Exclusion of m6A from splice-site proximal regions by the exon junction complex dictates m6A topologies and mRNA stability

Anna Uzonyi, David Dierks, Ronit Nir, Oh Sung Kwon, Ursula Toth, Isabelle Barbosa, Cindy Burel, Alexander Brandis, Walter Rossmanith, Hervé Le Hir, et al.

Exclusion of m6A from splice-site-proximal regions by the exon junction complex dictates m6A topologies and mRNA stability

Anna Uzonyi[1], David Dierks[1], Ronit Nir[1], Oh Sung Kwon[2,3], Ursula Toth[4], Isabelle Barbosa[2], Cindy Burel[2], Alexander Brandis[5], Walter Rossmanith[4], Hervé Le Hir[2], Boris Slobodin[1,6*] & Schraga Schwartz[1*]

[1]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 7630031, Israel
[2]Institut de Biologie de l'Ecole Normale Supérieure (IBENS), Ecole Normale Supérieure, CNRS, INSERM, Université PSL, 75005 Paris, France.
[3]Current address: BioResearch center, Samsung biologics, 300, Songdo bio-daero, Yeonsu-gu, Incheon 21987, South Korea.
[4]Center for Anatomy & Cell Biology, Medical University of Vienna, 1090 Vienna, Austria
[5]Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot 7630031, Israel
[6]Department of Biochemistry, Rappaport Faculty of Medicine, Technion - Israel Institute of Technology, Haifa 31096, Israel.

[*] - Corresponding authors:

Boris Slobodin - boris.sl@technion.ac.il
Schraga Schwartz - schwartz@weizmann.ac.il

Summary

N6-methyladenosine (m6A), a widespread destabilizing mark on mRNA, is non-uniformly distributed across the transcriptome, yet the basis for its selective deposition is unknown. Here, we propose that m6A deposition is not selective. Instead, it is exclusion-based: m6A consensus motifs are methylated by default, *unless* they are within a window of ~100 nt from a splice junction. A simple model which we extensively validate, relying exclusively on presence of m6A motifs and exon-intron architecture, allows *in-silico* recapitulation of experimentally-measured m6A profiles. We provide evidence that exclusion from splice junctions is mediated by the exon-junction complex (EJC), potentially via physical occlusion, and that previously observed associations between exon-intron architecture and mRNA decay are mechanistically mediated via m6A. Our findings establish a mechanism coupling nuclear mRNA splicing and packaging with the covalent installation of m6A, in turn controlling cytoplasmic decay.

Introduction

N6-methyladenosine (m6A) is the most widespread modification on mRNA, present at roughly 0.2-0.4% of all adenosines [1]. m6A is installed at a consensus motif, whose core is typically represented as a DRACH motif (D=A/G/T/ R=A/G, H=A/C/T), which also extends into adjacent nucleotides [2,3]. Yet, a major feature of m6A distribution, not accounted for by sequence, is the strong positional enrichment of m6A within genes. m6A was originally described as being enriched within atypically long internal exons and near stop codons [4,5]. A subsequent study suggested that rather than being associated with the stop codon, m6A was associated with the last intron-exon junction within a gene [6]. This prompted multiple studies to explore a potential impact of methylation on splicing, whereby only weak effects were typically observed [7–9]. In parallel, alternative potential mechanisms underlying the biased distribution of m6A have been proposed, invoking chromatin marks biased towards the ends of genes and potential involvement of the transcription termination machinery [6,10,11]. However, none of the proposed mechanisms accounts for the enrichment of m6A near stop codons, nor near last exons nor within long internal exons. The highly biased distribution of m6A has thus remained enigmatic, and the mechanism underlying it largely unknown. Given that m6A plays a well-established role in directing cytoplasmic degradation of mRNA [12–15], elucidating the forces shaping m6A is pivotal for obtaining a complete understanding of the rules governing mRNA stability.

mRNA levels are shaped by production and degradation. Intriguingly, diverse studies have consistently found that mRNA degradation rates correlate most strongly with ‘exon density’, a ratio between the number of exons within the coding region and the coding region length [16–20], explaining nearly 30% of the variability in degradation rates between different genes [16]. While this observation suggests a strong connection between exon-intron architecture and RNA stability, its mechanistic basis is unknown. Exon-intron architecture of genes has also been associated with additional features of genes including RNA processing, export, and translation [21–26]. The underlying mechanisms are poorly understood, whereby a key gap in our understanding is how the history of splicing in the nucleus is encoded to subsequently drive cytoplasmic fate.

In addition to intron excision, splicing also has a considerable impact on mRNP particle structure. Assembled during splicing and deposited upstream of exon-exon junctions, the Exon Junction Complex (EJC) is a stable molecular mark of splicing [27,28]. The core of this multimeric complex comprises the RNA helicase eIF4A3, the hetero-dimer MAGOH-Y14 and MLN51. The EJC core accompanies mRNAs to the cytoplasm, and serves as a binding platform for a dozen of more peripheral factors. Recent structural and biochemical studies revealed that pre-translational mRNPs adopt a compact and elongated rod-like structure [29–31] suggesting that RBPs including EJCs contribute to mRNA packaging. Functionally, the EJC has been found to link pre-mRNA splicing to downstream events including mRNA transport, localization, translation and surveillance [28,32], yet the underlying mechanisms remain incompletely understood, in particular given that the EJCs are disassembled during mRNA translation [33].

Here, we explore the basis for the region-specific distribution of m6A. We reveal that the transcriptome-wide selectivity in m6A deposition is not inclusion-based, but rather exclusion based. We suggest that sites harboring an m6A consensus sequence are methylated by default, *unless* they are within a window of up to ~100 nt from a splice junction, in which case their formation is inhibited by the EJC. We further provide evidence that m6A - installed in a splice-junction sensitive manner - serves as the mechanism connecting exon-intron architecture with cytoplasmic decay. Our findings thus causally link mRNA splicing and EJC deposition with the covalent installation of m6A in the nucleus, which - in turn - subsequently controls the mRNA cytoplasmic stability.

Results

Dissection of sequence determinants of m6A deposition

M6A is strongly enriched at long internal exons and last exons, near stop codons, yet the m6A consensus ('DRACH') motif is not enriched in these regions **(Fig. S1A)**, suggesting additional sources of specificity. To systematically assay the determinants governing m6A specificity, we established a massively parallel reporter assay, allowing us to interrogate the methylation status of 7584 different sequences, each 101-bp in length. These sequences were designed to include (1) WT and point-mutated DRACH motifs embedded within high-confidence m6A sites residing within human or mouse genes, (2) systematic mutations of each of the positions along a methylated sequence, (3) systematic shifting of the relative position of DRACH motifs along sequences, (4) synthetic sequences comprising varying number of DRACH motifs **(Fig. 1A)**. All sequences were cloned as a pool downstream of *GFP* and the resulting plasmids were transfected into HEK293T cells, following which m6A was monitored via m6A-IP and targeted amplification **(Fig. 1B)**. Each sequence was assigned an enrichment score, capturing the fold-change in read-count in m6A-IP samples in comparison to an 'input' control. Enrichment scores were highly reproducible internally (Pearson R (R_p) = 0.95, **Fig. S1B**), as evaluated on the basis of 150 sequences that had been integrated twice into the pool with different barcodes, and were also reproducible between biological replicates (R_p = 0.82, **Fig. S1C**).

Reassuringly, enrichment scores were considerably higher in WT sequences than in mutated counterparts **(Fig. 1C)**. In addition, an analysis of the systematically mutated series allowed to de-novo *functionally* reconstruct the requirement for a DRACH motif, which has previously been characterized primarily based on *associative* analysis **(Fig. 1D)**. However, these analyses failed to uncover any new requirement for methylation: Sequences were methylated at roughly uniform levels, regardless of their relative location within a sequence **(Fig. S1D)**, and no sequences appeared to be required beyond the positions harboring the m6A consensus motif **(Fig. 1E, S1E)**. Perhaps most surprisingly, we found that even the completely synthetic sequences were enriched in a manner correlating with the number of synthetic DRACH motifs installed in them **(Fig. 1F)**. These experiments thus suggested that under the surveyed contexts, a DRACH motif is not only required for methylation but may potentially also be sufficient.

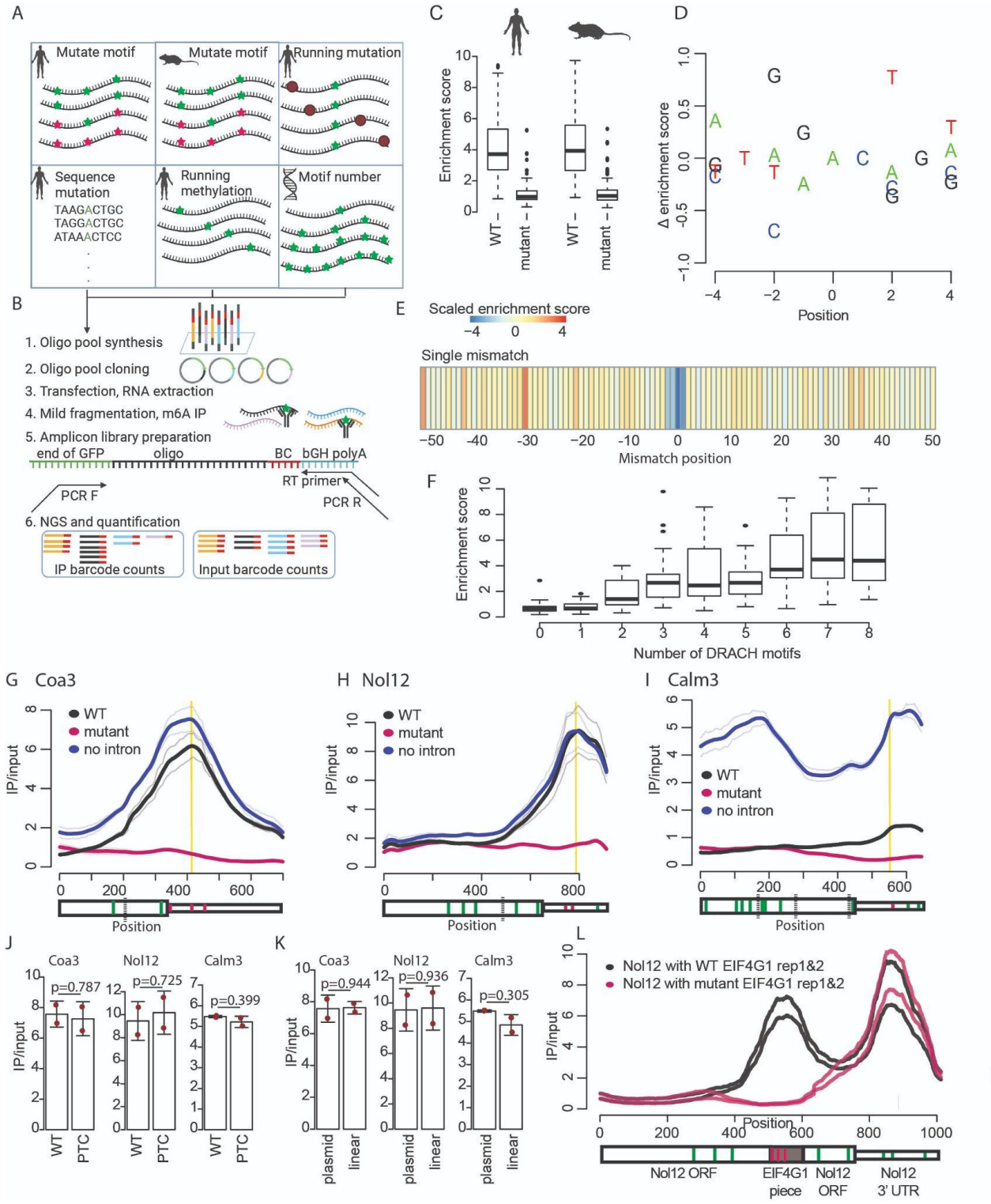

Figure 1. Interrogation of factors controlling m6A specificity. **A)** Depiction of the library subsets analyzed in Fig. 1 and Fig. S1. **B)** Experimental pipeline of the m6A measurement of the oligo library pool (see Methods). **C)** Enrichment score of WT and mutant sequences on the basis of 200 human and 200 mouse sequences. Outliers with an enrichment score over 10 are not shown. Box plots correspond to the median, Q1 and Q3, whiskers mark Q1-1.5 IQR and Q3+1.5 IQR. **D)** Delta (Δ) enrichment scores of the motif sequence set, on the basis of permutation of the

extended 9-mer consensus motif in four human mut-secondary sequences. The methylated A corresponds to position 0. Each value is a mean of all sequences with a certain base at a given position. **E)** Heatmap of enrichment z-scores in the running single nucleotide point mutation series. **F)** m6A Enrichment (Y axis) as a function of the number of synthetic DRACH motifs, ranging from 0 to 8. Each box is based on 18 synthetic sequences with different extended 9-mer consensus motifs. **G)** Top: IP/input ratio over the *Coa3* coding region and 3' UTR. Thin lines - measurements (two replicates). Thick lines - average with a 10% Loess fitting. Bottom: thick box - CDS, thin box - 3' UTR. Black dotted lines - exon-intron junctions. Green and red stripes - eligible DRACH motifs, where the red ones were mutated in the mutant constructs and the green ones were left unchanged. Vertical yellow stripe - main DRACH motif. **H-I)** Measurements for Nol12 and Calm3, as in G. **J)** IP/input ratios at the methylated site across each of the three constructs in the WT sequences (lacking an intron) in comparison to counterparts into which a premature termination codon (PTC) was introduced. All constructs lack introns. Red dots mark individual measurements, the height of the bars mark the average and the whiskers represent the standard deviation of the mean. P-values for two-tailed Student's t-test are shown. **K)** IP/input ratios at the methylated site across each of the three constructs in the WT constructs (lacking an intron) which were transfected as plasmids vs linear DNA into cells, represented as in (J) . **L)** IP/input coverage of constructs with a 101 nt long *EIF4G1* derived piece (WT or DRACH depleted mutant) cloned within the open reading frame (ORF) of the intronless *Nol12* gene. Based on two biological replicates. The gene model (bottom) is depicted as in panel G.

We next speculated that the basis for m6A selectivity might reside within more distal elements, such as the reading frame or intron-exon architecture. Motivated by the enrichment of m6A in internal and last exons, we sought to assess whether RNA splicing is required for methylation. We selected three mouse genes with well-supported m6A sites in the stop codon vicinity and with a relatively short last intron: *Coa3*, *Nol12*, and *Calm3*. The exons encoding each of the genes, in addition to the last intron (preceding the methylation site) were cloned into plasmids either in their wild-type form ('WT'), or with point-mutation of the suspected methylation motifs ('point-mut'). In addition, the same sequences were cloned without introns ('no-intron'). In one of the 3 cases (*'Calm3'*) the introns were sufficiently short to enable us to include three introns. Application of m6A-seq2 [15] to all 9 constructs, following their transfection into human MCF7 cells, revealed that the 'WT' constructs recapitulated the 'stop-codon' proximal enrichment of the endogenous constructs **(Fig. 1G-I, S1F)**, which was dramatically reduced in the point-mutated counterparts **(Fig. 1G-I, S1G)**. Similar profiles were observed for transiently expressed constructs and stably integrated counterparts **(Fig. S1H)**. However, no reduction was observed in the intron-less constructs; indeed, increased methylation levels were observed in two out of the three cases **(Fig. 1G-I, S1G)**. These results thus indicate that splicing is not a prerequisite for methylation. Given the association of m6A peaks with stop codons, we next assessed whether an intact reading frame was required for methylation. Introduction of a premature stop-codon into all three 'no-intron' constructs gave rise to nearly identical methylation levels as their 'WT' counterparts **(Fig. 1J)**, ruling out a requirement for a reading frame. We further found that methylation profiles were similar when DNA was transfected as a (circular) plasmid, or in a linear form, following PCR-amplification from the plasmid **(Fig. 1K, S1I),** suggestive of a relative indifference to chromatin properties. Finally, we found that introduction of a 101-bp methylated region from the 3' UTR of *EIF4G1* into the gene body of the *Nol12* no-intron construct led to strong methylation of the *EIF4G1* site **(Fig. 1L)**. This result demonstrates that presence within a 3' gene region is not a requirement for methylation, and - once again - pointed at DRACH motifs potentially being sufficient for inducing methylation, rendering the question of the determinants giving rise to the region-specific deposition of m6A all the more enigmatic.

<u>An m6A exclusion zone in the proximity of splice junctions</u>

In the above experiment we noted that removal of introns led to increased m6A signal **(Fig. 1G-I, S1G).** This was particularly pronounced in the case of the *Calm3* construct. Whereas in the intron-harboring construct enrichment was apparent only in the 3' UTR, in the intron lacking counterpart substantial enrichment was apparent over the beginning of the CDS, in a region harboring a high density of DRACH sites **(Fig. 1I)**. The enrichment pattern within the intron-less construct which appeared to mirror DRACH density led us to wonder whether it could be predicted in-silico simply on the basis of DRACH sites along it. To explore this, we established a simple model of m6A formation ('m6Apred-1') according to which every 'eligible' m6A site is methylated at a fixed level. Eligible m6A sites were defined as ones harboring any of the seven most prevalent consensus sequences (GGACC, AGACA, TGACT, AGACT, GAACT, GGACA, GGACT) [34]. To allow a direct comparison between our model and m6A-seq, in which the signal of an m6A site is spread as a 'peak' over a window surrounding the methylated site, we modeled each predicted m6A site as a gaussian centered over a 200-bp window surrounding the methylation site. The predicted methylation level of each position was defined as the sum of predicted signal overlapping that position (**Fig. 2A-top**). Remarkably, the m6Apred-1 predicted methylation profile for *Calm3* displayed an excellent correlation with measurements in the intron-less context (Spearman R_s=0.89, P<2.2e-16) (**Fig. 2A-bottom**), but no correlation with the corresponding measurements in the intron-harboring context (R_s=-0.06, P=0.1). Of note, the *in-silico* reconstituted methylation profile in *Calm3* is based on 11 different eligible DRACH motifs distributed along the length of this gene, all of which are predicted to undergo methylation by this model.

