

mRNA ageing shapes the Cap2 methylome in mammalian mRNA

<https://doi.org/10.1038/s41586-022-05668-z>

Vladimir Despic¹ & Samie R. Jaffrey^{1✉}

Received: 10 February 2022

Accepted: 16 December 2022

Published online: 1 February 2023

 Check for updates

The mRNA cap structure is a major site of dynamic mRNA methylation. mRNA caps exist in either the Cap1 or Cap2 form, depending on the presence of 2'-*O*-methylation on the first transcribed nucleotide or both the first and second transcribed nucleotides, respectively^{1,2}. However, the identity of Cap2-containing mRNAs and the function of Cap2 are unclear. Here we describe CLAM-Cap-seq, a method for transcriptome-wide mapping and quantification of Cap2. We find that unlike other epitranscriptomic modifications, Cap2 can occur on all mRNAs. Cap2 is formed through a slow continuous conversion of mRNAs from Cap1 to Cap2 as mRNAs age in the cytosol. As a result, Cap2 is enriched on long-lived mRNAs. Large increases in the abundance of Cap1 leads to activation of RIG-I, especially in conditions in which expression of RIG-I is increased. The methylation of Cap1 to Cap2 markedly reduces the ability of RNAs to bind to and activate RIG-I. The slow methylation rate of Cap2 allows Cap2 to accumulate on host mRNAs, yet ensures that low levels of Cap2 occur on newly expressed viral RNAs. Overall, these results reveal an immunostimulatory role for Cap1, and that Cap2 functions to reduce activation of the innate immune response.

The mRNA cap structure is a critical component of eukaryotic mRNAs but varies between transcripts based on its methylation state. Each cap structure comprises a *N*⁷-methylguanosine (m⁷G) cap at the 5' end of mRNA, adjacent to the first transcribed nucleotide^{3–5}. During mRNA biogenesis, the first nucleotide of all mRNAs becomes 2'-*O*-methylated (N_m) by cap methyltransferase 1 (CMTR1) to form a Cap1-modified mRNA terminus (m⁷G-ppp-N_m)^{6,7}. Upon mRNA export into the cytosol, a subset of Cap1 mRNAs is subjected to 2'-*O*-methylation of the ribose on the second nucleotide by cap methyltransferase 2 (CMTR2) to form Cap2-modified mRNA 5' ends (m⁷G-ppp-N_m-N_m)^{7–9} (Fig. 1a). However, the basis for this selectivity in mRNA cap methylation and the function of Cap2 are unclear.

Cap2 appears to have important regulatory roles in higher organisms, as deletion of the *Cmtr2* gene causes preweaning lethality in mice¹⁰ and affects growth and proliferation of *KRAS*-driven lung cancer cells¹¹. As Cap2 methylation is an abundant non-constitutive mRNA modification², these data suggest that Cap2 exerts important cellular functions via controlling a subset of mRNAs.

The identity of Cap2-modified mRNAs and the rules that govern the specificity of Cap2 methylation have remained unknown since the original discovery of Cap2 nearly 50 years ago. This is largely due to the lack of methods for transcriptome-wide mapping of Cap2 methylation.

Using a set of new methods to quantify and map Cap2, here we reveal unique features and function of this epitranscriptomic modification. Our quantitative transcriptome-wide map of Cap2 reveals an unexpected topology of the Cap2 methylome in mammalian transcriptomes, with essentially all mRNAs susceptible to Cap2 methylation. Cap2 methylation specificity is largely guided by mRNA age. We found

that this age-directed mechanism for Cap2 methylation provides a novel strategy to distinguish viral RNA from self RNA and to control the activation of the innate immune response.

Cap2 levels vary between cell types

As Cap2 is a non-constitutive mRNA modification^{1,2}, it has a potential for dynamic regulation in different cellular contexts and conditions. However, current methods for measuring Cap2 dynamics are cumbersome, as they rely on metabolic mRNA radiolabelling and chromatographic separation of Cap1 and Cap2 from the nuclease digests of mRNA¹. These methods utilize RNase T2 to selectively release the cap structures from mRNA. RNase T2 liberates the cap structures from mRNA as it cleaves all phosphodiester bonds in RNA except those after N_m (ref. 12) (Fig. 1b). Thus, RNase T2 digestion of Cap1 mRNA releases m⁷G-ppp-N_m-N, whereas the digestion of Cap2 mRNA releases m⁷G-ppp-N_m-N_m-N (Fig. 1b). We refer to these m⁷G-linked two-nucleotide-long or three-nucleotide-long fragments as 'cap tags', as they directly indicate the Cap1 or Cap2 status of an mRNA, respectively.

To quantify the levels and dynamics of Cap2 in any cellular sample, we developed CapTag-seq. In CapTag-seq, mRNAs are first subjected to enzymatic m⁷GDP removal (decapping), and the resulting 5'-monophosphorylated mRNA ends are ligated to a 5' adapter composed of N_m, thus rendering it resistant to RNase T2 (Fig. 1c). The resulting 5' adapter-linked mRNAs are fully digested with RNase T2, leaving the cap tags attached to the 5' adapter. A conversion of the 5' adapter-linked cap tags into cDNA libraries allows quantification of the cap tags by next-generation sequencing (Fig. 1c; see Methods).

¹Department of Pharmacology, Weill Cornell Medicine, Cornell University, New York, NY, USA. ✉e-mail: srj2003@med.cornell.edu

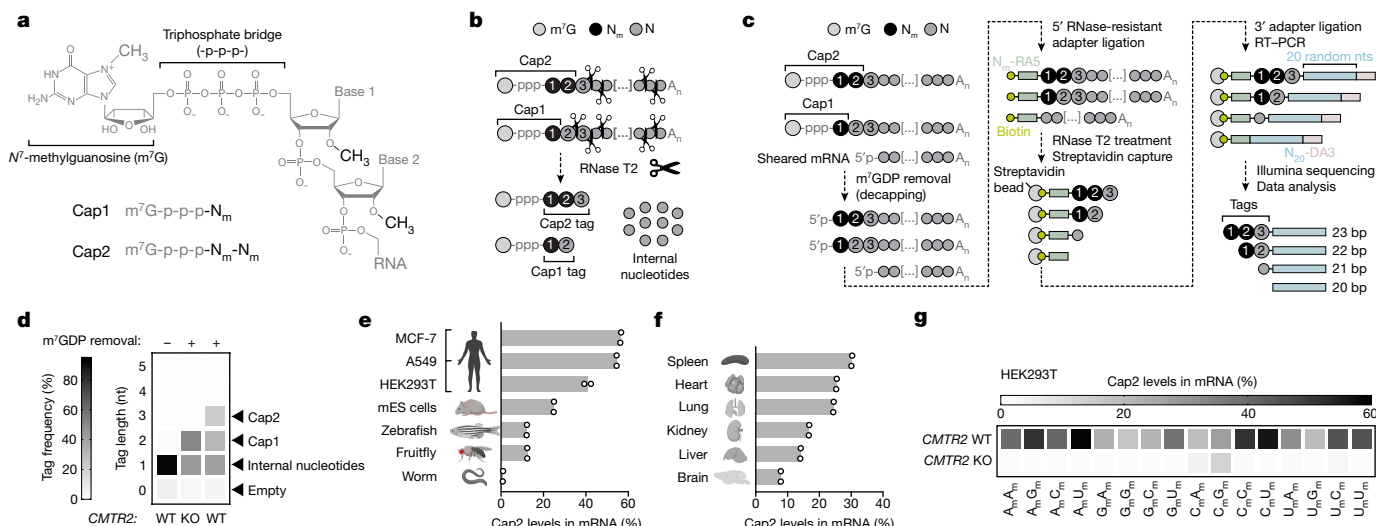


Fig. 1 | Cap2 is a variable mRNA modification found in diverse sequence contexts. **a**, Structure of Cap1-modified and Cap2-modified 5' mRNA termini. **b**, RNase T2 (scissors) cleaves RNA only after non-2'-*O*-methylated nucleotides. RNase T2 selectively releases Cap1 and Cap2 from mRNA in the form of two-nucleotide-long and three-nucleotide-long cap tags. **c**, CapTag-seq. mRNAs are decapped and a biotinylated 5' end adapter composed of 2'-*O*-methylated nucleotides (N_m -RA5) is ligated to the first mRNA nucleotide. The ligated mRNAs are digested with RNase T2, resulting in the formation of the 5' adapter-linked cap tags. The released 5' adapter-linked cap tags are captured on streptavidin beads and ligated to the 3' end DNA adapter containing 20 randomized nucleotides (N_{20} -DA3). After RT-PCR and next-generation sequencing, read length identifies Cap1 (22 base pair (bp)) and Cap2 (23 bp)

cDNA origins. Reads that are 21 bp long reflect 5' adapter ligations to uncapped, sheared mRNA. Reads that are 20 bp long represent adapter dimers. **d**, CapTag-seq specifically detects Cap1 and Cap2 mRNA species in bulk mRNA. CapTag-seq was conducted on mRNA extracted from *CMTR2* WT and KO HEK293T cells with (+) or without (-) the m^7 GDP removal step. A heatmap of the length-based tag frequency is shown for each sample. Average of $n = 2$ biological replicates. **e**, Cap2 levels in bulk mRNA from different organisms and mammalian cell types. Average of $n = 2$ biological replicates. **f**, Cap2 levels in bulk mRNA from different mouse tissues. Average of $n = 2$ biological replicates. **g**, A heatmap showing Cap2 levels on each of the 16 possible m^7 G-proximal dinucleotides in *CMTR2* WT and KO HEK293T cell transcriptomes. Average of $n = 2$ biological replicates.

In addition, the identity of nucleotides in the cap tags can reveal sequence contexts associated with Cap2 methylation.

We first assessed whether CapTag-seq specifically detects Cap1 and Cap2 mRNAs. CapTag-seq libraries prepared from HEK293T cells comprised almost exclusively three-nucleotide-long, two-nucleotide-long and one-nucleotide-long cap tags (Fig. 1d). The lack of longer cap tags demonstrates the specificity and efficiency of the RNase T2 cleavage step. In control mRNA samples that were not decapped, only the one-nucleotide tags remained (Fig. 1d), indicating that these tags probably reflect contamination from sheared mRNA with uncapped 5' ends that can be subjected to 5' adaptor ligation. They are unlikely to derive from Cap0 mRNAs (that is, no 2'-*O*-methylation at the first and second nucleotides), as Cap0 mRNAs are not detected in orthogonal biochemical assays^{1,2}.

To further confirm that the three-nucleotide-long cap tags derive from Cap2 mRNAs, we performed CapTag-seq on *CMTR2* knockout (KO) HEK293T cells (Extended Data Fig. 1a). Here we found a near-complete loss of the three-nucleotide-long cap tags (Fig. 1d), thus confirming their Cap2 origin and simultaneously validating Cap2 depletion in *CMTR2* KO HEK293T cells. CapTag-seq was quantitative, as mixtures of mRNA from *CMTR2*-overexpressing wild-type and *CMTR2* KO HEK293T cells at predefined ratios produced the expected Cap2 stoichiometries (Extended Data Fig. 1b).

We found that relative Cap2 abundance varied considerably between mammalian cell lines, with as low as approximately 25% in mouse embryonic stem (mES) cells and 40% in HEK293T cells, and as high as 54% in A549 cells and 56% in MCF-7 cells (Fig. 1e and Extended Data Fig. 1c). Of note, Cap2 was nearly absent in the *Caenorhabditis elegans* transcriptome (0.9%) and present at low levels in fruitfly and zebrafish (approximately 12%) mRNA (Fig. 1e). Cap2 was also highly variable in mouse tissues, ranging from approximately 8% in the mouse brain to 30% in the mouse spleen (Fig. 1f and Extended Data Fig. 1d). These

results show that Cap2 levels vary widely between organisms, tissues and cell types, suggesting regulation of Cap2 methylation in different cellular contexts.

Cap2 resides within diverse sequences

We next used CapTag-seq to determine whether Cap2 is associated with specific m^7 G-proximal nucleotides. To test whether certain sequences are selectively methylated by *CMTR2*, we examined Cap2 levels on all 16 possible m^7 G-proximal dinucleotides. In all cell types, we found that all 16 dinucleotides can be subjected to Cap2 methylation (Fig. 1g and Extended Data Fig. 1e). However, the extent of Cap2 methylation differed between individual dinucleotides (Fig. 1g and Extended Data Fig. 1e). Unexpectedly, we found that the sequences with the highest Cap2 methylation differed slightly among mouse tissues (Extended Data Fig. 1f). This was surprising as the sequence preferences of *CMTR2* would be expected to be identical regardless of the tissue. Overall, these data suggest that the pattern of Cap2 in the transcriptome may not be driven simply based on the sequence preferences of *CMTR2*, but, as will be shown below, is instead dictated in combination with other mechanisms.

Together, these data indicate that Cap2 is an abundant and promiscuous mRNA modification. The tissue-specific enrichment of Cap2 in different sequence contexts suggests that some mRNAs may have higher levels of Cap2 than other mRNAs, raising the possibility that Cap2 may have specific regulatory roles or functions.

CLAM-Cap-seq creates cDNA-mRNA chimeras

To understand pathways and processes that may be regulated by Cap2 methylation, we wanted to identify Cap2-modified mRNAs. N_m has been mapped at internal sites within ribosomal RNA based on

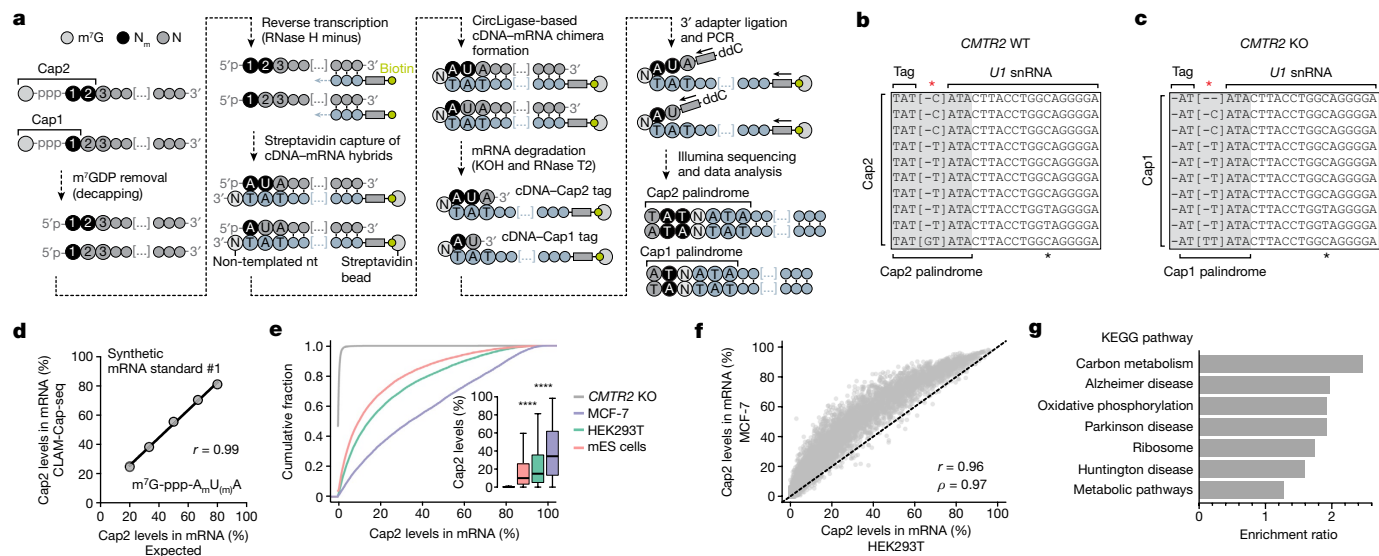


Fig. 2 | CLAM-Cap-seq identifies Cap2 methylation enrichment on specific mRNAs. **a**, In CLAM-Cap-seq, mRNAs are decapped and reverse transcribed using a 5'-biotinylated primer. Non-templated nucleotides added to the 3' cDNA ends by reverse transcriptase are indicated. cDNA-mRNA hybrids are captured on streptavidin beads. mRNA and cDNA ends are ligated by CircLigase to form a cDNA-mRNA chimera. The mRNA is degraded by KOH and RNase T2, leaving two-nucleotide-long and three-nucleotide-long cap tags attached to the cDNA. A DNA adapter is ligated to the cap tags, enabling the conversion of the cDNA-cap tag chimera into sequencing libraries. Sequencing reads begin with a palindrome, comprising the reverse complement of the cap tag followed by the transcript sequence. **b, c**, *U1* snRNA CLAM-Cap-seq PCR sequences obtained from *CMTR2* WT (**b**) and KO (**c**) HEK293T cells. The red asterisk indicates non-templated nucleotides, and the black asterisk denotes a mutation in the *U1* snRNA. **d**, CLAM-Cap-seq-predicted Cap2 stoichiometry plotted against the

expected Cap2 stoichiometry of a luciferase mRNA standard starting with A_mU_mA . Pearson's r is shown. $n = 2$ technical replicates. **e**, Cumulative distribution plot of Cap2 stoichiometry in TSN isoforms from mES cells ($n = 12,773$), *CMTR2* WT ($n = 12,776$) and KO ($n = 16,102$) HEK293T cells, and MCF-7 cells ($n = 10,630$). Each box shows the first quartile, median and third quartile, and whiskers represent 1.5 \times interquartile ranges. Two-sided Wilcoxon signed-rank test, **** $P < 0.00001$. **f**, Cap2 stoichiometry is correlated between human cell lines. Cap2 stoichiometry for a shared set of TSN isoforms in HEK293T and MCF-7 cells was plotted. Pearson's (r) and Spearman's (ρ) are shown. Average of $n = 4$ (HEK293T) and $n = 2$ (MCF-7) biological replicates. $n = 7,187$ TSN isoforms. **g**, Gene enrichment analysis performed for highly Cap2-modified mRNAs from HEK293T cells. The bar graph shows the enrichment ratio of the KEGG pathways associated with highly Cap2-modified mRNAs relative to all mRNAs with measurable Cap2 stoichiometry. FDR < 5%.

the ability of N_m to stochastically stall certain reverse transcriptases and induce termination of cDNA synthesis¹³. However, N_m -induced terminations are often inconsistent and influenced by sequence contexts surrounding modified nucleotides¹⁴. We therefore sought to develop a novel method for detection and quantification of Cap2 in individual mRNAs.

Although RNase T2 selectively releases Cap1 and Cap2 cap tags from mRNA, it destroys the sequence information of the mRNA from which the cap tag originated (Fig. 1b). To overcome this, we developed CLAM-Cap-seq (CircLigase-assisted mapping of caps by sequencing), a method which entails generating cDNA that is physically attached to the cap tag of its template mRNA. By creating a chimera between the cDNA and the cap tag, each cDNA sequence contains a record of the cap status of the original mRNA.

In CLAM-Cap-seq, decapped mRNA is first reverse transcribed to generate a cDNA-mRNA hybrid. Next, the cDNA 3' end is ligated to the first 5' mRNA nucleotide to create a cDNA-mRNA chimera (Fig. 2a). Subsequently, RNase T2 removes the entire mRNA except the cap tag, which remains covalently attached to the cDNA. Finally, a DNA adapter is ligated to the cap tag to enable the conversion of cDNA-cap tags into a sequencing library (Fig. 2a; see Methods). The resulting sequencing reads begin with 'palindromes' that reflect the cap tag of the mRNA followed by the first cDNA nucleotides, which are the reverse complement of the cap tag (Fig. 2a). Overall, CLAM-Cap-seq physically couples a remnant of the mRNA in the form of the cap tag to the cDNA, thus revealing the Cap1 or Cap2 status for each mRNA in the transcriptome.

Although the ligation of cDNA to mRNA has not been described, we considered the possibility that this could be achieved with CircLigase,

which has previously been used to ligate annealed DNA strands¹⁵. We tested CircLigase-assisted formation of chimeric cDNA-cap tags on *U1* small nuclear RNA (snRNA), which is nearly 100% Cap2-modified in various cell types¹⁶. First, we confirmed successful generation of cDNA-cap tags by performing PCR across the cDNA-cap tag ligation junction (Extended Data Fig. 2a). Sequencing of individual PCR products revealed the expected Cap2 palindromes at the beginning of the reads derived from *U1* snRNA extracted from wild-type cells (Fig. 2b). By contrast, only Cap1 palindromes were observed in *CMTR2* KO cells, demonstrating the specificity of this method (Fig. 2c). Overall, these data show that CLAM-Cap-seq selectively generates cDNA-cap tag chimeras that report the RNA identity as well as its Cap1 or Cap2 status.

In the centre of the palindromes, we routinely observed 0, 1 or 2 nucleotides (Fig. 2b,c), probably added to the cDNA ends by the terminal nucleotidyl-transferase activity of the reverse transcriptase¹⁷. Despite the presence of non-templated nucleotides, the complementary sequences within the palindrome are easily identified and demarcate the exact cap tag region (Fig. 2b,c). Thus, the cap tag can be extracted from the palindrome to determine whether the original mRNA was Cap1 or Cap2.

We next tested the ability of CLAM-Cap-seq to accurately predict Cap2 stoichiometry in mRNA. To test this, we mixed Cap1-modified and Cap2-modified luciferase mRNAs in specific molar ratios and spiked them into cellular poly(A)⁺ RNA. We found that CLAM-Cap-seq accurately predicted Cap2 stoichiometry present in the mixture of synthetic luciferase mRNA standards (Fig. 2d and Extended Data Fig. 2b). Collectively, these data demonstrate that CLAM-Cap-seq can accurately measure Cap2 stoichiometry in mRNA.

CLAM-Cap-seq reveals the Cap2 methylome

We next identified Cap2-modified mRNAs in mES, HEK293T and MCF-7 cells. Gene transcription can initiate at different locations within a gene promoter, giving rise to transcript isoforms that differ in the start nucleotides (transcript-start nucleotide (TSN) isoforms)^{18,19}. To identify Cap2 stoichiometry for each TSN isoform, rather than genes, we first identified TSN isoforms in the transcriptome by transcription start site-sequencing (TSS-seq)^{20,21} (Extended Data Fig. 2c–e).

Next, we performed CLAM-Cap-seq on mRNA extracted from mES, HEK293T and MCF-7 cells and calculated Cap2 stoichiometry for each TSN isoform identified by TSS-seq. CLAM-Cap-seq produced sufficient read coverage for prediction of Cap2 stoichiometry in 10,630–12,776 TSN isoforms, spread across 3,745–4,267 genes in the examined cell lines. CLAM-Cap-seq datasets showed low variability of Cap2 measurements across replicates (Extended Data Fig. 2f,g).

