Comprehensive profiling of *N*^{6}-methyladnosine (m^{6}A) readouts reveals novel m^{6}A readers that regulate human embryonic stem cell differentiation

Zhou Huang[1,2,3,4,5,#], Rucong Liu[1,2,3,6,#], Zibaguli Wubulikasimu[3,7,#], Wanqing Zhao[1,2,8,#], Jiaqi Huang[2,4], Jiaxuan Wang[3], Rui Fan[1,2], Wei Kong[2,4], Qinghua Cui[1,2], Yang Li[3,*], Yuan Zhou[1,2,9,*]

[1] Department of Biomedical Informatics, School of Basic Medical Sciences, Peking University, Beijing 100191, China.

[2] State Key Laboratory of Vascular Homeostasis and Remodeling, Peking University, Beijing 100191, China.

[3] Department of Cell Biology, School of Basic Medical Sciences, Peking University Stem Cell Research Center, Peking University, Beijing 100191, China.

[4] Department of Physiology and Pathophysiology, School of Basic Medical Sciences, Peking University, Beijing 100191, China.

[5] Cancer Center, Daping Hospital, Army Medical University, Chongqing 400042, China.

[6] National Institute of Geriatrics, National Health Commission, Beijing Hospital, Beijing 100730, China.

[7] State Key Laboratory of Female Fertility Promotion, Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing 100191, China.

[8] Shenzhen Key Laboratory of Cardiovascular Disease, Fuwai Hospital Chinese Academy of Medical Sciences, Shenzhen 518057, China.

[9] Department of Integrated Traditional Chinese and Western Medicine, School of Basic Medical Sciences, Peking University, Beijing 100191, China.

* Correspondence authors: Yuan Zhou (zhouyuanbioinfo@hsc.pku.edu.cn) and Yang Li (liyang@hsc.pku.edu.cn)

# Contributed equally.

Abstract

N^6-methyladenosine (m^{6}A) methylation has emerged as a prevalent RNA modification that extensively impacts various physiological and pathological processes via various post-transcriptional readout effects in mammals. High-throughput methylome profiling has outlined the landscape of m^{6}A modification sites, but their downstream readouts require comprehensive investigation. To this end, we systematically assessed the effects of m^{6}A on mRNA half-life, translation efficiency, and alternative splicing across five cell lines (A549, HEK293T, HUVEC, JURKAT, and human embryonic stem cells (hESCs)) using actinomycin D-disrupted temporal transcriptome sequencing, ribosome sequencing, and ultra-high-depth transcriptome sequencing, respectively. Our analysis, coupled with the integration of public and re-profiled m^{6}A methylome data, revealed high cell type specificity in m^{6}A readouts where m^{6}A level alone is insufficient to predict m^{6}A readouts. Machine learning models focused on the RNA binding protein (RBP) binding context of m^{6}A sites demonstrated substantial predictive ability of m^{6}A readouts while prioritizing putative m^{6}A readers from their informative RBP features. Four novel m^{6}A readers (DDX6, FUBP3, FXR2, and L1TD1) were identified and validated through m^{6}A RNA pull-down assays and transcriptome-wide RBP binding site mapping. Notably, FUBP3, FXR2 and L1TD1 were found to regulate hESC differentiation without impairing self-renewal, underscoring their critical roles in stem cell biology. Together, this study bridges the gap in understanding m^{6}A functional readouts and lays the groundwork for future research on m^{6}A-mediated stem cell fate decisions.

Introduction

The mammalian transcriptome is extensively regulated by a variety of chemical modifications, among which N^6-methyladenosine (m^6A) has been proven as one of the most prevalent modifications [1-3]. The on-and-off state and quantification of m^6A methylation are closely related to the reversible enzymatic reaction system behind. Among them, METTL3 is able to independently methylate substrate RNA *in vitro*, and *METTL3* knockdown leads to a significant reduction of transcriptome m^6A methylation level, together highlighting METTL3 as the core component of m^6A writer complex [4]. On the other hand, the m^6A modification can be removed by eraser proteins such as FTO and ALKBH5 [5]. Through sophisticated procedures executed by the m^6A writers and erasers, a complicated epitranscriptome is formed. To date, near 500,000 human single-nucleotide m^6A sites has been identified [6]. Comprehensive m^6A profiling by MeRIP-seq has revealed distinct but partly shared m^6A profiles across different cell types [7,8]. Nonetheless, considering the different transcriptome context among different cell types, direct and systematic measure of the regulatory effects of m^6A is necessary. Unfortunately, there remains a lack of sizable high-throughput profiling of m^6A's effects, and consequently, to what extent the m^6A regulatory effects are shared or distinct between different cell types is not clear.

The m^6A modification is also recognized as the top RNA modification with the greatest number of known recognizing partners, i.e. m^6A readers. This comprehensive and sophisticated reader toolkit gives rise to the diversity and complexity of the m^6A regulatory effects, i.e. m^6A readouts. Although with controversial about functional specificity [9,10], it is recognized that several key m^6A readers play vital roles in gene expression regulations. For example, YTHDF1 would promote the translation of mRNAs with m^6A modification, while YTHDF2 would reduce the half-life of m^6A target transcripts [11,12]. Nuclear YTHDC1 could recruit pre-mRNA splicing factor SRSF3 and inhibit the binding of SRSF10 to mRNAs to influence exon inclusion events in alternative splicing [13]. IGF2BP family proteins constitute a new clade of m^6A readers that inhibit mRNA degradation and enhance translational efficiency [14]. Nonetheless, disruption of any one of these m^6A readers cannot completely abolish even single type of regulatory effects of m^6A [11-14], suggesting the existence of undiscovered m^6A readers. Indeed, by leveraging mass spectrometry, dozens of proteins have shown favorable presence in the m^6A-associated proteome, although the direct m^6A recognition ability of these

putative m^{6}A readers need further validations [15]. Until recent, novel m^{6}A readers have been unveiled. For example, ELAVL1 (HuR) preferentially binds elements in the 3'-UTR that are enriched for AU or U, thereby regulating mRNA stability [16]. TARDBP (TDP43) is identified as a m^{6}A-binding protein that regulate neurodegeneration [17].

As the key epitranscriptomic factor, m^{6}A also exert substantial impacts on key properties of pluripotent stem cells. In fact, early demonstrations of m^{6}A regulatory function on embryonic stem cells (ESCs) constitute the pioneer cases supporting the functional importance of m^{6}A modification. In 2014, Wang et al. reported that *Mettl3* inhibition decreases the expression of pluripotency-associated genes (e.g., *Nanog* and *Sox2*) but up-regulates the expression of developmental markers (e.g., *Sox17*), and ultimately promotes mouse ESCs (mESCs) differentiation [29]. In induced pluripotent stem cells (iPSCs), Chen et al. also showed that increased m^{6}A formation promotes cell reprogramming to pluripotency [18]. However, there are also inconsistent results, for example, in 2014, Batista et al. found that depletion of mouse and human *Mettl3* promoted self-renewal of mESCs and human ESCs (hESCs) by prolonging *Nanog* expression upon differentiation. Subsequently, Geula et al. showed that deletion of m^{6}A could lead to either defected or accelerated differentiation depending on the pluripotent state of the mESCs, i.e. naive or primed [19]. These results imply complexity in the functional readout of the m^{6}A tags in ESCs. Besides, various m^{6}A readers also play essential roles in regulating ESC pluripotency. Knockout of *YTHDF* genes inhibits mESCs exit from pluripotency as evidenced by poor differentiation during teratoma and embryoid body formation [20]. Deletion of YTHDC1 initiates cellular reprogramming to a 2-cell like state[21]. These studies emphasize the importance of the m^{6}A modification in regulating the self-renewal and differentiation process of ESCs. However, which and how m^{6}A readers exert the specific regulatory effects of hESC differentiation has not been sufficiently investigated.

Together, m^{6}A has emerged as the key gene expression regulatory machinery but its regulatory effects and mechanisms have not been fully elucidated. Whether m^{6}A showed similar or distinct readouts across different cell types? Are there new readers behind the regulatory effects of m^{6}A. And if so, how could these m^{6}A readers influence basic stem cell functions including self-renewal and differentiations? To this end, we comprehensively profiled the three main layers of m^{6}A readouts (mRNA half-life, translational efficiency and alternative splicing) across five common cell types of

different origin (including cancerous A549 and JURKAT, and non-cancerous hESC, HEK293T and human umbilical vein endothelial cells (HUVECs)). Besides, quantitative profiling of m^6A methylation by GLORI, if applicable, was also performed. To exploit these resources, machine learning modeling of the m^6A readouts was conducted, where putative m^6A readers were suggested based on the interpretation of model features. Finally, we validated a few of these putative m^6A readers and unveiled their specific impacts on hESC differentiation.

Methods

Cell culture

HEK293T and A549 cell lines were cultured in DMEM medium (Gibco, catalog#C11995500BT) supplemented with 10% fetal bovine serum (FBS, Gemini, catalog#900-108) and antibiotics. JURKAT cell line was maintained in RPMI 1640 medium (Gibco, catalog#C11875500BT) with 10% FBS (Gemini, catalog#900-108) and antibiotics. hESCs were cultivated in serum-free, defined mTeSR1 medium (STEMCELL, catalog#85850) and grown on plates coated with growth-factor-reduced Matrigel. hESCs were passaged with non-enzymatic dissociation by 0.5 mM EDTA (Beyotime, catalog#C0196-100ml). HUVECs were isolated as described [22]. The clean vein of the umbilical cord was first lavaged with 0.25% trypsin at 37°C for 15 min. HUVECs were collected and suspended in Endothelial Cell Medium (ECM, ScienCell, catalog#1001). Isolated HUVECs were cultured between passage 4–6 in ECM for experiment. Cells were grown at 37°C with 5% CO_2.