To explore whether DRACH-motif distribution along genes was sufficient for predicting methylation patterns in intron-less genes transcriptome-wide, we applied m6Apred-1 to ~20,000 human gene models, and compared its predictions against m6A signal from human A549 cells [34]. To explore agreement between predicted and measured methylation levels *within* genes, we established a 'per-base similarity score' (PBSS), capturing the per-base Spearman correlation between predicted and measured (IP/input) profiles. Consistently with our above results, we found that variability within genes was well-captured by m6Apred-1 in intron-lacking genes (median PBSS=0.5, n=64), but not in intron harboring ones (median PBSS=0.21, n=4,238) **(Fig. 2B)**. Also variability *between* genes, assessed via correlating the mean predicted vs measured m6A level per genes, were well captured in intron-lacking genes (R_s=0.56), but poorly in intron-harboring counterparts (R_s=0.17) **(Fig. 2C)**. These results thus suggest that in the absence of introns, m6A accumulation to a large extent mirrors the distribution of methylation consensus sequences, and that introns have an inhibitive impact on m6A formation.

To explore the possibility of an inhibitive effect of introns on methylation, we examined the distribution of 81,518 methylation sites, identified in human cells at single-nucleotide resolution by six m6A-miCLIP experiments [35], with respect to the intron-exon junction. Remarkably, we found m6A depleted from both the 5' and 3' ends of long internal exons, as well as from the 5' of last exons, plateauing at a distance of ~200 nt from splice junctions **(Fig. 2D, S2A-D)**. The size of this 'm6A exclusion zone' was fixed, and did not depend on the length of the internal exon **(Fig. 2D, S2B)**. Similar results were obtained on the basis of 25,879 m6A sites identified via

miCLIP in mouse [36] **(Fig. S2C),** as well as upon analysis DART-seq data [37] **(Fig. S2D)**. These results thus suggest the relative depletion of methylation sites from splice junctions, extending previous findings [14].

To establish that distance from the junction causally impacts m6A formation, we selected sixteen 239-bp regions harboring a prominent methylation site originating from different human and mouse genes **(Fig. 2E)**. For each of these methylation sites, we designed a series of 58 variants in which we systematically shifted the position of either a single or two consecutive methylation consensus motifs in 4-nt increments. These oligos were synthesized as a pool, and cloned into the last exon of the Calm3 construct, such that the relative distance of the methylation consensus motif from the last exon/intron junction spanned the range from 30 to 258 bp **(Fig. S2E-G)**. Remarkably, a continuous and pronounced increase in methylation was observed with increased distance from the splice junction across nearly every target (**Fig. 2F,G,S2H**), confirming the causal role of distance from splice site in controlling methylation levels. Further support was obtained on the basis of a series of perturbations conducted on the Calm3 gene, involving modulation of the distance between the *Calm3* methylation site and the last intron-exon junction. We found that increasing the distance between the methylation site and the splice junction gave rise to increased relative levels, up to a distance of 225 bp as measured both by m6A-seq **(Fig. 2H)** and by SCARLET [38] **(Fig. S2I).** We further confirmed that introduction of a 5-bp deletion at the donor site of intron 3, which led to intron 3 retention **(Fig. 2I, S2J)** thereby increasing the distance of the methylated DRACH motif from a splice junction, gave rise to increased m6A levels, whereas retention of intron 1 led to a much more reduced effect **(Fig. 2I).**

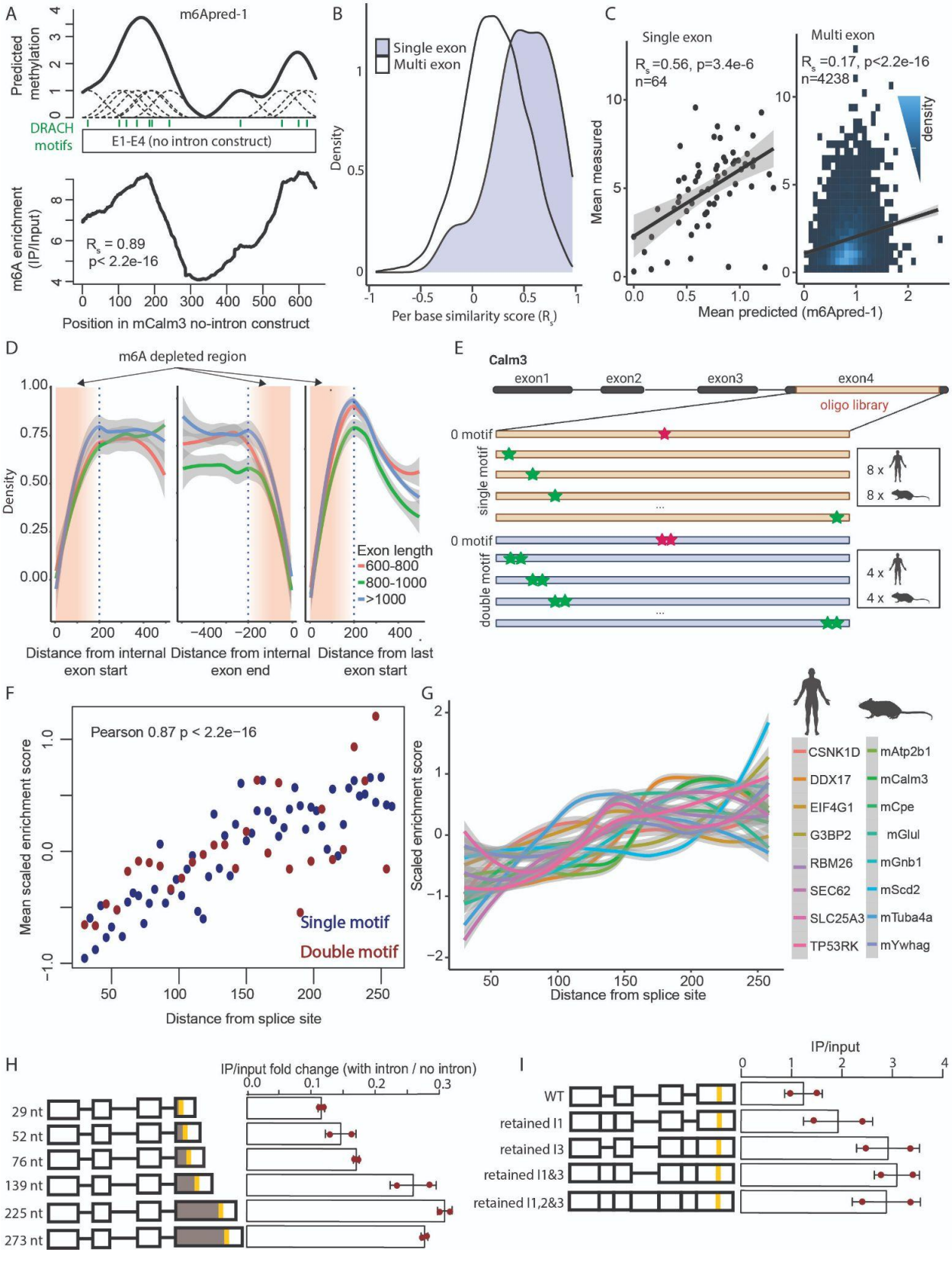
A
m6Apred-1
Predicted methylation
DRACH motifs
E1-E4 (no intron construct)
m6A enrichment (IP/Input)
Rs = 0.89
p< 2.2e-16
Position in mCalm3 no-intron construct
B
Single exon
Multi exon
Density
Per base similarity score (Rs)
C
Single exon
Rs =0.56, p=3.4e-6
n=64
Multi exon
Rs =0.17, p<2.2e-16
n=4238
density
Mean measured
Mean predicted (m6Apred-1)
D
m6A depleted region
Density
Exon length
600-800
800-1000
>1000
Distance from internal exon start
Distance from internal exon end
Distance from last exon start
E
Calm3
exon1
exon2
exon3
exon4
oligo library
0 motif
single motif
8 x
8 x
0 motif
double motif
4 x
4 x
F
Pearson 0.87 p < 2.2e−16
Mean scaled enrichment score
Single motif
Double motif
Distance from splice site
G
Scaled enrichment score
Distance from splice site
CSNK1D
DDX17
EIF4G1
G3BP2
RBM26
SEC62
SLC25A3
TP53RK
mAtp2b1
mCalm3
mCpe
mGlul
mGnb1
mScd2
mTuba4a
mYwhag
H
IP/input fold change (with intron / no intron)
29 nt
52 nt
76 nt
139 nt
225 nt
273 nt
I
IP/input
WT
retained I1
retained I3
retained I1&3
retained I1,2&3

Figure 2. Proximity to intron-exon junctions inhibits m6A formation. **A)** (top) Depiction of m6Apred-1 predictions on the basis of the *Calm3* gene. Every eligible DRACH motif (depicted as green bars) is modeled as a Gaussian, whereby the sum of all Gaussians overlapping a position defines its predicted methylation level. (bottom) Measured IP/input values along the *Calm3* gene are shown. The indicated correlation reflects the per-base correlation between the predicted (top) and measured (bottom) levels of methylation along the *Calm3* gene. **B)** Distribution of per-base similarity scores within single-exon and multi-exon genes. **C)** Correlation between the mean predicted levels of methylation (mean m6Apred-1 values along gene bodies) and corresponding measurements (IP/Input in A549 cells), shown for both single-exon and multi-exon genes. **D)** Frequency of detected m6A sites as a function of distance from the start (left panel) or end (middle panel) of internal exons, or from the start of the last exon (right panel). Blue dashed line marks 200 bp distance from the nearest exon-intron junction. **E)** Schematic of the long oligo pool planning. Oligo pool containing a single motif and a double motif set, with DRACH motifs moved from the beginning to the end of the sequence was cloned to the last exon of *Calm3*. The sequences are based on human and mouse sequences of the indicated numbers (right boxes). Green stars mark DRACH motifs, red stars mark mutated DRTCH motifs. **F)** Correlation of the mean enrichment score of different sequences with the relative position compared to the exon-intron junction. Enrichment scores were scaled within the sequences. Pearson correlation coefficient and p-value are indicated. **G)** Plot showing the scaled enrichment score for each of the natural (human and mouse) single motif constructs. **H)** Left panel: schematic of constructs with different distances between the last splice junction and the strong 3' m6A motif, on the basis of the *Calm3* gene. Vertical yellow lines mark the position of the quantified DRACH motif. The distance between the quantified 3' UTR DRACH motif and the last splice junction is marked in gray. Right panel: Bar plot comparing the methylation levels of the 3' UTR DRACH site of the WT *Calm3* construct, and corresponding constructs with reduced or increased distance between the last splice junction and the quantified site. For the quantification of input and IP signal, only reads beginning at most 10 bp upstream of the modification site were used, to avoid contribution of signal from a more upstream consensus motif (see Methods). Quantification shows the ratio of enrichment between the intron harboring and intronless constructs. Red dots mark individual measurements, the height of the bars mark the average and the whiskers represent the standard deviation of the mean. **I)** Left panel: schematic of constructs retaining introns 1, 3, 1&3 or all, on the basis of the *Calm3* gene. All graphical parameters are according to panel E. Right panel: Bar plots comparing the methylation levels of the 3' UTR DRACH site of the WT *Calm3* construct, as well as corresponding constructs with retention of intron 1, 3, 1&3 or 1,2&3. Red dots mark individual measurements, the height of the bars mark the average and the whiskers represent the standard deviation of the mean. For the quantification, 5' read starts up to 40 nt upstream were summed.

Transcriptome-wide m6A topologies are predictably driven via exclusion from splice junctions

The above results prompted us to explore the following model: rather than assuming that methylation is *off* by default and seeking elements driving its specificity, we considered an opposite model, in which methylation is *on* by default on all DRACH motifs *unless* in proximity to a splice junction. To assess whether a consensus sequence and a minimal distance from a splice junction are sufficient for *in-silico* reconstitution of m6A landscapes, we developed 'm6Apred-2', incorporating two considerations: (1) Every 'eligible' m6A site is methylated (as in 'm6Apred-1'), (2) unless it is within an 'm6A exclusion zone', a region spanning a fixed size ('ExclusionZoneSize') from the exon/intron junction (**Fig. 3A-top, S3A-C**). While the region depleted of m6A was roughly 200 nt long, the extent of depletion dropped continuously with distance from the junction **(Fig. 2D)**; Hence, the 'ExclusionZoneSize' was set at 100 nucleotides, roughly midway within the depleted region **(Fig. 2D)**. Of note, for intronless genes, m6Apred-1 and m6Apred-2 predictions are identical. Encouragingly, when we applied m6Apred-2 to the intron-harboring version of *Calm3*, the constraint on distance from a junction only allowed 3 (of the 11) DRACH motifs to undergo methylation, all of which in the last (long) exon. This resulted in a high agreement with experimental measurements, now also in an

intron-harboring context (R_s=0.81, **Fig. 3A-bottom**). Experimentally measured m6A profiles in mNol12 and mCoa3 changed only mildly between the intron-harboring and lacking constructs, consistent with only mild differences in predicted profiles, given that they were each cloned with only a single intron (**Fig. S3B-C)**. To validate this model more broadly, we selected 33 intron-harboring human genes, predicted to have different m6A-profiles in the presence of introns (on the basis of m6Apred-2) in comparison to the absence of introns (on the basis of m6Apred-1). We then overexpressed cDNAs encoding each of these genes and measured m6A profiles via m6A-seq2. As anticipated, we found that the (intron-less) cDNA m6A profiles correlated well with m6Apred-1 predictions but poorly with m6Apred-2 predictions. In contrast, the m6A profiles of genomically encoded counterparts were well predicted by m6Apred-2, but poorly with m6Apred-1 **(Fig. 3B,C, S3D)**. The same (exonic) sequence thus gets methylated in a differential - and predictable - manner, depending on the presence and whereabouts of introns.

Application of m6APred-2 to 20,102 human gene-models yielded 262,326 predicted methylation sites, with a mean of ~13 predicted sites per gene. These sites were distributed over 76,599 predicted 'peaks' (whereby each peak was defined as a consecutive stretch of bases with a predicted methylation signal >0). An average peak comprised 3.4 methylated sites, and spanned 373 nt. We obtained similar results across 20,954 mouse genes, where an average of 12.8 sites per gene distributed across 68,893 peaks.

To evaluate the performance of this model, we assessed its ability to capture m6A features at three different resolutions: individual m6a *sites*, m6A levels within and between *genes*, and *meta-gene* features. To validate the ability of m6Apred-2 to accurately detect individual m6A sites, we sought to investigate what fraction of the predicted sites could be confirmed on the basis of experimental data. We compiled an assembly of high-confidence m6A sites detected via distinct experimental single-nucleotide resolution methods (miCLIP [35], SAC-seq [39], miCLIP2 [40], DART-seq [37] and GLORI [41]). We then assessed the overlap between the predicted methylation status and the experimental measurements across 7,268 genes that had been pre-filtered based on expression-levels. While all sets of measurements provided substantial support to predictions by our model **(Fig. 3D)**, the most compelling results were observed via the recently established GLORI approach, which experimentally validates ~62% of all sites that we predict to undergo methylation, whereas only 6.3% of sites that we predict not to undergo methylation are experimentally supported by the GLORI dataset. Combining all datasets, 71% of the predicted methylated sites were experimentally validated, in comparison to 12% of the predicted negative ones **(Fig. 3D),** lending considerable support to the predictive power of m6Apred-2.

Next, we explored the ability of the model to predict diversity in methylation *between* genes. We observed a striking correlation (R_s=0.67) between experimentally measured m6A levels in A549 cells [34] and mean m6Apred-2 predictions **(Fig. 3E, top left panel)**. Importantly, this correlation coefficient dropped substantially (R_s=0.17) when predicted values were compared against m6A measurements in A549 cells which had been pre-treated with siRNAs targeting WTAP **(Fig. 3E, top right panel)**, a critical member of the methyltransferase complex [34]. We made essentially

identical observations in mouse embryonic stem cells, where we observed excellent agreement between predicted methylation densities and experimentally measured counterparts (R_s=0.71, **Fig. 3E, bottom left panel**), which were nearly entirely abolished in METTL3 knockout cells (R_s=0.1, **Fig. 3E, bottom right panel**). Consistent results were observed in an analysis of additional human and mouse cell lines **(Fig. S3E)**. Thus, roughly 45-50% of the variability in experimentally-measured m6A levels at the gene levels are captured by our model.

m6Apred-2 also performed well in de-novo recapitulating experimentally-measured m6A distribution *within* genes, with genes displaying remarkably high similarity scores between measured and predicted values **(Fig. 3F)**. Roughly 15% of the genes displayed excellent similarity scores (R>0.8) across gene bodies - in such cases, measured m6A profiles are nearly perfectly recapitulated using the simple, above-defined model **(Fig. 3F)**. Another ~26% of the cases displayed correlations between 0.6-0.8, and ~22% displayed correlations ranging from 0.4-0.6. Examples for genes displaying varying similarity scores are displayed in **Fig. 3G**. Median similarity scores in an intron-containing context were 0.54, marking a dramatic improvement over m6Apred-1 predictions that achieved a median score of 0.21 **(Fig. 2B)**. Genes with lower methylation levels generally showed substantially poorer similarity scores **(Fig. 3F, S4A)**, which is anticipated given that in lowly-methylated genes the relative levels of signal (true methylation) to background (non-specific binding) is low. We also found that M6Apred-2 tended to underestimate m6A signal in 5' UTR regions **(Fig. S4B,C)**, likely due to a substantial portion of m6A enrichment in the 5' UTR not reflecting m6A, but instead m6Am, installed by a different machinery at non-DRAC related sites [34,42–44].