The specificity of CLAM-Cap-seq is supported by the lack of Cap2 reads in CLAM-Cap-seq datasets from *CMTR2* KO HEK293T cells (Fig. 2e). Furthermore, the Cap2 stoichiometry was low (1–11%) in nuclear poly(A)⁺ RNAs compared with cytosolic mRNAs (2–84%), consistent with the formation of Cap2 in the cytosol⁸ (Extended Data Fig. 2h). The overall levels of Cap2 methylation and the Cap2 dinucleotide preferences identified by CLAM-Cap-seq were highly similar to those identified by CapTag-seq (Extended Data Fig. 2i,j).

To biochemically validate the levels of Cap2 measured using CLAM-Cap-seq, we developed an approach to measure Cap2 levels on specific cellular mRNAs. In this method, named CapOligo-PAGE, a self-splinting DNA oligo is selectively ligated to the radiolabelled 5' end of an mRNA of interest. Following RNase T2-mediated removal of the mRNA, the self-splinting oligo remains attached to the radiolabelled cap tags. A subsequent PAGE electrophoresis resolves DNA oligo-attached cap tags to establish the Cap1 or Cap2 status of the mRNA of interest (Extended Data Fig. 3a–c; see Methods). Using CapOligo-PAGE, we confirmed that Cap2 was present in *OAT* and *RPS12* mRNAs, with markedly higher Cap2 stoichiometry in *RPS12* mRNA (Extended Data Fig. 3d,e), as predicted by CLAM-Cap-seq (Extended Data Fig. 2h). Thus, the Cap2 stoichiometry measurements obtained with CLAM-Cap-seq are consistent with our orthogonal assay for determination of Cap1 and Cap2 status in individual mRNAs.

We next assessed the overall stoichiometry of Cap2 throughout transcriptomes. In all three cell lines, we detected transcripts with Cap2 stoichiometries that ranged from 0% to approximately 100% (Fig. 2e). Of note, the overall level of Cap2 varied substantially between cell types, with MCF-7 cell transcripts exhibiting threefold higher Cap2 methylation than mRNAs of the mES cell transcriptome (Fig. 2e). To understand the basis for the differences in Cap2 levels in different cell types, we compared Cap2 stoichiometry at the shared set of TSN isoforms between HEK293T and MCF-7 cells (Fig. 2f). We found that TSN isoforms with low Cap2 stoichiometry in HEK293T cells also exhibited low levels of Cap2 in MCF-7 cells. Similarly, high Cap2 stoichiometry TSN isoforms from HEK293T cells were also highly Cap2-modified in MCF-7 cells (Fig. 2f). The major difference was that the Cap2 stoichiometry of each TSN isoform was proportionally higher in MCF-7 cells (Fig. 2f). Thus, rather than distinct Cap2 epitranscriptomes in each cell type, we found that Cap2 levels at all TSN isoforms are largely correlated in these two human cell lines.

We next used gene enrichment analysis to determine whether Cap2-modified mRNAs are linked to specific cellular pathways. These analyses showed that transcripts with high Cap2 stoichiometry were significantly enriched in gene sets associated with general metabolic pathways and other housekeeping functions in all three cell lines (Fig. 2g and Extended Data Fig. 4). Thus, Cap2 enrichment on specific mRNA cohorts appears to be a conserved phenomenon in mammalian cells.

Overall, these data suggest that Cap2 could influence the function of specific groups of genes and thus regulate cell functions.

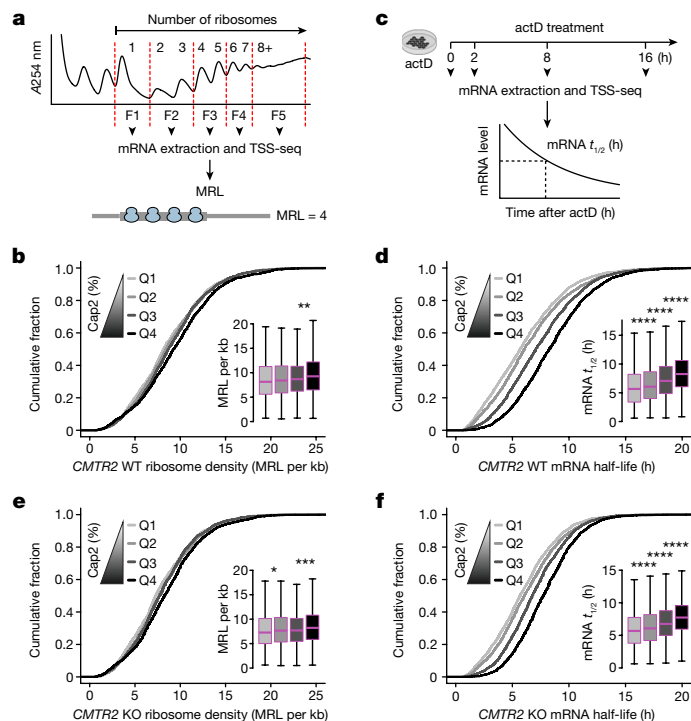


Fig. 3 | Cap2 methylation is highly enriched on long-lived mRNAs. **a**, Polysome-TSS-seq. Translated mRNAs were isolated from different polysome fractions (F1–F5) and subjected to TSS-seq. The average number of ribosomes bound to each TSN isoform was calculated, termed mean ribosome load (MRL). **b**, Cumulative distribution plot of the ribosome density (MRL per kb) for TSN isoforms binned into equally sized quartiles of increasing Cap2 stoichiometry (Q1–Q4) in *CMTR2* WT HEK293T cells. $Q_n = 1,782$. **c**, ActD-TSS-seq. mRNA was extracted from cells treated with actinomycin D (actD) for 0, 2, 8 and 16 h. Purified mRNA was subjected to TSS-seq for half-life measurements of each TSN isoform. **d**, Cumulative distribution plot of the half-lives for TSN isoforms binned into equally sized quartiles of increasing Cap2 stoichiometry in *CMTR2* WT HEK293T cells. $Q_n = 2,115$. **e**, Analysis of mRNA translation, as in **b**, using *CMTR2* KO HEK293T cells. $Q_n = 1,938$. **f**, Analysis of mRNA half-lives, as in **d**, using *CMTR2* KO HEK293T cells. $Q_n = 2,150$. In **b**, **d**–**f**, average of $n = 2$ biological replicates. Two-sided Wilcoxon signed-rank test, * $P < 0.01$, ** $P < 0.001$, *** $P < 0.0001$ and **** $P < 0.00001$. Each box shows the first quartile, median and third quartile, and whiskers represent 1.5 \times interquartile ranges.

Cap2 is enriched on long-lived mRNAs

As Cap2 is added in the cytosol, we asked whether Cap2 is associated with cytoplasmic mRNA processing events, such as translation or mRNA stability. Current approaches for measuring mRNA translation and stability generate sequencing reads that cannot be linked to specific TSN isoforms. To measure translation and stability of every Cap2-modified TSN isoform, we first developed methods for TSN-specific translation and stability analyses.

To specifically assess translation of TSN isoforms, we combined polysome profiling with TSS-seq (polysome-TSS-seq). The abundance of each TSN isoform along the sucrose gradient can be used to calculate the average number of actively translating ribosomes bound to each TSN isoform (mean ribosome load (MRL))²² (Fig. 3a).

We quantified TSN isoforms in five polysome fractions of HEK293T cells using TSS-seq. As a control, we showed that mRNAs known to be highly and lowly translated were found in the expected positions within the sucrose gradient (Extended Data Fig. 5a). Next, we calculated the MRL for each TSN isoform in two replicates (Extended Data Fig. 5b,c).

To determine whether Cap2 is associated with translation efficiency, we stratified all TSN isoforms into equally sized quartiles based on

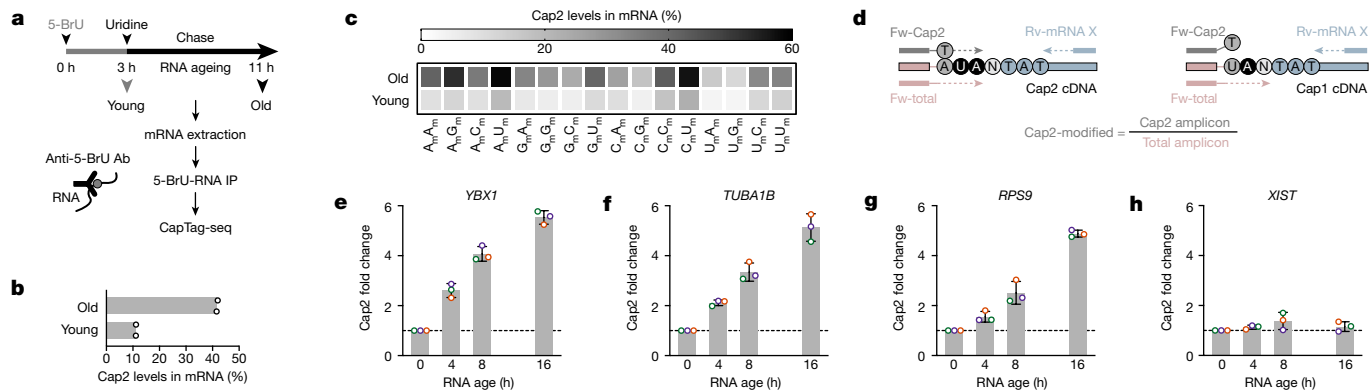


Fig. 4 | Cap2 is an epitranscriptomic mark of mRNA age. **a**, Transcripts were metabolically labelled using a 3 h 5-BrU pulse and chased in uridine-containing media for 0 h and 8 h to allow isolation of young and old RNA, respectively. BrU-containing mRNAs of different age were immunopurified using an anti-5-BrU antibody (Ab) and Cap2 levels were measured using CapTag-seq. IP, immunoprecipitation. **b**, Cap2 levels in young and old mRNA. Average of $n = 2$ biological replicates. **c**, Cap2 levels on each of the 16 possible m⁷G-proximal dinucleotides in young and old transcriptomes. Average of $n = 2$ biological replicates. **d**, CLAM-Cap-qPCR, a transcript-specific Cap2 methylation measurement method. Following generation of cDNA-cap tags attached to the

3' end DNA adapter, Cap2 levels were measured by qPCR using a gene-specific reverse (Rv-mRNA X) and Cap2 tag-specific forward (Fw-Cap2) primer. Total (Cap1 and Cap2) cDNA-cap tags were amplified in parallel by qPCR with a primer pair comprising Rv-mRNA X and a primer that anneals to the DNA adapter (Fw-total). **e-h**, Fold change in Cap2 methylation as a function of mRNA age for *YBX1* (**e**), *TUBA1B* (**f**) and *RPS9* (**g**) mRNAs and for the nuclear non-coding RNA *XIST* (**h**). The level of Cap2 methylation at 0 h of the uridine chase was set to 1 ($t_0 = 1$, dashed line). $n = 3$ biological replicates. Data shown as mean \pm s.d. Coloured data points denote measurements from the same experiment.

their Cap2 stoichiometry. When we examined the distribution of the ribosome density (MRL per kilobase of the open reading frame (MRL per kb)) for each Cap2 stoichiometry quartile, we observed slightly higher ribosome density in highly Cap2-modified transcripts relative to the mRNAs with the lowest levels of Cap2 (median Δ ribosome density ($Q_4 - Q_1$) = -1 MRL per kb; Fig. 3b and Extended Data Fig. 5d-f). Overall, these data suggest that Cap2 shows a slight enrichment on mRNAs with increased translation efficiency.

We next asked whether Cap2 is associated with mRNA stability. To calculate the half-lives of each TSN isoform, we quantified TSN isoform abundance by TSS-seq at different time points after transcriptional shut off with actinomycin D (actD-TSS-seq) (Fig. 3c and Extended Data Fig. 5g). TSN isoform half-lives were quantified in two replicates (Extended Data Fig. 5h,i).

When we compared the distribution of half-lives between mRNAs of different Cap2 stoichiometry, we observed a large difference in the half-lives and abundance of mRNAs with the highest and lowest Cap2 levels (median $\Delta t_{1/2}$ ($Q_4 - Q_1$) = 2.5 h; Fig. 3d and Extended Data Fig. 5j-m).

A similarly strong positive relationship between Cap2 methylation and mRNA stability was also observed in mES and MCF-7 cells (Extended Data Fig. 5n-t), suggesting a conserved relationship between Cap2 and mRNA stability across different cell types and species.

Overall, these data show that Cap2 methylation is enriched on mRNAs with high translation and stability, with a much more prominent association between Cap2 methylation and mRNA stability.

Cap2 does not confer high mRNA stability

We next wanted to determine whether Cap2 directly influences mRNA translation. We therefore measured mRNA translation in *CMTR2* KO HEK293T cells by polysome-TSS-seq (Extended Data Fig. 6a). We noticed that *CMTR2* KO HEK293T cells exhibited slower growth and overall reduced translation based on the puromycin incorporation and polysome profiling (Extended Data Fig. 6b-d). Despite the global reduction in translation, we asked whether *CMTR2* depletion causes a selective decrease in the translation efficiency of mRNAs with the highest Cap2 stoichiometry. Consistent with the overall impairment of growth and translation (Extended Data Fig. 6b-d), all transcripts exhibited a reduction in translation after *CMTR2*

depletion independent of their Cap2 status (median Δ ribosome density = -0.8 MRL per kb; Extended Data Fig. 6e). However, the slight difference in translation efficiency between the high and low Cap2-modified mRNAs observed in *CMTR2* WT cells remained largely unchanged in the *CMTR2* KO cells (Fig. 3e). These data suggest that *CMTR2* depletion does not regulate the translation capacities of transcripts based on their Cap2 stoichiometry.

We then wanted to determine whether Cap2 confers long half-lives to mRNA. To test this, we measured changes in mRNA half-lives ($\Delta t_{1/2}$) in *CMTR2* KO HEK293T cells with actD-TSS-seq (Extended Data Fig. 6f,g). Here we observed that highly Cap2-modified transcripts exhibited a subtle decrease in half-lives in *CMTR2* KO cells (Extended Data Fig. 6h). However, the large difference in mRNA stability observed between the high and low Cap2-modified mRNAs in *CMTR2* WT cells remained in the *CMTR2* KO cells (Fig. 3f and Extended Data Fig. 6i). These data suggest that the mild stabilizing effect of Cap2 on mRNA does not explain the unusual longevity of Cap2-marked mRNAs.

Cap2 levels increase with mRNA age

As Cap2 methylation does not confer long half-life to an mRNA, we considered the possibility that long half-life might cause high Cap2 methylation. Long-lived mRNAs persist longer in cells and therefore may have more time to acquire high Cap2 methylation during their cytoplasmic lifetime.

To test this, we used BruChase to capture mRNAs of increasing age²³. In this approach, cells were pulsed with 5-bromouridine (5-BrU) for 3 h to label newly synthesized RNA. Then, the cells were chased in uridine-rich media to allow for ageing of BrU-labelled transcripts. BrU-containing RNA was immunopurified from mRNA with an anti-5-BrU antibody at 0 and 8 h (Fig. 4a).

We next performed CapTag-seq on young (0 h) and old (8 h) mRNA. We found that young mRNA was 11% Cap2-modified, whereas old mRNA exhibited approximately 42% Cap2 methylation (Fig. 4b), suggesting that Cap2 methylation increases throughout the transcriptome as transcripts age in the cytosol. All 16 m⁷G-proximal dinucleotides exhibited an increase in Cap2 methylation over time (Fig. 4c), indicating the generality of age-dependent increases in Cap2 levels across all mRNA sequence contexts.

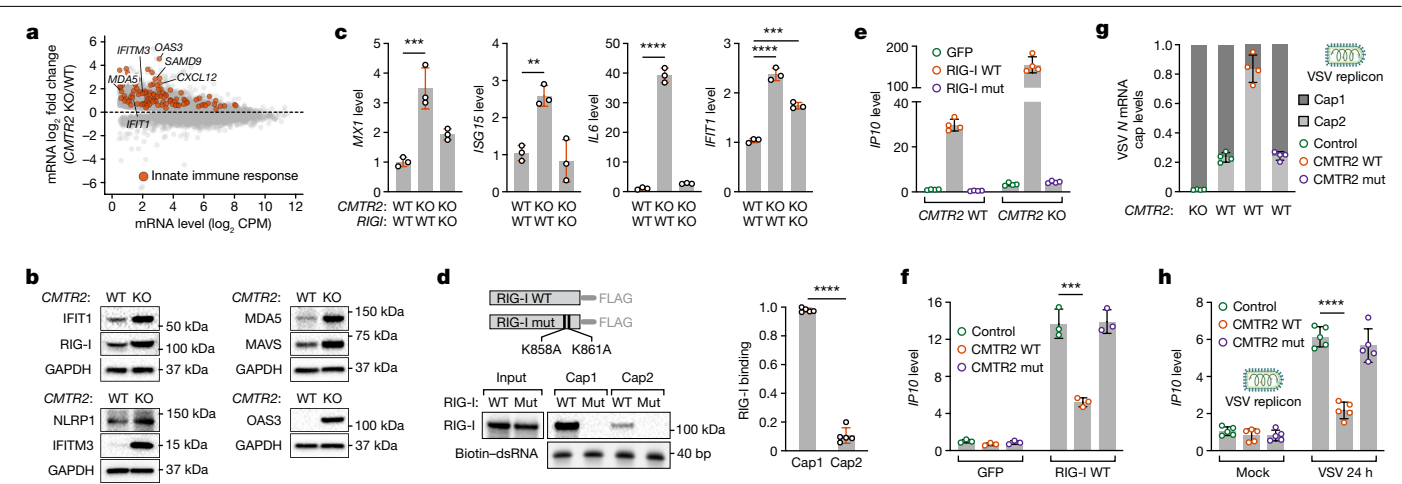


Fig. 5 | Cap2 methylation suppresses RIG-I activation and virus-induced innate immune response. **a**, RNA-seq-based measurements of gene expression changes upon *CMTR2* depletion. $n = 4$ biological replicates. $n = 9,507$ genes. $P < 0.05$. CPM, counts per million. **b**, Expression of antiviral proteins in *CMTR2* WT and KO HEK293T cells. GAPDH was the loading control. **c**, Expression of innate immune response genes in *CMTR2* KO A549 cells requires RIG-I. Transcript levels were quantified by RT-qPCR. $n = 3$ biological replicates. **d**, HEK293T lysates expressing RIG-I or the RIG-I-K858A/K861A cap-binding mutant were incubated with immobilized Cap1 and Cap2 double-stranded RNAs (dsRNAs). The pulled-down RIG-I protein was detected by western blot. The bars show quantification from $n = 5$ pull-downs. Two-tailed unpaired Student's *t*-test with Welch's correction, **** $P < 0.00001$. **e**, *CMTR2* WT and KO HEK293T cells were transfected with GFP, RIG-I-WT or RIG-I mutant (mut) from **d**. *IP10* expression was measured 48 h post-transfection by RT-qPCR.

$n = 4$ biological replicates. **f**, *CMTR2* WT HEK293T cells were transfected with control, *CMTR2* (*CMTR2* WT) and *CMTR2* W85A (*CMTR2* mut) constructs for 24 h. Each sample was further transfected with GFP or RIG-I-WT. *IP10* expression was measured by RT-qPCR 48 h after the second transfection. $n = 3$ biological replicates. **g**, HEK293T *CMTR2* KO cells expressing control and *CMTR2* WT cells expressing control, *CMTR2* WT or *CMTR2* mut for 24 h were either mock-infected or VSV-infected. Cap2 methylation on viral *N* mRNA was assessed with CLAM-Cap-qPCR 24 h post-infection. The stacked bar plot shows the fraction of Cap1-modified and Cap2-modified VSV *N* mRNA. The Cap1 fraction was calculated as $1 - \text{Cap2 fraction}$. $n = 4$ biological replicates. **h**, *CMTR2* WT HEK293T cells were treated as in **g**. *IP10* expression was assessed 24 h post-infection by RT-qPCR. $n = 5$ biological replicates. In **c-h**, data shown as mean \pm s.d. In **c, f, h**, one-way ANOVA with Bonferroni multiple comparison test was used; ** $P < 0.01$, *** $P < 0.001$ and **** $P < 0.0001$.

To examine the levels of Cap2 on individual transcripts as they age, we developed CLAM-Cap-quantitative PCR (CLAM-Cap-qPCR), an amplification-based method for measurement of Cap2 levels on low-abundance mRNAs. CLAM-Cap-qPCR involves preparation of cDNA-cap tag chimeras with a DNA adapter ligated to the 3' end of the cap tag according to the CLAM-Cap-seq protocol. Next, the levels of Cap2 are measured in an mRNA of interest by qPCR using a transcript-specific primer and a primer that hybridizes to the DNA adapter and the first nucleotide of the three-nucleotide-long cap tag that is unique to the Cap2 form of the mRNA (Fig. 4d). To determine the total abundance of the mRNA, a parallel qPCR is performed using a primer that hybridizes only to the 3' adapter, but not to any portion of the cap tag (Fig. 4d). We confirmed the accuracy of this method using luciferase mRNA standards with known Cap2 stoichiometries (Extended Data Fig. 6j). Overall, this approach allowed us to calculate the Cap2 levels in an mRNA of interest.

We performed CLAM-Cap-qPCR on three long-lived mRNAs: *YBX1*, *TUBA1B* and *RPS9*. In each case, we observed that the Cap2 levels continuously increased as each mRNA aged (Fig. 4e-g). As a control, we examined the nuclear RNA *XIST*²⁴, which is not expected to encounter *CMTR2*. Indeed, we found largely unchanged Cap2 levels on *XIST* for the entire duration of the uridine chase (Fig. 4h). Overall, these data suggest that Cap2 methylation represents a dynamic mRNA modification that continuously increases throughout the mRNA lifetime.

CMTR2 depletion induces antiviral genes

We next wanted to understand the functional implication of Cap2 methylation. To test this, we analysed gene expression changes in *CMTR2* KO HEK293T cells using RNA-seq. We found that downregulated genes were enriched in translation and RNA processing pathways, whereas markedly upregulated genes were related to the innate immune response

and inflammatory pathways (Fig. 5a and Extended Data Fig. 7a). Among upregulated genes, we noticed numerous interferon-stimulated genes (ISGs), a gene group that is transcriptionally induced after cell infection with viruses, as well as other pathogens²⁵. We confirmed the induction of ISGs at the protein and RNA level (Fig. 5b and Extended Data Fig. 7b-g). Similar effects were seen in *CMTR2* KO A549 cells (Fig. 5c and Extended Data Fig. 7h-j).