METTL3 knockdown, transfection and flow cytometry

The doxycycline-controlled Tet-Off/On gene expression systems were used for regulating the activity of METTL3 in cell lines. The sequence targeting *METTL3* were synthesized and inserted into the lentivirus vector GV307 (TetIIP-TurboRFP-MCS-Ubi-TetR-IRES-Puromycin) and the lentiviruses were purchased from GeneChem (Shanghai, China). Then according to the characteristics of different cell lines, lentiviruses transfection was performed using the HiTransG A/P developed by GeneChem. Cells transfected with shRNA lentivirus or negative control lentivirus were cultured for 48 h. The target sequence of *METTL3* was CAAGGAACAATCCATTGTT and the negative control sequence

136 of insertion was TCTCGCTTGGGCGAGAGTAAG. Other information of the oligo synthesis was
137 included in **Supplementary Table S1**. After the doxycycline induction for 48 h, the transfected cells
138 expressed red fluorescence (RFP) for roughly estimating the overall transfection efficiency. Cell
139 samples were resuspended in MACS buffer for fluorescence-activated cell sorting (FACS, BD
140 FACSAria™ III) to obtain high RFP expression transfected cells.

141

142 **RNA extraction**

143 Total RNA was isolated from cell samples using TRIzol (Invitrogen, catalog#15596018) following
144 the manufacturer's instruction. Briefly, cells were lysed with TRIzol reagent at room temperature
145 until the cell deposits disappear and chloroform was added. After centrifuge for 15 min at 12,000×g
146 at 4°C, the aqueous phase was transferred to the new EP tube, mixed with isopropanol and stand at
147 room temperature for 10~15 min. Then centrifuge for 10 min at 12,000×g at 4°C, the supernatant was
148 discarded and the RNA precipitate was washed with 75% ethanol. After centrifuge for 5 min at
149 7,500×g at 4°C and dry, RNA was dissolved with RNase-free water and its concentration was
150 measured by Thermo Scientific NanoDrop.

151

152 **Real-time qPCR for mRNA quantification**

153 Total RNA was extracted from transected cells scribed in 'RNA-extraction', then reverse transcribed
154 using PrimeScript RT Master Mix (TaKaRa, catalog#RR036B). All primers used for RT-qPCR were
155 listed in **Supplementary Table S2**. All RT-qPCRs were run on Bio-Rad CFX384 real-time PCR
156 detection system using PowerUp SYBR Green Master Mix (ThermoFisher Scientific,
157 catalog#A25742). GAPDH was used as internal control to normalize the data across different samples.

158

159 **Western blot**

160 Cell samples were collected, washed with ice-cold PBS and fully lysed with RIPA lysis buffer
161 containing protease inhibitors (Beyotime, catalog#P0013C) on ice. After centrifuge for 30 min at
162 12,000×g at 4°C, supernatant was collected and protein amounts were quantified using the BCA
163 Protein Assay Kit (Beyotime, catalog#P0012S). After boiling for 10 min, equal amounts of protein

were loaded and separated by 10% SDS-PAGE. After electrophoresis, they were transferred to PVDF membranes (Merck-Millipore, catalog#ISEQ00010). Membranes were blocked with 5% non-fat milk (Beyotime, catalog#P0216-1500g) for 1 h and then incubated at 4°C overnight with the primary antibodies. After incubating 1 h with secondary antibodies at room temperature, images were obtained using the Gel Doc EZ Imager (Bio-Rad). GAPDH was used as the loading control. All antibodies used for western blot were listed in **Supplementary Table S3**.

High-throughput profiling of m^6A readouts

Unless otherwise stated, the m^6A readouts were measured by comparing shControl (normal m^6A level) versus shMETTL3 (disrupted m^6A level) to intuitively depict the regulatory readouts in presence of m^6A. The sequencing service related to m^6A readout estimations was provided by Novogene Biotech (Tianjin, China). The experimental and computational pipelines of each readout category is described as follows:

a. mRNA half-life readout. Cells transfected with *METTL3* shRNA or negative control shRNA were re-seeded into 6 well plates. After the doxycycline induction for 48 h, actinomycin D was added to 2 μg/mL at 0 h, 1 h, 3 h and 6 h before total RNA extraction. Before construction of the library with mRNA sample, ERCC RNA spike-in control mix (Invitrogen, catalog# 4456740) was added to each sample following the manufacturer's instruction. NEBNext Ultra RNA Library Prep Kit for Illumina (NEB #E7770) and NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB #E7760) were used with 400 ng of total RNA for the construction of sequencing libraries. ERCC spike-in were added to total RNA before library construction, following the protocol recommended by ERCC spike-in kit manufacturer. RNA libraries were prepared for sequencing using standard Illumina protocols. Two biological replicates were prepared for each time point. For each sample collected at 1 h, 3 h and 6 h, a 15G amount of paired-end sequencing reads was obtained. For 0 h sample, we required 60G amount of sequences to ensure ultra-deep RNA sequencing that is necessary for systematic alterative splicing assay. This differences in data size were properly controlled by the aforementioned ERCC spike-ins. More specifically, quality control, adapter cutting and low-quality sequence removal were performed by FastQC, fastp and FASTX toolkits. The clean reads were mapped to human genome (hg38) by HISAT2 software (v2.2.1) and the expression values were calculated by featureCounts command of RSubread package with suggested parameter of the software. The

absolute quantification of each sample was performed by using ERCC RNA spike-ins as the reference, following Wang et al. [11]. The mRNA half-life was also estimated following the method modified from [11]. The degradation rate of RNA k was estimated by:

$$\log_2(A_t/A_0 \cdot \delta) = -kt$$

where t is transcription inhibition time (h), A_t and A_0 represent mRNA quantity at time t and time 0, δ is a regulating factor in case the RNA:ERCC ratio is prominently deviated from the expectation. k values of 3 h and 6 h were calculated: time 3 h versus time 0 h, and time 6 h versus time 0 h. A 1.5-fold change of half-life threshold was applied for subsequent analysis.

b. translational efficiency readout. The translational efficiency was measured by ribosome profiling sequencing (Ribo-seq). The reported protocols [23] was adapted with the following modifications: Before cells collection, cycloheximide (CHX, HARVEYBIO, catalog#C21865-100mg) was added to the culture media at 100 μg/mL for 5 min. Eighty million cells from each group were collected, rinsed in ice-cold PBS with 100 μg/mL CHX. After centrifuge for 5 min at 500×g at 4°C, cells were resuspended in cell freezing medium with 100 μg/mL CHX for dry ice transportation. Two replicates were prepared and the reads were processed and analyzed with the similar pipeline to the above part a. To comply with the Ribo-seq convention, the R1 reads were used for subsequent analysis. The Xtail software (v1.1.5) was utilized to measure the changes of translational efficiency with default parameter [24]. To ensure sufficient number of genes for subsequent analysis, a relaxed thresholds of 1.5-fold change and $P < 0.05$ were applied. We also considered public Ribo-seq data from GEO database for the machine learning modeling described later, including those from HeLa (GSE117299), HepG2 (GSE121952), Huh7 (GSE155447) and MOLM13 (GSE98623). These data were analyzed with the same pipeline.

c. alternative splicing readout. We exploited the ultra-deep sequencing of 0 h samples as described in the above part a to measure alternative splicing events. Two replicates were prepared and the reads were processed and analyzed with the similar pipeline to the above part a. The alternative splicing events and exon inclusion ratio (Psi) were calculated by rMATS-turbo (v4.1.0, default parameter) [25]. A change of exon inclusion ratio (delta-Psi) no less than 0.1, together with FDR < 0.05, were selected as the significant alternative splicing event.

d. Functional analysis. The functional enrichment of Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), WikiPathways and MSigDB functional gene sets were performed by

224 ClusterProfiler with default parameter and cutoffs [26]. Gene dependency score was obtained from
225 DepMap database [27].

226 High-throughput profiling of m^6A methylation sites

227 ***a. GLORI sequencing.*** Recently, Liu et al. developed GLORI, an antibody-independent sequencing
228 technique for single-base quantitative detection of m^6A methylation [28]. The main idea of GLORI is
229 to find a catalytic system of glyoxal and nitrite through chemical reaction combination screening,
230 which efficiently deaminate unmethylated adenosines to inosines (A-to-I), but not methylated ones.
231 We used this state-of-the-arts technique for the quantitative profiling of m^6A methylation in A549,
232 hESC and JURKAT. The GLORI-identified m^6A sites in HEK293T was directly obtained from the
233 original study [28]. HUVEC was not included because of the required large amount of input cells by
234 GLORI sequencing service. GLORI sequencing service was provided by Cloud-Seq Biotech
235 (Shanghai, China). Total RNA from the samples was extracted using TRIzol, and RNA was purified
236 according to the MEGA clear transcription purification kit (Thremo fisher, catalog#AM1908). RNA
237 was incubated with fragmentation buffer (NEB, catalog#E6150S) for 4 min at 94°C and glyoxal
238 solution and DMSO were added, and the mixture was incubated for 30 min at 50°C in a preheated
239 thermal cycler. minutes, and at the end of the incubation, the tubes were brought to room temperature,
240 and then saturated H_3BO_3 (Sigma-Aldrich, catalog#B0394) solution was added. The incubation was
241 continued at 50 °C for 30 min, and was then added to the pre-configured deamination buffer (750
242 mM $NaNO_2$ (Sigma-Aldrich, 31443), 40 mM MES (pH 6.0), and 10 μL of glyoxal solution for a total
243 of 8 h of incubation. Finally, RNA was purified by ethanol precipitation, and sequencing libraries
244 were constructed by GenSeq Low Input RNA Library Prep Kit (GenSeq, Inc.) following the process
245 of the instructions provided by the manufacturer. The constructed sequencing libraries were subjected
246 to quality control and quantification by BioAnalyzer 2100 system (Agilent Technologies, USA),
247 followed by 150 bp paired-end sequencing.