Finally, at the metagene level, we found that the predicted m6A profiles recapitulated the relative depletion of m6A from highly expressed genes **(Fig. S4D),** reflecting the generally more compact structures of highly expressed genes comprising shorter exons and shorter 3' UTRs [45]. An additional hallmark of m6A distribution, the peak in the last exon, was also well-captured by m6Apred-2 **(Fig. S5A)**. We noted, however, that the sharp decline in the last exon, which is typically observed in experimentally-derived m6A metagene plots **(Fig. S5A left),** was only partially predicted by m6Apred-2. We hypothesized that methylation might be occluded not only from the vicinity of splice-junctions but also from polyadenylation sites, potentially due to masking by the megadalton polyadenylation and transcription termination machineries. To explore this hypothesis, we generated a series of constructs where the m6A site was embedded at varying distances from the polyadenylation site ranging from 25 to 200 nt. We found that reducing the distance to the polyadenylation site led to continuously decreasing methylation levels **(Fig. S5B)**. We also noted that 3' annotations tended to considerably overestimate 3' UTRs lengths, likely because cell-type agnostic annotations are often based on the longest isoform; such overestimates of 3' lengths give rise to predicted methylations in deep 3' UTR regions that do not occur in practice **(Fig. S5C)**. To assess whether the rapid decline of the peak at the terminal exon could be explained by these two factors, we first reannotated gene ends based on the input data (**Methods**) and then updated m6Apred-2 to not allow methylations to occur at motifs adjacent to the polyadenylation site. This was sufficient to predict both the last exon peak and the rapid decline **(Fig. S5A right)**, reminiscent of the one observed experimentally **(Fig. S5A left)**. These findings suggest that while the peak in the last exon is

due to m6A exclusion from splice junctions, the decline may partially be a consequence of exclusion of m6A also from polyadenylation sites.

Collectively, these findings establish that transcriptome-wide m6A topologies in human and mouse are to a large extent shaped by exclusion, primarily from splice-junction proximal regions and potentially also from the vicinity of polyadenylation sites.

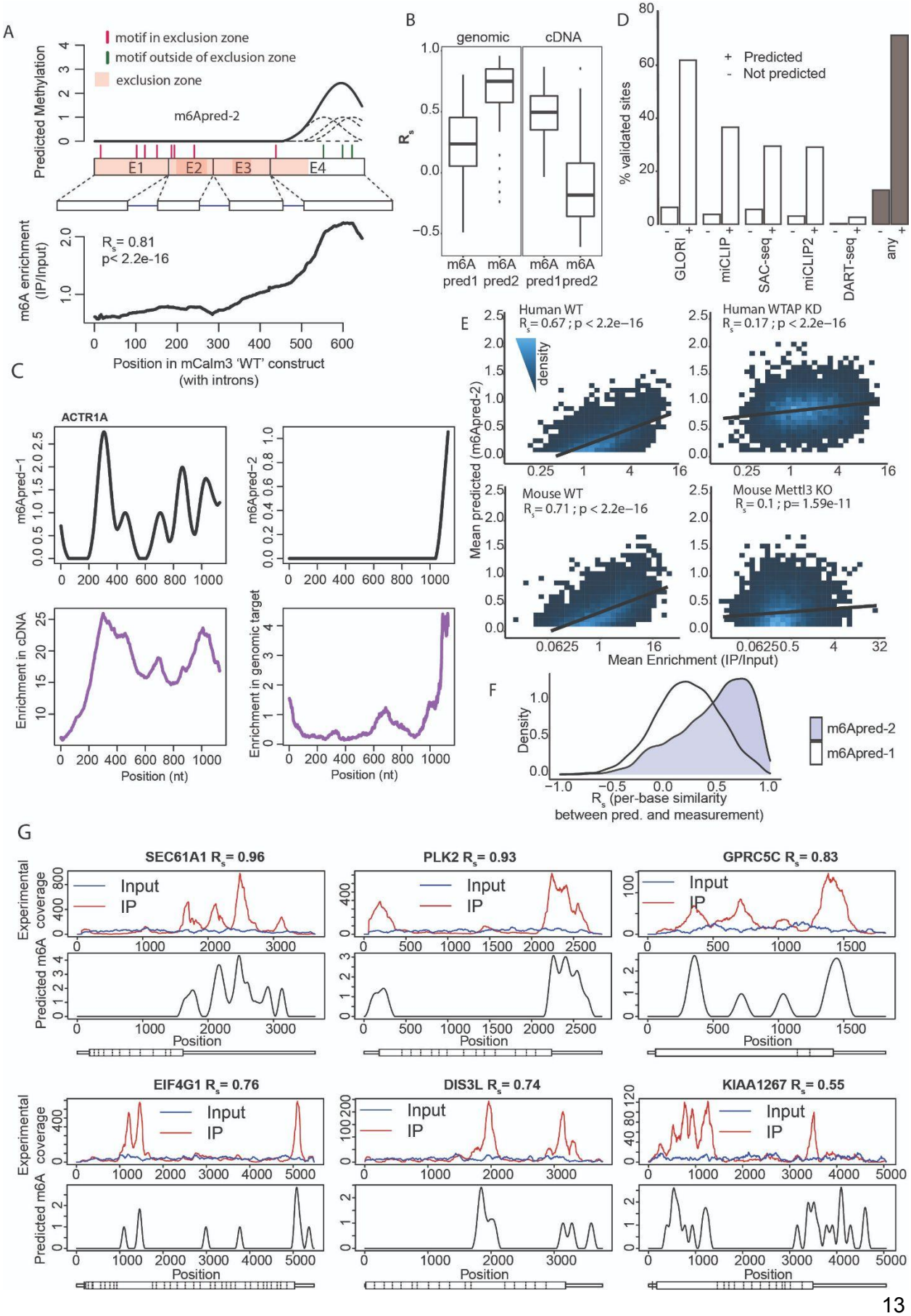
A
motif in exclusion zone
motif outside of exclusion zone
exclusion zone
m6Apred-2
Predicted Methylation
E1
E2
E3
E4
m6A enrichment (IP/Input)
Rs= 0.81
p< 2.2e-16
Position in mCalm3 'WT' construct (with introns)
B
genomic
cDNA
Rs
m6A pred1
m6A pred2
D
+ Predicted
- Not predicted
% validated sites
GLORI
miCLIP
SAC-seq
miCLIP2
DART-seq
any
C
ACTR1A
m6Apred−1
m6Apred−2
Enrichment in cDNA
Enrichment in genomic target
Position (nt)
E
Human WT
Rs= 0.67 ; p < 2.2e−16
Human WTAP KD
Rs= 0.17 ; p < 2.2e−16
density
Mean predicted (m6Apred-2)
Mouse WT
Rs= 0.71 ; p < 2.2e−16
Mouse Mettl3 KO
Rs= 0.1 ; p= 1.59e-11
Mean Enrichment (IP/Input)
F
Density
m6Apred-2
m6Apred-1
Rs (per-base similarity between pred. and measurement)
G
SEC61A1 Rs= 0.96
PLK2 Rs= 0.93
GPRC5C Rs= 0.83
EIF4G1 Rs= 0.76
DIS3L Rs= 0.74
KIAA1267 Rs= 0.55
Experimental coverage
Input
IP
Predicted m6A
Position

Figure 3. M6A topologies can be de-novo predicted based on motif and exon-intron architecture. **A)** (top) Development of m6Apred-2, relying on eligible DRACH motifs and exon-intron architecture, illustrated on the basis of the *Calm3* gene. Eligible m6A sites are depicted as lines, drawn in red if they are within an m6A exclusion zone (within 100 bp of an exon-intron junction, depicted as a transparent box in peach) or green otherwise. Each site is modeled as a gaussian, and the sum of the gaussians at each position define its predicted levels of methylation. (bottom) Measurements of m6A (IP/input) for the 'WT' *Calm3* construct (harboring agreement). The depicted correlation is the per-base correlation between the measured levels and the predicted ones (top). **B)** Boxplot of m6Apred-1 and m6Apred-2 correlation with measured m6A for 33 expressed cDNA constructs and their endogenous counterparts**.** For cDNA-based predictions the correlations are performed across the ORF, for genomic counterparts the correlations are across the gene-bodies. **C)** An example of observed and predicted methylation in cDNA constructs and endogenous counterparts. **D)** Predicted methylation levels are confirmed by experimental approaches. For each of ~120,000 sites across ~7000 genes, the fraction of sites within gene-bodies supported by different experimental approaches is plotted, binned by whether the site is predicted to undergo methylation or not. The right-most bars (in grey) reflect an analysis performed on all sites identified by at least one of the five methods. **E)** Agreement between predicted (m6Apred-2) and measured (IP/input) m6A values in human A549 cells (top) and mouse embryonic stem cells (bottom), in both WT cells (left) and cells depleted of methylation components, as indicated (right). Spearman correlations and P values are shown. **F)** Distribution of per-base similarity scores between predicted and measured (A549 cells) m6A levels on the basis of m6Apred-1 (white) and m6Apred-2 (purple). **G)** Predicted and measured m6A levels across five human genes. Top panels display the IP and input signal and bottom panels display the m6Apred-2 based prediction along the gene. The per-base correlation between predicted and measured values is shown at the top of each panel.

Exon-intron architecture dictates RNA stability via m6A

We next sought to explore the extent to which m6Apred-2 was predictive of mRNA stability. We found both in human and in mouse that predicted m6A levels correlated strongly with RNA stability (R_s = -0.5 and -0.47, respectively), close to the experimentally measured counterparts (R_s = -0.57 and -0.5, respectively) **(Fig. 4A-B).** These correlations with half-life between either the predicted m6A level or the experimentally measured one are both abolished in METTL3 KO mESCs **(Fig. 4C)**. To ensure the robustness of these associations, we obtained mRNA stability measurements from 34 different studies in human and mouse (assembled in [19]), revealing consistent negative correlations between half-lives and both predicted and measured m6A levels **(Fig. S5D,E)**. Further supporting the ability of our model to capture mRNA stability, we found that mRNA deadenylation rates, available in mouse 3T3 cells [46], correlate well with predicted m6A levels (R_s=0.42), close to levels achieved with measured ones (R_s=0.52) **(Fig. 4D)**, consistent with findings that binding of m6A by YTH-domain harboring proteins results in deadenylase recruitment [47].

The strong dependency of m6A on exon-intron architecture prompted us to explore whether m6A was the missing link connecting exon-intron architecture and mRNA stability [16–19]. We found that exon density correlates strongly with both measured (R_s=-0.58) and predicted (R_s=-0.54) m6A levels **(Fig. 4E)**, suggesting that exon density may be an indirect metric capturing m6A derived signal. To causally dissect this correlation, we explored the relationship between exon density and mRNA stability in WT and METTL3 KO mESCs. We recapitulated a strong correlation between mRNA half-life and exon density in WT cells (R_s=0.48). However, this correlation was nearly entirely abolished in METTL3 KO mESCs (R_s=0.11) **(Fig. 4F).** To gain further support for these findings, we measured mRNA stability in HEK293T cells acutely

depleted of m6A using STM2457 [48], an METTL3 inhibitor, and in mock-treated controls **(Fig. S6A)**, and observed reduced association between m6A and mRNA stability **(Fig. S6B)** as well as between mRNA stability and exon-density upon METTL3 inhibition **(Fig. S6C)**. These findings thus suggest that m6A - by virtue of being depleted from exon-rich regions - serves as the mechanism connecting exon-intron architecture to mRNA decay.

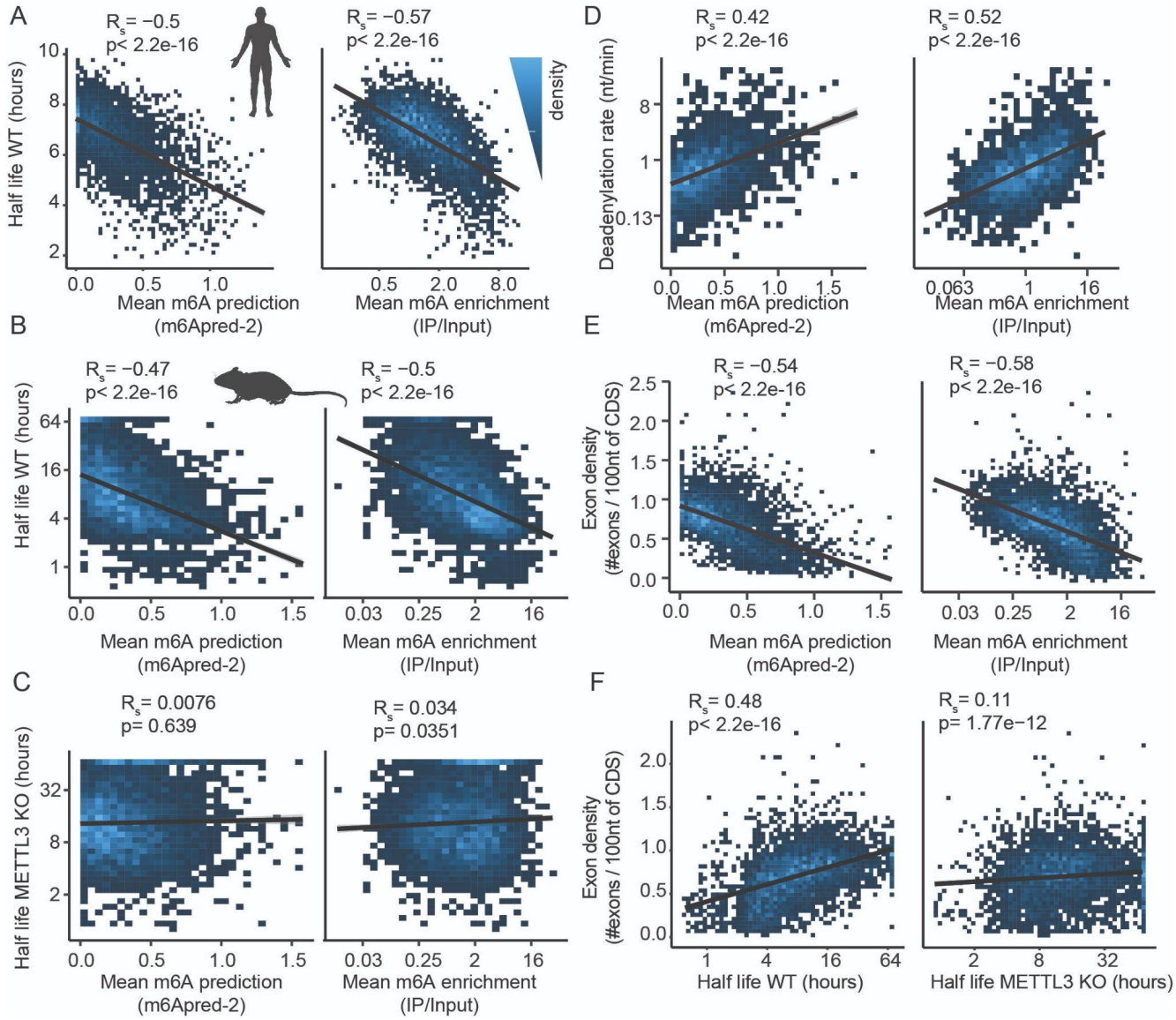

Figure 4. Exon-intron architecture drives mRNA stability through m6A. **A) (left)** Correlation between predicted m6A levels (m6Apred-2) and log-transformed mRNA stability rates in human cells. mRNA stability rates in MCF cells were obtained from [49]. **(right)** Correlation between measured m6A levels in A549 cells and mRNA stability. **B)** Plots as in (A), in mESCs on the basis of predictions and measurements in mESCs. **C)** Corresponding plots of predicted and measured m6A levels in mESCs against decay rates measured in METTL3 knock-out mESCs. **D)** Correlation between predicted and measured m6A levels in mESCs against deadenylation rates measured in mouse 3T3 cells. **E)** Correlation of exon density with predicted (left) and measured (right) m6A levels in mESCs. **F)** Correlation of exon density with decay rate in WT and METTL3 KO mESCs. Mouse half-life values were capped at 72 h.

Depletion of m6A from splice-junctions is mediated via the EJC

Given the tight association of the exon-junction complex (EJC) at exon-exon boundaries, we hypothesized that this complex might mediate loss of m6A from junction-proximal consensus motifs. Under this model, elimination of the EJC should lead to accumulation of m6A at eligible motifs within otherwise excluded zones. To investigate this possibility, we introduced a dTAG degron [50] downstream of Y14, a core EJC component, leading to efficient depletion over a 24-hour time course **(Fig. 5A,B)**. Remarkably, application of m6A-seq2 followed by peak-calling to samples harvested along this time course revealed a progressive and dramatic accumulation of m6A within - but not outside - of exclusion zones. Specifically, when we ordered all identified m6A peaks based on distance from the splice junction, a pronounced induction of m6A levels with Y14 depletion was observed up to a distance of ~200 nt, whereas little changes occurred in m6A intensities for sites more distal from splice junctions **(Fig. 5C, 5E, 5F)**. Consistently, the fraction of peaks within short exons grew substantially over the course of Y14 depletion, with only 0.59% of the called peaks residing within short exons at 0 hours in comparison to 21.89% at 24 hours following depletion **(Fig. 5D)**. Thus, the depletion of m6A from short exons and from junction-proximal regions is reversed with progressive EJC removal, consistent with a model that the EJC prevents their methylation. Furthermore, an analysis of gene-level m6A revealed a pronounced increase in m6A levels across multi-exon genes, and yet - critically - not in single-exon genes, which should not be affected by EJC loss **(Fig. 5G, S6D).** Finally, we were also able to confirm a dramatic induction of bulk m6A levels following EJC loss, via mass-spectrometry of digested mRNA derived nucleic acids **(Fig. 5H)**.