The magnitude of the ISG induction was comparable to that of the induction seen after wild-type HEK293T cells were treated with interferon (750 U ml^{-1}) for 6 h (Extended Data Fig. 7k). The induction of ISGs was sufficient to sensitize cells to diverse inflammatory signals (Extended Data Fig. 7l-p).

Overall, these data suggest that the lack of, and thus the increase in Cap1 mRNA levels, leads to the activation of pathways that are normally triggered in response to viral RNA. As *CMTR2* KO cells are not exposed to an exogenous virus, the induction of the innate immune response may be mediated by the change in cap methylation of endogenous RNAs.

Cap2 suppresses RIG-I activation

The effects of Cap2 loss may be mediated by a specific RNA-binding protein whose RNA-binding activity is affected by Cap2 methylation. RIG-I is a well-established sensor of foreign RNA that becomes activated upon binding to the triphosphate bridge of m⁷G-capped RNAs^{26,27}. RIG-I is markedly activated by Cap0 RNA due to the low nanomolar affinity (2 nM) of RIG-I for Cap0 RNA^{27,28}.

Methyl modifications of RNA impair the binding of capped RNA to RIG-I. A 2'-*O*-methylation at the first nucleotide in Cap1 RNA leads to markedly reduced binding affinity (425 nM)²⁷. 2'-*O*-methylation at the second nucleotide also reduces RIG-I activation²⁶. In this study, the RNA only contained a single methyl modification at the second position, and

not a dual methylation as seen in Cap2. Therefore, it remains unclear whether dual methylation in Cap2 would further reduce the binding affinity of Cap1 RNA to RIG-I. Overall, these studies raise the possibility that some of the cellular effects of CMTR2 depletion may be mediated by RIG-I.

To address this, we first asked whether RIG-I activity contributes to the induction of ISGs in *CMTR2* KO cells. We generated *CMTR2*, *RIG-I* double KO A549 cells and monitored the expression of several ISGs by qPCR with reverse transcription (RT-qPCR) (Fig. 5c and Extended Data Fig. 7h,j). We found that RIG-I depletion markedly reduced the expression of ISGs in *CMTR2* KO cells, demonstrating RIG-I involvement in the induction of the innate immune response in these cells.

Next, we directly measured the effect of Cap2 on the RNA-binding capacity of RIG-I. We incubated equal amounts of streptavidin-immobilized Cap1 and Cap2 double-stranded RNAs with cell lysates expressing FLAG-tagged RIG-I. Consistent with previous studies²⁷, RIG-I readily bound to Cap0 RNA (Extended Data Fig. 8a). RIG-I also showed lower, but clear binding to Cap1 RNA. However, RIG-I exhibited markedly lower interaction with Cap2 RNA (Fig. 5d and Extended Data Fig. 8a,b). Neither Cap1 nor Cap2 RNA bound to RIG-I K858A/K861A, a RIG-I mutant that cannot interact with the triphosphate bridge of the cap structure²⁹ (Fig. 5d). These results confirm the cap-dependent nature of the observed interactions. Overall, these data demonstrate that the dual methylation in Cap2 RNAs further reduces the ability of Cap1 RNAs to bind RIG-I.

Although Cap1 is not thought to be an activating ligand for RIG-I, we wanted to know whether high levels of Cap1 could activate RIG-I in cells. To measure RIG-I activation, we overexpressed RIG-I in HEK293T cells, which led to the induction of *IP10* (Fig. 5e), a frequently used marker for RIG-I activation^{28,30}. However, overexpression of RIG-I in *CMTR2* KO cells, whose transcriptome is solely composed of Cap1 RNA, resulted in markedly increased levels of *IP10* (Fig. 5e). These data suggest that Cap1 RNA can activate RIG-I if it is present at high levels.

Of note, the observed activation of RIG-I depends on the RNA cap structure as overexpression of RIG-I K858A/K861A failed to induce *IP10* (Fig. 5e). Expression of RIG-I constructs remained similar in all tested samples (Extended Data Fig. 9a).

To determine whether the low level of activation of RIG-I in wild-type HEK293T cells was due to Cap1 mRNAs (approximately 60% of mRNAs; see Fig. 1e), we overexpressed CMTR2 to convert Cap1 to Cap2 mRNAs before RIG-I overexpression. Here we found that RIG-I expression led to markedly reduced *IP10* levels relative to control RIG-I-expressing cells (Fig. 5f). Conversely, RIG-I activation was not reduced in cells expressing CMTR2 W85A, a mutant that cannot methylate the cap structure⁷ (Fig. 5f). RIG-I expression remained similar across all tested samples (Extended Data Fig. 9b). Overall, these data suggest that the function of the dual methylation in Cap2 is to further reduce the binding of Cap1 to RIG-I, thus decreasing the immunostimulatory effects seen at high concentrations of Cap1.

We confirmed previous studies showing that transfected Cap1 RNA is unable to activate RIG-I probably due to the low level of transfected RNA and its weak binding affinity (Extended Data Fig. 9c). However, the transcriptome-wide increase in Cap1 levels due to CMTR2 depletion may be sufficient to achieve Cap1 concentrations needed to activate RIG-I.

Cap2 in viral RNA impairs host response

Although our data demonstrate that Cap2 suppresses the ability of endogenous Cap1 RNA to induce the innate immune response, the purpose of a slow, time-dependent methylation of Cap1 to Cap2 is unclear. We considered the possibility that slow Cap2 methylation allows the host cell to detect and respond to rapidly replicating viral Cap1 RNAs before the viral RNA acquires high levels of Cap2 over time.

To test this model, we used vesicular stomatitis virus (VSV), an RNA virus that triggers the expression of antiviral genes, in part through RIG-I activation³¹. Although VSV RNA acquires Cap1 by virally encoded enzymes, it utilizes host CMTR2 to achieve low Cap2 levels³². Therefore, we used CMTR2 overexpression to enhance Cap2 methylation efficiency in VSV RNA. In this way, we could determine whether the normally low Cap2 methylation efficiency enables efficient induction of the innate immune response by VSV.

To test this, we first infected control HEK293T cells with propagation-incompetent VSV³³ (Extended Data Fig. 9d). Consistent with previous studies^{34,35}, VSV infection resulted in the rapid accumulation of viral RNA to an amount comparable to the entire mRNA transcriptome of mock-infected cells (Extended Data Fig. 9e,f). Viral transcripts were predominantly Cap1-modified (Fig. 5g and Extended Data Fig. 9g,h). This increase in Cap1 levels was associated with the expected induction of *IP10* mRNA (Fig. 5h). Thus, the VSV-induced increase in Cap1 may contribute to the activation of the innate immune response.

We next attempted to increase Cap2 methylation efficiency of viral RNA by overexpressing CMTR2. Overexpression of CMTR2 markedly increased Cap2 stoichiometry on viral RNA, resulting in the concomitant reduction in Cap1 levels (Fig. 5g and Extended Data Fig. 9g,h). In CMTR2-overexpressing cells, VSV infection failed to induce *IP10* mRNA to the levels seen in control infected cells (Fig. 5h). This was due to cap-mediated methylation, as CMTR2 W85A failed to suppress *IP10* induction (Fig. 5g,h and Extended Data Fig. 9g,h). Overall, these data suggest that cells maintain low efficiency of Cap2 methylation to prevent viral RNA from acquiring Cap2 and evading host defence mechanisms.

As a control, we asked whether CMTR2 overexpression nonspecifically suppresses the innate immune response. To test this, we activated the innate immune response using noncapped RIG-I ligands, such as triphosphorylated double-stranded RNA and poly(I:C)³⁶. Although HEK293T cells express RIG-I protein at low levels (Fig. 5b), they have previously been shown to respond to double-stranded RNA ligands in a RIG-I-dependent manner³⁷. Consistent with this, we found that both stimuli readily induced *IP10* mRNA expression (Extended Data Fig. 9i). However, CMTR2 overexpression did not suppress their effects on *IP10* mRNA levels (Extended Data Fig. 9i). These data are consistent with a model in which increased expression of CMTR2 selectively suppresses antiviral responses induced by Cap1 RNAs.

Discussion

Cap2 was discovered nearly 50 years ago as one of the five major methyl modifications that decorate mRNA along with m⁷G, N⁶,2'-O-dimethyladenosine (m⁶A_m), Cap1 and N⁶-methyladenosine (m⁶A). Despite the high prevalence of Cap2 in the transcriptome, Cap2 is the last major unmapped nucleotide modification. Using CLAM-Cap-seq, we generated a transcriptome-wide map of Cap2, which revealed strong Cap2 enrichment on long-lived mRNAs, occurring as a result of mRNA age-guided Cap2 deposition in the transcriptome. Rather than controlling mRNA processing events such as translation or stability, a major function of Cap2 is to further suppress the ability of endogenous RNAs to activate the innate immune response. Mechanistically, the dual methylation in Cap2 acts to prevent Cap1 from binding to RIG-I, thus suppressing the autoimmune potential of endogenous Cap1 RNA. We also show that slow, time-dependent accumulation of Cap2 in mRNAs represents a cellular adaptation used to cloak host RNAs from activating RIG-I. Simultaneously, slow Cap2 methylation reduces the likelihood that rapidly replicating viral RNA acquires Cap2 and thus evades recognition by host cell defences.

The cell requires mechanisms to distinguish self from non-self RNA. Methylation of mRNA caps is one of the major mechanisms to mark self mRNAs. Cap0 RNAs that lack ribose methylation are high-affinity ligands and potent activators of RIG-I²⁷. Small amounts of Cap0 can

therefore activate the innate immune response. The presence of a single methyl group in Cap1 is sufficient to reduce the ability of RNA to activate RIG-I. Although the binding affinity of Cap0 to RIG-I is 2 nM, Cap1 still binds at an affinity of 425 nM (ref. 27). We found that large amounts of Cap1 RNA, achieved by depletion of CMTR2 or by viral infection, can provide sufficient levels of Cap1 to activate RIG-I. Thus, even though Cap1 only activates RIG-I weakly, the large amount of Cap1 in the cell, coupled with the induction of RIG-I expression in response to viral infection, can make Cap1 RNA an important agonist of RIG-I signalling.

Previous studies have shown that methylation at either the first or second position of RNA is sufficient to reduce activation of RIG-I²⁶. We showed that the dual methylation in Cap2 functions to reduce the ability of Cap1 RNA to bind to and activate RIG-I. Although our study identifies RIG-I as an ‘anti-reader’ of Cap2, *RIG-I*-KO did not completely reduce the induction of the ISGs after CMTR2 depletion. Other proteins, such as IFITs^{38,39} or other foreign RNA sensors such as MDA5 (ref. 40), may also be sensitive to Cap2 methylation and therefore regulate aspects of mRNA biology.

Of note, some viruses have acquired CMTR2 homologues, such as Mimivirus and African swine fever virus⁹. The CMTR homologue in vaccinia virus has been proposed to methylate the first, second and possibly other nucleotides⁴¹. Thus, viral CMTR homologues may have broader methylation functions than previously recognized, including Cap2 methylation, which may help to them to evade host responses.

The levels of Cap1 and Cap2 vary in different cells and tissues, which correlate in part with CMTR2 levels (Extended Data Fig. 1c,d). Although higher Cap1 levels may be deleterious for cells, we found that these cells typically exhibit lower RIG-I expression, which may reduce the autoinflammatory potential of high levels of Cap1 RNA (Extended Data Fig. 10a–c). Cells may adjust the levels of CMTR2 to influence their cellular responses to either host or viral RNAs, and aberrant levels of Cap1 may contribute to diseases that are linked to excessive activation of the innate immune response (Extended Data Fig. 10d).

In addition to suppressing the immunostimulatory effects of cellular RNAs, Cap2 may affect mRNAs in other ways. Our global analyses showed small but clear effects of Cap2 on mRNA translation or stability. Cap2 may thus shape gene expression, particularly for long-lived transcripts, to further enhance their stability in cells and to increase their protein expression. In addition, some RNAs may have a stronger dependence on Cap2 for preventing their interaction with RIG-I, and thus may have a more important role in suppressing the innate response when Cap2-modified. Other RNAs may contribute to the biology of Cap2, including snRNAs¹⁶. Finally, Cap2 has been linked to neuronal functions in *Drosophila*⁴¹, suggesting different roles of Cap2 in lower organisms.

We developed a suite of tools to profile and measure Cap2 in transcriptomes and on specific mRNAs of interest. Measurements of Cap1 and Cap2 cap tags using CapTag-seq reveals the overall Cap2 prevalence in bulk mRNA, whereas CLAM-Cap-seq involves creating cDNA–cap tag chimeras that directly link the Cap2 status of an mRNA to the cDNA that is generated from it. Along with newly developed biochemical and PCR-based methods, the Cap1 or Cap2 state of any mRNA of interest can be readily measured.

Our data show that Cap2 is fundamentally different from other epi-transcriptomic mRNA modifications. First, Cap2 methylation is a dynamic process, as it continuously accumulates throughout the lifetime of an mRNA in the cytoplasm. This contrasts other mRNA modifications such as m⁶A, which are largely ‘written’ in the nucleus and thus exhibit little or no potential for dynamics once the mRNA has left the nucleus. Although CMTR2 exhibits slight preferences for methylation of some m⁷G-proximal sequences, it does not require strictly defined sequence motifs. Furthermore, the primary purpose of Cap2 is not to regulate mRNA fates, as is seen with m⁶A⁴², but is instead to lower the overall burden of Cap1 RNAs on the intracellular defence mechanisms designed to fight against invading pathogens.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-022-05668-z>.

- Adams, J. M. & Cory, S. Modified nucleosides and bizarre 5'-termini in mouse myeloma mRNA. *Nature* **255**, 28–33 (1975).
- Furuichi, Y. et al. Methylated, blocked 5' termini in HeLa cell mRNA. *Proc. Natl Acad. Sci. USA* **72**, 1904–1908 (1975).
- Cho, E. J., Takagi, T., Moore, C. R. & Buratowski, S. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **11**, 3319–3326 (1997).
- McCracken, S. et al. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* **11**, 3306–3318 (1997).
- Moteki, S. & Price, D. Functional coupling of capping and transcription of mRNA. *Mol. Cell* **10**, 599–609 (2002).
- Belanger, F., Stepinski, J., Darzynkiewicz, E. & Pelletier, J. Characterization of hMTR1, a human Cap1 2'-O-ribose methyltransferase. *J. Biol. Chem.* **285**, 33037–33044 (2010).
- Smietanski, M. et al. Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nat. Commun.* **5**, 3004 (2014).
- Langberg, S. R. & Moss, B. Post-transcriptional modifications of mRNA. Purification and characterization of cap I and cap II RNA (nucleoside-2'-)-methyltransferases from HeLa cells. *J. Biol. Chem.* **256**, 10054–10060 (1981).
- Werner, M. et al. 2'-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. *Nucleic Acids Res.* **39**, 4756–4768 (2011).
- Dickinson, M. E. et al. High-throughput discovery of novel developmental phenotypes. *Nature* **537**, 508–514 (2016).
- Cai, H. et al. A functional taxonomy of tumor suppression in oncogenic KRAS-driven lung cancer. *Cancer Discov.* **11**, 1754–1773 (2021).
- Motorin, Y. & Marchand, V. Detection and analysis of RNA ribose 2'-O-methylations: challenges and solutions. *Genes (Basel)* **9**, 642 (2018).
- Maden, B. E., Corbett, M. E., Heeney, P. A., Pugh, K. & Ajuh, P. M. Classical and novel approaches to the detection and localization of the numerous modified nucleotides in eukaryotic ribosomal RNA. *Biochimie* **77**, 22–29 (1995).
- Motorin, Y., Muller, S., Behm-Ansmant, I. & Branlant, C. Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol.* **425**, 21–53 (2007).
- Dorsett, Y. et al. HCoDES reveals chromosomal DNA end structures with single-nucleotide resolution. *Mol. Cell* **56**, 808–818 (2014).
- Krogh, N., Kongsbak-Wismann, M., Geisler, C. & Nielsen, H. Substoichiometric ribose methylations in spliceosomal snRNAs. *Org. Biomol. Chem.* **15**, 8872–8876 (2017).
- Wulf, M. G. et al. Non-templated addition and template switching by Moloney murine leukemia virus (MMLV)-based reverse transcriptases co-occur and compete with each other. *J. Biol. Chem.* **294**, 18220–18231 (2019).
- Consortium, F. et al. A promoter-level mammalian expression atlas. *Nature* **507**, 462–470 (2014).
- Reyes, A. & Huber, W. Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues. *Nucleic Acids Res.* **46**, 582–592 (2018).
- Yamashita, R. et al. Genome-wide characterization of transcriptional start sites in humans by integrative transcriptome analysis. *Genome Res.* **21**, 775–789 (2011).
- Tsuchihara, K. et al. Massive transcriptional start site analysis of human genes in hypoxia cells. *Nucleic Acids Res.* **37**, 2249–2263 (2009).
- Sample, P. J. et al. Human 5' UTR design and variant effect prediction from a massively parallel translation assay. *Nat. Biotechnol.* **37**, 803–809 (2019).
- Paulsen, M. T. et al. Use of Bru-seq and BruChase-seq for genome-wide assessment of the synthesis and stability of RNA. *Methods* **67**, 45–54 (2014).
- Brown, C. J. et al. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **71**, 527–542 (1992).
- Schneider, W. M., Chevillotte, M. D. & Rice, C. M. Interferon-stimulated genes: a complex web of host defenses. *Annu. Rev. Immunol.* **32**, 513–545 (2014).
- Wang, Y. et al. Structural and functional insights into 5'-ppp RNA pattern recognition by the innate immune receptor RIG-I. *Nat. Struct. Mol. Biol.* **17**, 781–787 (2010).
- Devarkar, S. C. et al. Structural basis for m⁷G recognition and 2'-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proc. Natl Acad. Sci. USA* **113**, 596–601 (2016).
- Schubert-Wagner, C. et al. A conserved histidine in the RNA sensor RIG-I controls immune tolerance to Ni²⁺-O-methylated self RNA. *Immunity* **43**, 41–51 (2015).
- Takahashi, K. et al. Nonself RNA-sensing mechanism of RIG-I helicase and activation of antiviral immune responses. *Mol. Cell* **29**, 428–440 (2008).
- Brownell, J. et al. Direct, interferon-independent activation of the CXCL10 promoter by NF-κB and interferon regulatory factor 3 during hepatitis C virus infection. *J. Virol.* **88**, 1582–1590 (2014).
- Kato, H. et al. Cell type-specific involvement of RIG-I in antiviral response. *Immunity* **23**, 19–28 (2005).
- Moyer, S. A. & Banerjee, A. K. In vivo methylation of vesicular stomatitis virus and its host-cell messenger RNA species. *Virology* **70**, 339–351 (1976).
- Berger Rentsch, M. & Zimmer, G. A vesicular stomatitis virus replicon-based bioassay for the rapid and sensitive determination of multi-species type I interferon. *PLoS ONE* **6**, e25858 (2011).
- Lodish, H. F. & Porter, M. Translational control of protein synthesis after infection by vesicular stomatitis virus. *J. Virol.* **36**, 719–733 (1980).

35. Neidermyer, W. J. Jr & Whelan, S. P. J. Global analysis of polysome-associated mRNA in vesicular stomatitis virus infected cells. *PLoS Pathog.* **15**, e1007875 (2019).
36. Kato, H. et al. Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-1 and melanoma differentiation-associated gene 5. *J. Exp. Med.* **205**, 1601–1610 (2008).
37. Wienert, B., Shin, J., Zelin, E., Pestal, K. & Corn, J. E. In vitro-transcribed guide RNAs trigger an innate immune response via the RIG-I pathway. *PLoS Biol.* **16**, e2005840 (2018).
38. Kumar, P. et al. Inhibition of translation by IFIT family members is determined by their ability to interact selectively with the 5'-terminal regions of cap0-, cap1- and 5'ppp-mRNAs. *Nucleic Acids Res.* **42**, 3228–3245 (2014).
39. Abbas, Y. M. et al. Structure of human IFIT1 with capped RNA reveals adaptable mRNA binding and mechanisms for sensing N1 and N2 ribose 2'-O methylations. *Proc. Natl Acad. Sci. USA* **114**, E2106–E2115 (2017).
40. Züst, R. et al. Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat. Immunol.* **12**, 137–143 (2011).
41. Hausmann, I. U. et al. CMTr cap-adjacent 2'-O-ribose mRNA methyltransferases are required for reward learning and mRNA localization to synapses. *Nat. Commun.* **13**, 1209 (2022).
42. Zaccara, S., Ries, R. J. & Jaffrey, S. R. Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* **20**, 608–624 (2019).

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

Methods

Cell culture

HEK293T, A549 and MCF-7 cells were purchased from the American Type Culture Collection. mES cells were obtained as a kind gift from J. Hanna's laboratory. HEK293T and MCF-7 cells were maintained in DMEM (11995065, Gibco) supplemented with 10% FBS and 100 U penicillin–streptomycin (15140148, Gibco). A549 cells were grown in Ham's F-12K medium (21127022, Gibco) supplemented with 10% FBS and 100 U penicillin–streptomycin. mES cells were cultured in Knockout DMEM (10829018, Gibco) supplemented with 15% heat-inactivated FBS, 100 U penicillin–streptomycin, 1× GlutaMAX (35050061, Gibco), 55 μM β-mercaptoethanol, 1× MEM non-essential amino acid solution (11140076, Gibco), 10³ U ml⁻¹ LIF (ESG1107, ESGRO), 3 μM CHIR99021 (72052, STEMCELL Technologies) and 1 μM PD0325901 (72182, STEMCELL Technologies). All cell types were grown in sterile cell culture incubators at 37 °C and 5% CO₂. Cell lines were not authenticated. All cell types tested negative for mycoplasma contamination. Mycoplasma contamination was routinely tested with Hoechst staining.

Generation of KO cell lines

To generate *CMTR2* KO HEK293T cells, 6 × 10⁵ HEK293T cells were seeded in a single well of a six-well cell culture plate. The next day, cells were transfected with 1 μg FTSJD1 double nickase plasmid (SC-412604-NIC, Santa Cruz Biotechnology) using 2 μl LipoD239 transfection reagent (SL100668, SignaGen Laboratories). After 36 h, GFP-positive cells were sorted by flow cytometry and seeded into a single well of a 12-well cell culture plate. After 24 h, GFP-positive cells were treated with 5 μg ml⁻¹ puromycin for 36 h. The remaining viable cells were washed twice with PBS and grown until 30–50% confluency was reached. Single-cell clones were isolated and screened for *CMTR2* depletion using western blot. To generate *CMTR2* KO and *CMTR2RIG-I* double KO A549 cells, 2.5 × 10⁵ A549 cells were seeded in a single well of a six-well plate. The following day, cells were transfected with either 2.5 μg FTSJD1 double nickase plasmid or a mixture of 1.5 μg FTSJD1 and 1.5 μg RIG-I (SC-400812-NIC, Santa Cruz Biotechnology) double nickase plasmids using Lipofectamine 3000 transfection reagent (L3000001, Invitrogen) according to the manufacturer's instructions. The procedure for isolation of single-cell KO clones was performed as described for HEK293T cells.