248 ***b. GLORI data analysis.*** Quality control of raw reads was first performed using Q30. Then cutadapt
249 software (v1.9.3) was used for read de-multiplex and removal of low-quality reads. PCR duplicates
250 were removed based on Unique Molecular Identifiers (UMI) tags. Clean reads were aligned to the
251 reference genome (hg38) and BAM files were obtained using STAR (v2.7) and bowtie2 (v2.2.4)
252 software. To ensure acceptable m^6A site detection rate, we performed two rounds of sequencing, and
253 the aligned BAM files from these two rounds were combined by samtools (v1.12). Finally, the m^6A

sites were called and quantified by GLORI-tools [28] with parameter "--combine --rvs_fa -c 1 -C 0 -r 0.1 -p 0.05 -adp 0.05". Finally, only m^6A sites that showed modification rate > 0.1 for both replicates were retained [28].

c. Public MeRIP-seq methylation profiles. Currently, GLORI methylation profiles are not available for most cell lines. To obtain m^6A sites for other cell types, we exploited widely used MeRIP-seq-based m^6A methylation profiles [1,2]. To ensure comparable results, raw sequences of MeRIP-seq studies were obtained from GEO (as specified in **Supplementary Table S4**). The adapter trimming and quality control was performed by fastp, and the reads were aligned to human genome (hg38) by HISAT2. Finally, exomePeak2 was used to call m^6A peaks [6,29]. To obtain single-nucleotide m6A sites, we compared m6A peaks with known single nucleotide m^6A sites from m6A-Atlas V2.0 database [6] and GLORI technique by using bedtools (v2.31.0). Sites following the consensus m^6A motif DRACH were retained.

Machine learning modeling of readouts

a. Obtaining RNA binding protein (RBP) binding site annotations: Our m^6A readout prediction models were heavily relied on transcriptome-wide annotation of RBP binding sites. CLIP-seq and its derivates constitute the major source of such annotation. We obtained CLIP-seq (and derivates)-based RBP binding sites from POSTAR2 database [30]. We did not consider POSTAR3 because there was no RBP binding site score in the downloadable data of POSTAR3. Moreover, to make our annotations up-to-date, we obtained and re-analyzed raw CLIP-seq data from GEO, and adopted a pipeline similar to that of POSTAR database to process these data. Briefly, raw reads were filtered and trimmed by fastp and cutadpat software, and aligned to human genome (hg38) by using bowtie (v1.3.0, for short reads) or STAR (for longer reads). Samtools was used to process the aligned reads and the RBP binding regions were called by Piranha (v1.2.1) with recommended parameters [31]. We also obtained eCLIP-based annotation from ENCODE projects [32]. We noted that the default merging of biological replicates in ENCODE would result in very limited binding sites for some RBPs. To this end, we also used bedtools to obtain the union of RBP binding regions derived from both replicates to enhance the sensitivity and transcriptome coverage. Finally, RBP site annotations from non-CLIP techniques were directly obtained from the corresponding GEO accession. The summary of RBP binding site profiles were available in **Supplementary Table S5**.

b. Feature encoding and model training: For predicting half-life and translational efficiency readouts, eleven set of features were considered: 1) The score of RBP binding site with the nearest distance on the genome; 2) The score of RBP binding site with the nearest distance on the transcripts; 3) The minimum RBP-to-m^{6}A-site distance among different m^{6}A sites of the same gene on the genome; 4) The minimum RBP-to-m^{6}A-site distance on the transcripts; 5) The maximum RBP-to-m^{6}A-site distance on the genome; 6) The maximum RBP-to-m^{6}A-site distance on the transcripts; 7) The site to gene topology (e.g. minimum distance to start codon); 8) Tissue expression level, as measured by GTEx project [33]; 9) Gene dependency score, as record by DepMap database [27]; 10) Nearest distance to a 6-mer on transcript; 11) Maximum counts of a specific k-mer in ±100 nt range flanking known m^{6}A sites. As for the prediction of exon inclusion ratio alterations, we used the annotation of the closest m^{6}A site to the exon boundary to encode features (if any within 10kb of exon boundary). Since only one representative m^{6}A site was considered for each exon, no minimum/maximum operation of RBP features among multiple m6A sites on the same gene was required. Besides, as an exon-level prediction, gene features like GTEx tissue expression and DepMap gene dependency score were also not included.

c. Machine learning model: We exploited XGBoost [34] to model the readouts for each readout and each cell type. XGBoost has been successfully implemented in both bioinformatic prediction and clinical application [35,36]. Ten-fold cross-validation to train and test the model, and the overall performance was measure by receiver operating characteristic (ROC) curve and area under ROC curve (AUC). A heustic search of best combination of feature set was performed to measure the AUC gain of the feature set. And for RBP features, the feature importance score derived from the XGBoost models.

RNA pull-down

The RNA oligonucleotides [ss-A: 5'-CGUCUCGGACUCGGACUGCU-3'; ss-m^{6}A: 5'-CGUCUCGG(m^{6}A)CUCGG(m^{6}A)CUGCU-3'] with the same sequences of ss-A and ss-m^{6}A were synthesized by Rui Biotech (Beijing, China). RNA pull-down assay was conducted by the Pierce™ Magnetic RNA-Protein Pull-Down Kit (Thermo Fisher Scientific, catalog#20164). The included Thermo Scientific Pierce RNA 3' Desthiobiotinylation Kit (catalog#20163) was used to label the target RNA. Next, biotin-labeled RNA oligonucleotides containing adenosine or m^{6}A (50 pmol) were

314 immobilized onto 50 μL streptavidin magnetic beads in binding buffer at room temperature for 15-
315 30 min. RNA-conjugated streptavidin beads were then incubated with 200 μg total proteins from
316 target cells in binding buffer in a final volume of 100 μL at 4°C for 30-60 min. RNA-protein complex-
317 containing beads were washed three times with wash buffer. After extensive washing, RNA-binding
318 protein complexes were boiled and eluted in 1 × SDS loading buffer to separate on 10% SDS-PAGE
319 gels for western blot analysis.

320

321 **RBP binding site profiling by HyperTRIBE**

322 ***a. HyperTRIBE plasmids construction and transfection.*** HyperTRIBE is a promising high-
323 throughput sequencing technique for identifying potential binding site of an RBP. Leveraging
324 plasmid fusing target RBP with hyperactive ADAR, the RNA editing signals close to the RBP binding
325 site can be detected at transcriptome level through conventional RNA sequencing [37]. Here, the
326 hyperactive ADAR containing TRIBE plasmids were constructed by Rui Biotech (Beijing, China).
327 The control plasmid, mammalian expression plasmid, containing mCherry ADAR control and p2A
328 GFP reporter can be obtained from Addgene (Plasmid #154786). As for RBP-TRIBE plasmid, using
329 standard restriction enzyme-based cloning to clone known and putative m^6A reader RBPs (YTHDF1,
330 DDX6, FUBP3, FXR2 and L1TD1) into TRIBE plasmid. Sanger sequencing was used to confirm
331 proper insertion and then control and RBP-ADAR plasmids were infected into hESCs by transfection
332 reagent jetOPTIMUS (Polyplus, catalog#101000051). After transient transfection into target cells for
333 24 h, the cells containing RBP-ADAR fusion or ADAR alone proteins were sorted by GFP and
334 mCherry fluorescence and using TRIzol reagent to extract the total mRNA from the sorted cells.