If the EJC prevents methylation from occurring at junction proximal sequences, this gives rise to two predictions. First, the ability to predict m6A landscapes *in silico* on the basis of m6Apred-2 - which takes into account exon-intron architecture - should continuously *decrease* with progressive Y14 depletion. Second, the ability to predict m6A landscapes *in silico* on the basis of m6Apred-1, which only takes into account motif availability but not exon-intron architecture, should continuously *increase* with progressive depletion. Both trends were evident, based on analyses capturing m6A variability both within **(Fig. 5I)** and between genes **(Fig. S6E).** Thus, upon depletion of a critical EJC component, m6A landscapes gradually cease to be governed by proximity to splice-junctions, leading to the predictable emergence of m6A sites within m6A exclusion zones.

Collectively, these findings strongly implicate the EJC in restricting methylation near splice sites, serving as the link connecting exon-intron architecture to m6A deposition, and thereby to mRNA stability.

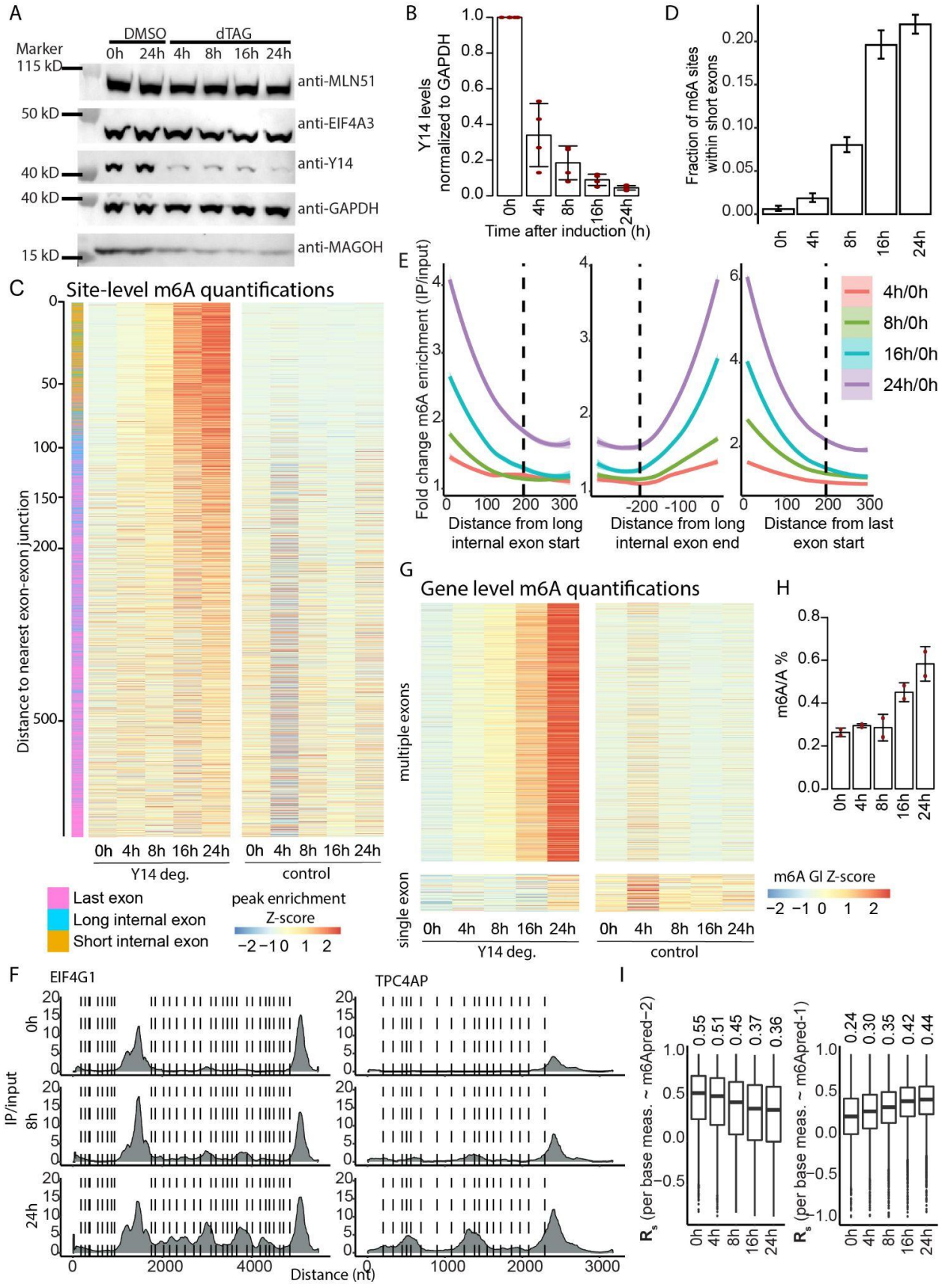
A
DMSO
dTAG
Marker
115 kD
0h 24h 4h 8h 16h 24h
anti-MLN51
50 kD
anti-EIF4A3
40 kD
anti-Y14
40 kD
anti-GAPDH
15 kD
anti-MAGOH
B
Y14 levels normalized to GAPDH
Time after induction (h)
D
Fraction of m6A sites within short exons
E
Fold change m6A enrichment (IP/input)
4h/0h
8h/0h
16h/0h
24h/0h
Distance from long internal exon start
Distance from long internal exon end
Distance from last exon start
C
Site-level m6A quantifications
Distance to nearest exon-exon junction
Y14 deg.
control
Last exon
Long internal exon
Short internal exon
peak enrichment Z-score
G
Gene level m6A quantifications
multiple exons
single exon
m6A GI Z-score
H
m6A/A %
F
EIF4G1
TPC4AP
IP/input
Distance (nt)
I
R_s (per base meas. ~ m6Apred−2)
R_s (per base meas. ~ m6Apred−1)

Figure 5. Inducible target specific depletion of exon-junction complex member Y14 leads to increased methylation in multi-exonic genes, in short exons. **A)** Representative Western blot images of Y14 degron and control DMSO treated time points, with antibodies against different EJC members and control GAPDH. Uncut gels are presented as Fig. S6G. **B)** Quantification of Western blot band intensities of Y14 normalized to GAPDH in the Y14 degron time points, based on four biological replicates. **C)** Heatmap of row-scaled m6A site scores of all Y14 degradation and control samples analyzed for unique m6A-sites that were previously detected by de-novo m6A peak calling (see Materials and methods) based on all Y14 degradation samples (0h, 4h, 8h, 16h and 24). m6A-sites were ordered by the distance to the closest exon-exon junction (distances indicated in black, left) and labeled by the type of exon (small internal exon; <200 bases, long internal exon; > 200 bases and last exon. **D)** Barplot depicting the ratio of m6A-sites that were depicted in small internal exons compared to any other type of exon. Error bars show the 95% confidence interval of the test of equal proportions. **E)** Meta fold-change m6A-IP/Input of timepoints after the induction of Y14 degradation (4h, 8h, 16h, 24h) compared to timepoint 0h. Analysis was performed for 300 bases of the 5' beginning (left), 300 bases before the 3' end (center) of long internal exons (>600 bases) and the 5' beginning of last exons (right). Calculated fold-changes are based on the median m6A-IP/Input score for 10-nt long bins per time point. **F)** Heatmap showing the row-scaled m6A gene index (m6A-GIs) of all the samples of the Y14 degradation dataset for genes annotated with > 1 exon (multiple exons, top) and single exon genes (bottom). **G)** m6A IP/ Input coverage (y-axis) per base in transcript space (x-axis) based on three samples of the Y14 degradation dataset (0h, 8h, 24h) for two representative genes *EIF4G1* (left) and TPC4AP (right). Dashed lines mark exon-exon boundaries. **H)** m6A/A percentage of the indicated time points of the Y14 degron samples, measured by LC-MS/MS. **I) (left)** Distribution of per-base correlations between m6Apred-2 predictions of m6A levels and experimentally measured counterparts (capturing the ability of m6Apred-2 to predict m6A variability **within** genes), plotted at indicated time points following Y14 depletion. **(right)** As in the left panel, but on the basis of m6Apred-1 predictions.

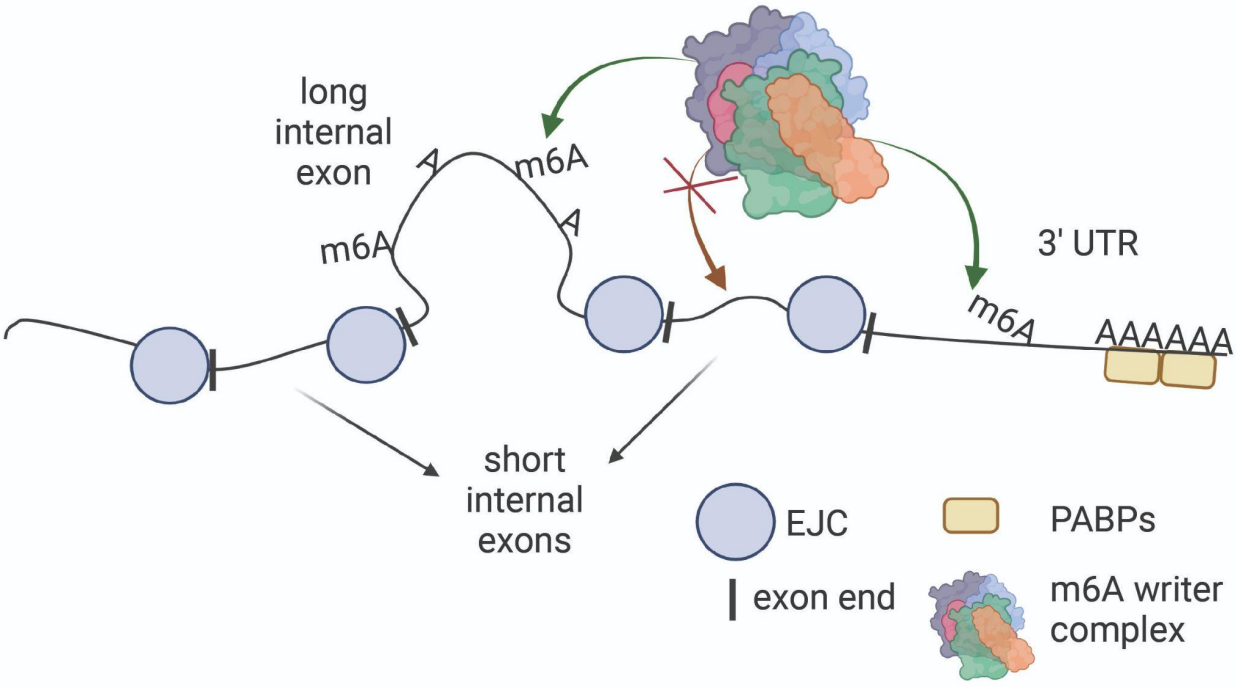

Figure 6. Model proposing the physical exclusion of methylation marks near splice junctions by the EJC (see **Discussion**). PABP: poly(A) binding protein.

Discussion

Here we propose a model for how specificity in m6A deposition is achieved. We propose that m6A deposition is dictated by sequence motifs and exon-intron architecture, and demonstrate that the depletion from splice site proximity is mediated via the EJC. In unraveling the link between m6A and exon-intron architecture, we also identify a mechanism linking exon-intron architecture to mRNA stability. Our findings suggest that deposition of m6A is radically different from tRNA and rRNA modifications: While the latter are selectively installed at specific positions, m6A is installed in a non-selective manner, governed by exclusion. Rather than being present on only a handful of sites per gene, our model predicts that m6A is installed at hundreds of thousands of sites transcriptome-wide, with an average of ~13 m6A sites per gene, which is consistent with most recent measurements of m6A on the basis of GLORI, reporting on average ~10 sites per gene [41]. Our findings have important bearings on the biogenesis, regulation and function of m6A, on our understanding of the determinants and regulation of mRNA stability, as well as on our understanding of a new critical role played by the EJC. Our findings further offer an attractive lens through which to re-examine additional contexts in which the cytoplasmic fate of an mRNA (e.g. localization, translation) were found to be associated with nuclear splicing [21–26,51,52].

How does the EJC prevent methylation from occurring at splice-site proximal consensus motifs? One possibility is that EJCs physically occlude m6A consensus motifs, thereby preventing access of the megadalton RNA methylation machinery [53] **(Fig. 6)**. The 8-nt footprint of EJCs [54] is insufficient to account for the depletion across a 200 bp window. This large exclusion zone most likely results from tight association of EJCs with other RBPs including members of the SR family [27,55] with a footprint of 30-150 nt [27]. Under this scenario, methylation occurs following EJC deposition, and hence as a consequence of splicing [28]. This is consistent with studies finding m6A to be strongly depleted from introns [9,14,56]. We do note, however, that m6A was found to be present at near steady-state levels already on chromatin-associated RNA [14] and was reported by one study on nascent RNA [57] and in another in introns [58], albeit at low levels, which would be inconsistent with the posited model.

Our findings have widespread implications on the regulation of m6A and of mRNA stability. The majority of exons are constitutively spliced most of the time. As such, mechanisms tied to splicing are similarly expected to be constitutive, to a considerable extent. This is consistent with the relative robustness of m6A profiles across different cell and tissue types [34,59] as well as with recent findings that mRNA stability is also highly correlated across diverse cell types [19]. Nonetheless, exon-intron architecture can be modulated via different forms of alternative splicing or via alternative polyadenylation, which can alter the proximity of m6A consensus sites to splice junctions, thereby impacting methylation and, in turn, mRNA stability. Alternative polyadenylation events are particularly interesting to consider, as it has been observed in diverse contexts that longer isoforms are less stable. This reduced stability was attributed to abolishment of microRNA binding sites, though this abolishment was shown to only account for a minor portion of this destabilization [60,61]. Our results offer a fresh interpretation for these observations, given that longer 3’ UTRs give rise to broad methylation-permissive regions

(vertebrate 3' UTR typically lack introns), which increase the methylation load of a gene and would therefore be predicted to reduce stability.

It is interesting to speculate about the 'logic' of coupling mRNA degradation to absence of splice junctions. One possibility is that such coupling came about in order to inhibit gene expression from retrotransposons. Such elements are pervasive in the human genome, they lack introns, and their expression has the potential to create havoc. Retrotransposons are epigenetically silenced, at the transcriptional level, via 5mC on DNA - indeed, 5mC is speculated to have evolved as a mechanism for their silencing [62,63]. m6A, which - similar to m5C - we now find to be installed by default, and which also imposes silencing but at the RNA level, may serve as a post-transcriptional counterpart. This idea resonates with recent literature implicating m6A in the silencing of transposable elements [64–68].

Limitations of the Study

Our model exhibits a substantial, yet still incomplete, agreement with experimentally measured methylation levels, explaining roughly ~50% of the variability in m6A levels between genes, and ~25% of the variability within a median gene. This incomplete agreement is likely attributed to several factors. First, a strength - but also weakness - of our model is its simplicity: it relies on binary, rather than probabilistic, decisions regarding the eligibility of motifs and splice junction-distances. Second, the experimental measurements on which we rely are limited, and subject to biases due to antibody promiscuity [3,69] and incomplete sensitivity [2]. Finally, as indicated above, genomic annotations - on which the model heavily relies - are an additional source of error. More quantitative models, based on improved and more quantitative m6A measurements and improved annotations, will allow an updated estimate of the fraction of m6A deposition not accounted for by our findings, and potentially allow discovery of additional levels shaping m6A deposition.

Our findings provide a unifying framework, quantitatively and causally connecting exon-intron architecture, the EJC, m6A deposition and mRNA stability. We anticipate that this study will open up future explorations into the logic of tying mRNA stability to exon-intron architecture, and into physiological and pathological contexts in which regulation of EJC deposition, via modulation of splicing, regulates mRNA stability by controlling the deposition of m6A.

Acknowledgements

We are grateful to Ghil Jona (Bacteriology & Genomic Repository Unit, Weizmann Institute of Science) for sharing dozens of plasmids from the ORFeome collection, to Joe Georgeson for conceiving the cDNA overexpression experiment, and to Quentin Alasseur for adapting the dTAG degron strategy.

SS is funded by the Israel Science Foundation (913/21) and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant no. 714023 and 101000970). S.S. is the incumbent of the Robert Edward and Roselyn Rich Manson Career Development Chair in Perpetuity. H.L.H. was supported by the Agence Nationale de la Recherche (ANR-17-CE12-0021 and ANR-21-CE12-0041), by Fondation pour la

Recherche Médicale (FRM EQU202003010226), and by continuous financial support from the Centre National de Recherche Scientifique, the Ecole Normale Supérieure and the Institut National de la Santé et de la Recherche Médicale, France. Anna Uzonyi is grateful to the Azrieli Foundation for the award of an Azrieli Fellowship.

Author contributions

B.S, A.U. and S.S. conceived the project. B.S., A.U. and R.N performed most of the experiments. A.U., D.D, and S.S performed the computational analyses. S.S. implemented the m6A prediction models. S.S. wrote the manuscript with input from all authors. O.S.K., C.B., I.B. and H.L.H. performed the Y14 degron experiments. U.T. and W.R acquired the SCARLET measurements. A.B. performed the mass spec experiments.

Declaration of interests

The authors declare no competing interests.

STAR Methods

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Schraga Schwartz (schwartz@weizmann.ac.il).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- NGS data have been deposited at Gene Expression Omnibus (GEO) and is publicly available as of the date of publication with accession number GSE204980.
- The code for m6Apred-1 and m6Apred-2 is included as supplementary code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Experimental model and subject details

HEK293T (female), MCF7 (female) and NIH3T3 (gender unknown) cells are from ATCC. These cell lines were cultured in DMEM (Gibco) supplemented with 10% FBS and 1% Penicillin and Streptomycin, at 37 °C.