Animal maintenance and procedures

All animals used in this study were maintained in compliance with Weill Cornell Medicine Institutional Animal Care and Use Committee (IACUC) protocols. Eight-week-old CL57B/L wild-type female mice were purchased from Charles River Laboratory and housed in standard cages with unrestricted supplies of water and food with 14 h light–10 h dark cycle at 18–23 °C and 40–60% humidity. Sixteen-week-old female mice were dissected for isolation of the brain, liver, kidney, lung, heart and spleen. Upon isolation, organs were stored in TRIzol reagent at –80 °C until further use. The zebrafish wild-type AB line was used in this study. Embryos were maintained in E3 medium at 28 °C and staged as previously described⁴³. Forty-eight-hour-old zebrafish embryos were collected in TRIzol reagent and stored at –80 °C until further use.

Total RNA isolation

Total RNA was isolated from cells using TRIzol reagent according to the manufacturer's instructions unless otherwise stated. For isolation of total RNA from mouse tissues, TRIzol-submerged tissues were homogenized with high-impact zirconium 1.5-mm beads (D1032-15, Benchmark Scientific) twice at 50 Hz for 3 min using TissueLyser II (Qiagen). For isolation of total RNA from zebrafish samples, embryos were dissolved in TRIzol reagent and passed three times through a 21-gauge needle.

mRNA extraction

Total RNA (20 μg) was diluted in 75 μl nuclease-free water. Oligo(dT)₂₅ magnetic beads (50 μl; bed volume; S1419S, NEB) were resuspended in 75 μl 2× mRNA binding buffer (40 mM Tris-HCl pH 7.5, 1 M LiCl, 2 mM EDTA and 0.1% Triton X-100) and mixed with total RNA. Samples were heated at 65 °C for 5 min, placed on ice for 3 min and incubated at room temperature for 10 min with constant rotation. Beads were washed twice with 150 μl mRNA wash buffer (20 mM Tris-HCl pH 7.5, 150 mM LiCl, 1 mM EDTA and 0.01% Triton X-100) and resuspended in 75 μl mRNA elution buffer (10 mM Tris-HCl pH 7.5). Samples were heated at 75 °C for 2.5 min and placed on ice for 2 min. To ensure pure mRNA isolation, a second round of poly(A)⁺ RNA purification was conducted. 75 μl 2× mRNA binding buffer was mixed with the beads from the previous step, followed by incubation for 10 min at room temperature with constant rotation. Beads were then washed once with 150 μl mRNA wash buffer and resuspended in 25 μl water. To allow for the final mRNA elution from the beads, samples were heated at 75 °C for 3 min and the supernatant was collected. The extracted mRNA was further purified using a Zymo Research Clean and Concentrator-5 (RCC-5) column (R1013, Zymo Research). In the experiments in which a higher or lower amount of total RNA was used, mRNA was isolated by upscaling or downscaling of the mRNA isolation reagents.

CapTag-seq

mRNA (2 μg) was treated with 25 U Quick CIP (M0525L, NEB) in a 30-μl reaction for 30 min at 37 °C. The reaction was cleaned using a Zymo RCC-5 column and mRNA was eluted with 20 μl water. m⁷GDP was removed from mRNA 5' termini (decapping) using 5 U Cap-Clip acid pyrophosphatase (C-CC15011H, CELLSCRIPT) in a 20-μl reaction for 1 h at 37 °C. The mRNA decapping reaction was cleaned with Zymo RCC-5 column and mRNA was eluted in 11.5 μl water. A biotinylated 5' adapter composed of 2'-O-methylated nucleotides (biotin-N_m-RA5) was ligated to the 5'-monophosphorylated mRNA ends using 60 U T4 RNA ligase 1 (M0437M, NEB) in the following 30 μl reaction mixture: 3 μl 10× RNA ligase buffer, 10 μl decapped mRNA, 3 μl 10 mM ATP, 1 μl 30 μM biotin-N_m-RA5, 10.5 μl 50% PEG-8000, 0.5 μl 40 U μl⁻¹ RNaseOUT and 2 μl 30 U μl⁻¹ T4 RNA ligase 1 for 4 h at 25 °C. To successfully remove non-ligated biotin-N_m-RA5, the ligation reaction was cleaned three times with Zymo RCC-5 column following the manufacturer's protocol for isolation of RNA longer than 200 nucleotides. Ligated mRNA was eluted in the third cleanup step in 40 μl water. Next, ligated mRNA was digested to completion with 5 U RNase T2 (GE-NUC00400-02, MoBiTec) in 30 mM Na-acetate pH 4.5 overnight at 37 °C. The resulting adapter-linked cap tags were captured on M-280 streptavidin Dynabeads (20 μl bed volume, Invitrogen) in 300 μl binding buffer (10 mM Tris-HCl pH 7.5, 300 mM NaCl and 0.05% Triton X-100) for 30 min at room temperature. Beads were washed twice with 500 μl high-salt buffer (10 mM Tris-HCl pH 7.5, 2 M NaCl and 0.1% Triton X-100), twice with 500 μl binding buffer, and twice with 500 μl low-salt buffer (10 mM Tris-HCl pH 7.5, 50 mM NaCl and 0.025% Triton X-100). The removal of the 2',3'-cyclic phosphate from the 3' end of cap tags was conducted using 10 U T4 PNK (M0201L, NEB) in a 20 μl dephosphorylation buffer (100 mM Na-acetate pH 6.0, 10 mM MgCl₂ and 5 mM DTT) for 20 min at 23 °C. Beads were washed twice with 500 μl binding buffer, and twice with 500 μl low-salt buffer. Next, the preadenylated DNA adapter with a 20-nucleotide randomized region (N₂₀-DA3) was ligated to the dephosphorylated 3' ends of cap tags with 200 U T4 RNA ligase 2, truncated KQ (M0373L, NEB) in a 20 μl ligation mix (2 μl 10× RNA ligase buffer, 1 μl 20 μM DDA3, 7 μl water, 1 μl 200 U μl⁻¹ T4 RNA ligase 2, truncated KQ, 8.5 μl 50% PEG-8000 and 0.5 μl 40 U μl⁻¹ RNaseOUT) for 4 h at 27 °C. The non-ligated N₂₀-DA3 was removed with 25 U yeast 5'-deadenylase (M0331S, NEB) and 15 U RecJf (M0264S, NEB) for 45 min at 30 °C. Beads were washed twice with 500 μl high-salt buffer, twice with 500 μl binding buffer and twice with 500 μl low-salt buffer. For cDNA preparation,

Article

beads were first resuspended in 12 μl reverse transcription annealing mix (1 μl 10 μM reverse transcription primer (RTP), 4 μl 250 mM Tris-HCl pH 8.3, 5 μl 300 mM KCl and 2 μl water), heated at 90 $^{\circ}\text{C}$ for 2 min, and slowly cooled down to 25 $^{\circ}\text{C}$ with a rate of 0.1 $^{\circ}\text{C s}^{-1}$. Following the RTP annealing, 7.25 μl RT mixture (4 μl water, 1 μl 0.1 M DTT, 1 μl 10 mM dNTPs (each), 1 μl 60 mM MgCl_2 and 0.25 μl 40 U μl^{-1} RNaseOUT) and 0.75 μl SuperScript III (18080051, Invitrogen) were added to the annealing mix. The reaction was incubated for 30 min at 55 $^{\circ}\text{C}$, and heat inactivated for 10 min at 75 $^{\circ}\text{C}$. PCR amplification of cDNA was performed with 2 \times Phusion High Fidelity PCR master mix with HF buffer (M0531L, NEB). The resulting PCR products were purified twice with 1.8 \times PCR volume AMPure XP beads (A63881, Beckman Coulter). Amplified cDNA libraries were sequenced on Illumina instruments in a single-end or paired-end modes.

CapTag-seq data analysis

Low-quality sequencing reads were filtered out and the 3' end adapter (DA3) was trimmed off using Flexbar v2.5 (ref. 44). Duplicated reads were removed using the pyFastqDuplicateRemover.py script within the pyCRAC package⁴⁵. Next, the 20-bp-long randomized region was removed from the 3' end of sequencing reads using the UNIX cut command. The remaining portion of the reads represents RNase T2-released tags. Tag length distribution, quantification of the two-nucleotide and three-nucleotide-long cap tags from Cap1 and Cap2 mRNAs, respectively, and Cap1 and Cap2 m⁷G-proximal sequences of cap tags were obtained using basic UNIX commands.

TSS-seq

mRNA (200 ng) was treated with 25 U Quick CIP in a 30- μl reaction for 30 min at 37 $^{\circ}\text{C}$. The reaction was cleaned using a Zymo RCC-5 column and mRNA was eluted in 17 μl water. m⁷GDP was removed from the mRNA 5'-termini using 5 U Cap-Clip acid pyrophosphatase in a 20- μl reaction for 1 h at 37 $^{\circ}\text{C}$. The decapped mRNA was isolated by mixing the decapping reaction with 44 μl RNAClean XP beads (A63987, Beckman Coulter). After 15 min of incubation at room temperature, the beads were washed twice with 80% ethanol and mRNA was eluted from the beads in 10 μl water for 5 min. Next, a biotinylated 5' RNA adapter with the eight-nucleotide-long unique molecular identifiers (biotin-RA5-UMIs) was ligated to the 5'-monophosphorylated mRNA ends with 60 U RNA ligase I in the 30- μl reaction as described in CapTag-seq. Ligated mRNA was purified by mixing the ligation reaction with 44 μl 1.5 M NaCl and 31 μl RNAClean XP beads. After 10 min of incubation at room temperature, beads were washed twice with 80% ethanol and RNA was eluted from the beads in 20 μl water for 5 min. Eluted RNA was further mixed with 28 μl RNAClean XP beads and incubated for 10 min at room temperature. Beads were washed twice with 80% ethanol and RNA was eluted as before in 20 μl water. Next, the ligated mRNA was subjected to a partial alkaline-based fragmentation by mixing 20 μl mRNA with 5 μl 240 mM NaHCO_3 and 5 μl 360 mM Na_2CO_3 . The fragmentation mixture was incubated for 11 min at 60 $^{\circ}\text{C}$ and immediately placed on ice. Fragmented RNA was extracted using a Zymo RCC-5 column, eluted in 20 μl water and stored at -80 $^{\circ}\text{C}$ until further use. The following day, ligated mRNA 5' ends were captured on M-280 streptavidin Dynabeads (20 μl bed volume) in 100 μl binding buffer for 30 min at room temperature. Next, the beads were washed twice with 500 μl high-salt buffer, twice with 500 μl binding buffer and twice with 500 μl low-salt buffer. Dephosphorylation of the 3' ends of RNA fragments, ligation of the preadenylated DNA adapter (DA3) and reverse transcription steps were conducted as described in the CapTag-seq protocol. For the reverse transcription reaction, samples were incubated for 45 min at 55 $^{\circ}\text{C}$, followed by heat inactivation for 10 min at 75 $^{\circ}\text{C}$. cDNA was PCR amplified using 2 \times Phusion High Fidelity PCR master mix with HF buffer and the resulting PCR products were purified twice with 0.9 \times PCR volume AMPure XP beads. Amplified cDNA libraries were sequenced on the Illumina instrument in a paired-end mode.

TSS-seq data analysis

Low-quality sequencing reads were filtered out and the 3' end adapter (DA3) was trimmed off using Flexbar v2.5 (ref. 44). Duplicated reads were removed using the pyFastqDuplicateRemover.py script within the pyCRAC package v1.3.2 (ref. 45). Only the read R1 was considered for further analysis. Eight-nucleotide-long randomized region (UMIs) was removed from the 5' end of sequencing reads using seqtk. UMI-free reads were then shortened from the 3' end to the universal length of 25 bp using seqtk. The first nucleotide of the processed, 25-bp-long reads represents the TSN of an mRNA. The processed reads were first aligned to the *Drosophila melanogaster* genome (dm6) using Bowtie v1.2.3 (ref. 46). The remaining, unmapped reads were then aligned to the human (h38) or mouse (mm10) genome using Bowtie v1.2.3. The 5' end read coverage (representing a TSN coverage) per each genomic position was obtained using BEDTools v2.28.0 (ref. 47). The Ensembl gene annotation file was obtained from Ensembl⁴⁸ and annotated gene starts were extended by 250 bp. The aligned reads were annotated with BEDTools using the modified Ensembl gene annotation file. For each sample, 5' end read coverage per each genomic position was normalized to the total number of mapped reads (mapped reads per million (RPM)). Genomic positions with the normalized 5' end coverage < 2 RPM were discarded. A genomic position with the maximum coverage within an annotated gene was identified as the major TSN isoform of a gene. All other genomic positions within that gene with normalized 5' end coverage more than 10% of the maximum (major) TSN were considered as alternative TSN isoforms for that gene.

ActD-TSS-seq

CMTR2 WT or *CMTR2* KO HEK293T cells (1.8×10^6) were seeded in a 6-mm cell culture dish. The day after seeding, cells were treated with 5 $\mu\text{g ml}^{-1}$ actD to block synthesis of new transcripts. Total RNA was extracted from cells after 0, 2, 8 and 16 h of actD treatment. Total RNA of each sample was spiked-in with 40 ng *D. melanogaster* poly(A)⁺ RNA (636222, Takara). Cellular poly(A)⁺ RNA was extracted using oligo(dT)₂₅ magnetic beads. mRNA (200 ng) from each time point of actD treatment was subjected to the TSS-seq protocol as described above.

ActD-TSS-seq data analysis

Sequencing reads were processed as described above with minor modifications.

The annotated TSN isoforms with expression more than 4 RPM at 0 h of actD treatment were considered for calculations of mRNA half-lives. The TSN isoform expression at each time point of actD treatment was first normalized to the total number of reads mapping to the *D. melanogaster* genome in each sample. The normalized TSN expression was then transformed with a sample-specific scale factor determined from the general mRNA decay rate in HEK293T cells (see Extended Data Figs. 5g and 6f). The scaling factors were as follows: for 0 h actD = 1, for 2 h actD = 0.77, for 8 h actD = 0.56 and for 16 h actD = 0.41. The fully normalized TSN isoform expression values at each time point after actD treatment were then used for the calculation of the mRNA decay rates with a one-phase decay model using the *drm* function in the *drcR* package⁴⁹. TSN isoform half-lives were calculated as: $t_{1/2} = \ln 2/k$, where k represents the TSN isoform-specific decay constant derived from the *drm* function.

Polysome-TSS-seq

CMTR2 WT or *CMTR2* KO HEK293T cells (8×10^6) were seeded onto a 150-mm cell culture dish. The following day, cells were treated with 100 $\mu\text{g ml}^{-1}$ cycloheximide (CHX; C7698, Sigma) for 5 min at 37 $^{\circ}\text{C}$. Cell dishes were placed in ice, washed once with 10 ml ice-cold PBS supplemented with 100 $\mu\text{g ml}^{-1}$ CHX (PBS + CHX), and scraped in 8 ml PBS + CHX. Cells were pelleted at 300g for 3 min at 4 $^{\circ}\text{C}$, supernatant aspirated, and cell pellet was resuspended in 0.5 ml ice-cold polysome

extraction buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 1% Triton X-100, 5 mM MgCl₂, 2 mM DTT, 100 μg ml⁻¹ CHX and 1× Halt protease and phosphatase inhibitor cocktail (78440, Thermo Scientific)). Cells were left on ice for 5 min to lyse. The cell lysis process was facilitated by passing cells through a 21-gauge needle. Cell lysate was centrifuged for 5 min at 16,000g at 4 °C, the supernatant was transferred to a new 1.5-ml tube and snap frozen in liquid nitrogen. Frozen cell lysates were stored at -80 °C until further use. Frozen cell lysates were thawed on ice. Lysate (500 μl) was layered on top of the 10–50% linear sucrose gradient prepared with polysome extraction buffer. The gradient was centrifuged for 2 h at 36,000g in a SW-41 Ti swinging bucket rotor at 4 °C. Polysome fractionation was performed using an automated fraction collector (BioComp) with a continuous monitoring of the 254-nm absorbance. Fractions corresponding to a one, two–three, four–five, six–seven and more than eight ribosomes were collected manually. *Drosophila* poly(A)⁺ RNA (3 μl of 2.5 ng μl⁻¹) was spiked-in to each fraction to account for differences in RNA extraction between fractions. Next, an equal volume of TRIzol LS was added to each isolated polysome fraction along with 15 mM EDTA (final concentration). Tubes were briefly vortexed and left at room temperature for 30 min. Total RNA was isolated as instructed by the TRIzol LS manufacturer's protocol and precipitated using isopropanol. mRNA was extracted from the total RNA of each polysome fraction with oligo(dT)₂₅ magnetic beads. Poly(A)⁺ RNA (200 ng) from each fraction was subjected to the TSS-seq protocol as described above.

Polysome–TSS-seq data analysis

Sequencing reads were processed as described in the actD-TSS-seq data analysis with minor modifications. The genomic positions with 5' end read coverage of more than 1 RPM in each polysome fraction were considered. Genomic positions that passed this expression threshold were then annotated to the TSN isoforms identified in the input sample (0 h actD-TSS-seq libraries) using BEDTools. The expression (RPM) of the annotated TSN isoforms was further normalized to the total number of reads mapping to the *D. melanogaster* genome in each polysome fraction. Following the normalization, an average number of ribosomes bound to each TSN isoform was calculated as follows: total abundance of a TSN isoform (N_{total}) was calculated as the sum of the TSN isoform expression levels (N) in each of the five isolated polysome fractions (F_n): $N_{total} = N_{F1} + N_{F2} + N_{F3} + N_{F4} + N_{F5}$, where F denotes polysome fraction number. Total number of ribosomes (R_{total}) associated with each TSN isoform was calculated as $R_{total} = f_1 \times N_{F1} + f_2 \times N_{F2} + f_3 \times N_{F3} + f_4 \times N_{F4} + f_5 \times N_{F5}$, where f denotes the number of ribosomes present in each polysome fraction ($f_1 = 1, f_2 = 2.5, f_3 = 4.5, f_4 = 6.5$ and $f_5 = 12$). The average number of ribosomes bound to each TSN isoform (MRL) was then calculated as $MRL = R_{total} / N_{total}$. Translation efficiency of each TSN isoform was defined as the average ribosome density on each TSN isoform. Ribosome density (MRL per kb) was calculated as MRL per kilobase of the mRNA open reading frame.

CLAM-Cap-seq

mRNA (1 μg) was partially fragmented in a 60-μl reaction containing 40 mM NaHCO₃ and 60 mM Na₂CO₃ for 8.5 min at 60 °C. Seven independent fragmentation reactions were combined (7 μg poly(A)⁺ RNA in total) and collectively purified using a Zymo RCC-5 column. Fragmented mRNA was eluted from the column in 31 μl water. Next, the internal mRNA fragments were 5' end phosphorylated with 10 U T4 PNK in two separate 30-μl reactions (15 μl mRNA, each) for 30 min at 37 °C. Two phosphorylation reactions were combined and cleaned as described before. mRNA was eluted from the column in 21 μl water. Next, 5'-phosphorylated internal mRNA fragments were removed with 1 U Terminator 5' phosphate-dependent exonuclease (TER51020, Lucigen) in two separate 15-μl reactions (10 μl mRNA, each) for 1 h at 30 °C. The remaining m⁷G-protected 5' mRNA fragments were purified by mixing two combined Terminator reactions (30 μl total) with 45 μl

RNAclean XP beads. The mixtures were left at room temperature for 15 min, followed by two bead washes with 80% ethanol. RNA was eluted from the beads in 16 μl water for 5 min. Enriched 5' mRNA fragments were subjected to m⁷GDP removal with 7.5 U Cap-Clip acid pyrophosphatase in a 20-μl reaction for 1 h at 37 °C. The decapping reaction was mixed with 36 μl RNAclean XP beads and 84 μl 100% ethanol, and left at room temperature for 15 min. The beads were washed twice with 80% ethanol, and decapped mRNA fragments were eluted from the beads in 11 μl water for 5 min. Next, the decapped mRNA fragments were incubated in the annealing mix (10 μl decapped mRNA fragments, 1 μl 10 μM biotin-N₆-DA3-RTP, 2 μl 500 mM Tris-HCl pH 8.3, and 2 μl 750 mM KCl) at 90 °C for 1.5 min and cooled down quickly to 4 °C. Following N₆-DA3-RTP annealing, 4.25 μl RT mix (2 μl 0.1 M DTT, 1 μl 10 mM dNTPs (each), 1 μl 60 mM MgCl₂ and 0.25 μl 40 U μl⁻¹ RNaseOUT) was added to the annealing mix along with 0.75 μl 200 U μl⁻¹ Maxima reverse transcriptase RNase H minus (EP0752, Thermo Scientific). The reverse transcription reaction was incubated at 25 °C for 7.5 min, followed by a 10-min incubation at 52 °C and a 30-min incubation at 57 °C. Reactions were cooled down to 4 °C, mixed with M-280 streptavidin Dynabeads (10-μl bed volume) in 100 μl binding buffer and incubated at room temperature for 30 min. The beads were washed twice with 500 μl binding buffer and twice with 500 μl low-salt buffer. Next, the beads were resuspended in 18.5 μl 1× CircLigase mix (2 μl 10× CircLigase buffer, 4 μl 5 M betaine, 1 μl 50 mM MnCl₂, 1 μl 1 mM ATP and 10.5 μl water), and 1 μl 100 U μl⁻¹ CircLigase II was added. The CircLigase reaction was incubated for 8 h at 60 °C with occasional pipetting of the beads. The reaction was then resuspended in 100 μl binding buffer containing fresh M-280 streptavidin Dynabeads (5-μl bed volume) and incubated for 30 min at room temperature. Beads were then washed once with 500 μl high-salt buffer, twice with 500 μl binding buffer and once with 500 μl low-salt buffer. Next, the beads were resuspended in 50 μl 0.2 M KOH and incubated overnight at 37 °C. The following day, HCl was added to the solution to obtain the pH of 7.5, and new M-280 streptavidin Dynabeads (10-μl bed volume) were added. Sample was incubated for 30 min at room temperature, followed by bead washes with 500 μl high-salt buffer, twice with 500 μl binding buffer and once in 500 μl low-salt buffer. Beads were then incubated for 2.5 h at 37 °C in a 50-μl reaction containing 30 mM Na-acetate pH 4.5 and 5 U RNase T2. Beads were then washed once with 500 μl high-salt buffer, twice with 500 μl binding buffer and twice with 500 μl low-salt buffer. The removal of the 2',3'-cyclic phosphate from the 3' ends of cap tags and the ligation of the preadenylated DNA adapter with UMIs (DA5-UMIs) was performed as described in the CapTag-seq protocol. After ligation, the beads were washed twice with 500 μl high-salt buffer, twice with 500 μl binding buffer and twice with 500 μl low-salt buffer. Beads were then resuspended in 20 μl cDNA buffer (10 mM Tris-HCl pH 7.5) and used for subsequent PCR amplification of cDNA–cap tags with 2× Phusion High Fidelity PCR master mix with HF buffer. The resulting PCR products were purified twice with 0.9× PCR volume AMPure XP beads. Purified cDNA libraries were sequenced on Illumina instrument in a paired-end mode.