335 ***b. TRIBE sequencing and analysis.*** RNA-sequencing for cells transfected with HyperTRIBE
336 plasmid were conducted by Novogene Biotech (Tianjin, China). Messenger RNA was purified from
337 total RNA using poly-T oligo-attached magnetic beads. After fragmentation, the first strand cDNA
338 was synthesized using random hexamer primers followed by the second strand cDNA synthesis. After
339 library quality control, quantified libraries were pooled and sequenced on Illumina NovaSeq with
340 paired-end 150-bp reads, providing sequencing data at 40G per library. According to Biswas et al.'s
341 protocol [38], firstly, cutadapt was used to remove the adaptor and low-quality reads. Clean reads were
342 then aligned to the reference genome (hg38) by using STAR software. Picard (v3.3.0) was used to
343 remove duplicates and sort the output BAM files. To identify editing sites, we used Bullseye, a set of

344 Perl scripts which was originally developed by Flamand et al. for analysis of editing events induced
345 by transfected plasmids [9]. BAM files were parsed to generate coverage matrices at each position in
346 the genome, excluding duplicate and multi-mapped reads. Editing sites for HyperTRIBE in individual
347 samples were then called by comparing A2I mutation rates at all positions within annotated exons
348 (hg38) with samples expressing ADAR alone. Sites edited at rates between 5% and 99%, edited to
349 higher levels than control cells (1.5 × for A2I), with at least two mutations, and that did not overlap
350 annotated SNP were kept for further analysis. The final list of sites was identified as those found in
351 three replicates and with an average editing rate of at least 5% in all samples where sites were called.
352 To identify the binding motif of the editing sites, we broadened each site by 500 bp in each direction
353 and then imported into HOMER software [39] for *de novo* motif identification. It is worth mentioning
354 that because we are interested in identification of m^6A-binding proteins, we further focused on the
355 proteins that have motifs containing 'RRACH', which represents the most common m6A consensus
356 sequence. To investigate the binding preference of the RBPs near m6A sites, BAM alignments were
357 further processed using deepTools [40]. Heatmaps and profiles were generated using the plotHeatmap
358 and plotProfile script from the deepTools suite.

359

360 **CRISPR-mediated *FUBP3*, *FXR2* and *L1TD1* knock out (KO)**

361 The CRISPR-Cas9 knock-out vector, containing cassettes for GFP, Cas9 endonuclease, and the
362 double guide RNA (gRNA) moiety designed to target FUBP3, FXR2 and L1TD1, was custom
363 synthesized by Guangzhou Regenverse Therapeutics Co., Ltd. (Guangzhou, China). After plasmid
364 transfection for 48 h, GFP+ hESCs were sorted using flow cytometry, and plated at low density for
365 single-colony isolation. Selected single colonies were tested by western blot for loss of protein.
366 Mutations were validated by sequencing products of PCR amplification of the regions flanking the
367 targeting sites.

368

369 **Cell proliferation assays**

370 ***a. Cell counting kit 8 (CCK8) assay.*** Human embryonic stem cells (hESCs) were digested into single
371 cell suspension using Accutase (Merck-Millipore, catalog#SCR005) and the number of cells were
372 counted by automated cell counter (RWD, catalog#C100). After counting, the cells were seeded in

373 Matrigel-coated 96-well with a cell density of 2000 cells per well. After appropriate cultivate time,
374 10% CCK8 (Dojindo, catalog#CK04) was added into the culture medium and incubate cells for 2 h,
375 then optical density value at 450 nm was detected using a microplate reader. Cells were detected in
376 every 24 h.

377 ***b. Alkaline phosphatase staining assay.*** For ALP (alkaline phosphatase) staining, hESCs were
378 digested into single cell suspension using Accutase and the number of cells were counted by
379 automated cell counter. After counting, the cells were seeded in Matirgel-coated 12-well with the cell
380 density of 1.2×10^5 cells per well and incubated in mTeSR1 medium for 3 days. Then the cells were
381 fixed with 4% paraformaldehyde for 3–5 min followed by washing with PBS. Next, the samples were
382 added with BCIP/NBT staining solution (Beyotime, catalog#C3206) for 15 min in dark at room
383 temperature. The usage of regents was followed with the manufacturer's recommendations. The cell
384 images were taken after washing the cells with PBS.

385

386 **hESCs differentiation assays**

387 ***a. Endodermal differentiation.*** To initiate endodermal cell differentiation, hESCs were cultured on
388 Matrigel-coated 6-well plate in RPMI1640 basal medium supplemented with B27 (Gibco,
389 catalog#12587-010) and 100 ng/ml Activin A (Peprotech, catalog#AF-120-14E-10UG). After cell
390 differentiation inducing for 3 days, endoderm markers were detected by flow cytometry and RT-
391 qPCR analysis.

392 ***b. Hematopoietic mesodermal differentiation.*** For Hematopoietic mesodermal differentiation,
393 hESCs were cultured to 60% density on Matrigel-coated plates. Using Accutase to digest cells for 3-
394 5 min, the cells were resuspended by blowing with mTeSR1 and seeded into ultra-low attachment 6-
395 well plates with the cell density of 2×10^5 cells per well. Cells were cultured in a hypoxic incubator
396 (5% CO_2, 5% O_2) for 24 h, until they formed to embryoid bodies (EB). To induce mesodermal
397 differentiation, STEMdiff APELTM2 (STEMCELL, catalog#05270) was used as the basal medium,
398 and the differentiation medium was configured by adding corresponding cell factors according to the
399 Kaufman hematopoietic differentiation protocol [41]. After cell differentiation inducing for 6 days,
400 mesodermal markers were detected by flow cytometry and RT-qPCR analysis.

c. Neural ectodermal differentiation. Monolayer culture protocol was used to initiate neural ectodermal differentiation. hESCs were plated onto Matrigel-coated 6-well plates with the cell density of 2×10^5 cells/cm^2, and then cultured in STEMdiff Neural Induction Medium (STEMCELL, catalog#08581) added with SMADi. After cell differentiation inducing for 3 days, neural ectodermal markers were detected by flow cytometry and RT-qPCR analysis.

Statistical analysis

Unless otherwise stated, bar charts represent the mean ± standard deviation of the mean (SD), and the results for the two groups were compared using a two-tailed unpaired Student's *t*-test. A *P*-value smaller than 0.05 was considered significant unless stated differently, and the exact degree of significance as indicated by asterisks is stated in the legends. Statistical significance was presented as $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$; ns indicates $P > 0.05$.

Results

Systematic profiling of m^{6}A methylation readouts across different cell types

The overall design of this study is illustrated in **Figure 1A**. Briefly, we perturbed m^{6}A in five representative cell lines through knockdown of core m^{6}A writer *METLL3*. The efficiency of the knockdown was verified by qRT-PCR and Western blot (**Supplementary Figure S1)**. By procedure, we could measure the mRNA half-life, translation efficiency, and alternative splicing readouts of m^{6}A by comparing the post-transcriptional gene expression dynamics of m^{6}A-normal cells (i.e. shControl cells) versus m^{6}A-disrpted cells (i.e. shMETTL3 cells). Actinomycin D-treated temporal RNA-seq, Ribo-seq and ultra-deep RNA-seq were adopted to measure the three main categories of m^{6}A readouts, respectively. We introduced a cost-effective experiment design by performing ultra-deep RNA-seq on the untreated (0h) cells, which could also serve as the start point for temporal RNA-seq as well as the input sample for Ribo-seq, with the discrepancy in sequencing library size controlled by ERCC spike-ins (see Methods).

The mRNA half-life was firstly measured by fitting the gene expression to the mRNA decay curve along the time points. For all cell types, a decrease of absolute mRNA quantification along the time points were observed, demonstrating substantial mRNA degradation with the transcription process inhibited by Actinomycin D. Nonetheless, the decay rate of the mRNA half-life could be diverged between different m^6A states, where for most of the cell types, an overall down-regulation of mRNA half-lives in m^6A normal cells (shControl cells) could be observed (**Figure 1B**), which is consistent with previous literature [11,42]. Nonetheless, the overlap of half-life down-regulated genes is limited between different cell types, indicating a high cell-type specificity of the half-life readout of m^6A (**Figure 1C**). It is possible that the limited overlap between down-regulated genes is simply caused by the threshold in measuring the readout. Thus, we performed enrichment for different functional gene sets and pathways to check if different cell types could show consistent changes at the pathway level. Indeed, gene sets representing fundamental biological functions such as mitochondrial oxidative phosphorylation, glycolysis and DNA repairs were shared by multiple cell types, indicating a consensus regulation by m^6A in these basic cellular functions (**Figure 1D-E**). On the other hand, some pathways were only enriched in particular cell type, and these pathways are likely to be associated the functions of specific cell type. For example, multiple oncogenic pathways were enriched in lung cancer cell line A549, whereas cytoskeletal and endothelial cell migration functions were enriched in blood vessel-derived HUVEC cells (**Figure 1F-G**).

By querying to the public MeRIP-seq methylation profiles from the corresponding cell type, we confirmed that most (>90%) of the half-life down-regulated genes were the direct target genes of m^6A methylation (**Figure 1H**). To precisely describe the relationship between m^6A modification levels and half-life changes, we evaluated the m^6A methylation levels of target genes in different cells using GLORI technique [28]. The quantitative m^6A map of four out of the five cell types were included, where A549, hESC and JURKAT data were newly measured in this study. HUVEC was not included in GLORI analysis because of its proliferation capacity is insufficient to accumulate required amount of input cells. The m^6A methylation sites identified by GLORI were all enriched near the stop codon (**Supplementary Figure S2A**), which is a typical topology feature of m^6A sites on coding genes [2]. Our GLORI experiments could capture m^6A sites with medium-to-high modification ratios, although low rate sites were relatively underrepresented because of the reduced sequencing depth due to the cost concerns (**Supplementary Figure S2B**). Besides, the m^6A sites and

genes were partially shared between different cell types, and noticeable overlaps between m^6A target genes identified by GLORI and those from MeRIP-seq could be observed, confirming that the main m^6A target genes were captured (**Supplementary Figure S2C-E**). Finally, >80% of the half-life down-regulated genes harbored a GLORI-derived m^6A site (**Supplementary Figure S2F**), enabling a direct correlation analysis between the absolute quantification of m^6A sites and the mRNA half-life fold-change of their target genes. Significant correlations between m^6A modification rates and the half-life changes in most cells were observed, suggesting the half-life regulation would partly rely on the methylation level of the target genes. Nonetheless, the correlations were generally weak (**Figure 1I**), suggesting that other factors, such as the genomic context of m^6A sites, may also have significant contributions.