For the Y14 dTAG experiments, HEK293T and HEK293T Y14-HA-dTAG cells were maintained in DMEM high glucose (Dutscher) supplemented with 10% of fetal bovine serum (Sigma-Aldrich), and 1% of penicillin and streptomycin (Sigma-Aldrich).

The commonly used V6.5 [70] mESC line (male) was kept in DMEM (Gibco) supplemented with 1% Penicillin-Streptomycin, 1 mM L-glutamine, 1% non-essential amino acids, 20% high-grade fetal bovine serum (Biological Industries), beta-mercaptoethanol and 10 µg recombinant leukemia inhibiting factor, at 37 °C. Cells were kept on tissue culture plates covered by gelatin

(0.2%), in co-culture with in-house generated, radiation-inactivated mouse embryonic fibroblasts.
No cell line authentication was performed.

Method details

Cloning of the short oligo library pool

The short oligo library pool (related to **Fig. 1A-F**) was synthesized by Twist Biosciences and cloned into a SNRPN-GFP plasmid as detailed in [71]. The full list of sequences in the pool can be found in Table S2.

Design of the short oligo library pool

For the wild type sequences, 200 human and 200 mouse genes were selected based on an annotated methylation site in several miCLIP data sets, aiming to keep a variability of different sequence motifs, as well as different origins of the sites, such as internal exons, end of CDS and 3' UTR. A 50 bp region on both sides of the methylated adenosine was selected based on the hg19 and mm9 genome annotations, leading to a total of 400 sequences with 101 nt length. For the mut-main set, the adenosine at position 51 was mutated to a T. For the mut-secondary set, all As in a DRACH motif, except for position 51 were mutated to a T. For the mut-all set, all As in DRACH motifs were mutated to a T.

The 'Running mutation' series is based on 10 human mut-secondary sequences. In this series, 1, 3 or 5 consecutive bases were mutated to their complement, base-by-base, from the first to the last of the 101 nucleotides. The 'Sequence mutation' series is based on 4 human mut-secondary sequences. In this set, the extended, 9-mer consensus motif in the middle of the sequence was permuted to all possible combinations of NTNRACNGN, with the methylated A in the middle. The 'Running methylation' set is based on ten human mut-secondary sequences. In this set, the extended 9-mer motif was moved from the beginning to the end of the 101 nt sequence, base-by-base. The 'Motif number' seq is based on 18 different 9-mer m6A consensus motifs: ATAAACAGT, ATAAACCGT, ATAAACTGT, ATAGACAGT, ATAGACCGT, ATAGACTGT, ATGAACAGT, ATGAACCGT, ATGAACTGT, ATGGACAGT, ATGGACCGT, ATGGACTGT, ATTAACAGT, ATTAACCGT, ATTAACTGT, ATTGACAGT, ATTGACCGT and ATTGACTGT. Each motif was placed in the 101 nt sequence in its wild type form n times (0-8) and in its point mutant (T instead of A in the middle) form 8-n times, at random locations. The remaining bases were filled up with random nucleotides maintaining a 50% GC content.

Each sequence was planned with a unique 10 nt barcode. Barcodes were designed to be devoid of 'AC' dinucleotides, to avoid any chance of undergoing methylation, and of purine triplets, as purine rich regions may be promiscuously immunoprecipitated by the anti-m6A antibody [3]. Furthermore, barcodes and the planned sequences were designed to be devoid of canonical poly-adenylation signals and the restriction sites used in the cloning strategy.

Cloning of the long, splicing oligo library pool

The long oligo pool (related to **Fig. 2E-G**) was synthesized by Twist Biosciences. The full list of sequences can be found in Table S3. Each sequence was planned with an 8 nucleotide

barcode, with the same considerations as above. Some of the sequences included in the library will be described elsewhere. The single motif set is based on 8 human and 8 mouse sequences. They contain a single 9-mer consensus motif which is moved in steps of four base pairs from the beginning to the end of the sequence. The double motif set is based on 4 human and 4 mouse sequences. Two consecutive DRACH motifs are moved together through the entire length of the sequence in steps of 8 base pairs. The length of the variable region of the constructs is 239 base pairs. With corresponding point mutated zero motif controls, as well as 100 sequences ordered with two different barcodes, the size of the library is 1996 different sequences. With the primers for PCR amplification and restriction sites the entire sequence length is 300 base pairs. This oligo pool was cloned into the pTwist CMV plasmid, containing the 'with intron' version of the *Calm3* gene (ordered from Twist Biosciences), with all DRACH motifs point mutated without disruption of the coding sequence. The majority of the last exon of *Calm3* was replaced with a library placeholder sample sequence for the backbone ordering. The library placeholder was designed to be in between two AvrII restriction sites. We first ensured that the last intron was properly spliced out, following which the placeholder was removed and replaced by the library pool. The library pool was cloned as above, with a traditional restriction-ligation based method. The oligo pool was digested with the restriction enzymes NheI (NEB) and BstBI (NEB), while the plasmid was digested with AvrII (Thermo Fisher XmaJI, complementary to NheI) and BstBI.

Cloning of the *Coa3*, *Nol12* and *Calm3* plasmids

DNA sequences encoding the mice *Nol12*, *Coa3* and *Calm3* genes were ordered from Twist Bioscience. These initial sequences included the full respective ORFs with either 1) the last intron of each gene (three introns for *Calm3*), 2) point mutations of adenosines suspected to be m6A-modified or, 3) no introns. These sequences were PCR-amplified and introduced between HindIII and NotI sites in the pcDNA5/FRT/TO plasmid (Thermo Fisher). For subsequent mutagenesis or modifications, the desired changes were introduced using restriction-free cloning. In brief, the desired changes were introduced using primers with partial complementarity to the original construct being modified.PCRs producing two parts of the new construct with complementary overhangs were conducted. After isolation of the correct products from agarose gels, both PCR products were combined and amplified using terminal primers. The final products were isolated from gels, digested with HindIII and NotI and ligated into a linearized pcDNA5/FRT/TO plasmid (Thermo Fisher). All PCRs were performed using Kapa HiFi DNA polymerase (Roche). Final constructs were partially sequenced using Sanger sequencing. All final sequences are described in Table S1.

Cloning of cDNA constructs

33 cDNA clones encoding full-length ORFs (Table S1) were selected from the 90/90 Human ORFeome V1 collection of Gateway Entry Vectors. The following criteria were used for selection of the cDNAs: (1) We ensured that the ORF sequence in the clone perfectly matched the ORF sequence as annotated in our assembly, (2) We filtered for genes with >5 exons with ORFs >800 nt, (4) We ensured that the endogenous genes were expressed above the 25% quartile of gene expression in HEK293T cells, to allow comparison of profiles in cDNA to endogenous counterparts, (5) We filtered for genes with experimentally measured m6A gene index

(m6A-GIs) >1.8, to ensure that they were endogenously modified somewhere, and (5) We filtered for genes with poor correlations between their predicted m6A profiles on the basis of m6Apred-1 and m6Apred-2 (r<0.3). The cDNAs were cloned into pLenti6 V5-DEST gateway vector (Invitrogen) that was linearized by FastDigest SalI (Thermo Scientific), using Gateway LR Clonase II (Invitrogen), according to the manufacturer's manual. The LR reactions were transformed into chemically competent Stbl3 bacteria (Invitrogen) and grown in 30 degrees.

Cell culture and transfections

For transfection, V6.5 mESCs were seeded on gelatin-covered 6-well plates (without MEF feeder cells) for transfection. *GFP-SLC25A3* plasmid was transfected with the TransIT-X2 reagent (Mirus) according to the manufacturer's instructions. Cells were collected 24 hours post-transfection.

For the library pool transfection, 0.5 million HEK293T cells were seeded in each well of a 6-well plate. 24 hours later cells were transfected with 2 µg plasmid DNA and 8 ul home-made PEI reagent. Cells were harvested 24 hours post transfection. Each 3 wells of the 6-well plate were merged to form two replicates.

For combined treatment with Actinomycin D and m6A inhibitor STM2457 (DC Chemicals, Shanghai, #DC53045), the cells were seeded into 60-mm dishes and grown for one day until 50% confluent. Then, STM2457 [48] compound was added at a final concentration of 5µM. After 6 hours of incubation, Actinomycin D was added (7.5µM) for the indicated time points and the cells were collected after a total treatment of 15 hours with STM2457. After lysis, RNA was extracted and RNA sequencing libraries were prepared according to [72].

Transfections of MCF7 cells were done with JetPrime (Polyplus) reagent, according to the supplier's instructions. Typically, the cells were seeded in 6-cm dishes (Corning) and on the next day transfected with 1-2µg of plasmids. The cells were harvested after additional 24 hours by scraping and subjected to RNA extraction. For stable expression of reporter genes, MCF7 cells bearing a genome-integrated Flp-In site were transfected with pcDNA5/FRT/TO-encoded genes and pOG44 recombinase in a 1:9 ratio. After 48 hours, the medium was supplemented with Hygromycin at 0.1mg/ml final concentration and grown for three weeks with weekly change of medium.

NIH3T3 cells were seeded on 60 mm plates 24 hours prior to transfection. Cells were transfected with pLenti6 cDNA constructs with Lipofectamine 2000 (Thermo Fisher Scientific) reagent according to the manufacturer's protocol. A total of 33 different cDNA plasmids were transfected in pools of 11-12, in duplicates.

Preparation and validation of Y14-HA-dTAG HEK293T cell

The dTAG degron (FKBP12F36V, [50]) was fused in frame into the C-terminus of Y14 in HEK293T using CRISPR-Cas9-mediated knock-in as previously described [73]. Briefly, 0.6 million of HEK293T cells were co-transfected using JET Prime (Polyplus, 101000001) with 0.25 µg of each of sgRNAs expressing plasmids and 1.5 µg of repair plasmid to fuse 3xHA tag and dTAG degron. 24 hr post transfection, cells were split in 10 cm dish, and 500 µg/ml of geneticin (G418,

InvivoGen) was added after 24 hr. CRISPR modification was screened by western blot with Y14 antibody (Santacruz, 4C4, sc-32312) and HA antibody (Sigma-Aldrich). The selected clones were transfected with Cre expressing plasmid with JET prime to eliminate neomycin resistant gene, and split into a 96 well plate with a single cell per well. Genomic DNA was purified from each clone by SureClean Plus (Bioline, BIO-37047) following manufacturer's instructions. Transgene integration and removal of Neomycin resistant gene were verified by genomic DNA PCR with the primer set F: 5'-TCTGTGTTTGCTGTGGTTGG-3', R: 5'-CCAGGTAACACAGCAAGACC-3'.

dTAG time-course treatment

In a 6 well plate, HEK293T and HEK293T Y14-HA-dTAG cells were cultivated at 70% confluency and treated with DMSO 1:1000 or 50 nM dTAG13 (Sigma-Aldrich) in DMSO, and harvested after 4, 8, 16 and 24 hours post-treatment.

For western blot, the following antibodies and dilutions were used: α-GAPDH (Cell Signaling, 14C10, 2118S, rabbit, 1:2000), α-eIF4A3 (Daguenet et al. 2012, 1:1000), α-Y14 (Santacruz, 4C4, sc-32312, H2416, mouse IgG2b, 1:500), α-MLN51 (Daguenet et al. 2012, 1:1000), α-MAGOH (E. Izaurralde, Max Planck Institute for Developmental Biology, Tuebingen, Germany, affinity purified from rabbit, 1:500).

LC-MS/MS for quantification of m6A/A

The LC-MC/MS sample preparation and analysis was performed as previously described [2].

SCARLET analyses for quantification of m6A/A

SCARLET analyses of mRNA samples (1 µg each) were performed as previously described [2], except that two-dimensional thin-layer chromatography (TLC) was used to finally resolve the 5'-monophosphate nucleosides (1st dimension solvent: 66% isobutyric acid, 0.25% ammonia; 2nd dimension solvent: 68% 2-propanol, 6.4% hydrochloric acid) [74]. The use of two-dimensional TLC resulted in superior separation of A and m6A from the "background" of other nucleosides.

RNA extraction, m6A IP and library preparation

RNA from mESC, NIH/3T3 and HEK293T cells was extracted with NucleoZOL (Macherey-Nagel); from MCF7 cells BIO TRI RNA reagent (Bio-lab) was used. RNA was poly-A selected with oligo dT-beads (Dynabeads mRNA DIRECT Kit, life tech). For all single-plasmid transfections, sample pooling, m6A immunoprecipitation and NGS library preparation was prepared according to the step-by-step m6A-seq2 protocol described in [15]. m6A IP and amplicon library preparation of the oligo library pool was performed on the basis of the same protocol with the following modifications: the 3' adapter ligation and pooling steps were omitted, the 5' adapter ligation step was omitted, reverse transcription was performed with a sequence specific RT primer (Table S1), library amplification was performed with sequence specific PCR primers containing Illumina sequencing adapters (Table S1). All NGS libraries were sequenced on the Illumina NovaSeq 6000 platform.

Datasets analyzed in this study

Measurements of m6A in WT A549 and in counterparts depleted of WTAP, METTL3, METTL14, and KIAA1429, as well as from mouse dendritic cells and embryonic fibroblasts, were obtained from [34] (GSE54365). Measurements of m6A in WT and METTL3 KO mESCs were obtained from [2], (GSE122961, replicate 1). mRNA stability in WT human MCF7 cells were obtained from [49] (GSE49831). mRNA stability in WT and METTL3 KO mESC cells were obtained from [14]. Deadenylation rates were obtained from [46].

Compilation of m6A catalogs at single nucleotide resolution: For the assembly of a high-confidence catalog of m6A sites in human cells employed in **Fig. 3D**, we collected m6A catalogs from the following sources: (1) A catalog of 81,519 sites identified across at least one of six miCLIP sites, assembled in [35], (2) A set of 12,672 m6A sites identified via DART-seq [37], (3) A set of 45,616 sites identified via SAC-seq, pooled from measurements conducted in HepG2, HEK293 and HeLa cells, which were downloaded from GSE162356, (4) A set of 36,552 sites identified in HEK293 cells using miCLIP2 [40], (5) A set of 170,240 sites identified in HEK293 cells via GLORI [41]. The UCSC 'LiftOver' tool was used to convert hg38 coordinates to hg19, when needed. We then generated a dataset of all eligible DRACH motifs within 7,300 gene-bodies (filtered based on expression level). Each motif was annotated with its predicted methylation status on the basis of m6Apred-2, following which it was merged with the assembly of m6A sites to allow assessment of experimentally-defined m6A status.

Alignment and calculation of IP/input enrichment scores

Oligo pools: Oligo pool data was analyzed with a custom R script. Enrichment scores were calculated as the ratio between the number of reads originating from a single sequence upon m6A-IP in comparison to the input and normalized to the median score of all human, mouse and synthetic constructs with no m6A consensus motif in the corresponding library.
To calculate the delta enrichment scores at each position for Fig. 1D, from the mean enrichment score of a given nucleotide we subtracted the mean enrichment score of any other nucleotide. These scores were averaged for all the different motif permutations for each position.

Constructs: Transcriptome wide m6A-seq2 datasets were mapped to mm9 or hg19 genome assembly and the transfected plasmid using STAR/2.7.9a [75]. Per-base read coverages were obtained with txtools (https://github.com/AngelCampos/txtools), and normalized to the total number of reads in the pool. Raw input and IP read counts were first normalized by the sum of reads across the entire corresponding pool of samples subjected to m6A-seq simultaneously (i.e. across all input samples or all IP samples, respectively). m6A enrichment levels were calculated as the fold-change in normalized coverage at the modified position between IP and input. For constructs with large insertions and deletions (in **Fig. 2H,I**), we sought to avoid contribution of signal originating from the insert (or from upstream intronic regions). In these analyses we therefore constrained the analysis of input and IP reads only to ones overlapping the sites and beginning at most n bases upstream of the methylated position. In the case of **Fig. 2H**, n=10 was employed, given the presence of a DRACH motif 10 nt upstream of the shortest construct (in which the targeted m6A site was 29 nt away from the splice junction). To quantify the constructs with intron retention **(Fig. 2I)**, n=40 was employed; In this case a longer distance could be used, allowing integration of more signal, without reaching the DRACH site.

Transcriptome-wide: m6A-seq measurements (IP and input) were aligned against the hg19 and mm9 genomes for human and mouse, respectively, using STAR/2.7.9a [75]. An in-house script was employed to map the genomic coordinates to transcriptomic annotations. For these annotations, we used canonical 'UCSC Known Genes' for human and mouse. Only reads fully matching a transcript structure were retained. Paired end reads were computationally extended in transcriptome space from the beginning of the first read to the end of its mate, and coverage in transcriptome-space was calculated for each nucleotide across all transcripts. For each base along each transcript, an IP/input enrichment score was calculated.

Implementation of m6Apred-1 and m6Apred-2 models

Gene models for roughly 20,000 'canonical' human and mouse genes were downloaded from the 'UCSC Known Genes' annotation table in the UCSC genome browser. For each gene we identified all 'eligible' DRACH motifs, and furthermore recorded their distance from the nearest exon-intron junction. In m6Apred-1, each of the motifs was considered methylated. In m6Apred-2, an eligible DRACH motif was only considered methylated if its distance to the nearest splice junction exceeded 100 nt. With the exception of analyses in **Fig. S5C**, also motifs within 100 nt from the transcript start or end site were considered non-eligible by m6Apred-2, based on the finding that m6A is also depleted in the vicinity of polyadenylation sites. To mimic the regional enrichment in m6A-seq, every site predicted to undergo methylation was modeled as a gaussian over a 200 bp region centered at the methylated site. The values along this gaussian were derived using the density function of a gaussian distribution (mean=0, sd=4), which were calculated for 100 values distributed at fixed intervals between 0 and 9 using the dnorm() function in R, and min-max normalized, to distribute between 0 and 1. The final predicted enrichment value at each position along the gene was defined as the sum of all signals (stemming from zero, one or potentially multiple gaussians) overlapping this position (see Supplementary code).