CLAM-Cap-seq data analysis

Low-quality sequencing reads were filtered out and the 3' end DNA adapter (DA3) was trimmed off using Flexbar v2.5 (ref. ⁴⁴). Duplicated reads were removed using the pyFastqDuplicateRemover.py script within the pyCRAC package v1.3.2 (ref. ⁴⁵). Following the removal of PCR duplicates, only the read R1 was considered for further analysis. The six-nucleotide-long UMI was first removed from the 5' end of sequencing reads using seqtk. Next, we determined Cap1 and Cap2 origins of the sequencing reads. To achieve that, 5' ends of processed reads were screened for the presence of Cap1 and Cap2 palindromes containing zero and one non-templated nucleotide using grep UNIX command. The palindrome search was focused on all 64 possible trinucleotide sequences that an mRNA may begin with (for example,

Article

AGA, CTT, and so on), except for the homopolymeric trinucleotide stretches (for example, AAA, GGG, CCC and UUU), as these can generate palindromes with ambiguous Cap1 and Cap2 assignment. Of note, as U-starting mRNAs are very rare in mammalian transcriptomes (see Extended Data Fig. 2d), they were also excluded from the Cap1 and Cap2 palindrome search. Next, the sequencing reads containing Cap1 and Cap2 palindromes were separated, followed by the removal of cap tags and a non-templated nucleotide with seqtk to obtain mappable reads whose starts represent the first mRNA nucleotide. Cap tag-free Cap1-derived and Cap2-derived reads were then trimmed at their 3' end to obtain the universal read length of 25 bp. After processing, Cap1 and Cap2 reads were separately mapped to the human (hg38) or mouse (mm10) genome using Bowtie v1.2.3. 5' end read coverage per each genomic position was assessed using BEDTools. The aligned reads were then annotated to the TSN isoforms identified via TSS-seq using BEDTools. Cap2 stoichiometry for each TSN isoform was calculated as the Cap2 read fraction of the total (Cap1 + Cap2) reads. Only TSN isoforms whose total (Cap1 + Cap2) CLAM-Cap-seq read coverage was higher or equal to 30 were considered for Cap2 stoichiometry calculations. TSN isoforms showing low variability in Cap2 stoichiometry between replicates (standard deviation of less than 5%) were used for all downstream analysis.

Gene enrichment analysis

For gene enrichment analysis (WebGestalt (<http://www.webgestalt.org/>)), Cap2 stoichiometry of a gene was calculated as a weighted Cap2 stoichiometry average of all TSN isoforms of that gene. In these calculations, the contribution of each TSN isoform to the average Cap2 stoichiometry for a gene was based on the relative expression level of each Cap2-modified TSN isoform. Genes with the highest and lowest Cap2 stoichiometry (top 20%) were identified. The overrepresentation analysis⁵⁰ for KEGG pathways was performed using all genes with measured Cap2 stoichiometry as a reference gene set. In the RNA-seq dataset, genes were ranked based on the level of the changes in their expression between *CMTR2* KO and *CMTR2* WT cells (\log_2 fold change) and subjected to the gene set enrichment analysis⁵⁰ for identification of gene sets enriched in specific Reactome pathways.

CapOligo-PAGE

Small or poly(A)⁺ RNA (3 μ g) was subjected to m⁷GDP removal with 5 U Cap-Clip acid pyrophosphatase in a 20- μ l reaction for 1 h at 37 °C. The decapped RNA was purified using a Zymo RCC-5 column and eluted in 20 μ l water. RNA 5' ends were dephosphorylated using 25 U Quick CIP in a 30- μ l reaction at 37 °C. RNA was purified as described above and eluted in 10 μ l water. RNA (7 μ l) from the previous step was radiolabelled in a 10- μ l reaction using 10 U T4 PNK and 1 μ l 10 mCi ml⁻¹ [³²P]-ATP for 30 min at 37 °C. The reaction was heat inactivated at 85 °C for 2 min. Next, 1 μ l 10 μ M 5'-biotinylated self-splinting DNA oligonucleotide was added to the reaction mix. The sample was heated at 85 °C for 4 min and left at room temperature for 20 min to allow DNA oligo annealing to the 5' end of the target RNA. Ligation mix (3.7 μ l; 0.56 \times PNK buffer, 0.15 mM ATP, 22.8% DMSO and 1.9 U μ l⁻¹ T4 DNA ligase (EL001I, Thermo Scientific)) was added to the reaction, mixed and incubated for 4 h at 37 °C. M-280 streptavidin Dynabeads (7- μ l bed volume) in 100 μ l binding buffer were added to the ligation mix and incubated for 30 min at room temperature with constant shaking. Next, the beads were washed twice with 500 μ l high-salt buffer, twice with 500 μ l binding buffer and twice with 500 μ l low-salt-buffer. After the last wash, beads were resuspended in 45 μ l 30 mM Na-acetate pH 4.5, and 4 μ g Monarch RNase A (T3018L, NEB) and 5 U RNase T2 were added. Samples were incubated overnight at 37 °C to completely degrade mRNA, leaving the cap tags attached to the self-splinting DNA oligonucleotide. The following day, the beads were washed twice with 500 μ l binding buffer and twice with low-salt buffer. The beads were then resuspended in 50 μ l 10 mM HCl and incubated for 30 min at 37 °C to open up the 2',3'-cyclic phosphate

at the 3' end of RNase-derived cap tags. Binding buffer (150 μ l) was added, and the pH of the reaction was adjusted with KOH to 7.5. M-280 streptavidin Dynabeads (5- μ l bed volume) were added to the mixture and incubated for 30 min at room temperature. Next, the beads were washed once with 500 μ l binding buffer and once with low-salt buffer. Beads were then resuspended in 10 μ l 1.1 \times CutSmart buffer and digested with 20 U BamHI HF (3136L, NEB) for 1.25 h at 37 °C. Beads were gently washed once with 150 μ l binding buffer and once with 150 μ l low-salt buffer. Beads were then resuspended in 10 mM Tris-HCl pH 7.5, heated at 85 °C for 7 min and supernatants were quickly collected into a fresh tube. The eluted samples containing DNA oligo-linked cap tags were mixed with 2 \times Novex TBE-urea sample buffer (LC6876, Invitrogen) and loaded onto a 20-cm-long 15% TBE-urea sequencing gel. Samples were run on the gel until the bromophenol blue dye reached the end of the gel. The resolved DNA oligo-linked cap tags were transferred onto a nylon membrane, UV crosslinked and exposed to the phosphor screen. Autoradiographs were developed using the FLA7000IP Typhoon Phosphorimager.

BruChase

HEK293T cells (5×10^6) were seeded in a 10-cm cell culture dish. The following day, cells were incubated in growing media containing 200 μ g ml⁻¹ 5-BrU (850187, Sigma) for 3 h to allow for metabolic labelling of newly synthesized transcripts. Then, BrU-containing media were withdrawn, and cells were chased in growing media containing 2 mM uridine (U3750, Sigma) for 0, 4, 8 and 16 h. Total RNA was extracted from cells at each time point of the uridine chase using TRIzol reagent and isopropanol precipitation. Next, 100 μ g total RNA for each time point was incubated with 10 μ g anti-bromodeoxyuridine antibody (MI-11-3, MBL Life Science) in 500 μ l IP buffer (5 mM Tris-HCl pH 7.5, 0.5 \times PBS, 0.05% Triton X-100, 1 mM EDTA and 40 U ml⁻¹ RNaseOUT) for 2 h at 4 °C with constant rotation. Pierce Protein A/G magnetic beads (30- μ l bed volume; 88802, Thermo Scientific) in 100 μ l IP buffer were added to the RNA-antibody mix and incubated for an additional 1 h at 4 °C with constant rotation. Beads were then washed four times with 700 μ l IP buffer. The immunopurified RNA was eluted from the beads via proteinase K treatment. Beads were mixed in 300 μ l proteinase K buffer (100 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.5% SDS and 10 mM EDTA) and 10 μ l 20 mg ml⁻¹ proteinase K solution (RNA grade; 25530049, Invitrogen), and incubated for 45 min at 50 °C. Next, the supernatant was collected and mixed with 300 μ l phenol:chloroform:IAA, 25:24:1, pH 6.6 (AM9730, Invitrogen) in a Phase Lock Gel Heavy tube (2302830, QuantaBio). The mixture was incubated for 5 min at 30 °C and centrifuged at 13,000g for 5 min. Following centrifugation, the aqueous phase was collected, and BrU-containing RNA was precipitated with 0.3 M Na-acetate pH 5.5, 2.5 \times sample volume 100% ethanol and 2 μ l 15 mg ml⁻¹ Glycoblue (AM9515, Invitrogen). Two independent immunopurifications were combined and subjected to the CLAM-Cap-seq protocol. DNA adapter-linked cDNA-cap tags were resuspended in 40 μ l water and subjected to the RT-qPCR analysis.

BruChase-CapTag-seq

HEK293T cells (9×10^6) were seeded in a 15-cm cell culture dish. The following day, cells were incubated in growing media containing 200 μ g ml⁻¹ 5-BrU (850187, Sigma) for 3 h to allow for metabolic labelling of newly synthesized transcripts. Then, BrU-containing media were withdrawn, and cells were chased in growing media containing 2 mM uridine (U3750, Sigma) for 0 h and 8 h. Total RNA was extracted from cells at both time points of the uridine chase using TRIzol reagent and isopropanol precipitation. mRNA was extracted from total RNA using oligo(dT) capture as described above. Isolated mRNA from three 15-cm cell culture dishes constituted a single biological replicate. BrU-labelled mRNA was immunopurified as described above. BrU-labelled mRNA (200 ng) from each time point of the uridine chase was subjected to the CapTag-seq procedure.

RNA-seq

Total RNA was extracted from 2×10^6 *CMTR2*WT or *CMTR2*KO HEK293T cells using TRIzol reagent. The RNA quality was assessed by Bioanalyzer analysis. Total RNA was spiked in with ERCC RNA spike-in mix 1 according to the manufacturer's instructions (4456740, Invitrogen). Total RNA (1 μ g) was used for RNA-seq library preparation with the NEBNext Ultra II RNA Library Prep kit for Illumina (E7770S, NEB). Ribosomal RNA was removed using the NEBNext rRNA Depletion kit (E6310L, NEB). The libraries were sequenced on the Illumina instrument in a paired-end mode. Four independent biological replicates were sequenced for each condition.

RNA-seq data analysis

Sequencing reads of low quality were discarded and reads shorter than 18 bp were removed. Ribosomal rRNA reads were removed using STAR aligner⁵¹. The remaining reads were then mapped to the human (hg38) protein-coding transcriptome using STAR aligner. Trimmed mean of M values (TMM) normalization, empirical Bayes estimate of the negative binomial dispersion, and measurement of the changes in gene expression (\log_2 fold change) were performed for all samples and replicates at the same time using edgeR⁵².

Western blotting and antibodies

Cell pellets were collected by centrifugation at 300g for 5 min. After a single wash in ice-cold PBS, cell pellets were resuspended in cell lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP-40, 0.1% SDS and 1 \times Halt protease and phosphatase inhibitor cocktail) and lysed on ice for 10 min. Cell lysis was facilitated by sonication with the following parameters: four times 5 s on, 10 s off at 10% amplitude (Branson). Cell lysates were cleared by a 5 min of centrifugation at 20,000g at 4 °C. Protein concentration in isolated cell lysates was determined using the Pierce BCA Protein Assay kit (23225, Thermo scientific). Protein (20–25 μ g) was loaded per lane onto a NuPAGE 4–12% Bis-Tris gel (NP0322BOX, Invitrogen). Resolved proteins were transferred onto a PVDF membrane and probed with an antibody recognizing a protein of interest. The following antibodies were used in this study: rabbit anti-FTSJD1 polyclonal antibody (1:500 dilution; PA5-61696, Invitrogen), rabbit anti-RIG-I (D14G6) monoclonal antibody (1:1,000 dilution; 3743T, CST), rabbit RIG-I/DDX58 (EPRI8629) monoclonal antibody (1:2,000 dilution; ab180675, Abcam), rabbit anti-MDA5 (D74E4) monoclonal antibody (1:1,000 dilution; 5321T, CST), rabbit anti-IFIT1 (D2X9Z) monoclonal antibody (1:1,000 dilution; 14769S, CST), rabbit anti-IFITM3 (D8E8G) XP monoclonal antibody (1:1,000 dilution; 59212T, CST), mouse anti-NLRP1 monoclonal antibody (1:1,000 dilution; 679802, BioLegend), rabbit anti-MAVS polyclonal antibody (1:1,000 dilution; 3993T, CST), rabbit anti-OAS3 polyclonal antibody (1:1,000 dilution; ab154270, Abcam), mouse anti-puromycin (12D10) monoclonal antibody (1:5,000 dilution; MABE343, Millipore Sigma), mouse anti-GAPDH monoclonal antibody (1:10,000; GT239, GeneTex), mouse anti- β -actin monoclonal antibody (1:4,000; AM4302, Thermo), horseradish peroxidase-conjugated donkey anti-rabbit IgG (1:5,000; NA934, Cytiva) and horseradish peroxidase-conjugated sheep anti-mouse IgG (1:5,000; NA931, Cytiva). For anti-RIG-I (D14G6), anti-NLRP1 and anti-MDA5 (D74E4) antibodies, membranes were incubated with 3% BSA in TBS supplemented with 0.1% Tween-20 (TBST). All other antibodies were used with 3% milk in TBST.

RT-qPCR

To remove potential DNA contamination, total RNA was first treated with 1 U DNase I (EN0521, Thermo Scientific) for 20 min at 37 °C and purified using a Zymo RCC-5 column. Total RNA (1–3 μ g) was reverse transcribed to cDNA with the SuperScript III First Strand synthesis system (18080051, Invitrogen) using either random hexamers (N8080127, Invitrogen) or oligo(dT)₂₀ (18418020, Invitrogen) as reverse transcription primers. The same amount of total RNA was used for directly compared

conditions. qPCR was performed using the iQ SYBR Green Supermix (1708880, Bio-Rad) with 150 nM primers in a 10- μ l reaction. The amplifications were conducted with the following protocol in all experiments: 95 °C for 10 min, 40 cycles of 95 °C for 15 s, 58 °C for 15 s and 68 °C for 20 s. The specificity of primer pairs was tested with melting curves at the end of the 40th amplification cycle. *GAPDH*-normalized gene expression was calculated and presented. Normalized gene expression values were set to 1 for control conditions with a propagation of variability across all samples and replicates.

Puromycin incorporation assay

*CMTR2*WT or *CMTR2*KO HEK293T cells (1.8×10^6) were seeded in a 6-mm cell culture dish. The following day, cells were incubated in growing media containing 0.8 μ g ml⁻¹ puromycin for 0, 30, 60 and 90 min. Cells were washed twice with ice-cold PBS and collected by centrifugation at 300g for 5 min. Cell lysate preparation and western blotting were performed as described above. The PVDF membrane with transferred proteins was probed with anti-puromycin antibody for the detection of nascent proteins. Following anti-puromycin western blot, the membrane was washed and stained with Amido Black Staining solution (A8181, Sigma-Aldrich) to ensure equal protein loading in all lanes.

MTT cell proliferation assay

Changes in cell growth upon *CMTR2* depletion were tested using MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) cell proliferation assay. *CMTR2* WT or *CMTR2* KO HEK293T cells (2×10^4) were seeded in a single well of a 12-well plate. The MTT assay was performed on cells at 1, 2 and 3 days following the initial seeding. For each time point, cells were washed once with pre-warmed PBS, followed by incubation in the solution containing 1:1 mixture of phenol red-free DMEM and MTT reagent (5 mg ml⁻¹ MTT (ab 146345, Abcam) for 3 h at 37 °C. MTT solvent (1.5 \times volume) (4 mM HCl and 0.1% NP-40 in isopropanol) was added to the cells, and formazan crystals were dissolved by pipetting. Samples were incubated for 15 min at room temperature and 570-nm absorbance read to estimate the number of viable cells. A well without seeded cells was used for background subtraction.

VSV infections

HEK293T cells (3.5×10^5) were seeded in a poly-D-lysine (A3890401, Gibco) coated well of a six-well plate. The following day, cells were transfected with 800 ng pcDNA4.0/TO-NeonGreen, pcDNA4/TO-NeonGreen-*CMTR2* or pcDNA4/TO-NeonGreen-*CMTR2* W85A plasmids using 1.2 μ l LipoD293T transfection reagent. Twenty-four hours after transfection, cells were infected with 1.25×10^8 propagation-incompetent VSV particles. After 24 h of the VSV infection, cells were washed once with PBS and collected in 1 ml TRIzol reagent for total RNA extraction. Isolated total RNA was treated with DNase I, purified and subjected to CLAM-Cap-qPCR and RT-qPCR analysis.

Poly(I:C) and 3p-hpRNA cell treatments

HEK293T cells (3.5×10^5) were seeded in a well of a six-well plate. The following day, cells were transfected with 800 ng pcDNA4.0/TO-NeonGreen, pcDNA4/TO-NeonGreen-*CMTR2* or pcDNA4/TO-NeonGreen-*CMTR2* W85A plasmids using 1.2 μ l LipoD293T transfection reagent. Twenty-four hours after transfection, cells were transfected with either 500 ng LMW poly(I:C) (InvivoGen) or 750 ng 3p-hpRNA (InvivoGen) using LyoVec transfection reagent (InvivoGen) according to the manufacturer's instructions. After 24 h (for poly(I:C)) and 8 h (for 3p-hpRNA), cells were collected in 1 ml TRIzol reagent for total RNA extraction. Isolated total RNA was treated with DNase I, purified and subjected to RT-qPCR analysis.

Plasmids

The *CMTR2* open reading frame was obtained by RT-PCR on mRNA extracted from HEK293T cells. The *CMTR2* open reading frame was

Article

cloned into the pcDNA4/TO-NeonGreen plasmid between the KpnI and XbaI restriction sites to obtain the N-terminal NeonGreen-tagged CMTR2 expression construct. The pcDNA4/TO-NeonGreen-CMTR2 W85A mutant construct was generated by PCR-based site-directed mutagenesis of the pcDNA4/TO-NeonGreen-CMTR2 plasmid. The pcDNA3.1(+)-FLAG-RIG-I plasmid was purchased from OriGene (OHU25414). The pcDNA3.1(+)-FLAG-RIG-I K858A/K861A, H830A and C829A mutant constructs were generated by PCR-based site-directed mutagenesis of the original pcDNA3.1(+)-FLAG-RIG-I plasmid. All generated plasmids are available from the lead contact on request.

Preparation of the biotinylated dsRNA oligos for RIG-I pulldown

Single-stranded Cap0-modified, Cap1-modified and Cap2-modified RNA oligonucleotides were provided by TriLink Biotechnologies. To remove any residual non-capped RNA, the oligos were treated with Terminator exonuclease for 1 h at 30 °C and purified using RNAClean XP beads. A complementary 5'-biotinylated RNA oligo was annealed to Cap0-modified, Cap1-modified and Cap2-modified RNA oligos in a 1:1:1 ratio in 50 µl annealing buffer (10 mM Tris-HCl pH 7.5, 100 mM NaCl and 1 mM EDTA). Biotinylated Cap0-terminated, Cap1-terminated or Cap2-terminated dsRNA (700 ng) was incubated with M-280 streptavidin Dynabeads (40 µl bed volume) in 300 µl binding buffer supplemented 1 U µl⁻¹ RNaseOUT at room temperature for 30 min with constant rotation. The RNA-bound beads were washed three times with 500 µl binding buffer, resuspended in 150 µl cell lysis buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 5 mM MgCl₂, 0.5% NP-40, 40 U ml⁻¹ RNaseOUT and 1× Halt protease and phosphatase inhibitor cocktail), and kept on ice.

RIG-I pulldown assay

HEK293T cells (4.8 × 10⁶) were seeded on a 10-cm Petri dish. The following day, cells were transfected with 4.8 µg pcDNA3.1(+)-FLAG-RIG-I, pcDNA3.1(+)-FLAG-RIG-I K858A/K858A, pcDNA3.1(+)-FLAG-RIG-I H830A or pcDNA3.1(+)-FLAG-RIG-I C829A plasmids using 10 µl LipoD293 transfection reagent. Forty-eight hours after transfection, cells were washed once with 5 ml ice-cold PBS and scraped in 8 ml ice-cold PBS. Cell pellets were collected by 5 min of centrifugation at 300g at 4 °C, supernatant was removed by aspiration and cells were resuspended in 500 µl cell lysis buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 5 mM MgCl₂, 0.5% NP-40, 40 U ml⁻¹ RNaseOUT and 1× Halt protease and phosphatase inhibitor cocktail). Cells were left to lyse on ice for 10 min. To facilitate cell lysis, cells were passed three times through a 21-gauge needle and incubated on ice for an additional 10 min. The lysates were cleared by centrifugation at 16,000g for 10 min and the supernatants were saved. Protein concentration was measured by the Pierce BCA Protein Assay kit according to the manufacturer's instructions. Protein lysate (0.75 mg) at the concentration of 1 mg ml⁻¹ was incubated with streptavidin-immobilized Cap0, Cap1 and Cap2 dsRNA for 45 min at room temperature with agitation. Following the incubation, the beads were washed four times with 750 µl cell lysis buffer and separated in two equal parts. RNA-bound proteins were eluted in 25 µl 1.5× NuPAGE LDS sample buffer (NP0007, Invitrogen) from one bead fraction. The eluted proteins were loaded onto the NuPAGE 4–12% Bis-Tris gel and run at 180 V for 60 min. The resolved proteins were transferred onto the PVDF membrane and western blot was performed using anti-RIG-I antibody. To show equal amount of RNA bait across different samples, the remaining part of the beads was subjected to the proteinase K treatment for RNA isolation. Isolated RNA was loaded onto the 20% TBE non-denaturing PAGE and run at 150 V for 2.5 h. RNA was visualized after the gel was stained with 1× SYBR Gold nucleic acid gel stain (S11494, Invitrogen).