Next, we profiled the translational efficiency readout of m^6A. In terms of translational efficiency, different cell types showed more balanced changes, with roughly comparable numbers of up- and down-regulated genes (**Figure 2A**). The overlap of genes showing translational efficiency changes was also very limited between different cell types, for both up- and down-regulated genes (**Figure 2B-C**). Some of the genes showed both half-life and translational efficiency changes (**Figure 2D and Supplementary Figure S3A-C**). In A549 cell lines, genes with half-life or translational efficiency changes had higher importance (DepMap dependency [27]) scores than background, and those with both types of changes ranked the highest; in JURKAT, higher importance scores of m^6A-regulated genes were also observed (**Figure 2E**). Similar to the observation in half-life readout analysis, some fundamental pathways are also the favored target of the translational efficiency regulations of m^6A (**Supplementary Figure S3D-I).** Finally, more than 90% of translational efficiency-change genes were also the m^6A methylation target genes, as suggested by the MeRIP-seq methylation profiles (**Figure 2F**). When considering GLORI-based methylation profiles, approximately 80% of genes showing translational efficiency changes could be targeted by m^6A (**Supplementary Figure S2G**), but GLORI-estimated methylation levels were only weakly correlated or non-significantly correlated with the fold change in translational efficiency (**Figure 2G**). We also noted that genes with up- or down-regulated translational efficiency did not always tend to have a m^6A site near the start codon (**Figure 2H**). These results indicate that neither m^6A level nor m^6A topology is sufficient to fully explain the translational efficiency readout of m^6A.

The ultra-deep sequencing enabled extensive detection of alternative splicing events, where the m^{6}A-mediated alternative splicing was mainly manifested as exon skipping, at both the event and the gene level (**Figure 3A**). Therefore, subsequent analyses of alternative splicing were focused on exon skipping. Generally, an increase of exon inclusion ratio (Psi) could be observed in m^{6}A-normal cells (**Figure 3B**), consistent with previous studies [13,43]. Genes showed altered exon inclusion were less shared between different cell types (**Figure 3C-D**). Basic functions like organelle organization and DNA repair are enriched for these genes (**Supplementary Figure S4**). However, it is noteworthy that only a small proportion of these exons had known m^{6}A methylation sites in their vicinity, suggesting that there might be a greater influence of indirect factors on the observed exon inclusion changes (**Figure 3E and Supplementary Figure S2H**). To reduce the influence of indirect factors, for the subsequent analysis we only considered exons with m^{6}A methylation sites within 10kb flanking the exon boundary. Among these exons, the m^{6}A-to-exon-boundary distances were only weakly or not significantly correlated with the changes in exon inclusion ratios (**Figure 3F**). In addition, GLORI-estimated methylation levels were only weakly correlated or not significantly correlated with changes in exon inclusion ratio (**Figure 3G**), suggesting that simple transcript topological features are not sufficient to predict the alternative splicing readout of m^{6}A.

Predicting m^{6}A readouts based on the RBP binding context

Given that m^{6}A methylation level and topology could not accurately predict m^{6}A functional readout, we paid more attention on the genomic context, especially the RBP binding context of the m^{6}A sites. To this end, we systematically collected and re-analyzed high-throughput RBP binding site data from public databases (**Supplementary Table S5**) to comprehensively evaluate the associations between m^{6}A sites and RBP binding sites. To ensure applicability, we used MeRIP-seq-based m^{6}A site annotation here (see Methods; **Supplementary Table S4**). The score of the closest RBP binding site and the genomic/transcript distance were calculated as the RBP features of an m^{6}A site. We also introduced the distribution features of m^{6}A sites in transcripts, tissue expression level, gene importance score and k-mer features to better characterize m^{6}A sites and their target genes. The above features were inputted into the XGBoost model to build the m^{6}A readout predictor, hereafter referred as Reading-m6A. Reading-m6A could deal with five readout categories: half-life down-regulation,

translational efficiency down-regulation and up-regulation, and exon inclusion ratio decrease or increase. The half-life up-regulation cases were not considered due to the limited number of this type of genes. Because of the obvious differences in m^6A readout between different cells, each cell type was modeled separately and the prediction accuracy of the model was assessed using ten-fold cross-validation. The results showed that the overall prediction performance of the model was acceptable, with an average AUROC of 0.8, especially for half-life and alternative splicing readouts (**Figure 4A-E**). To further test this modeling approach, we also introduced translational efficiency readout data derived from public Ribo-seq datasets. Four cell lines (HeLa, HepG2, Huh7 and MOLM13) were included. and the predictive performance of the model is equally acceptable in these cells except MOLM13 (**Fig. 4B, C**). Compared to other classical machine learning frameworks, XGBoost showed the best performance for most cases (**Supplementary Figure S5A**); and its performance is insensitive to hyperparameter optimization, with at most 0.01 AUC gain after hyperparameter optimization (**Supplementary Figure S5B**). We also checked whether considering GLORI-based m^6A sites, especially their m^6A modification rates, could result in better prediction performance. The results demonstrated that replacing MeRIP-seq-based m^6A site annotation with GLORI-based ones could benefit the performance in some cases but reduce the performance in other cases, even the highly comprehensive GLORI profile from HEK293T was adopted, while including m^6A modification rate features show marginal influence to the performance (**Supplementary Figure S5C**). Therefore, we kept MeRIP-seq-based m^6A site annotation as the primary source to build our models.

One of the main reasons we used the XGBoost framework for modeling was its good interpretability. We found that the RBP-based feature sets indeed ranked the best for their contribution to AUC. For half-life and translational efficiency readout predictions, the binding scores of RBP-binding sites near the m^6A site were the most important feature set (**Figure 4F-G** and **Supplementary Figure S5D**). While the transcript of genomic distance between m^6A and RBP-binding site ranked top in exon inclusion readout predictions (**Supplementary Figure S5E-F**). Gene topology features and tissue expression contributed to the performance to a much smaller extent. K-mer features showed only marginal or no contribution to AUC, implying the context features related to m^6A readouts were largely covered by current RBP binding site data. Moreover, further analysis of the importance scores of individual RBP features revealed that some known m^6A readers had high importance scores, while

549 other important RBP features suggested putative novel m^{6}A readers, for which we will be
550 experimentally validated in the next sections (**Figure 4H**).

551

552 **M^{6}A readout-related RBPs as the novel m^{6}A readers**

553 We noticed some RBPs with high feature importance in our m^{6}A readout models have been reported
554 as m^{6}A readers [4]. These RBPs included typical m^{6}A readers such as YTHDF1/2/3, YTHDC1/2 and
555 IGF2BPs as well as some recently discovered m^{6}A readers like TARDBP and ELAVL1 (**Figure 4H**).
556 To further screened novel m^{6}A-binding proteins, we used methylated single-stranded RNA bait (ss-
557 m^{6}A, with the consensus sequence GG(m^{6}A)CU) or unmethylated control RNA (ss-A) to perform
558 RNA pull-down assay, where several known m^{6}A readers (YTHDF1, YTHDC2, RBFOX2, TARDBP)
559 were included as the positive controls (**Figure 5A-B** and **Supplementary Figure S6**). The pull-down
560 screening was mainly performed in A549 cells, and the results showed that DDX6, FUBP3 and FXR2
561 could strongly bind to the methylated bait instead of the unmethylated control (**Figure 5A**), and their
562 binding to m^{6}A was also confirmed in hESCs (**Figure 5B**). Because L1TD1 has no expression in
563 A549 cells, we could only validate L1TD1 binding to methylated bait in hESCs only (**Figure 5B**).

564

565 We further investigated the transcriptome-wide binding preference of these candidate readers toward
566 m^{6}A sites by using TRIBE/HyperTRIBE technique [37,44,45]. TRIBE involves fusing the target RBP to
567 the catalytic domain of the RNA editing enzyme ADAR. When the fusion proteins are expressed in
568 target cells, the A-to-I editing at residues in proximity to RBP binding sites was catalyzed by ADAR,
569 providing a signal detectable in conventional RNA-seq (**Figure 5C**). HyperTRIBE exploits a mutated
570 version of ADAR domain to improve its activity thus enhance the A-to-I signals [37,45]. To demonstrate
571 the feasibility of HyperTRIBE approach, we firstly applied this technique to well-known m^{6}A reader
572 YTHDF1 in HEK293T cells (**Figure 5D**). After applying stringent thresholds and subtracting
573 negative controls (see Methods), over 100,000 sites targeted by YTHDF1-ADAR fusion protein were
574 observed. *De novo* motif analysis revealed that YTHDF1 significantly bind to the GGACU consensus
575 sequence, which is a sub-motif of the DRACH m^{6}A consensus motif (**Supplementary Figure S7A**).
576 Over 90% YTHDF1 target genes significantly overlapped with known m^{6}A-modified genes
577 identified by GLORI ($P < 1.0e-200$, Fisher's exact test; **Supplementary Figure S7B**). Moreover, the
578 majority of YTHDF1 target genes overlapped with those identified by the previous TRIBE assay [9]

(**Supplementary Figure S7C**), and with known m^{6}A target genes from MeRIP-seq assay [46] (**Supplementary Figure S7D**). These results together demonstrate the feasibility of HyperTRIBE.