For the analyses summarizing gene-level predicted or measured methylation levels, we calculated the mean signal over gene bodies. Gene bodies were defined as the entirety of the gene excluding TSS-proximal regions (first 200 bp of the gene) and distal 3' UTR regions (greater than 400 bp from the stop codon). These regions were excluded to avoid contribution from signal originating from m6Am at the beginning of genes [34,42–44] and overcome limitation in available annotations of 3' UTRs (see Results). The gene-level quantification of experimentally-measured m6A levels was calculated as the overall number of reads per gene-body in the IP experiment divided by the corresponding number in the input experiment. This metric draws on our previously developed m6A-gene index (GI) [15].

For assessing agreement between predicted and experimentally measured profiles, we only included protein-coding genes with an average coverage per base exceeding 5 reads in the Input sample, to ensure adequate expression for estimating enrichment levels. These filters typically allowed quantifications of ~5000-~10000 genes.

For a rough approximation of 3' termini on the basis of 'input' RNA-seq data, we used a strategy similar to the one employed in [3]. We scanned the 3' UTRs from 5' to 3', and defined the 3' UTR as the first point at which coverage dropped >3-fold with respect to median coverage within the open reading frame of the gene.

Analysis of Mettl3 inhibitor RNA degradation timepoints

Mettl3 inhibitor and actinomycin D treated samples were mapped to hg19 genome annotation using STAR/2.7.9a [75]. Total read count per gene was computed with bedtools multicov [76], with the samples ordered as non-treated 0 hours replicates, 3 hours replicates, 6 hours replicates, 9 hours replicates, followed by the treated samples ordered similarly. Genes with less than a sum of 100 reads for all inhibitor treated and control timepoints were filtered out. Read counts were normalized to library size. Read counts were log transformed and the decay rate was calculated as a linear model fit on all time point duplicate measurements for treatment and control. Genes that had an r^2 value of the fit below 0.5 were filtered out. Remaining decay rates were correlated with measured and predicted m6A levels, as well as exon density.

Analysis of Y14 degron timepoints m6A IP data

The reads were aligned with STAR/2.7.9a to the hg19 genome. For further analysis, the bam files were filtered for uniquely aligned reads and processed by the bam2Endreads Rscript [2] for RNA fragment coverage in transcript space, based on the UCSC hg19 canonical gene annotation. M6a-GI and m6A-site score1 calculation were performed according to [15]. High confidence m6A-sites were defined as de-novo detected m6A-peaks in which the summit has an absolute distance to the nearest DRACH motif with less than 5 bases (see below, De-novo m6A peak calling).

De-novo m6A peak calling

m6A peak calling was performed as in [3], with minor adaptations. Genes were filtered for sufficient coverage for m6A peak detection with a threshold of at least 5 reads median coverage in the input sample. For a 51 bases rolling window centered around each base of an annotated transcript, the m6A site-score 2 was determined, defined as the ratio of the mean m6A-IP coverage of the window against the median IP. An Enrichment-Window was defined when m6A-site score 2 exceeded 4 for 15 consecutive bases. For each detected Enrichment-Windows, the Winscore was calculated by dividing the IP-derived m6A site score 2 by the same metric based on the Input coverage. Enrichment-Windows with a Winscore > 2 were isolated and consecutive windows merged. The final m6A-peak criteria was an IP to Input coverage fold-change of > 3. High-confidence single-nucleotide m6A-sites were defined by the summit of an m6A-peak with a maximum absolute distance of 5 bases to the nearest adenosine of a DRACH motif.

Supplementary Figures

Supplementary figure 1.

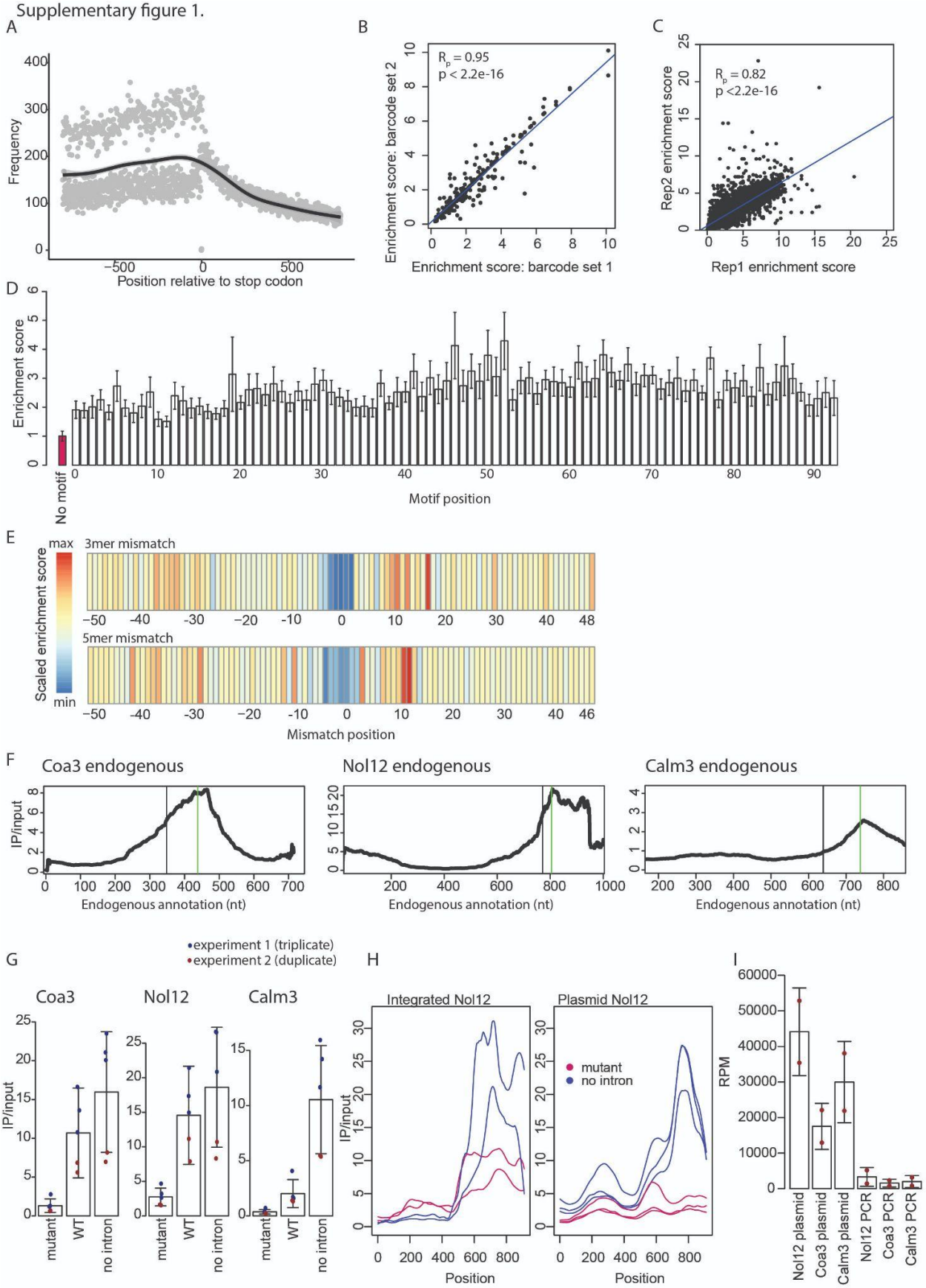

Figure S1. Related to Fig. 1. **A)** Metagene plot of the distribution of the 7 most common DRACH motifs (GGACC, AGACA, TGACT, AGACT, GAACT, GGACA, GGACT) in the human genome. Black line shows Loess fitted average.. **B)** Correlation of the enrichment score of 150 sequences planned with two different barcodes. Enrichment score was calculated by dividing the IP reads ratio mapped to a barcode by the input reads ratio mapped to a barcode (IP/input ratio), and normalized to the median IP/input ratio of all human, mouse and synthetic sequences with no DRACH motifs present in the pool. Pearson correlation coefficient and p-value are marked on the plot. **C)** Correlation of biological replicates of the MPRA assay. Pearson correlation coefficient and p-value are marked. **D)** Enrichment scores of the moving motif series. 9-mer m6A motif was moved base-by-base from the beginning to the end of the 101 nt long sequence. Bar plot depicts the mean and standard error of the mean. The value for the negative control sequences with no DRACH motif is marked in red. **E)** Scaled enrichment scores in the running 3mer and 5mer nucleotide point mutation, similarly to Fig. 1E. The mismatch position marks the first mutated base. **F)** IP/input ratios of the endogenous counterparts of the *Coa3*, *Nol12* and *Calm3* constructs. The main 3' UTR m6A motif of the synthetic constructs are marked with green, stop codon is marked with black vertical line. Only the regions which were part of the cloned constructs (without introns) are shown. X-axis presents the positions in the annotation of the endogenous transcript including 5' UTR. **G)** IP/input ratios at the position of the main DRACH motif of the *Coa3*, *Nol12* and *Calm3* constructs related to Fig. 1G-I. Red and blue dots mark individual measurements for separate two experiments, the height of the bars mark the average and the whiskers represent the standard deviation of the mean. **H)** Comparison of Nol12 expression from a genomic integration (left) and expressed from a plasmid (right) showing similar dynamic range of IP/input ratios. **I)** RPM expression levels of the *Nol12*, *Coa3* and *Calm3* constructs, transfected as circular plasmid or linear PCR product.

Supplementary figure 2.

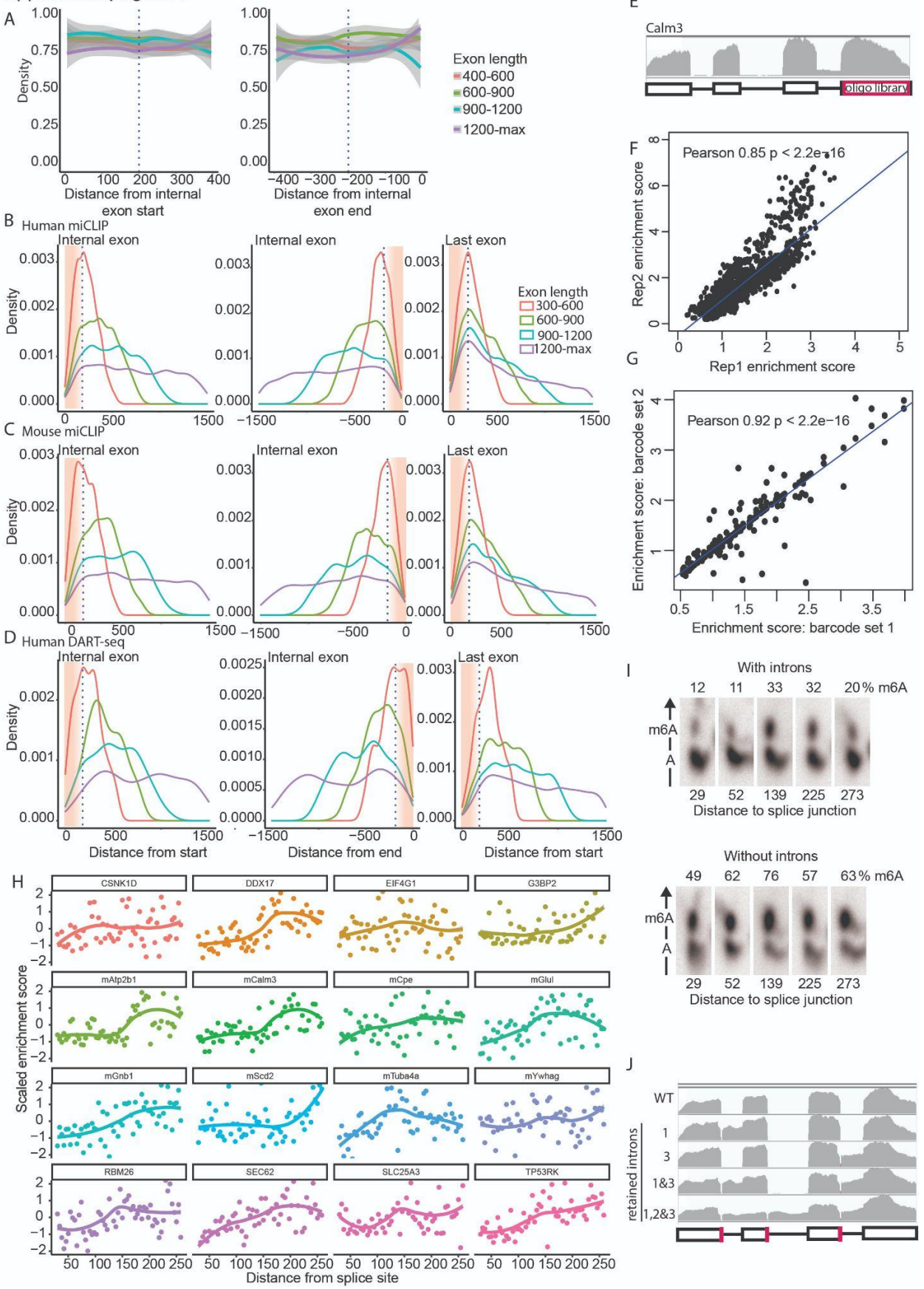

Figure S2. Related to Fig. 2. **A)** Density of DRACH motifs relative to splice junctions, shown by the distance from exon starts (left) or exon ends (right). **B)** Histogram of detected methylation sites in human miCLIP data based on exon length, according to Fig. 2D. While Fig. 2D only shows exons over 600 bp and the first/last 500 bp of exons, this figure shows all exons over 300 bp in a distance of up to 1500 bp from the splice junction. Left panel shows distance from the start of internal exons, right panel: distance from the end of internal exons, **C)** As in B), on the basis of mouse miCLIP data. **D)** As in B), on the basis of human bulk DART-seq data. **E)** IGV screenshot showing proper splicing of the *Calm3* gene with the oligo library place holder sequence. Exons are marked with boxes, introns with horizontal lines, the oligo library position is indicated in magenta. **F)** Correlation of biological replicates of the *Calm3*-splicing MPRA assay. Pearson correlation coefficient and p-value are marked. **G)** Correlation of the enrichment score of 100 sequences planned with two different barcodes. Pearson correlation coefficient and p-value are marked on the plot. **H)** Scaled enrichments scores relative to the distance from the splice junction for the single motif set, for all human and mouse constructs. Line indicates a Loess smoothing while individual data points are shown as well. **I)** SCARLET measurement of the Calm3 main m6A motif placed in different distances from the nearest splice junction in the intron-harboring (left) and intron-lacking (right) form. Segments containing A and m6A were cropped from 2D-TLCs. The chromatographic direction of the 2nd TLC dimension is indicated by the arrow on the left side. Uncut gels are presented in Fig. S6F. **J)** IGV coverage profiles of WT (intron harboring) *Calm3* construct, and corresponding constructs exhibiting retention of intron 1, 3, 1&3 or 1,2,&3, due to 5 nt deletions at the splice junction. Black boxes mark exons, horizontal lines mark introns, magenta marks the 5 nt deletion leading to intron retention.

Supplementary figure 3.

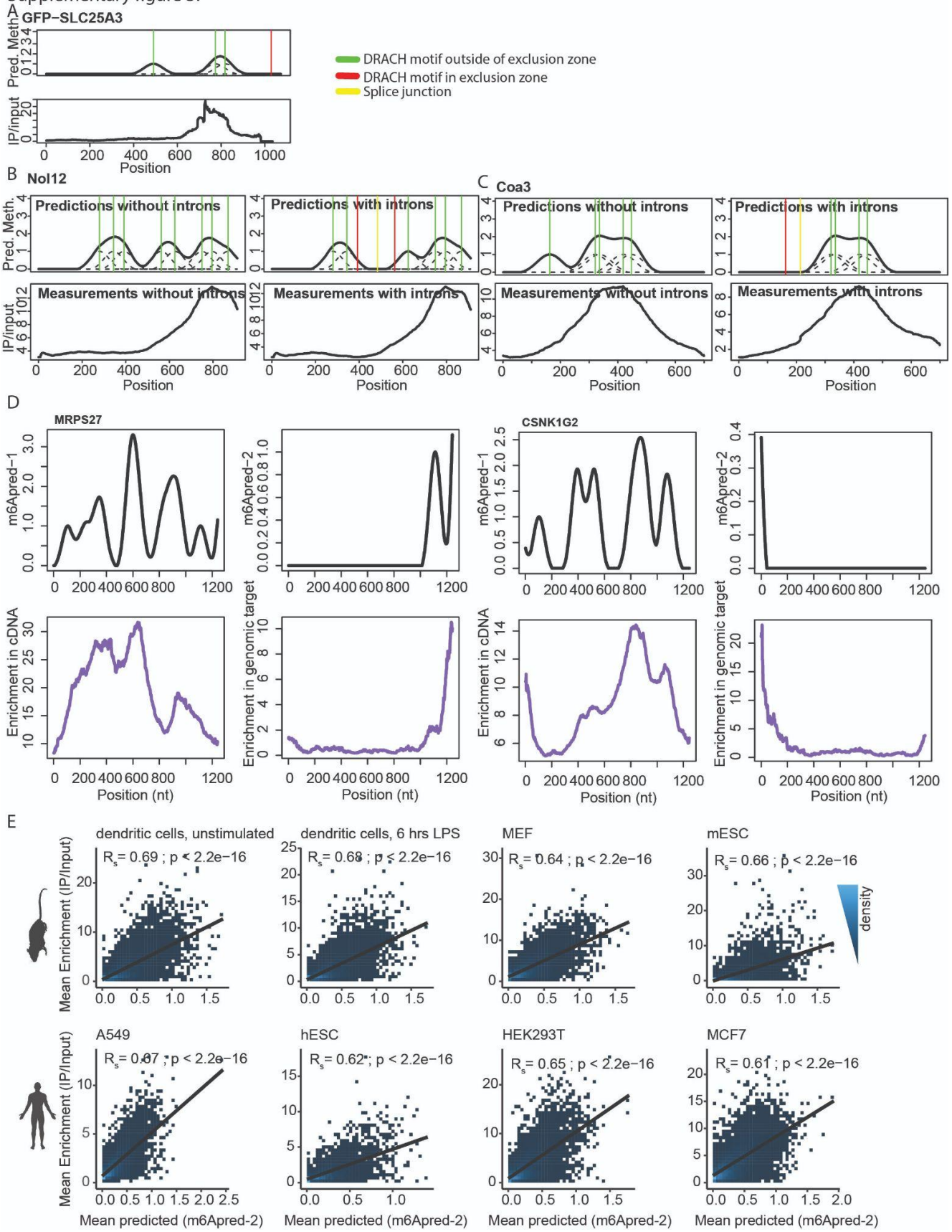

Figure S3. Related to Fig.3. **A)** Predicted (top) and measured (bottom) m6A levels of the *GFP-SLC25A3* construct. **B)** Predicted (top) and measured (bottom) m6A levels for the *Nol12* construct without introns (left) or with introns (right). For the no intron prediction m6Apred-1 was used, for the with intron prediction m6Apred-2 was used. Dashed lines mark single peak predictions. Red vertical lines mark m6A motifs in exclusion zones (not predicted by m6Apred-2), green lines mark m6A motifs outside the exclusion zone. Yellow lines mark splice junctions. **C)** As in B, for the Coa3 construct. **D)** Additional examples of observed and predicted methylation in cDNA constructs and endogenous counterparts, as in Fig. 3C. **E)** Agreement between predicted (m6Apred-2) and measured (IP/input) m6A values in mouse (top) and human (bottom) cell lines, as shown in Fig.3E.