CMTR2 and RIG-I co-overexpression in CMTR2 WT HEK293T cells

HEK293T cells (3.5 × 10⁵) were seeded in a single poly-D-lysine (A3890401, Gibco) coated well of a six-well plate. The following day, cells were trans-

ected with 800 ng pcDNA4.0/TO-NeonGreen, pcDNA4/TO-NeonGreen-CMTR2 or pcDNA4/TO-NeonGreen-CMTR2 W85A plasmids using 1.2 µl LipoD293T transfection reagent. After 24 h, cells were transfected with either 1 µg pcDNA3.1(+)-FLAG-GFP or pcDNA3.1(+)-FLAG-RIG-I. After 48 h, cells were washed once with PBS, scraped and collected by 5 min centrifugation at 300g at 4 °C. Cell pellets were divided into equal parts for protein and total RNA extraction.

RIG-I overexpression in CMTR2 WT and CMTR2 KO HEK293T cells

CMTR2 WT or CMTR2 KO HEK293T cells (3.5 × 10⁵) were seeded in a single well of a six-well plate. The following day, cells were transfected with 1 µg pcDNA3.1(+)-FLAG-GFP, pcDNA3.1(+)-FLAG-RIG-I or pcDNA3.1(+)-FLAG-RIG-I K858A/K861A constructs. After 48 h, cells were washed once with PBS, scraped and collected by 5 min of centrifugation at 300g at 4 °C. Cell pellets were divided into equal parts for protein and total RNA extraction.

Oligonucleotides

The sequences of the oligonucleotides used in this study are provided in Supplementary Table 3.

Quantification and statistical analysis

Quantitative and statistical methods are described above and in the figure legends. R v4.0.1, GraphPad Prism v9.0.1 and ImageJ 1.53a were used for all statistical analysis and data visualization. Figures were prepared using Graphic v3.1 and Adobe Illustrator v27.0.1. All statistical tests and *P* values are provided in Supplementary Table 2. Experimental results shown as representative blots were successfully replicated two or more times to ensure the reproducibility of the reported findings.

Reporting summary

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

Data availability

Raw and processed sequencing data can be accessed from the NCBI Gene Expression Omnibus under the accession number GSE196043. In addition, TSS-seq, CLAM-Cap-seq, polysome-TSS-seq and actD-TSS-seq processed data are provided as Supplementary Table 1. The following public datasets were used in this study: GSE49831, GSE99978, GSE69352, GSE141507 and GSE52662. The detailed description of the data processing can be found in the Methods section and within GSE196043. Any additional information for the analyses of the data reported in this study is available from the corresponding author on request. Source data are provided with this paper.

43. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253–310 (1995).
44. Dodt, M., Roehr, J. T., Ahmed, R. & Dieterich, C. FLEXBAR-Flexible barcode and adapter processing for next-generation sequencing platforms. *Biology (Basel)* **1**, 895–905 (2012).
45. Webb, S., Hector, R. D., Kudla, G. & Granneman, S. PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. *Genome Biol.* **15**, R8 (2014).
46. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
47. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
48. Yates, A. D. et al. Ensembl 2020. *Nucleic Acids Res.* **48**, D682–D688 (2020).
49. Ritz, C., Baty, F., Streibig, J. C. & Gerhard, D. Dose-response analysis using R. *PLoS ONE* **10**, e0146021 (2015).
50. Wang, J., Vasaike, S., Shi, Z., Greer, M. & Zhang, B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* **45**, W130–W137 (2017).
51. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
52. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
53. Lizio, M. et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol.* **16**, 22 (2015).

54. Schueler, M. et al. Differential protein occupancy profiling of the mRNA transcriptome. *Genome Biol.* **15**, R15 (2014).
55. Batista, P. J. et al. m⁶A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* **15**, 707–719 (2014).
56. Weber, R. A. et al. Maintaining iron homeostasis is the key role of lysosomal acidity for cell proliferation. *Mol. Cell* **77**, 645–655.e7 (2020).
57. Floor, S. N. & Doudna, J. A. Tunable protein synthesis by transcript isoforms in human cells. *eLife* **5**, e10921 (2016).
58. Herzog, V. A. et al. Thiol-linked alkylation of RNA to assess expression dynamics. *Nat. Methods* **14**, 1198–1204 (2017).
59. Wei, J. et al. Genome-wide CRISPR screens reveal host factors critical for SARS-CoV-2 infection. *Cell* **184**, 76–91.e13 (2021).

Acknowledgements We thank A. Mirza for helpful suggestions on the development of the CapOligo-PAGE; members of the Jaffrey laboratory for comments and suggestions throughout the duration of this project; J. Dittman and J. Cao laboratories for providing us with worm and zebrafish samples, respectively; the R. Schwartz laboratory for propagation-incompetent VSV;

members of the Genomics Core Facility at Weill Cornell Medicine; and A. McCaffrey and his colleagues at TriLink Biotechnologies for providing us with chemically synthesized Cap1 and Cap2 luciferase mRNAs, as well as short Cap1 and Cap2 RNA oligonucleotides. This work is supported by US National Institutes of Health grants S10OD030335, R35NS111631, RM1HG011563 and MH121072 (to S.R.J.).

Author contributions V.D. and S.R.J. designed the study. V.D. performed the experiments, analysed the data and prepared the figures. V.D. and S.R.J. wrote the manuscript.

Competing interests S.R.J. is an advisor to and owns equity in 858 Therapeutics. V.D. declares no competing interests.

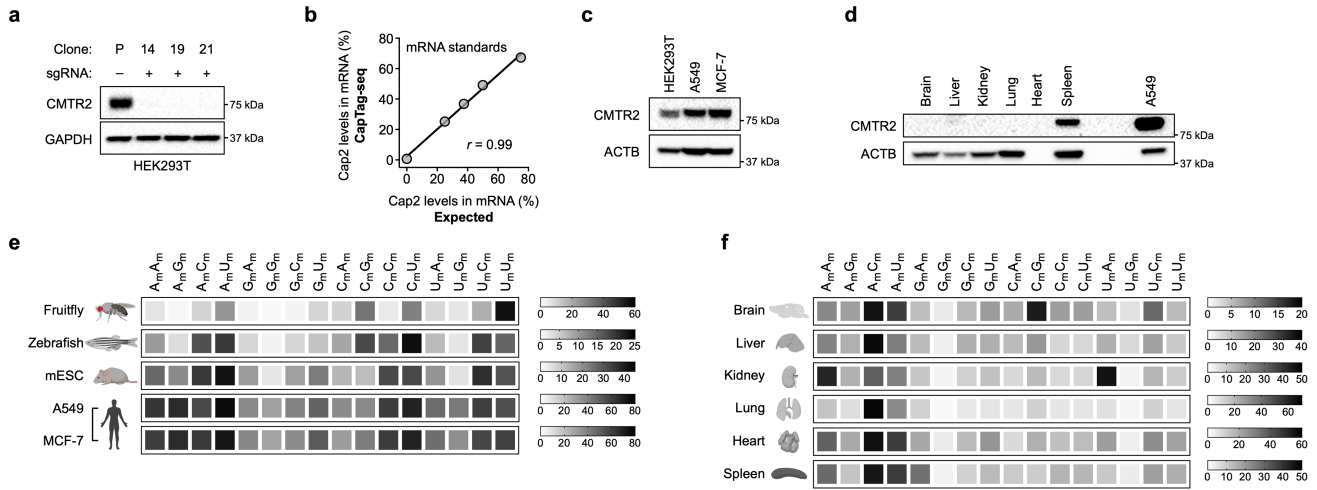
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-022-05668-z>.

Correspondence and requests for materials should be addressed to Samie R. Jaffrey.

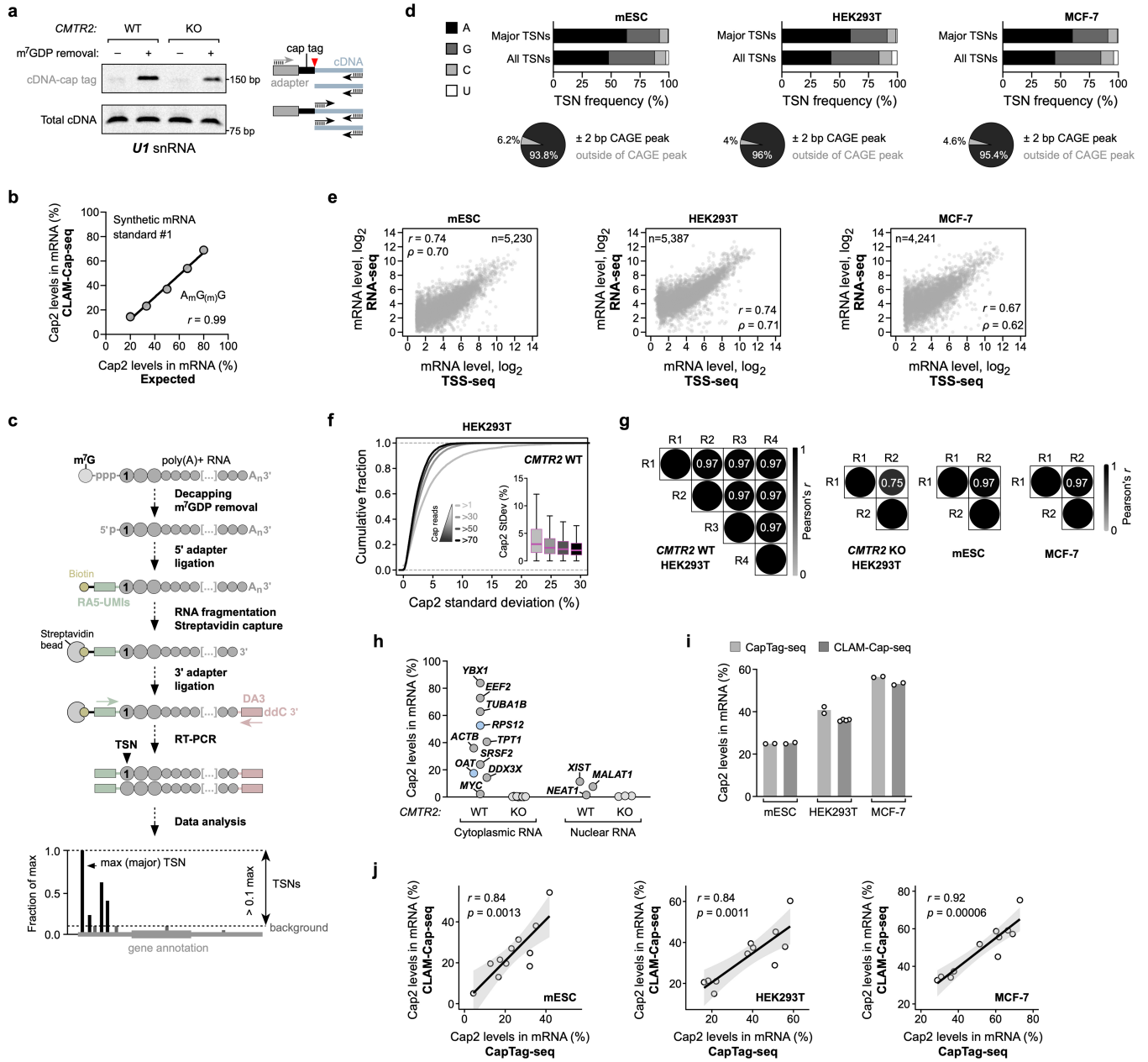
Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.

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



Extended Data Fig. 1 | Cap2 methylation residues in diverse sequence contexts. **a**, Representative western blot showing CMTR2 expression in the parental (P) and three *CMTR2* KO HEK293T cell clones (#14, #19, #21). GAPDH, loading control. *CMTR2* KO clone #21 was used for all experiments throughout the manuscript, unless otherwise indicated. sgRNA, single-guide RNA. **b**, CapTag-seq accurately measures Cap2 levels in bulk mRNA. CapTag-seq-determined Cap2 stoichiometry was plotted against the expected Cap2 stoichiometry of a predefined mixture of mRNAs from *CMTR2* KO and *CMTR2*-overexpressing cells in specific molar ratios. Pearson's (r), $n = 1$ experiment.

c, CMTR2 levels vary between human cell lines. Representative western blot showing CMTR2 expression in three human cell lines; HEK293T, A549 and MCF-7 cells. ACTB, loading control. **d**, CMTR2 is highly expressed in mouse spleen. Representative western blot showing CMTR2 expression in mouse organs. A549 cell lysate was used to identify mouse CMTR2 protein on the blot. ACTB, loading control. **e-f**, Heatmaps showing Cap2 levels across all 16 m⁷G-proximal dinucleotides in transcriptomes of different organisms, mammalian cell lines (**e**) and mouse tissues (**f**). Average of $n = 2$ biological replicates.



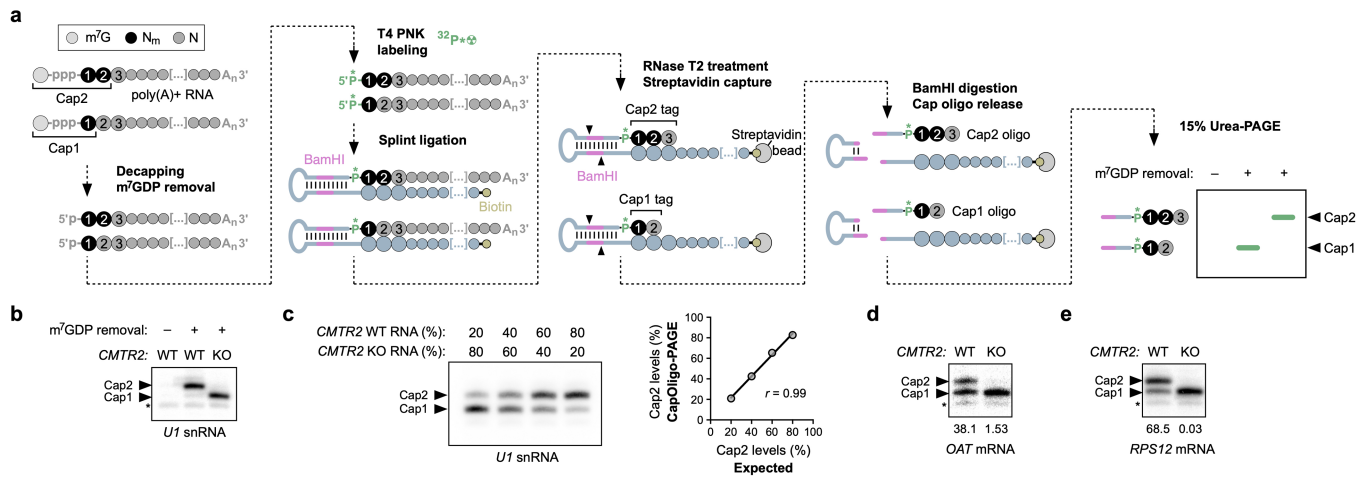
Extended Data Fig. 2 | See next page for caption.

Article

Extended Data Fig. 2 | CLAM-Cap-seq quantitatively measures Cap2

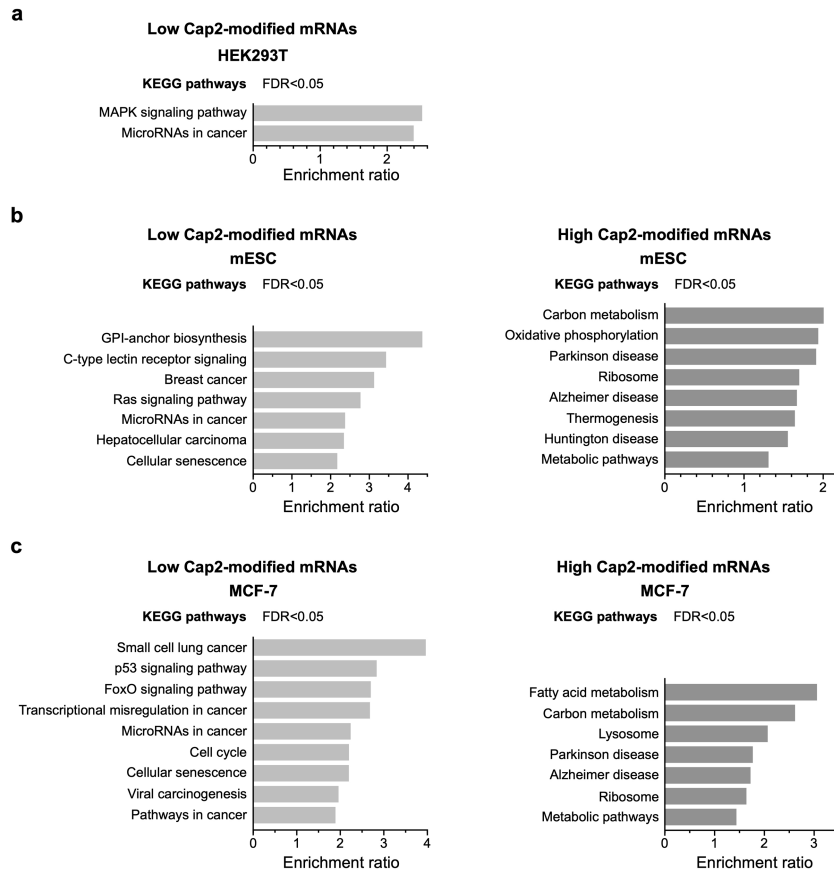
stoichiometry in individual mRNAs transcriptome-wide. **a**, CirLigase-assisted formation of cDNA-mRNA chimera requires m^7 GDP removal from the mRNA 5'-terminus. Fully Cap2- and Cap1-modified *U1* snRNAs from *CMTR2* WT and *CMTR2* KO HEK293T cells, respectively, were subjected to the CLAM-Cap-seq protocol. cDNA-cap tags were PCR-amplified with primers that surround cDNA-cap tag ligation junction (red arrow). PCR product representing cDNA-cap tag was detected only upon the enzymatic m^7 GDP removal from the *U1* snRNA 5'-end. The total amount of synthesized cDNA in all samples was determined by a parallel PCR using a primer pair that hybridizes to the internal portion of the U1 cDNA. Representative nondenaturing, 6% TBE gel pieces with the respective PCR products are shown. **b**, CLAM-Cap-seq accurately predicts Cap2 stoichiometry in mRNA. The CLAM-Cap-seq-determined Cap2 stoichiometry plotted against the expected Cap2 stoichiometry of a synthetic luciferase mRNA standard with m^7 G-proximal A_m-G_m-G sequence. The expected Cap2 stoichiometry refers to a predefined mixture of Cap1- and Cap2-terminated luciferase mRNAs in specific molar ratios. Pearson's (r), $n = 2$ technical replicates. **c**, TSS-seq, an approach for identification of the TSN (transcript-start nucleotide) mRNA isoforms. Poly(A)+ RNA is subjected to the enzymatic m^7 GDP removal (decapping). Following mRNA decapping, a biotinylated 5' RNA adapter with unique molecular identifiers (RA5-UMIs) is ligated to the first mRNA nucleotide. After a partial alkaline-based mRNA fragmentation, ligated mRNA 5'-ends are captured on streptavidin beads to perform a 3'-end DNA adapter ligation (DA3). Following reverse transcription, cDNA libraries are PCR amplified, sequenced, and sequencing data analyzed. After UMI removal from the 5'-end of sequencing reads, the first nucleotide of the UMI-free sequencing reads defines a TSN (a circle denoted with 1) of each mRNA in the transcriptome. A single nucleotide position in an annotated gene exhibiting the highest 5'-end read coverage is marked as the major, maximum TSN isoform. All other positions

within an annotated gene with 5'-end read coverage higher than 10% of the maximum TSN are considered as TSN isoforms produced from a gene. **d**, TSS-seq-identified TSN isoforms are consistent with previous TSN mapping methods. Stacked bar plots show the observed frequency of different TSNs in the indicated transcriptomes (A>G>C>>U). Pie charts below show the percent of TSNs from TSS-Seq found within and outside the established CAGE (cap analysis of gene expression) peaks⁵³. **e**, TSS-seq accurately measures mRNA expression. Correlation plots between TSS-seq-derived and RNA-seq-based calculations of mRNA expression in different cell lines⁵⁴⁻⁵⁶. Pearson's (r) and Spearman's (ρ). (**d-e**) Average of $n = 2$ for HEK293T and $n = 1$ for mES and MCF-7 cells. **f**, Cumulative distribution plot of the Cap2 stoichiometry variation between replicates (Cap2 standard deviation, StDev, %) for TSN mRNA isoforms binned based on the total number of CLAM-Cap-seq (Cap1 and Cap2) reads. $n = 4$ biological replicates. Each box shows the first quartile, median, and third quartile, whiskers represent $1.5 \times$ interquartile ranges. **g**, Reproducibility of the CLAM-Cap-seq. Pearson's correlation coefficient (r) of the CLAM-Cap-seq-predicated Cap2 stoichiometry across biological replicates was calculated and plotted for different cell lines. **h**, Cap2 stoichiometry for a selected set of cytoplasmic and nuclear poly(A)+ RNAs in *CMTR2* WT and *CMTR2* KO HEK293T cells was plotted. Cap2 stoichiometry of the major TSN isoform was considered. Blue data points indicate mRNAs whose Cap2 stoichiometry was biochemically validated (see Extended Data Fig. 3d,e). Average of $n = 4$ and $n = 2$ biological replicates for *CMTR2* WT and *CMTR2* KO HEK293T cells, respectively. **i**, Bar plot showing global Cap2 levels in cells based on CapTag-seq and CLAM-Cap-seq measurements. Average of $n = 2$ and $n = 4$ biological replicates. **j**, Correlation between CapTag-seq- and CLAM-Cap-seq-derived Cap2 measurements for m^7 G-proximal dinucleotides in three different mammalian cell lines. Pearson's (r), F test, grey area denotes 95% confidence interval for predictions from a linear model. Average of $n = 2$ and $n = 4$ biological replicates.