Then we applied HyperTRIBE to DDX6, FUBP3, FXR2 and L1TD1 proteins in hESCs. The enrichment heatmaps and profiles showed that DDX6, FUBP3, FXR2 and L1TD1-binding sites were all notably enriched near m^{6}A sites rather than non-m^{6}A sites (**Figure 5E**). The *de novo* motif analysis revealed that DDX6, FUBP3, FXR2 and L1TD1-binding regions significantly enriched for the DRACH-like motifs (**Figure 5F**). When comparing DDX6, FUBP3, FXR2 and L1TD1 target genes with the m^{6}A target genes (identified either by our GLORI profiling or by MeRIP-seq), significant overlaps could be observed (**Figure 5G** and **Supplementary Figure S7E-H**). Taken together, these results indicated that DDX6, FUBP3, FXR2 and L1TD1 are novel m^{6}A readers in hESCs.

Effect of m^{6}A readers on early differentiation in hESCs

The regulatory function of DDX6 in hESCs has been clarified [47], but the functions of FUBP3, FXR2 and L1TD1 in hESCs have not been fully elucidated. To investigate potential roles for FUBP3, FXR2 and L1TD1 in self-renewal and early differentiation of hESCs, we constructed *FUBP3*, *FXR2* and *L1TD1* knockout (KO) hESCs by using the CRISPR/Cas9 gene-editing approach. The validty of KO constructs were confirmed by Western blot and Sanger sequencing (**Supplementary Figure S8**). We first evaluated the effects of these m^{6}A readers on hESC self-renewal, but no significant differences in cell morphology between *FUBP3*, *FXR2* and *L1TD1* KO hESCs and wild-type hESCs were observed (**Figure 6A**). Besides, compared to wild-type hESCs, the protein expressions of core pluripotent markers (OCT4, SOX2 and NANOG) did not have significant changes in *FUBP3*, *FXR2* and *L1TD1* KO cells (**Supplementary Figure S9A-B**). RT-qPCR analysis also found no significantly differences between FUBP3, FXR2 and L1TD1 KO hESCs and wild-type hESCs in the expression of core pluripotent markers (**Supplementary Figure S9C**). These results revealed that FUBP3, FXR2 and L1TD1 had no significant impact on the expression of core pluripotent genes in hESCs. We also employed the CCK8 and ALP staining assays to investigate the effects of FUBP3, FXR2 and L1TD1 on stem cell proliferation. CCK8 assay showed that the growth rate of *L1TD1* KO and *FXR2* KO cells was changed at 24-48 h; but when extending to 72-120 h after cell inoculation, there was no significant difference in proliferation rate between *FUBP3*, *FXR2* and *L1TD1* KO hESCs and

wild-type hESCs (**Figure 6B**). The ALP staining assay also showed no obvious distinction in the number of colonies formed by *FUBP3*, *FXR2* and *L1TD1* KO hESCs in comparison with wild-type hESCs (**Figure 6C** and **Supplementary Figure S9D**), indicating that deletion of *FUBP3*, *FXR2* and *L1TD1* had no significant effect on the proliferation rate of hESCs. Together, these results indicated that m^6A readers FUBP3, FXR2 and L1TD1 had no prominent impact on self-renewal of hESCs.

Next, we investigated the effect of m^6A readers FUBP3, FXR2 and L1TD1 on early differentiation of hESCs by inducing cells into three different germ layers. For endoderm differentiation, Activin A treatment was used in this study. The cells were cultured in RPMI 1640 basic medium supplemented with B27 (**Figure 6D**). After treating with Activin A for 3 days, cells were collected to detect the expression of endodermal markers FOXA2 and SOX17 by flow cytometry and RT-qPCR. The flow cytometry results showed that $FOXA2^+SOX17^+$double positive cells were significantly increased in the *FUBP3* KO cultures compared with wild-type hESCs after endodermal induction for 3 days (**Figure 6E-F**). However, in FXR2 and L1TD1 KO cultures, the proportion of $FOXA2^+SOX17^+$double positive cells was significantly reduced, indicating that loss of *FUBP3* enhances the ability of hESCs to endodermal differentiation, while loss of *FXR2* and *L1TD1* inhibits the ability to differentiate into endoderm (**Figure 6E-F**). As for hematopoietic mesodermal differentiation, cells were seeded on low attachment plates to form EB spheres, and then cultured in STEMdiff APEL^{TM}2 medium supplemented with factors SCF, BMP4 and VEGF for 6 days (**Figure 6G**). Flow cytometry results showed that *FUBP3*, *FXR2* and *L1TD1* KO hESCs significantly reduced the proportion of mesodermal marker $CD34^+$, indicating that FUBP3, FXR2 and L1TD1 were essential for hematopoietic mesodermal differentiation of hESCs. Deletion of *FUBP3*, *FXR2* and *L1TD1* significantly inhibits the ability to differentiate into mesoderm (**Figure 6H-I**). As for neural ectodermal differentiation, STEMdiff^{TM} neural induction medium with SMADi were used to induce cells to differentiate into neural progenitor cells by monolayer culture protocol (**Figure 6J**). Consistent with the tendency of endoderm differentiation, after neural induction for 3 days, more $SOX1^+$positive cells were present in the *FUBP3* KO cultures compared to wild-type group (**Figure 6K-L**). In *FXR2* and *L1TD1* KO cultures, the proportion of $SOX1^+$positive cells was significantly declined, indicating that loss of *FUBP3* promotes the ability of hESCs to neural ectodermal differentiation, while loss of *FXR2* and *L1TD1* inhibits the ability to ectodermal differentiation (**Figure 6K-L**).

Furthermore, we detected the mRNA expression level of marker genes of three germ layer through RT-qPCR. Consistent with the flow cytometry results, in endoderm differentiation stage, the mRNA expression levels of FOXA2 and SOX17 were significantly increased in *FUBP3* KO cells but decreased in *FXR2* and *L1TD1* KO cells (**Figure 6M**). During the hematopoietic mesoderm differentiation stage, the mRNA expression levels of TBXT (also known as T) and CD34 were all significantly reduced in *FUBP3*, *FXR2* and *L1TD1* KO cells compared to wild-type group (**Figure 6N**). As for neural ectoderm differentiation stage, the mRNA expression levels of neural-ectoderm markers SOX1, PAX6 and NES (also known as NESTIN) were significantly increased in *FUBP3* KO cells but decreased in *FXR2* and *L1TD1* KO cells (**Figure 6O**). These results indicated that FUBP3 may inhibit the endoderm and ectoderm differentiation, but is important for the mesoderm differentiation. FXR2 and L1TD1 genes are crucial for the differentiation of all three germ layers, and absence may significantly affect the differentiation ability of hESCs. In conclusion, the above results revealed that m^6A readers play an important role in regulating the early differentiation of hESCs.

Discussion

In recent years, m^6A has emerged as a key RNA modification in various biological processes and diseases [4,48]. Thanks to significant advances in high-throughput sequencing technologies such as MeRIP-seq [1,2] and GLORI [28], systematic mapping of m^6A modification sites and their dynamic changes in different tissues, conditions and cell types has become feasible [3,6-8]. On the other hand, m^6A modification exhibits diverse post-transcriptional functional readouts, including changes in RNA half-life, in translation efficiency, and in alternative splicing [49]. Despite the diverse functional readouts of m^6A, there is a lack of systematic measurement and comparison of m^6A post-transcriptional functional readouts across multiple cell lines. In fact, an extensive search of the public omics database GEO could only yield a very limited number of cell lines where the m^6A methylome could be coupled with its readouts. For example, we obtained the m^6A profiles along with Ribo-seq data for only four cell lines (HepG2, HeLa, Huh7, and MOLM13) with acceptable quality, and these data were introduced as an expanding dataset for translational efficiency readout models. Even for

alternative splicing, which can be measured by RNA-seq, few datasets could meet the required sequencing depth for comprehensive alternative splicing identification. To this end, in this study we first selected five representative cell lines (A549, hESC, HEK293T, JURKAT and HUVEC) and performed temporal transcriptome sequencing, Ribo-seq and ultra-deep RNA sequencing on m^6A-normal (siControl) and m^6A-disrupted (siMETTL3) cells. Our high-throughput data provide a comprehensive resource for further exploration of the regulatory mechanism as well as the key regulatory targets of m^6A methylation. Indeed, machine learning modeling of the profiled m^6A readouts has successfully prioritized novel m^6A readers from the RBP features of the models, demonstrating the utility of m^6A readout profiles.

Noticeable cell type specificity of the m^6A readouts was observed. On the one hand, the m^6A methylation rate alone is not sufficient to explain this diversity. There are some overlaps of m^6A target genes between different cell types. However, the overlaps of m^6A readouts are limited except for some basic pathways. Furthermore, the correlations between GLORI-measured m^6A modification ratios with the post-transcriptional readouts were compromised, and the inclusion of m^6A modification ratio did not result in better prediction of m^6A readouts. Instead, we pay more attention on the association of RBP-binding context of m^6A sites and m^6A readouts. Analysis of our readout prediction models suggests that the prediction performance was largely contributed by RBP-based features, either the RBP binding site score near the m^6A site or the distance between RBP binding sites and m^6A modification sites. Interestingly, we found that there are few RBPs that significantly contribute to the predictions of readouts, including some well-known m^6A readers, and screened novel m^6A readers based on this candidate list. Together, our results emphasize the importance of m^6A readers, which form a complex regulatory context that controls how the m^6A modification is interpreted in cells. On the other hand, it is likely that cell type specificity is overestimated here, when compared to the entire repository of human cell types. To be honest, we intentionally selected five cell types from the different cell clades that show distinctive gene expression patterns to ensure representativeness of our data at an affordable cost. When focusing on a specific physiological or disease process, our readout analysis pipeline would be applied to similar cell lines to investigate the consensus m^6A readout targets and mechanisms.