Supplementary figure 4.

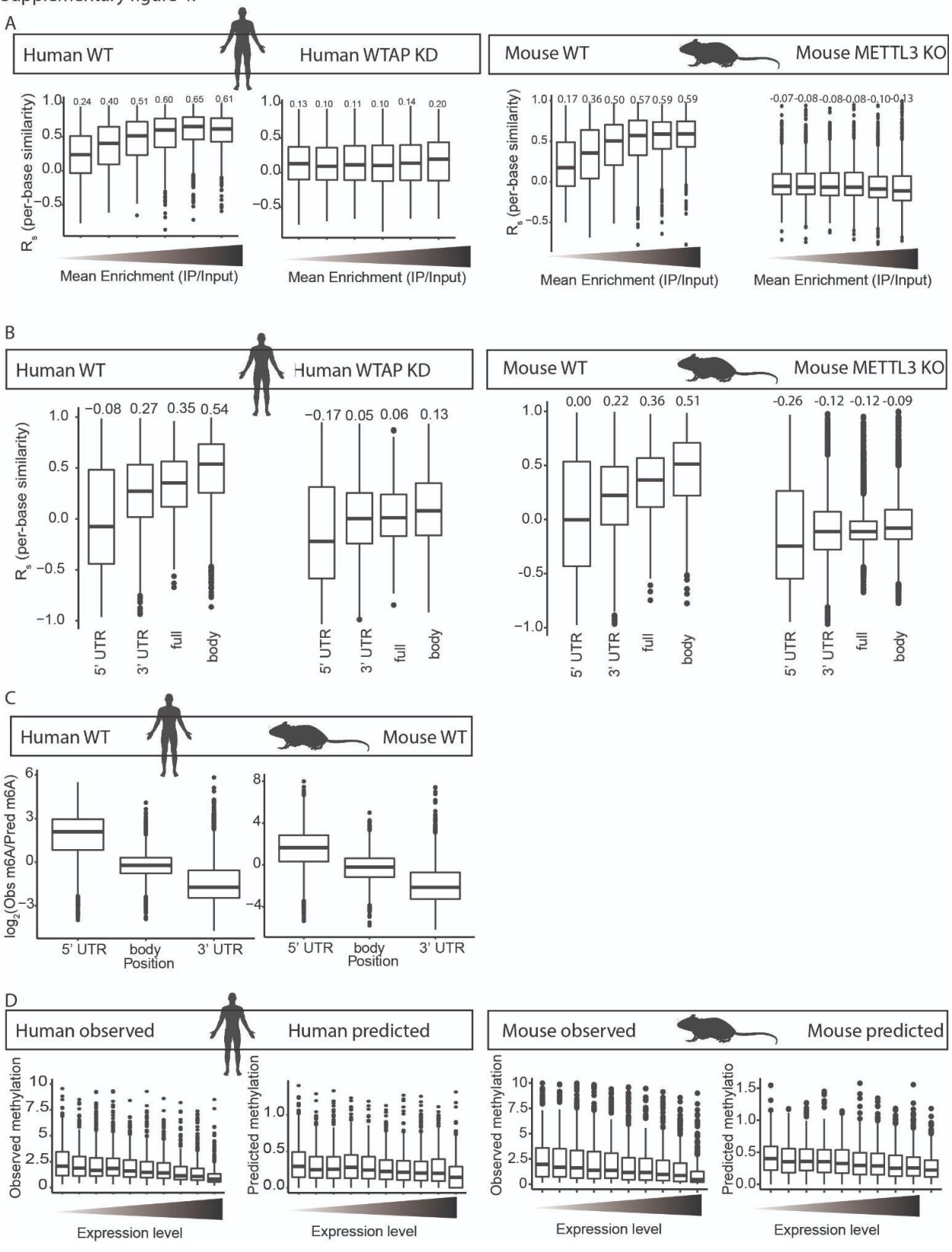

Figure S4. Related to Fig. 3. **A)** Boxplots depicting the distribution of the per-base similarity scores across human (left) and mouse (right) genes, binned into 6 equally sized groups on the basis of the experimentally measured m6A levels (IP/Input). Distributions are shown for both WT and WTAP knockdown (human) or Mettl3 knockout (mouse) samples. Median correlation coefficients of each box are written over it. **B)** Box plots of the m6Apred-2 model performance in human (left) and mouse (right) in different genetic regions. Y-axis shows Spearman correlation coefficient of per-base similarity. **C)** Box plots of observed versus m6Apred-2 predicted m6A in human (left) and mouse (right) in different genetic regions; 5' UTR, gene body and 3' UTR. **D)** Box plots showing the relation between expression level and experimentally observed methylation in human (left panel), m6Apred-2 predicted methylation in human (middle-left), experimentally observed methylation in mice (middle-right) and m6Apred-2 predicted methylation in mice (right). Box plots correspond to the median, Q1 and Q3, whiskers mark Q1-1.5 IQR and Q3+1.5 IQR. Genes are divided into ten equally sized bins according to their expression levels.

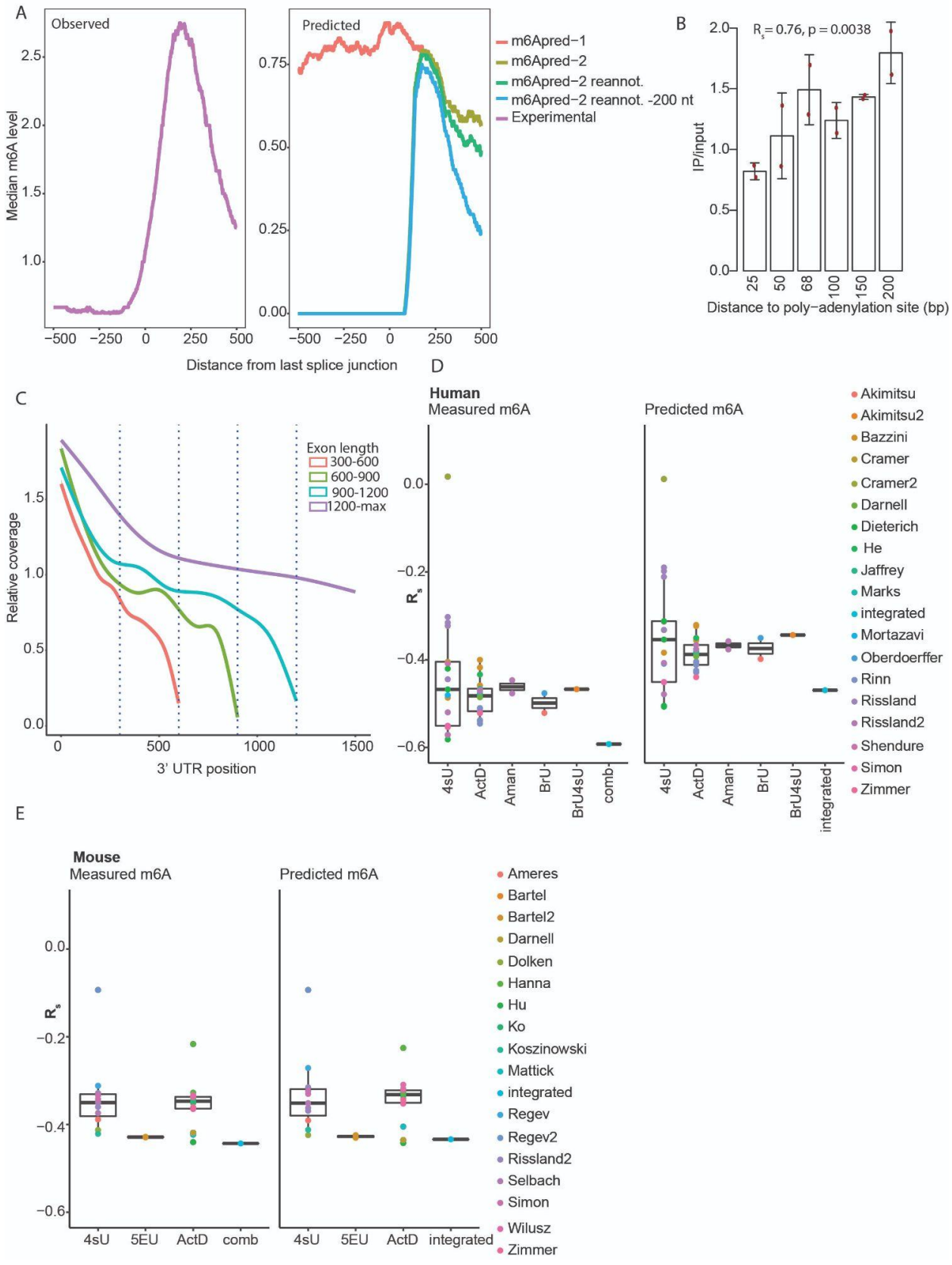
A
Observed
Predicted
Median m6A level
Distance from last splice junction
m6Apred−1
m6Apred−2
m6Apred−2 reannot.
m6Apred−2 reannot. -200 nt
Experimental
B
$R_s = 0.76, p = 0.0038$
IP/input
Distance to poly−adenylation site (bp)
C
Relative coverage
3' UTR position
Exon length
300-600
600-900
900-1200
1200-max
D
Human
Measured m6A
Predicted m6A
R_s
4sU
ActD
Aman
BrU
BrU4sU
comb
integrated
Akimitsu
Akimitsu2
Bazzini
Cramer
Cramer2
Darnell
Dieterich
He
Jaffrey
Marks
integrated
Mortazavi
Oberdoerffer
Rinn
Rissland
Rissland2
Shendure
Simon
Zimmer
E
Mouse
Measured m6A
Predicted m6A
R_s
4sU
5EU
ActD
comb
integrated
Ameres
Bartel
Bartel2
Darnell
Dolken
Hanna
Hu
Ko
Koszinowski
Mattick
integrated
Regev
Regev2
Rissland2
Selbach
Simon
Wilusz
Zimmer

Figure S5. Related to Fig. 3 and Fig. 4. **A)** Metagene profiles depicting median m6A signal intensity either on the basis of experimental measurements (left) obtained in A549 cells, or based on computational predictions either by (1) m6Apred-1 - note that in this case the decline reflects the decreased density of eligible consensus motifs, (2) m6Apred-2, prohibiting m6A from junction-proximal regions but not from regions adjacent to the transcript end, (3) m6Apred-2, but not allowing methylation to occur beyond the transcript termination site as based on re-annotation of transcript ends using the 'input' data, and (4) as in (3) but prohibiting methylation in the 200 most terminal nucleotides of the re-annotated transcripts. Metagene plots were derived for ~6000 genes, filtered to both ensure adequate expression levels and requiring an annotated last exon >500 nt to ensure that each of the genes spans the entirety of the region depicted in the X axis. **B)** M6A IP/input ratio as a function of the distance of the Calm3 3' m6A site to the poly-adenylation site. Experiment was based on two biological replicates, the red dots show individual measurements, the error bars show the standard deviation of the mean. **C)** Metagene coverage plots along the last annotated exon within genes (harboring the 3' UTR), binned based on annotated length of the last exon. Substantial drops in coverage are observed long before the annotated ends, suggesting that the actual 3' UTR is typically shorter than its annotated version. **D)** mRNA stability measurements of 16 human studies (collected in [19]) correlated against measured and predicted m6A levels. **E)** Same as D), for 18 mouse studies.

Supplementary figure 6.

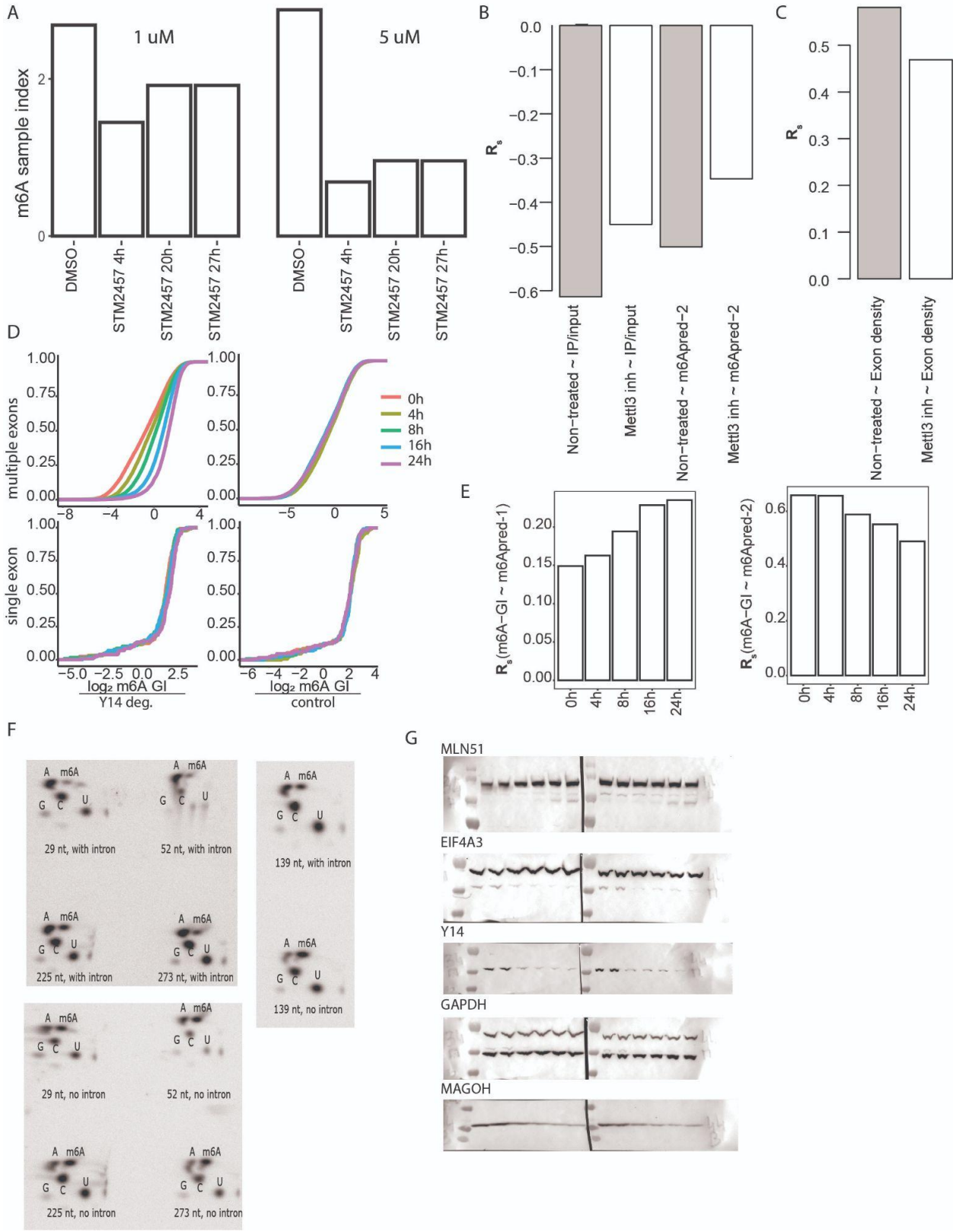

Figure S6. Related to Fig. 4. And Fig. 5. **A)** Barplots showing m6A sample indices [15], quantifying the relative amounts of m6A per sample, of DMSO or Mettl3 inhibitor (STM2457, 1μM (left panel) and 5μM (right panel)) treated samples

at the indicated time points. **B)** Barplot indicating Spearman correlation coefficients of the degradation rate of actD treated, and in parallel Mettl3 inhibitor treated or non-treated (DMSO treated) HEK293T cells with measured m6A levels and predicted m6A levels. **C)** Same as **B**, showing correlation with exon density. **D)** Cumulative plot of the log2 transformed m6A-GIs from Fig. 5G. **E)** (left) Correlation between experimentally measured gene-level m6A levels and m6Apred-1 predicted counterparts (capturing the ability of m6Apred-1 to predict m6A variability **between** genes), plotted at indicated time points following Y14 depletion. (right) As on the top panel, but using m6Apred-2 based predictions. **F)** Uncut original SCARLET images. **G)** Uncut original Western blot gels related to Fig. 5A.