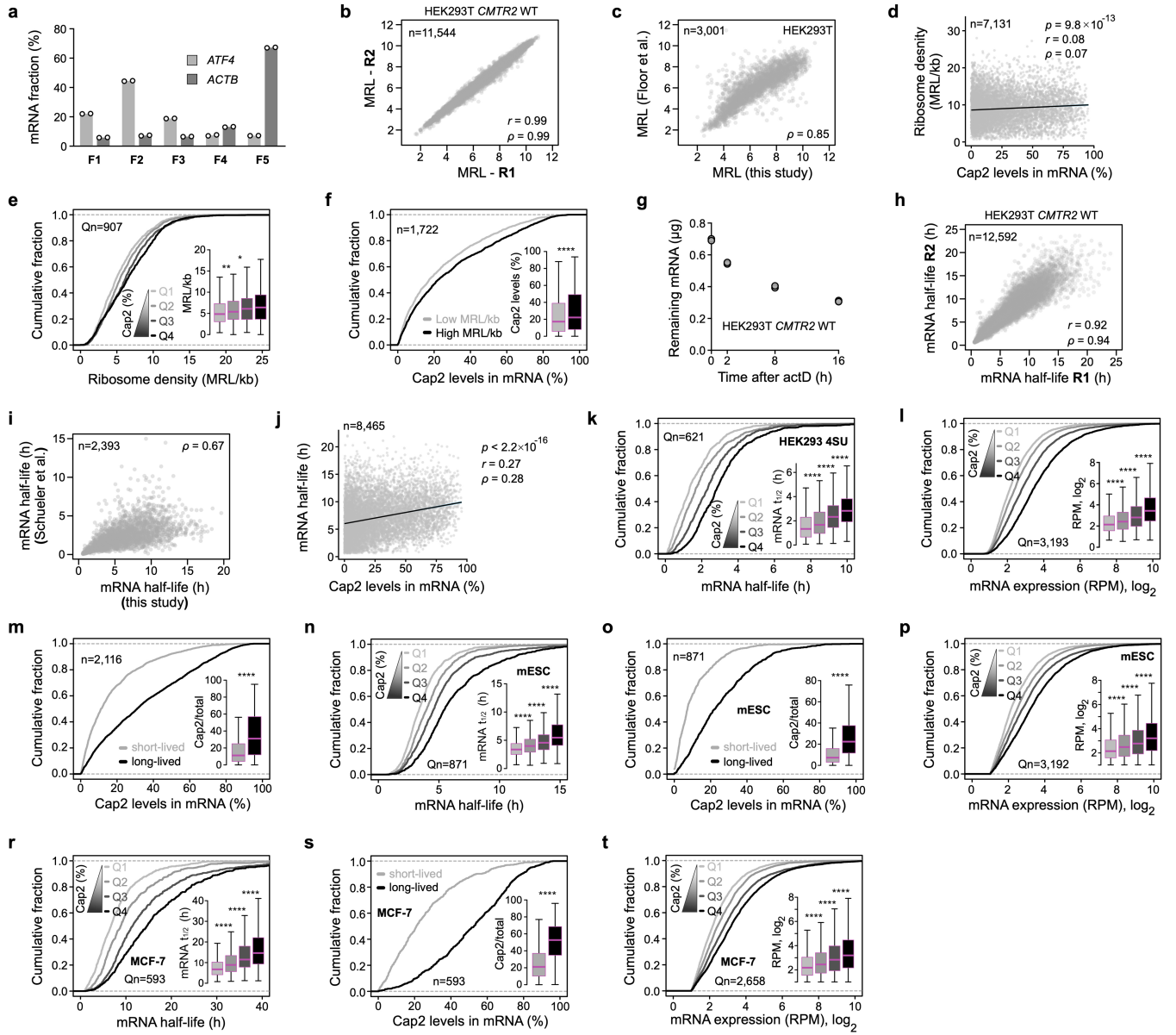
Extended Data Fig. 3 | A biochemical assessment of Cap2 stoichiometry in individual mRNAs with CapOligo-PAGE. a, CapOligo-PAGE, a biochemical assay for Cap2 measurements in individual transcripts. A mixture of Cap1- and Cap2-terminated mRNAs is subjected to enzymatic m⁷GDP removal, followed by T4 PNK 5'-end radiolabeling with [³²P]-ATP (green P with an asterisk). Next, a 5'-biotinylated self-splinting DNA oligonucleotide (blue) containing BamHI restriction site (pink) is ligated to the first radiolabeled mRNA nucleotide of a target transcript. Following splint ligation of the DNA oligonucleotide, an mRNA of interest is captured on streptavidin beads. A subsequent RNase-A and RNase T2-mediated mRNA removal creates self-splinting DNA-cap tag chimeras. BamHI cleaves the self-splinting DNA oligo, thus liberating radiolabeled Cap1 and Cap2 cap tags in a form of short oligonucleotides that differ by a single nucleotide in length. The released cap tag-containing DNA oligos are separated on the 15% sequencing TBE-Urea PAGE, transferred onto the nylon membrane, and visualized by autoradiography. **b**, CapOligo-PAGE specifically detects Cap2

and Cap1 in individual transcripts. Fully Cap2- and Cap1-modified *U1* snRNAs from *CMTR2* WT and *CMTR2* KO cells, respectively, were subjected to the CapOligo-PAGE protocol. Representative autoradiograph with resolved oligonucleotides containing radiolabeled cap tags is shown. Asterisk indicates a background band detected in all samples, including the sample in which the decapping step was omitted. **c**, CapOligo-PAGE accurately measures Cap2 stoichiometry in individual transcripts. Small RNA from *CMTR2* WT and *CMTR2* KO cells were mixed in predefined molar ratios and subjected to CapOligo-PAGE. Representative autoradiograph is shown. A plot shows linear regression of the expected and CapOligo-PAGE-derived Cap2 stoichiometries. Pearson's (r), $n = 1$ experiment. **d-e**, CapOligo-PAGE reveals Cap2 levels on cellular mRNAs. Representative autoradiographs showing Cap1 and Cap2 status of the **(d)** *OAT* and **(e)** *RPS12* mRNAs in *CMTR2* WT and *CMTR2* KO HEK293T cells. Asterisk, background band.



Extended Data Fig. 4 | Cap2 is enriched on a conserved set of specific mRNAs in mammalian cell lines. **a**, Gene enrichment analysis for lowly Cap2-modified mRNAs in HEK293T cells. **b-c**, Gene enrichment analysis for lowly and highly Cap2-modified mRNAs in **(b)** mESC and **(c)** MCF-7 cells. Bar graphs show the enrichment ratio of the KEGG pathways associated with highly

(dark grey) or lowly (light grey) Cap2-modified mRNAs relative to all mRNAs with measurable Cap2 stoichiometry. The enrichment analyses were performed using the overrepresentation analysis (ORA) method with a false discovery rate (FDR) threshold of 5%.

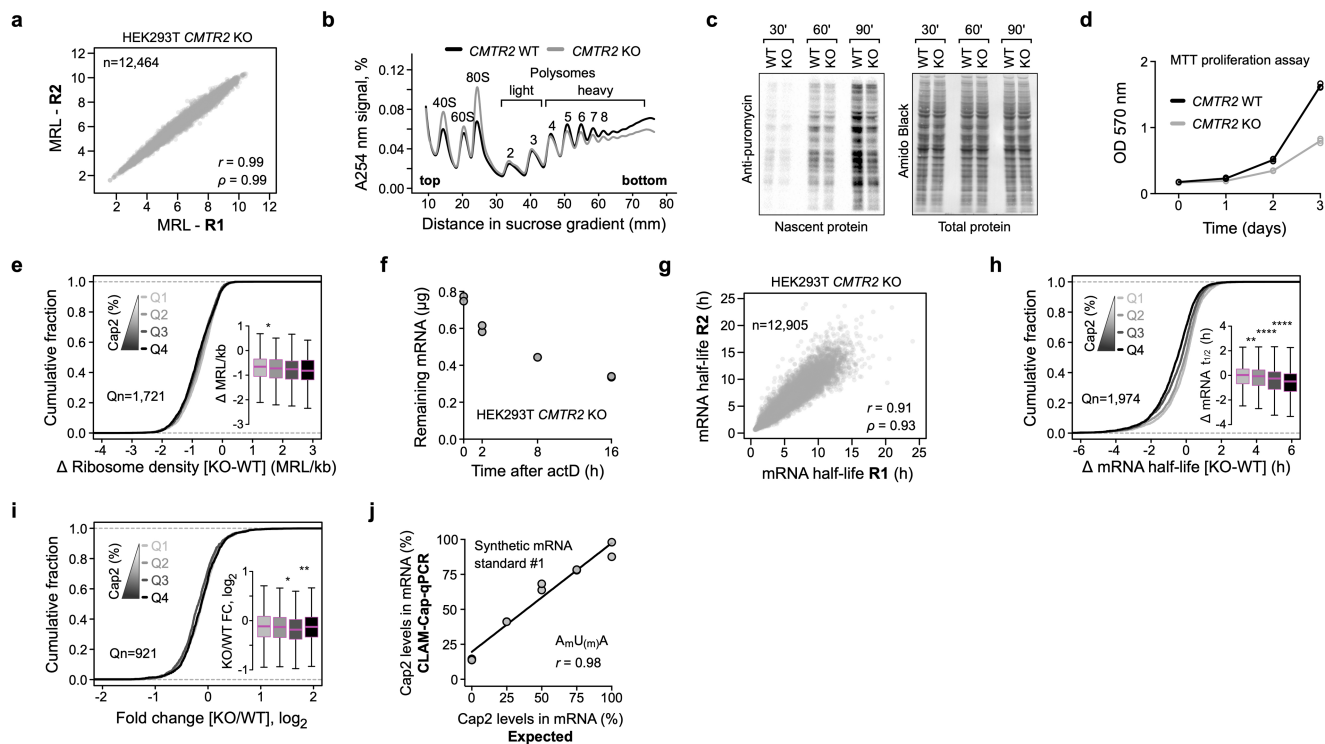


Extended Data Fig. 5 | See next page for caption.

Article

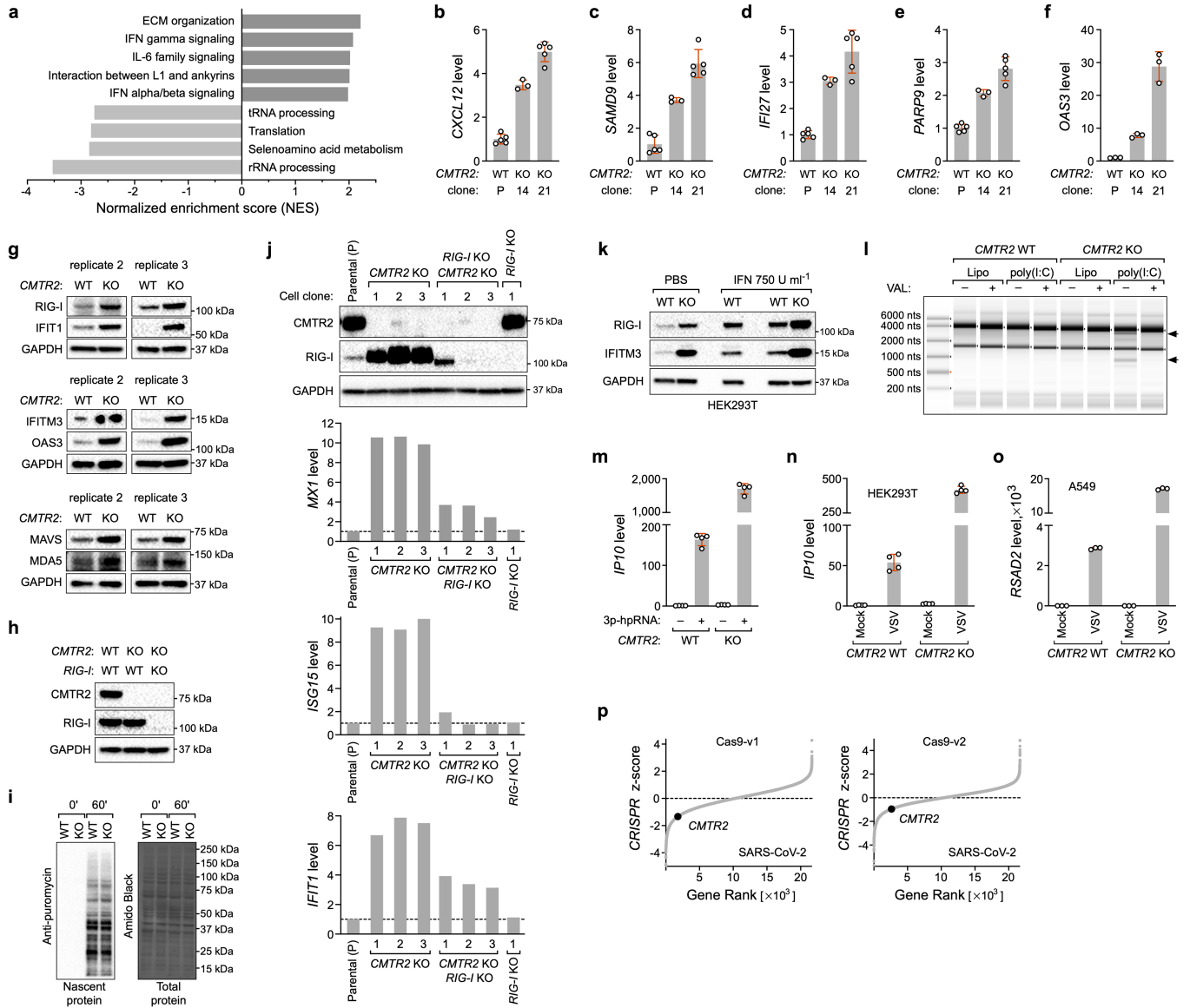
Extended Data Fig. 5 | Cap2 methylation is highly enriched on long-lived mRNAs. **a**, Bar graph showing the percent of *ACTB* and *ATF4* mRNAs found in each polysome fraction (F1-5) of the linear sucrose density gradient. Average of $n = 2$ biological experiments. **b**, Reproducibility of the translation measurements in *CMTR2* WT cells. The mean ribosome load (MRL) measurements of TSN isoforms from two independent polysome-TSS-seq experiments plotted against each other. Pearson's (r) and Spearman's (ρ). **c**, Polysome-TSS-seq-derived MRL measurements are consistent with the previously published mRNA translation dataset. An average MRL of all TSN isoforms for a gene plotted against gene-level MRL calculated from previously published polysome-RNA-seq HEK293T data⁵⁷. Spearman's (ρ). **d**, Cap2 methylation is poorly correlated with translation. Scatter plot showing correlation between Cap2 stoichiometry and ribosome density (MRL/kb) for each TSN isoform in HEK293T cells. Pearson's (r) and Spearman's (ρ). F test, light blue area denotes 95% confidence interval for predictions from a linear model. **e**, Cap2 methylation is enriched on highly translated mRNAs. Cumulative distribution plot of the ribosome densities (MRL/kb) for HEK293T-expressed mRNAs⁵⁷ binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. **f**, Highly translated mRNAs exhibit subtle Cap2-enrichment. Cumulative distribution plot of the Cap2 stoichiometry for the TSN isoforms binned into equally sized quartiles (Q) of increasing ribosome density (MRL/kb). Only the first and fourth quartiles were plotted. Average of $n = 2$ biological replicates. **g**, Global mRNA decay rate in *CMTR2* WT cells. The remaining amount of poly(A)+ RNA isolated from 20 μ g total RNA at different time points following transcription block by actinomycin D, actD, was plotted. $n = 2$ biological replicates. **h**, Reproducibility of the mRNA half-life measurements in *CMTR2* WT cells. TSN isoform half-lives derived from two independent actD-TSS-seq experiments plotted against each other. Pearson's (r) and Spearman's (ρ). **i**, ActD-TSS-seq-derived mRNA half-lives are consistent with previous mRNA half-life measurements. An average half-life of all TSN isoforms for a gene plotted against previously reported gene-level half-lives in HEK293 cells⁵⁴. Spearman's (ρ). **j**, Cap2 methylation is well-correlated with mRNA half-lives. Scatter plot showing the correlation between Cap2 stoichiometry and half-lives for each TSN isoform in HEK293T cells. Pearson's (r) and Spearman's (ρ). F test, light blue area denotes 95% confidence interval for predictions from a

linear model. **k**, Cap2 methylation is enriched on long-lived mRNAs. Cumulative distribution plot of the half-lives for HEK293-expressed mRNAs⁵⁴ binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. **l**, Cap2 methylation is enriched on highly abundant mRNAs in HEK293T cells. Cumulative distribution plot of the expression levels (RPM; reads per million, \log_2) for TSN isoforms in HEK293T cells binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. Average of $n = 2$ biological replicates. **m**, Long-lived mRNAs exhibit strong Cap2-enrichment in HEK293T cells. Cumulative distribution plot of the Cap2 stoichiometry for the TSN isoforms binned into equally sized quartiles (Q) of increasing half-lives. Only the first and fourth quartiles were plotted. Average of $n = 2$ biological replicates. **n**, Highly Cap2-modified mRNAs exhibit long half-lives in mESCs. Cumulative distribution plot of the half-lives for mESC-expressed mRNAs⁵⁸ binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. **o**, Long-lived mRNAs exhibit high Cap2 stoichiometry in mESCs. Cumulative distribution plot of the Cap2 stoichiometry for mESC-expressed mRNAs⁵⁸ binned into equally sized quartiles (Q) of increasing half-lives. Only the first and fourth quartiles were plotted. **p**, Cap2 methylation is enriched on highly abundant mRNAs in mESCs. Cumulative distribution plot of the expression levels (RPM; reads per million, \log_2) for TSN isoforms in mESCs binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. $n = 1$ experiment. **r**, Highly Cap2-modified mRNAs exhibit long half-lives in MCF-7 cells. Cumulative distribution plot of the half-lives for MCF-7-expressed mRNAs⁵⁴ binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. **s**, Long-lived mRNAs exhibit high Cap2 stoichiometry in MCF-7 cells. Cumulative distribution plot of the Cap2 stoichiometry for MCF-7-expressed mRNAs⁵⁴ binned into equally sized quartiles (Q) of increasing half-lives. Only the first and fourth quartiles were plotted. **t**, Cap2 methylation is enriched on highly abundant mRNAs in MCF-7 cells. Cumulative distribution plot of the expression levels (RPM; reads per million, \log_2) for TSN isoforms in MCF-7 cells binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. $n = 1$ experiment. **(e-f, k-t)** Each box shows the first quartile, median, and third quartile, whiskers represent 1.5 \times interquartile ranges. Two-sided Wilcoxon signed-rank test; * $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$, **** $p < 0.00001$.



Extended Data Fig. 6 | Cap2 methylation does not confer high translation and stability to mRNA. **a**, Reproducibility of the MRL measurements in *CMTR2* KO HEK293T cells. TSN isoform MRLs derived from two independent polysome-TSS-seq experiments plotted against each other. Pearson's (r) and Spearman's (ρ). **b**, A loss of Cap2 causes global reduction in translation in HEK293T cells. Polysome profiles of *CMTR2* WT and *CMTR2* KO HEK293T cells. The percent of the RNA signal (A254 nm, absorbance at 254 nm) observed at different positions across the linear (10-50%) sucrose gradient was plotted. $n = 1$ experiment. **c**, A loss of Cap2 results in the reduced synthesis rate of nascent proteins in HEK293T cells. *CMTR2* WT and *CMTR2* KO HEK293T cells were incubated in the growing media containing low levels of puromycin ($0.8 \mu\text{g ml}^{-1}$) for the indicated time periods to label nascent proteins. The left panel represents the western blot showing levels of puromycin-labeled proteins synthesized in the respective cell lines. The right panel shows the amido black-stained membrane from the left panel to indicate equal protein loading in all lanes. **d**, *CMTR2* KO cells exhibit slower growth rate. A plot showing *CMTR2* WT and *CMTR2* KO cell growth rate over time. The y-axis indicates the absorbance signal at 570 nm that is directly proportional to the number of viable, growing cells. $n = 3$ biological replicates. **e**, Cap2 loss does not lead to selective change in the translation of highly Cap2-modified transcripts. Cumulative distribution plot of changes in the ribosome density (MRL/kb) between *CMTR2* KO and *CMTR2* WT HEK293T cells for TSN isoforms binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. Average of $n = 2$ biological replicates. **f**, Global mRNA decay rate in *CMTR2* KO cells. The remaining amounts of poly(A)+ RNA isolated from 20 μg total RNA at different time points following transcriptional block by actD were plotted. $n = 2$ biological replicates. **g**, Reproducibility of the mRNA half-life measurements in *CMTR2* KO HEK293T cells. TSN isoform half-lives derived from two independent actD-TSS-seq experiments plotted against each other. Pearson's (r) and Spearman's (ρ). **h**, Cap2 loss leads to mild, but selective change in the half-lives of highly Cap2-modified transcripts. Cumulative distribution plot of changes in the half-lives between *CMTR2* KO and *CMTR2* WT HEK293T cells for TSN isoforms binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. Average of $n = 2$ biological replicates. **i**, Cap2 loss does not lead to selective change in the expression of highly Cap2-modified transcripts. Cumulative distribution plot of the RNA-seq-determined fold changes in gene expression between *CMTR2* KO and *CMTR2* WT HEK293T cells for genes binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. $n = 4$ biological replicates. **j**, CLAM-Cap-qPCR accurately detects differences in Cap2 levels in mRNA. CLAM-Cap-qPCR-determined Cap2 levels were plotted against the expected Cap2 stoichiometry of a synthetic luciferase mRNA standard. The expected Cap2 stoichiometry refers to Cap1- and Cap2-terminated luciferase mRNAs mixed in specific molar ratios. A linear regression line with the Pearson's (r). $n = 2$ technical replicates. (**e**, **h**, **i**) Each box shows the first quartile, median, and third quartile, whiskers represent $1.5 \times$ interquartile ranges. Two-sided Wilcoxon signed-rank test; * $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$, **** $p < 0.00001$.