References

1. Perry, R.P., Kelley, D.E., Friderici, K., and Rottman, F. (1975). The methylated constituents of L cell messenger RNA: evidence for an unusual cluster at the 5’ terminus. Cell *4*, 387–394.

2. Garcia-Campos, M.A., Edelheit, S., Toth, U., Safra, M., Shachar, R., Viukov, S., Winkler, R., Nir, R., Lasman, L., Brandis, A., et al. (2019). Deciphering the “m6A Code” via Antibody-Independent Quantitative Profiling. Cell *178*, 731–747.e16.

3. Schwartz, S., Agarwala, S.D., Mumbach, M.R., Jovanovic, M., Mertins, P., Shishkin, A., Tabach, Y., Mikkelsen, T.S., Satija, R., Ruvkun, G., et al. (2013). High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. Cell *155*, 1409–1421.

4. Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature *485*, 201–206.

5. Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., and Jaffrey, S.R. (2012). Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3’ UTRs and near Stop Codons. Cell *149*, 1635–1646.

6. Ke, S., Alemu, E.A., Mertens, C., Gantman, E.C., Fak, J.J., Mele, A., Haripal, B., Zucker-Scharff, I., Moore, M.J., Park, C.Y., et al. (2015). A majority of m6A residues are in the last exons, allowing the potential for 3′ UTR regulation. Genes Dev. *29*, 2037–2053.

7. Xiao, W., Adhikari, S., Dahal, U., Chen, Y.-S., Hao, Y.-J., Sun, B.-F., Sun, H.-Y., Li, A., Ping, X.-L., Lai, W.-Y., et al. (2016). Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. Mol. Cell. 10.1016/j.molcel.2016.01.012.

8. Lence, T., Akhtar, J., Bayer, M., Schmid, K., Spindler, L., Ho, C.H., Kreim, N., Andrade-Navarro, M.A., Poeck, B., Helm, M., et al. (2016). m6A modulates neuronal functions and sex determination in Drosophila. Nature. 10.1038/nature20568.

9. Wei, G., Almeida, M., Pintacuda, G., and Coker, H. (2021). Acute depletion of METTL3 implicates N6-methyladenosine in alternative intron/exon inclusion in the nascent transcriptome. Genome.

10. Yang, X., Liu, Q.-L., Xu, W., Zhang, Y.-C., Yang, Y., Ju, L.-F., Chen, J., Chen, Y.-S., Li, K., Ren, J., et al. (2019). m6A promotes R-loop formation to facilitate transcription termination. Cell Research *29*, 1035–1038. 10.1038/s41422-019-0235-7.

11. Abakir, A., Giles, T.C., Cristini, A., Foster, J.M., Dai, N., Starczak, M., Rubio-Roldan, A., Li, M., Eleftheriou, M., Crutchley, J., et al. (2019). N 6 -methyladenosine regulates the stability of RNA:DNA hybrids in human cells. Nat. Genet. *52*, 48–55.

12. Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. Nature *505*, 117–120.

13. Zaccara, S., and Jaffrey, S.R. (2020). A Unified Model for the Function of YTHDF Proteins in Regulating m6A-Modified mRNA. Cell *181*, 1582–1595.e18.

14. Ke, S., Pandya-Jones, A., Saito, Y., Fak, J.J., Vågbø, C.B., Geula, S., Hanna, J.H., Black, D.L., Darnell, J.E., Jr, and Darnell, R.B. (2017). m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. Genes Dev. *31*, 990–1006.

15. Dierks, D., Garcia-Campos, M.A., Uzonyi, A., Safra, M., Edelheit, S., Rossi, A., Sideri, T., Varier, R.A., Brandis, A., Stelzer, Y., et al. (2021). Multiplexed profiling facilitates robust m6A quantification at site, gene and sample resolution. Nat. Methods *18*, 1060–1067.

16. Spies, N., Burge, C.B., and Bartel, D.P. (2013). 3′ UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. Genome Research *23*, 2078–2090. 10.1101/gr.156919.113.

17. Agarwal, V., and Shendure, J. (2020). Predicting mRNA Abundance Directly from Genomic Sequence Using Deep Convolutional Neural Networks. Cell Rep. *31*, 107663.

18. Sharova, L.V., Sharov, A.A., Nedorezov, T., Piao, Y., Shaik, N., and Ko, M.S.H. (2009). Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. DNA Res. *16*, 45–58.

19. Agarwal, V., and Kelley, D.R. The genetic and biochemical determinants of mRNA degradation rates in mammals. 10.1101/2022.03.18.484474.

20. Clark, M.B., Johnston, R.L., Inostroza-Ponta, M., Fox, A.H., Fortini, E., Moscato, P., Dinger, M.E., and Mattick, J.S. (2012). Genome-wide analysis of long noncoding RNA stability. Genome Res. *22*, 885–898.

21. Le Hir, H., Nott, A., and Moore, M.J. (2003). How introns influence and enhance eukaryotic gene expression. Trends Biochem. Sci. *28*, 215–220.

22. Buchman, A.R., and Berg, P. (1988). Comparison of intron-dependent and intron-independent gene expression. Mol. Cell. Biol. *8*, 4395–4405.

23. Proudfoot, N.J., Furger, A., and Dye, M.J. (2002). Integrating mRNA processing with transcription. Cell *108*, 501–512.

24. Luo, M.J., and Reed, R. (1999). Splicing is required for rapid and efficient mRNA export in metazoans. Proc. Natl. Acad. Sci. U. S. A. *96*, 14937–14942.

25. Dwyer, K., Agarwal, N., Pile, L., and Ansari, A. (2021). Gene Architecture Facilitates Intron-Mediated Enhancement of Transcription. Front Mol Biosci *8*, 669004.

26. Matsumoto, K., Wassarman, K.M., and Wolffe, A.P. (1998). Nuclear history of a pre-mRNA determines the translational activity of cytoplasmic mRNA. The EMBO Journal *17*, 2107–2121. 10.1093/emboj/17.7.2107.

27. Singh, G., Kucukural, A., Cenik, C., Leszyk, J.D., Shaffer, S.A., Weng, Z., and Moore, M.J. (2012). The Cellular EJC Interactome Reveals Higher-Order mRNP Structure and an EJC-SR Protein Nexus. Cell *151*, 915–916.

28. Le Hir, H., Saulière, J., and Wang, Z. (2016). The exon junction complex as a node of post-transcriptional networks. Nat. Rev. Mol. Cell Biol. *17*, 41–54.

29. Adivarahan, S., Livingston, N., Nicholson, B., Rahman, S., Wu, B., Rissland, O.S., and Zenklusen, D. (2018). Spatial Organization of Single mRNPs at Different Stages of the Gene Expression Pathway. Mol. Cell *72*, 727–738.e5.

30. Metkar, M., Ozadam, H., Lajoie, B.R., Imakaev, M., Mirny, L.A., Dekker, J., and Moore, M.J. (2018). Higher-Order Organization Principles of Pre-translational mRNPs. Mol. Cell *72*, 715–726.e3.

31. Khong, A., and Parker, R. (2018). mRNP architecture in translating and stress conditions reveals an ordered pathway of mRNP compaction. J. Cell Biol. *217*, 4124–4140.

32. Woodward, L.A., Mabin, J.W., Gangras, P., and Singh, G. (2017). The exon junction complex: a lifelong guardian of mRNA fate. Wiley Interdiscip. Rev. RNA *8*. 10.1002/wrna.1411.

33. Gehring, N.H., Lamprinaki, S., Hentze, M.W., and Kulozik, A.E. (2009). The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay. PLoS Biol. *7*, e1000120.

34. Schwartz, S., Mumbach, M.R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G.G., Mertins, P., Ter-Ovanesyan, D., Habib, N., Cacchiarelli, D., et al. (2014). Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. Cell Rep. *8*, 284–296.

35. Chen, K., Wei, Z., Zhang, Q., Wu, X., Rong, R., Lu, Z., Su, J., de Magalhães, J.P., Rigden, D.J., and Meng, J. (2019). WHISTLE: a high-accuracy map of the human N6-methyladenosine (m6A) epitranscriptome predicted using a machine learning approach. Nucleic Acids Res. 10.1093/nar/gkz074.

36. Linder, B., Grozhik, A.V., Olarerin-George, A.O., Meydan, C., Mason, C.E., and Jaffrey, S.R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat. Methods. 10.1038/nmeth.3453.

37. Tegowski, M., Flamand, M.N., and Meyer, K.D. (2022). scDART-seq reveals distinct m6A signatures and mRNA methylation heterogeneity in single cells. Mol. Cell. 10.1016/j.molcel.2021.12.038.

38. Liu, N., Parisien, M., Dai, Q., Zheng, G., He, C., and Pan, T. (2013). Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA *19*, 1848–1856.

39. Hu, L., Liu, S., Peng, Y., Ge, R., Su, R., Senevirathne, C., Harada, B.T., Dai, Q., Wei, J., Zhang, L., et al. (2022). m6A RNA modifications are measured at single-base resolution across the mammalian transcriptome. Nat. Biotechnol. *40*, 1210–1219.

40. Körtel, N., Rücklé, C., Zhou, Y., Busch, A., Hoch-Kraft, P., Sutandy, F.X.R., Haase, J., Pradhan, M., Musheev, M., Ostareck, D., et al. (2021). Deep and accurate detection of m6A RNA modifications using miCLIP2 and m6Aboost machine learning. Nucleic Acids Res. *49*, e92.

41. Liu, C., Sun, H., Yi, Y., Shen, W., Li, K., Xiao, Y., Li, F., Li, Y., Hou, Y., Lu, B., et al. (2022). Absolute quantification of single-base m6A methylation in the mammalian transcriptome using GLORI. Nat. Biotechnol. 10.1038/s41587-022-01487-9.

42. Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A.V., Patil, D.P., Linder, B., Pickering, B.F., Vasseur, J.-J., Chen, Q., et al. (2016). Reversible methylation of m6Am in the 5′ cap controls mRNA stability. Nature *541*, 371–375.

43. Akichika, S., Hirano, S., Shichino, Y., Suzuki, T., Nishimasu, H., Ishitani, R., Sugita, A., Hirose, Y., Iwasaki, S., Nureki, O., et al. (2019). Cap-specific terminal N 6-methylation of RNA by an RNA polymerase II-associated methyltransferase. Science *363*. 10.1126/science.aav0080.

44. Keith, J.M., Ensinger, M.J., and Moss, B. (1978). HeLa cell RNA (2‘-O-methyladenosine-N6-)-methyltransferase specific for the capped 5‘-end of messenger RNA. J. Biol. Chem. *253*, 5033–5039.

45. Eisenberg, E., and Levanon, E.Y. (2003). Human housekeeping genes are compact. Trends Genet. *19*, 362–365.

46. Eisen, T.J., Eichhorn, S.W., Subtelny, A.O., Lin, K.S., McGeary, S.E., Gupta, S., and Bartel, D.P. (2020). The Dynamics of Cytoplasmic mRNA Metabolism. Mol. Cell *77*, 786–799.e10.

47. Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., and Wu, L. (2016). YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. Nat. Commun. *7*, 1–11.

48. Yankova, E., Blackaby, W., Albertella, M., Rak, J., De Braekeleer, E., Tsagkogeorga, G., Pilka, E.S., Aspris, D., Leggate, D., Hendrick, A.G., et al. (2021). Small-molecule inhibition of METTL3 as a strategy against myeloid leukaemia. Nature *593*, 597–601.

49. Schueler, M., Munschauer, M., Gregersen, L.H., Finzel, A., Loewer, A., Chen, W., Landthaler, M., and Dieterich, C. (2014). Differential protein occupancy profiling of the mRNA transcriptome. Genome Biol. *15*, R15.

50. Nabet, B., Roberts, J.M., Buckley, D.L., Paulk, J., Dastjerdi, S., Yang, A., Leggett, A.L., Erb, M.A., Lawlor, M.A., Souza, A., et al. (2018). The dTAG system for immediate and target-specific protein degradation. Nat. Chem. Biol. *14*, 431–441.

51. Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J. *20*, 4987–4997.

52. Zuckerman, B., Ron, M., Mikl, M., Segal, E., and Ulitsky, I. (2020). Gene Architecture and Sequence Composition Underpin Selective Dependency of Nuclear Export of Long RNAs on NXF1 and the TREX Complex. Mol. Cell *79*, 251–267.e6.

53. Bokar, J.A., Shambaugh, M.E., Polayes, D., Matera, A.G., and Rottman, F.M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. RNA *3*, 1233–1247.

54. Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Séraphin, B., and Le Hir, H. (2005).

The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. Nat. Struct. Mol. Biol. *12*, 861–869.

55. Saulière, J., Murigneux, V., Wang, Z., Marquenet, E., Barbosa, I., Le Tonquèze, O., Audic, Y., Paillard, L., Roest Crollius, H., and Le Hir, H. (2012). CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. Nat. Struct. Mol. Biol. *19*, 1124–1131.

56. Chen-Kiang, S., Nevins, J.R., and Darnell, J.E., Jr (1979). N-6-methyl-adenosine in adenovirus type 2 nuclear RNA is conserved in the formation of messenger RNA. J. Mol. Biol. *135*, 733–752.

57. Xu, W., He, C., Kaye, E.G., Li, J., Mu, M., Nelson, G.M., Dong, L., Wang, J., Wu, F., Shi, Y.G., et al. (2022). Dynamic control of chromatin-associated m6A methylation regulates nascent RNA synthesis. Mol. Cell *82*, 1156–1168.e7.

58. Louloupi, A., Ntini, E., Conrad, T., and Ørom, U.A.V. (2018). Transient N-6-Methyladenosine Transcriptome Sequencing Reveals a Regulatory Role of m6A in Splicing Efficiency. Cell Rep. *23*, 3429–3437.

59. Liu, J. ’e, Li, K., Cai, J., Zhang, M., Zhang, X., Xiong, X., Meng, H., Xu, X., Huang, Z., Peng, J., et al. (2020). Landscape and Regulation of m6A and m6Am Methylome across Human and Mouse Tissues. Mol. Cell *77*, 426–440.e6.

60. Mayr, C., and Bartel, D.P. (2009). Widespread shortening of 3’UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell *138*, 673–684.

61. Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., and Burge, C.B. (2008). Proliferating cells express mRNAs with shortened 3’ untranslated regions and fewer microRNA target sites. Science *320*, 1643–1647.

62. Yoder, J.A., Walsh, C.P., and Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. *13*, 335–340.

63. Deniz, Ö., Frost, J.M., and Branco, M.R. (2019). Regulation of transposable elements by DNA modifications. Nat. Rev. Genet. *20*, 417–431.

64. Chelmicki, T., Roger, E., Teissandier, A., Rucli, S., Dossin, F., Dura, M., Fouassier, C., Lameiras, S., and Bourc’his, D. (2020). m6A RNA methylation regulates the fate of endogenous retroviruses. 2020.03.24.005488. 10.1101/2020.03.24.005488.

65. Liu, J., Dou, X., Chen, C., Chen, C., Liu, C., Xu, M.M., Zhao, S., Shen, B., Gao, Y., Han, D., et al. (2020). N6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. Science *367*, 580–586.

66. Liu, J., Gao, M., He, J., Wu, K., Lin, S., Jin, L., Chen, Y., Liu, H., Shi, J., Wang, X., et al. (2021). The RNA m6A reader YTHDC1 silences retrotransposons and guards ES cell identity. Nature *591*, 322–326.

67. Chen, C., Liu, W., Guo, J., Liu, Y., Liu, X., Liu, J., Dou, X., Le, R., Huang, Y., Li, C., et al. (2021). Nuclear m6A reader YTHDC1 regulates the scaffold function of LINE1 RNA in mouse ESCs and early embryos. Protein Cell *12*, 455–474.

68. Xu, W., Li, J., He, C., Wen, J., Ma, H., Rong, B., Diao, J., Wang, L., Wang, J., Wu, F., et al. (2021). METTL3 regulates heterochromatin in mouse embryonic stem cells. Nature *591*, 317–321.

69. McIntyre, A.B.R., Gokhale, N.S., Cerchietti, L., Jaffrey, S.R., Horner, S.M., and Mason, C.E. Limits in the detection of m6A changes using MeRIP/m6A-seq. 10.1101/657130.

70. Yonatan Stelzer, A., Shakti Shivalila, C., Soldner, F., Markoulaki, S., Stelzer, Y., and Jaenisch, R. (2015). Tracing Dynamic Changes of DNA Methylation at Single-Cell Resolution In Brief Tracing Dynamic Changes of DNA Methylation at Single-Cell Resolution. Cell *163*, 218–229.

71. Uzonyi, A., Nir, R., and Schwartz, S. (2022). Cloning of DNA oligo pools for in vitro expression. STAR Protocols *3*, 101103.

72. Shishkin, A.A., Giannoukos, G., Kucukural, A., Ciulla, D., Busby, M., Surka, C., Chen, J., Bhattacharyya, R.P., Rudy, R.F., Patel, M.M., et al. (2015). Simultaneous generation of many RNA-seq libraries in a single reaction. Nat. Methods *12*, 323–325.

73. Hocq, R., Paternina, J., Alasseur, Q., Genovesio, A., and Le Hir, H. (2018). Monitored eCLIP: high accuracy mapping of RNA-protein interactions. Nucleic Acids Res. *46*, 11553–11565.

74. Grosjean, H., Droogmans, L., Roovers, M., and Keith, G. (2007). Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates. Methods Enzymol. *425*, 55–101.

75. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

76. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. BIOINFORMATICS APPLICATIONS NOTE *26*, 841–842.