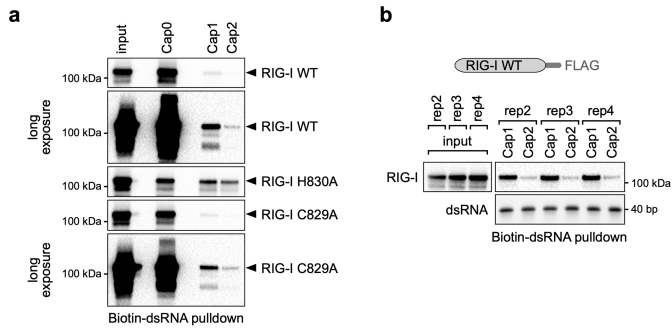
f, Global mRNA decay rate in *CMTR2* KO cells. The remaining amounts of poly(A)+ RNA isolated from 20 μg total RNA at different time points following transcriptional block by actD were plotted. $n = 2$ biological replicates. **g**, Reproducibility of the mRNA half-life measurements in *CMTR2* KO HEK293T cells. TSN isoform half-lives derived from two independent actD-TSS-seq experiments plotted against each other. Pearson's (r) and Spearman's (ρ). **h**, Cap2 loss leads to mild, but selective change in the half-lives of highly Cap2-modified transcripts. Cumulative distribution plot of changes in the half-lives between *CMTR2* KO and *CMTR2* WT HEK293T cells for TSN isoforms binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. Average of $n = 2$ biological replicates. **i**, Cap2 loss does not lead to selective change in the expression of highly Cap2-modified transcripts. Cumulative distribution plot of the RNA-seq-determined fold changes in gene expression between *CMTR2* KO and *CMTR2* WT HEK293T cells for genes binned into equally sized quartiles (Q) of increasing Cap2 stoichiometry. $n = 4$ biological replicates. **j**, CLAM-Cap-qPCR accurately detects differences in Cap2 levels in mRNA. CLAM-Cap-qPCR-determined Cap2 levels were plotted against the expected Cap2 stoichiometry of a synthetic luciferase mRNA standard. The expected Cap2 stoichiometry refers to Cap1- and Cap2-terminated luciferase mRNAs mixed in specific molar ratios. A linear regression line with the Pearson's (r). $n = 2$ technical replicates. (**e**, **h**, **i**) Each box shows the first quartile, median, and third quartile, whiskers represent $1.5 \times$ interquartile ranges. Two-sided Wilcoxon signed-rank test; * $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$, **** $p < 0.00001$.



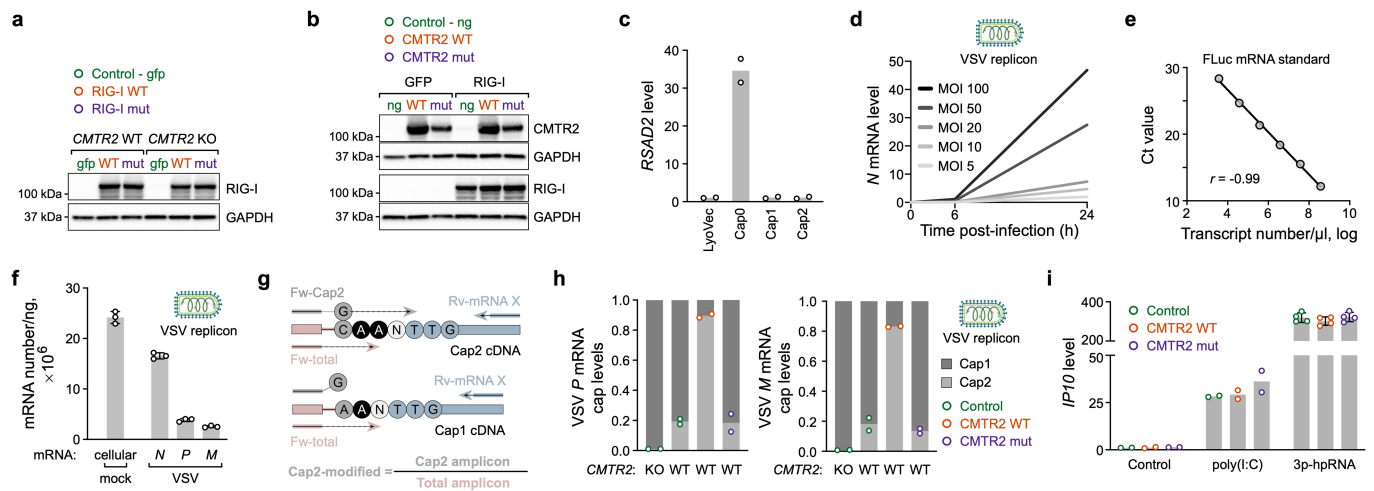
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Loss of Cap2 causes activation of the innate immune response. **a**, Cap2 loss causes the upregulation of genes related to the innate immune response and inflammatory pathways. Bar graph shows the normalized enrichment score (NES) of the Reactome pathways for differentially regulated genes in *CMTR2*KO HEK293T cells. Dark grey, upregulated genes. Light grey, downregulated genes. The analysis was conducted using the gene set enrichment analysis (GSEA) method with a false discovery rate (FDR) threshold of 5%. **b-f**, Cap2 depletion causes the upregulation of the interferon-stimulated genes (ISGs). *GAPDH*-normalized expression level for the selected set of ISGs measured in *CMTR2*WT and *CMTR2*KO HEK293T cells by RT-qPCR. n = 3 and n = 5 biological replicates, mean ± s.d. **g**, Prevention of Cap2 formation induces the upregulation of antiviral proteins. Western blots showing the expression levels of the selected antiviral proteins in *CMTR2*WT and *CMTR2*KO HEK293T cells. *GAPDH*, loading control. **h**, Generation of the *CMTR2*KO and *CMTR2*, *RIG-I* double KO A549 cell lines. Representative western blot showing *CMTR2* and *RIG-I* expression in WT, *CMTR2*KO and *CMTR2*, *RIG-I* double KO A549 cells. *GAPDH*, loading control. **i**, A loss of Cap2 has no effect on the synthesis rate of nascent proteins in A549 cells. *CMTR2*WT and *CMTR2*KO A549 cells were incubated in the growing media containing low levels of puromycin (1 µg ml⁻¹) for 90 min to label nascent proteins. The left panel represents a western blot showing levels of puromycin-labeled proteins synthesized in the respective cell lines. The right panel shows the amido black-stained membrane from the left panel to indicate equal protein loading in all lanes. **j**, *CMTR2* loss causes the induction of the innate immune response in A549 cells across different cell clones in a *RIG-I*-dependent manner. Western blots showing *CMTR2* and *RIG-I* expression levels in multiple knockout cell clones. *GAPDH*, loading control. Bar plots below show *GAPDH*-normalized expression of *MX1*, *ISG15* and *IFIT1* mRNAs determined using RT-qPCR. RNA samples were extracted from the respective clones at the time of the western blot screens in search for knockout cell clones. n = 1 experiment. *CMTR2*KO clone #1, *CMTR2*, *RIG-I* double KO clone #3

were selected for all downstream experiments. Note that the magnitude of the ISG induction decreases in the selected clones over several cell passages (Fig. 5c and Extended Data Fig. 7h), indicating A549 cell adaptation to chronic inflammation caused by permanent *CMTR2* depletion. **k**, Antiviral protein induction caused by *CMTR2* depletion is comparable to interferon (IFN) treatment in HEK293T cells. Representative western blot shows *RIG-I* and *IFITM3* expression in *CMTR2*WT and *CMTR2*KO cells treated with PBS or IFN for 6 h at 750 U ml⁻¹. *GAPDH*, loading control. **l**, Poly(I:C) treatment of *CMTR2*KO cells induces RNase L activation. Representative Bioanalyzer trace of total RNA extracted from *CMTR2*WT and *CMTR2*KO HEK293T cells pretreated with RNase L inhibitor (valoneic acid dilactone, VAL) followed by mock (Lipo) or poly(I:C) transfection for 6 h. rRNA cleavage products appeared only in *CMTR2*KO cells transfected with poly(I:C). rRNA cleavage products disappeared when cells were pretreated with RNase L inhibitor. **m**, *CMTR2*KO HEK293T cells are sensitive to treatment with triphosphorylated hairpin RNA (3p-hpRNA). *GAPDH*-normalized *IP10* transcript levels in *CMTR2*WT and *CMTR2*KO HEK293T cells transfected with 3p-hpRNA for 8 h. n = 4 biological replicates, mean ± s.d. **n**, *CMTR2*KO HEK293T cells are sensitive to viral infection. *GAPDH*-normalized *IP10* transcript levels in *CMTR2*WT and *CMTR2*KO HEK293T cells 24 h after the infection with the propagation-incompetent VSV replicon. n = 4 biological replicates, mean ± s.d. **o**, *CMTR2*KO A549 cells are sensitive to viral infection. *GAPDH*-normalized *RSAD2* transcript levels in *CMTR2*WT and *CMTR2*KO A549 cells 24 h after the infection with the propagation-incompetent VSV replicon. n = 3 biological replicates, mean ± s.d. **p**, *CMTR2* depletion sensitizes cells to SARS-CoV-2 infection. Plots showing the gene rank with respect to the effect of gene knockout (CRISPR z-score) on cell growth upon SARS-CoV-2 infection⁵⁹. Negative CRISPR z-score indicates negative effect of SARS-CoV-2 infection on cell growth, whereas positive score denotes positive effect of SARS-CoV-2 infection on cell growth upon the knockout of each gene.

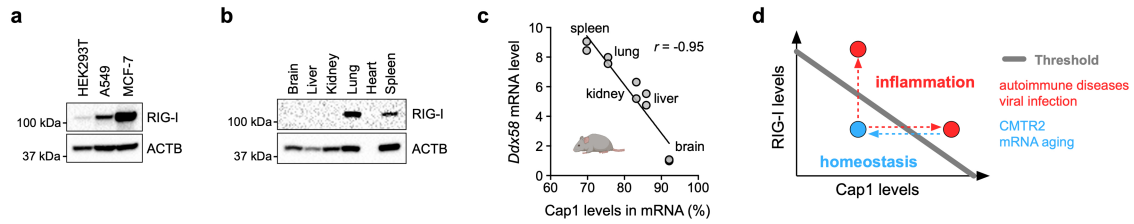


Extended Data Fig. 8 | Cap2 methylation inhibits RIG-I binding to capped RNA. **a**, Equal amounts of biotinylated Cap0, Cap1, and Cap2 double-stranded, dsRNAs were immobilized on streptavidin beads and incubated with HEK293T cell lysates expressing FLAG-tagged RIG-I wild-type (RIG-I WT), RIG-I H830A and RIG-I C829A mutant proteins. Representative western blots show the amount of the respective RIG-I proteins recovered by each capped RNA bait. **b**, Western blot showing RIG-I levels from three independent pulldowns with biotinylated Cap1 and Cap2 dsRNAs. A SYBR Gold-stained nondenaturing 20% TBE gel below shows equal amounts of streptavidin-immobilized dsRNA baits used in each pulldown experiment.



Extended Data Fig. 9 | Cap2 methylation suppresses RIG-I activation and VSV-induced innate immune response. **a**, Representative western blot showing RIG-I levels in *CMTR2* WT and *CMTR2* KO HEK293T cells transfected with constructs encoding FLAG-tagged GFP (Control - gfp), RIG-I wild-type (RIG-I WT) and RIG-I K858A/K861A mutant (RIG-I mut) proteins. GAPDH, loading control. **b**, Representative western blots showing CMTR2 and RIG-I levels in cells transfected with constructs encoding either FLAG-tagged GFP or RIG-I wild-type proteins, and plasmids expressing NeonGreen (Control - ng), NeonGreen-tagged CMTR2 wild-type (*CMTR2* WT) or *CMTR2* W85A (*CMTR2* mut) proteins. GAPDH, loading control. **c**, Cap0 RNA transfection results in the induction of the innate immune response. *GAPDH*-normalized *RSAD2* mRNA levels in wild-type A549 cells transfected with transfection reagent only (LyoVec), and equal amounts (1.5 micrograms) of Cap0, Cap1, and Cap2 dsRNA were plotted. Average of $n = 2$ biological replicates. **d**, VSV RNA efficiently replicates in HEK293T cells. Line plot showing RT-qPCR-measured VSV RNA levels 0, 6 and 24 h of VSV infection at the indicated multiplicity of infection (MOI). *GAPDH*-normalized VSV *N* mRNA levels were plotted. $n = 1$ experiment. **e**, Correlation plot of the predefined number of the firefly luciferase (FLuc) transcripts and RT-qPCR-derived Ct values. To obtain FLuc mRNA standard curve, a $10\times$ serial dilutions of FLuc mRNA standard were added to the fixed amount of total RNA extracted from HEK293T cells and quantified by RT-qPCR. Pearson's (r). $n = 2$ technical replicates. **f**, Viral RNA accumulates at high levels in VSV-infected cells. Bar plot showing the absolute number of *N*, *P*, and *M* VSV mRNAs per ng total RNA. Absolute number of viral transcripts was determined using the FLuc mRNA standard curve from Extended Data Fig. 9e. The number of endogenous mRNAs in mock-infected cells was determined based on the three independent measurements of oligo(dT)-isolated mRNA from fixed amount of total RNA.

The calculations of the absolute mRNA numbers were based on the average size of mRNA in human cells of 2.5 kb. $n = 3$ biological replicates, mean \pm s.d. **g**, A schematic of the Cap2 methylation measurement in VSV mRNAs by CLAM-Cap-qPCR. VSV mRNAs start with $A_m A_m C m^7 G$ -proximal sequence. Thus, after the creation of the DNA adapter-ligated cDNA-cap tags, VSV mRNAs Cap2 levels were measured by qPCR using a *N*, *P*, or *M* mRNA-specific reverse primer (Rv-mRNA X, blue arrow) and a forward primer that hybridizes to the DNA adapter and the first C nucleotide of the CAA three-nucleotide cap tag of the Cap2 form of VSV mRNAs (Fw-Cap2, grey arrow). Total abundance of (Cap1 and Cap2) cDNA-cap tags was determined via parallel qPCR using a primer pair comprising a VSV mRNA-specific reverse primer and a primer that only anneals to the DNA adapter and not to any portion of the VSV mRNA cap tags (Fw-total, pink arrow). **h**, *CMTR2* overexpression leads to increased Cap2 levels in viral *P* and *M* mRNAs. *CMTR2* KO HEK293T cells expressing Control, and *CMTR2* WT HEK293T cells expressing Control, *CMTR2* WT or *CMTR2* mut were infected with propagation-incompetent VSV replicon. Twenty-four h post-infection, RNA was collected, and CLAM-Cap-qPCR was conducted on viral *P* and *M* mRNAs. Stacked bar plot shows the fraction of Cap1- and Cap2-modified VSV *P* and *M* mRNAs. Cap1-modified fraction was inferred from Cap2 measurements as $1 - \text{Cap2}$ fraction. Average of $n = 2$ biological replicates. **i**, *CMTR2* overexpression does not suppress the innate immune response triggered by non-capped RNA ligands. Bar plot showing *IP10* mRNA levels in poly(I:C)- and 3p-hpRNA-stimulated HEK293T cells expressing Control, *CMTR2* WT or *CMTR2* mut proteins. *GAPDH*-normalized *IP10* mRNA expression levels. Average of $n = 2$ biological replicates for Control and 3p-hpRNA treatments. $n = 4$ biological replicates for poly(I:C) treatment, mean \pm s.d.



Extended Data Fig. 10 | Cap1 RNA levels anticorrelate with RIG-I expression.

a, RIG-I levels vary between human cell lines. Representative western blot showing RIG-I expression in three human cell lines; HEK293T, A549 and MCF-7 cells. ACTB, loading control. **b**, RIG-I levels vary between mouse organs. Representative western blot showing RIG-I expression in mouse organs. ACTB, loading control. **c**, *Ddx58* mRNA expression anticorrelates with Cap1 levels in mouse organs. Scatter plot showing strong negative correlation between Cap1 levels and *Ddx58* mRNA expression level determined via CapTag-seq and RT-qPCR, respectively. *Ddx58* mRNA expression was normalized to the expression of *Eef2* housekeeping mRNA. $n = 2$ biological replicates. *Ddx58* mRNA encodes for RIG-I protein. **d**, Schematic of how RIG-I and cap methylation

regulate the activation of the innate immune response. In this model, cellular homeostasis is achieved by balancing Cap1 and RIG-I levels. Regulation of Cap2 methylation via CMTR2 expression and mRNA ageing contributes to the establishment of this balance, allowing cell homeostasis. Low levels of Cap1 can induce RIG-I activation only if RIG-I levels are high. However, when Cap1 levels are elevated, even small amounts of RIG-I are sufficient to mediate Cap1 effects and lead to inflammation. Notably, Cap1 levels can increase with CMTR2 depletion or when large amounts of new viral RNA are synthesized, leading to activation of RIG-I. Additionally, situations associated with increased RIG-I expression, such as viral infection and autoimmune diseases, may further sensitize cells to Cap1 RNAs of the viral or host origins.

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our [Editorial Policies](#) 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.

n/a Confirmed

- 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

Data analysis https://github.com/lh3/seqtk), Bowtie v1.2.3, BEDTools v2.28.0., STAR aligner v2.7.
For further data processing and quantification: GraphPad Prism v9.0.1, R packages (edgeR v3.36.0), R packages (drc v3.0-1), WebGestalt (<http://www.webgestalt.org/>).

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All raw and processed data generated in this study are available at the NCBI Gene Expression Omnibus (GEO) with GSE196043 accession number. Published datasets used in this study are also available at the NCBI GEO with the following accession numbers: GSE69352, GSE49831, GSE99978, GSE141507, GSE52662. Genome build: human (hg38), mouse (mm10), Drosophila (dm6).

FANTOM5 CAGE-seq:

human: https://fantom.gsc.riken.jp/5/datafiles/reprocessed/hg38_latest/extra/CAGE_peaks/mouse: https://fantom.gsc.riken.jp/5/datafiles/reprocessed/mm10_latest/extra/CAGE_peaks/

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.

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 | Sample size was chosen to ensure the reproducibility of the results. All experiments reported in this study were performed at least twice, but often three or more times unless otherwise indicated. |
| Data exclusions | No data were excluded. |
| Replication | All attempts to replicate experiments in this study were successful. Sequencing experiments were performed in two replicates at minimum, with the exception of TSS-seq dataset from mES and MCF-7 cells. The validity of the TSS-seq datasets was confirmed by comparison to the FANTOM5 CAGE-seq datasets. Replicates from sequencing experiments were in good agreement for each reported method in this study. Sequencing-based findings in this study were routinely cross-validated and replicated using different, orthogonal approaches. |
| Randomization | The random allocation is not relevant to our study; all experiments were performed under controlled experimental conditions. |
| Blinding | The investigators were not blinded. The experiments were performed and the results were analyzed by the experimenter. |

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 |
|-------------------------------------|---|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Antibodies |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

| n/a | Involved in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Antibodies

Antibodies used

Rabbit anti-FTSJD1 polyclonal antibody (Invitrogen, PAS-61696); 1:500 dilution in 3% milk-TBST
 Rabbit anti-RIG-I (D14G6) monoclonal antibody (Cell Signaling, 3743T); 1:1,000 in 3% BSA-TBST
 Rabbit RIG-I/DDX58 [EPR18629] monoclonal antibody (Abcam, ab180675); 1:2,000 in 3% milk-TBST
 Rabbit anti-MDA-5 (D74E4) monoclonal antibody (Cell Signaling Technology, 5321T); 1:1,000 in 3% BSA-TBST
 Rabbit anti-IFIT1 (D2X9Z) monoclonal antibody (Cell Signaling Technology, 14769S); 1:1,000 in 3% milk in TBST
 Rabbit anti-IFITM3 (D8E8G) XP monoclonal antibody (Cell Signaling Technology, 59212T); 1:1,000 in 3% milk-TBST
 Mouse anti-NLRP1 monoclonal antibody (Biolegend, 679802); 1:1,000 in 3% BSA-TBST
 Rabbit anti-MAVS polyclonal antibody (Cell Signaling Technology, 3993T); 1:1,000 in 3% milk-TBST
 Rabbit anti-OAS3 polyclonal antibody (Abcam, ab154270); 1:1,000 in 3% milk-TBST
 Mouse anti-GAPDH monoclonal antibody (GeneTex, GT239); 1:10,000 in 3% milk-TBST
 Mouse anti-Beta Actin monoclonal antibody (Thermo, AM4302); 1:4,000 in 3% milk-TBST
 anti-Bromodeoxyuridine antibody (MBL Life Science, MI-11-3); for IP 0.5-1 micrograms antibody per 10 micrograms RNA
 Mouse anti-puromycin (12D10) monoclonal antibody (Millipore Sigma, MABE343); 1:5,000 in 3% milk-TBST
 HRP-conjugated donkey anti-rabbit IgG (Cytiva, NA934); 1:5,000 in 3% milk-TBST or 3% BSA-TBST
 HRP-conjugated sheep anti-mouse IgG (Cytiva, NA931); 1:5,000 in 3% milk-TBST or 3% BSA-TBST

Validation

All antibodies used in this study were compared to previously published results and antibody manufacturers data.

Validation

Rabbit anti-FTSJD1 polyclonal antibody for CMTR2 was validated by western blots on protein samples extracted from CMTR2 knockout cells (Extended Data Figures 1&7), as well as through an overexpression of the NeonGreen-tagged CMTR2 (Extended Data Figure 9).

Antibody validation source for primary antibodies used in this study:

<https://www.cellsignal.com/products/primary-antibodies/rig-i-d14g6-rabbit-mab/3743>

<https://www.abcam.com/rig-iddx58-antibody-epr18629-ab180675.html>

<https://www.cellsignal.com/products/primary-antibodies/mda-5-d74e4-rabbit-mab/5321>

<https://www.cellsignal.com/products/primary-antibodies/ifu1-d2x9z-rabbit-mab/14769>

<https://www.cellsignal.com/products/primary-antibodies/ifu3-d8e8g-xp-rabbit-mab/59212>

<https://www.biolegend.com/fr-fr/products/purified-anti-nlrp1-antibody-12213>

<https://www.cellsignal.com/products/primary-antibodies/mavs-antibody/3993>

<https://www.abcam.com/oas3-antibody-ab154270.html>

<https://www.genetex.com/Product/Detail/GAPDH-antibody-GT239/GTX627408>

<https://www.thermofisher.com/antibody/product/beta-Actin-Antibody-clone-AC-15-Monoclonal/AM4302>

<https://www.mblbio.com/bio/g/dt/A/index.html?pcd=MI-11-3>

https://www.emdmillipore.com/US/en/product/Anti-Puromycin-Antibody-clone-12D10,MM_NF-MABE343

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

Mouse embryonic stem cells were a gift from J. Hanna and S. Geula (Weizmann Institute of Science). HEK293T (ATCC CRL-3216), MCF-7 (ATCC HTB-22), A549 (ATCC CRM-CCL-185).

Authentication

mES cells were previously described in Geula, et. al. 2015. HEK293T, MCF-7, and A549 cells were purchased from ATCC. The authors did not authenticate cell lines.

Mycoplasma contamination

All cell lines used in this study tested negative for mycoplasma contamination. Mycoplasma contamination was routinely tested with Hoechst staining.

Commonly misidentified lines (See [ICLAC](#) register)

None of the ICLAC cell lines were used.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

C. elegans, adult; Zebrafish, AB line, 48 h old larvae; Mice, C57BL/6, female, 16 weeks old

Wild animals

This study did not involve wild animals.

Field-collected samples

This study did not involve field-collected samples.

Ethics oversight

The mice and zebrafish species were maintained in compliance with Weill Cornell Medicine Institutional Animal Care and Use Committee (IACUC) protocols.

Note that full information on the approval of the study protocol must also be provided in the manuscript.