In this study, we screened DDX6, FUBP3, FXR2 and L1TD1 as the novel m^6A readers and assessed their roles in regulating pluripotency in hESCs. DDX6 is an RNA helicase belonging to the Dead Box family, which participates in many aspects of RNA metabolism including processing body (P-body) formation [50], mRNA storage and decay [51], microRNA pathways [52], and translational suppression [51]. DDX6 plays an essential role in regulating self-renewal and differentiation. Di Stefano et al. has reported that DDX6 is required for ESC exit from the pluripotent state by mediating the translational suppression of target mRNAs in P-bodies [47]. Unlike DDX6, the functions of FUBP3, FXR2 and L1TD1 in ESCs is less explored. FUBP3 is an RNA-binding protein that can bind to the far upstream element (FUSE) of specific RNA molecules, thereby regulating gene expression. In cancer, FUBP3 functions as an oncogene or tumor suppressor gene depending on the type and context of the cancer [53-55]. However, the role of FUBP3 in cell fate transitions remains largely unclear. Our data in the first time revealed that FUBP3 exerted important impact on early differentiation of three germ lays in hESCs. FXR2 is also an RNA-binding protein related to Fragile X syndrome. FXR2 are highly enriched in mammalian brains, and has important effects on neurogenesis by regulating the proliferation and differentiation of neural progenitors in mice [56]. In this study, we observed that FXR2 deletion in hESCs impaired early neural development, substantially expanding the scale of FXR2 in neural development regulation. L1TD1 gene was first identified from the expressed sequence tag libraries by Yamanaka group. L1TD1 is abundantly expressed in ESCs and rapidly downregulates upon differentiation [57]. L1TD1 was reported as a part of the pluripotency interactome network of OCT4, SOX2, and NANOG [58]. Interestingly, L1TD1 is dispensable for the maintenance of pluripotency in mESCs, despite its specific expression pattern [57]. We observed that LITD1 deletion had no significant effects on self-renewal in hESCs. Of note, hESCs with LITD1 showed significant defects in early differentiation of three germ layers, meaning that L1TD1 play an essential role in human early embryonic development. Our work uncovered the new roles of FUBP3, FXR2 and L1TD1 in stem cell fate control, and sheds light on post-translational regulation in pluripotency and the importance of RBPs in pluripotent cells.

Our study also has some limitations. Firstly, the scope of cell types included is not sufficient in comparison with the wide spectrum of human cell atlas. Secondly, recent studies have revealed new layers of m^6A readouts such as the regulatory interplay with chromatin, but this aspect was not covered in current study. Thirdly, the molecular mechanism of the identified m^6A readers have not

been in-depth investigated. Nonetheless, our study has provided useful resource to bridge the gap between m^6A methylome landscape and its readouts, and prioritizing interesting novel m^6A readers for stem cell fate control. We expected the future research could expand the range of cell types studied, improve data resources, and delve deeper into the molecular mechanisms underlying m^6A functional readouts.

Software availability

Reading-m6A online tool is accessible at http://www.rnanut.net/reading-m6a.

Data availability

The data can be accessed in Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo) with accessions of GSE179872 (m^6A readout profiling), GSE290408 (GLORI) and GSE287679, GSE288304 (HyperTRIBE).

Funding

This study was supported by the National Natural Science Foundation of China (32222020 and 32070658 to Y.Z., 32371533 to Y.L.), Michigan Medicine-PKUHSC Joint Institute for Translational and Clinical Research (BMU2023JI002 to Y.L.), PKU-Baidu Fund (2019BD014 to Q.C.).

Reference

1. 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. 10.1038/nature11112.

2. 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. 10.1016/j.cell.2012.05.003.

3. Xuan, J., Chen, L., Chen, Z., Pang, J., Huang, J., Lin, J., Zheng, L., Li, B., Qu, L., and Yang, J. (2024). RMBase v3.0: decode the landscape, mechanisms and functions of RNA modifications. Nucleic Acids Res *52*, D273-D284. 10.1093/nar/gkad1070.

4. Liu, W.W., Zheng, S.Q., Li, T., Fei, Y.F., Wang, C., Zhang, S., Wang, F., Jiang, G.M., and Wang, H. (2024). RNA modifications in cellular metabolism: implications for metabolism-targeted therapy and immunotherapy. Signal Transduct Target Ther *9*, 70. 10.1038/s41392-024-01777-5.

5. Fu, Y., Dominissini, D., Rechavi, G., and He, C. (2014). Gene expression regulation mediated through reversible m6A RNA methylation. Nat Rev Genet *15*, 293-306. 10.1038/nrg3724.

6. Liang, Z., Ye, H., Ma, J., Wei, Z., Wang, Y., Zhang, Y., Huang, D., Song, B., Meng, J., Rigden, D.J., and Chen, K. (2024). m6A-Atlas v2.0: updated resources for unraveling the N6-methyladenosine (m6A) epitranscriptome among multiple species. Nucleic Acids Res *52*, D194-d202. 10.1093/nar/gkad691.

7. Xiao, S., Cao, S., Huang, Q., Xia, L., Deng, M., Yang, M., Jia, G., Liu, X., Shi, J., Wang, W., et al. (2019). The RNA N(6)-methyladenosine modification landscape of human fetal tissues. Nat Cell Biol *21*, 651-661. 10.1038/s41556-019-0315-4.

8. Liu, J., 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 m(6)A and m(6)Am Methylome across Human and Mouse Tissues. Mol Cell *77*, 426-440 e426. 10.1016/j.molcel.2019.09.032.

9. Flamand, M.N., Ke, K., Tamming, R., and Meyer, K.D. (2022). Single-molecule identification of the target RNAs of different RNA binding proteins simultaneously in cells. Genes Dev *36*, 1002-1015. 10.1101/gad.349983.122.

10. Zaccara, S., and Jaffrey, S.R. (2020). A Unified Model for the Function of YTHDF Proteins in Regulating m(6)A-Modified mRNA. Cell *181*, 1582-1595.e1518. 10.1016/j.cell.2020.05.012.

11. 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. 10.1038/nature12730.

12. Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., and He, C. (2015). N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell *161*, 1388-1399. 10.1016/j.cell.2015.05.014.

13. 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 *61*, 507-519. 10.1016/j.molcel.2016.01.012.

14. Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B.S., Mesquita, A., Liu, C., Yuan, C.L., et al. (2018). Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat Cell Biol *20*, 285-295. 10.1038/s41556-018-0045-z.

15. Edupuganti, R.R., Geiger, S., Lindeboom, R.G.H., Shi, H., Hsu, P.J., Lu, Z., Wang, S.Y., Baltissen, M.P.A., Jansen, P., Rossa, M., et al. (2017). N(6)-methyladenosine (m(6)A) recruits and repels proteins to regulate mRNA homeostasis. Nat Struct Mol Biol *24*, 870-878. 10.1038/nsmb.3462.

16. Visvanathan, A., Patil, V., Arora, A., Hegde, A.S., Arivazhagan, A., Santosh, V., and Somasundaram, K. (2018). Essential role of METTL3-mediated m(6)A modification in glioma stem-like cells maintenance and radioresistance. Oncogene *37*, 522-533. 10.1038/onc.2017.351.

17. McMillan, M., Gomez, N., Hsieh, C., Bekier, M., Li, X., Miguez, R., Tank, E.M.H., and Barmada, S.J. (2023). RNA methylation influences TDP43 binding and disease pathogenesis in models of amyotrophic lateral sclerosis and frontotemporal dementia. Mol Cell *83*, 219-236.e217. 10.1016/j.molcel.2022.12.019.

18. Chen, T., Hao, Y.J., Zhang, Y., Li, M.M., Wang, M., Han, W., Wu, Y., Lv, Y., Hao, J., Wang, L., et al. (2015). m(6)A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. Cell Stem Cell *16*, 289-301. 10.1016/j.stem.2015.01.016.

19. Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A.A., Kol, N., Salmon-Divon, M., Hershkovitz, V., Peer, E., Mor, N., Manor, Y.S., et al. (2015). Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science *347*, 1002-1006. 10.1126/science.1261417.

20. Lasman, L., Krupalnik, V., Viukov, S., Mor, N., Aguilera-Castrejon, A., Schneir, D., Bayerl, J., Mizrahi, O., Peles, S., Tawil, S., et al. (2020). Context-dependent functional compensation between Ythdf m(6)A reader proteins. Genes Dev *34*, 1373-1391. 10.1101/gad.340695.120.

21. 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 m(6)A reader YTHDC1 silences retrotransposons and guards ES cell identity. Nature *591*, 322-326. 10.1038/s41586-021-03313-9.

22. Baudin, B., Bruneel, A., Bosselut, N., and Vaubourdolle, M. (2007). A protocol for isolation and culture of human umbilical vein endothelial cells. Nat Protoc *2*, 481-485. 10.1038/nprot.2007.54.

23. Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M., and Weissman, J.S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. Nat Protoc *7*, 1534-1550. 10.1038/nprot.2012.086.

24. Xiao, Z., Zou, Q., Liu, Y., and Yang, X. (2016). Genome-wide assessment of differential translations with ribosome profiling data. Nat Commun *7*, 11194. 10.1038/ncomms11194.

25. Wang, Y., Xie, Z., Kutschera, E., Adams, J.I., Kadash-Edmondson, K.E., and Xing, Y. (2024). rMATS-turbo: an efficient and flexible computational tool for alternative splicing analysis of large-scale RNA-seq data. Nat Protoc *19*, 1083-1104. 10.1038/s41596-023-00944-2.

26. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb) *2*, 100141. 10.1016/j.xinn.2021.100141.

27. Tsherniak, A., Vazquez, F., Montgomery, P.G., Weir, B.A., Kryukov, G., Cowley, G.S., Gill, S., Harrington, W.F., Pantel, S., Krill-Burger, J.M., et al. (2017). Defining a Cancer Dependency Map. Cell *170*, 564-576.e516. 10.1016/j.cell.2017.06.010.

28. Liu, C., Sun, H., Yi, Y., Shen, W., Li, K., Xiao, Y., Li, F., Li, Y., Hou, Y., Lu, B., et al. (2023). Absolute quantification of single-base m(6)A methylation in the mammalian transcriptome using GLORI. Nat Biotechnol *41*, 355-366. 10.1038/s41587-022-01487-9.

29. Meng, J., Lu, Z., Liu, H., Zhang, L., Zhang, S., Chen, Y., Rao, M.K., and Huang, Y. (2014). A protocol for RNA methylation differential analysis with MeRIP-Seq data and exomePeak R/Bioconductor package. Methods *69*, 274-281. 10.1016/j.ymeth.2014.06.008.

30. Zhu, Y., Xu, G., Yang, Y.T., Xu, Z., Chen, X., Shi, B., Xie, D., Lu, Z.J., and Wang, P. (2019). POSTAR2: deciphering the post-transcriptional regulatory logics. Nucleic Acids Res *47*, D203-d211. 10.1093/nar/gky830.

31. Uren, P.J., Bahrami-Samani, E., Burns, S.C., Qiao, M., Karginov, F.V., Hodges, E., Hannon, G.J., Sanford, J.R., Penalva, L.O., and Smith, A.D. (2012). Site identification in high-throughput RNA-protein interaction data. Bioinformatics *28*, 3013-3020. 10.1093/bioinformatics/bts569.

32. Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R., Blue, S.M., Chen, J.Y., Cody, N.A.L., Dominguez, D., et al. (2020). A large-scale binding and functional map of human RNA-binding proteins. Nature *583*, 711-719. 10.1038/s41586-020-2077-3.

33. GTEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science *369*, 1318-1330. 10.1126/science.aaz1776.

34. Chen, T., and Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery.

35. Li, J., He, S., Guo, F., and Zou, Q. (2021). HSM6AP: a high-precision predictor for the Homo sapiens N6-methyladenosine (m^6 A) based on multiple weights and feature stitching. RNA Biol *18*, 1882-1892. 10.1080/15476286.2021.1875180.

36. Zhou, J., You, D., Bai, J., Chen, X., Wu, Y., Wang, Z., Tang, Y., Zhao, Y., and Feng, G. (2023). Machine Learning Methods in Real-World Studies of Cardiovascular Disease. Cardiovasc Innov Appl *7*, CVIA.2023.0011.

37. Rahman, R., Xu, W., Jin, H., and Rosbash, M. (2018). Identification of RNA-binding protein targets with HyperTRIBE. Nat Protoc *13*, 1829-1849. 10.1038/s41596-018-0020-y.

38. Biswas, J., Rosbash, M., Singer, R.H., and Rahman, R. (2021). Protocol for using TRIBE to study RNA-protein interactions and nuclear organization in mammalian cells. STAR Protoc *2*, 100634. 10.1016/j.xpro.2021.100634.

39. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589. 10.1016/j.molcel.2010.05.004.

40. Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res *42*, W187-191. 10.1093/nar/gku365.

41. Zhu, H., and Kaufman, D.S. (2019). An Improved Method to Produce Clinical-Scale Natural Killer Cells from Human Pluripotent Stem Cells. Methods Mol Biol *2048*, 107-119. 10.1007/978-1-4939-9728-2_12.

42. Li, M., Zhao, X., Wang, W., Shi, H., Pan, Q., Lu, Z., Perez, S.P., Suganthan, R., He, C., Bjoras, M., and Klungland, A. (2018). Ythdf2-mediated m(6)A mRNA clearance modulates neural development in mice. Genome Biol *19*, 69. 10.1186/s13059-018-1436-y.

43. Kasowitz, S.D., Ma, J., Anderson, S.J., Leu, N.A., Xu, Y., Gregory, B.D., Schultz, R.M., and Wang, P.J. (2018). Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. PLoS Genet *14*, e1007412. 10.1371/journal.pgen.1007412.

44. McMahon, A.C., Rahman, R., Jin, H., Shen, J.L., Fieldsend, A., Luo, W., and Rosbash, M. (2016). TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins. Cell *165*, 742-753. 10.1016/j.cell.2016.03.007.

45. Xu, W., Rahman, R., and Rosbash, M. (2018). Mechanistic implications of enhanced editing by a HyperTRIBE RNA-binding protein. RNA *24*, 173-182. 10.1261/rna.064691.117.

46. Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M., Lujan, E., Haddad, B., Daneshvar, K., et al. (2014). m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell *15*, 707-719. 10.1016/j.stem.2014.09.019.

47. Di Stefano, B., Luo, E.C., Haggerty, C., Aigner, S., Charlton, J., Brumbaugh, J., Ji, F., Rabano Jiménez, I., Clowers, K.J., Huebner, A.J., et al. (2019). The RNA Helicase DDX6 Controls Cellular Plasticity by Modulating P-Body Homeostasis. Cell Stem Cell *25*, 622-638.e613. 10.1016/j.stem.2019.08.018.

48. Deng, X., Qing, Y., Horne, D., Huang, H., and Chen, J. (2023). The roles and implications of RNA m(6)A modification in cancer. Nat Rev Clin Oncol *20*, 507-526. 10.1038/s41571-023-00774-x.

49. Yang, Y., Hsu, P.J., Chen, Y.S., and Yang, Y.G. (2018). Dynamic transcriptomic m(6)A decoration: writers, erasers, readers and functions in RNA metabolism. Cell Res *28*, 616-624. 10.1038/s41422-018-0040-8.

50. Serman, A., Le Roy, F., Aigueperse, C., Kress, M., Dautry, F., and Weil, D. (2007). GW body disassembly triggered by siRNAs independently of their silencing activity. Nucleic Acids Res *35*, 4715-4727. 10.1093/nar/gkm491.

51. Coller, J., and Parker, R. (2005). General translational repression by activators of mRNA decapping. Cell *122*, 875-886. 10.1016/j.cell.2005.07.012.

52. Chu, C.Y., and Rana, T.M. (2006). Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. PLoS Biol *4*, e210. 10.1371/journal.pbio.0040210.

53. Gao, Q., Zhou, R., Meng, Y., Duan, R., Wu, L., Li, R., Deng, F., Lin, C., and Zhao, L. (2020). Long noncoding RNA CMPK2 promotes colorectal cancer progression by activating the FUBP3-c-Myc axis. Oncogene *39*, 3926-3938. 10.1038/s41388-020-1266-8.

54. Sharma, M., Anandram, S., Ross, C., and Srivastava, S. (2023). FUBP3 regulates chronic myeloid leukaemia progression through PRC2 complex regulated PAK1-ERK signalling. J Cell Mol Med *27*, 15-29. 10.1111/jcmm.17584.

55. Yao, Q., Xie, Y., Xu, D., Qu, Z., Wu, J., Zhou, Y., Wei, Y., Xiong, H., and Zhang, X.L. (2022). Lnc-EST12, which is negatively regulated by mycobacterial EST12, suppresses antimycobacterial innate immunity through its interaction with FUBP3. Cell Mol Immunol *19*, 883-897. 10.1038/s41423-022-00878-x.

56. Guo, W., Zhang, L., Christopher, D.M., Teng, Z.Q., Fausett, S.R., Liu, C., George, O.L., Klingensmith, J., Jin, P., and Zhao, X. (2011). RNA-binding protein FXR2 regulates adult hippocampal neurogenesis by reducing Noggin expression. Neuron *70*, 924-938. 10.1016/j.neuron.2011.03.027.

57. Iwabuchi, K.A., Yamakawa, T., Sato, Y., Ichisaka, T., Takahashi, K., Okita, K., and Yamanaka, S. (2011). ECAT11/L1td1 is enriched in ESCs and rapidly activated during iPSC generation, but it is dispensable for the maintenance and induction of pluripotency. PLoS One *6*, e20461. 10.1371/journal.pone.0020461.

58. Emani, M.R., Närvä, E., Stubb, A., Chakroborty, D., Viitala, M., Rokka, A., Rahkonen, N., Moulder, R., Denessiouk, K., Trokovic, R., et al. (2015). The L1TD1 protein interactome reveals the importance of post-transcriptional regulation in human pluripotency. Stem Cell Reports *4*, 519-528. 10.1016/j.stemcr.2015.01.014.

Figures

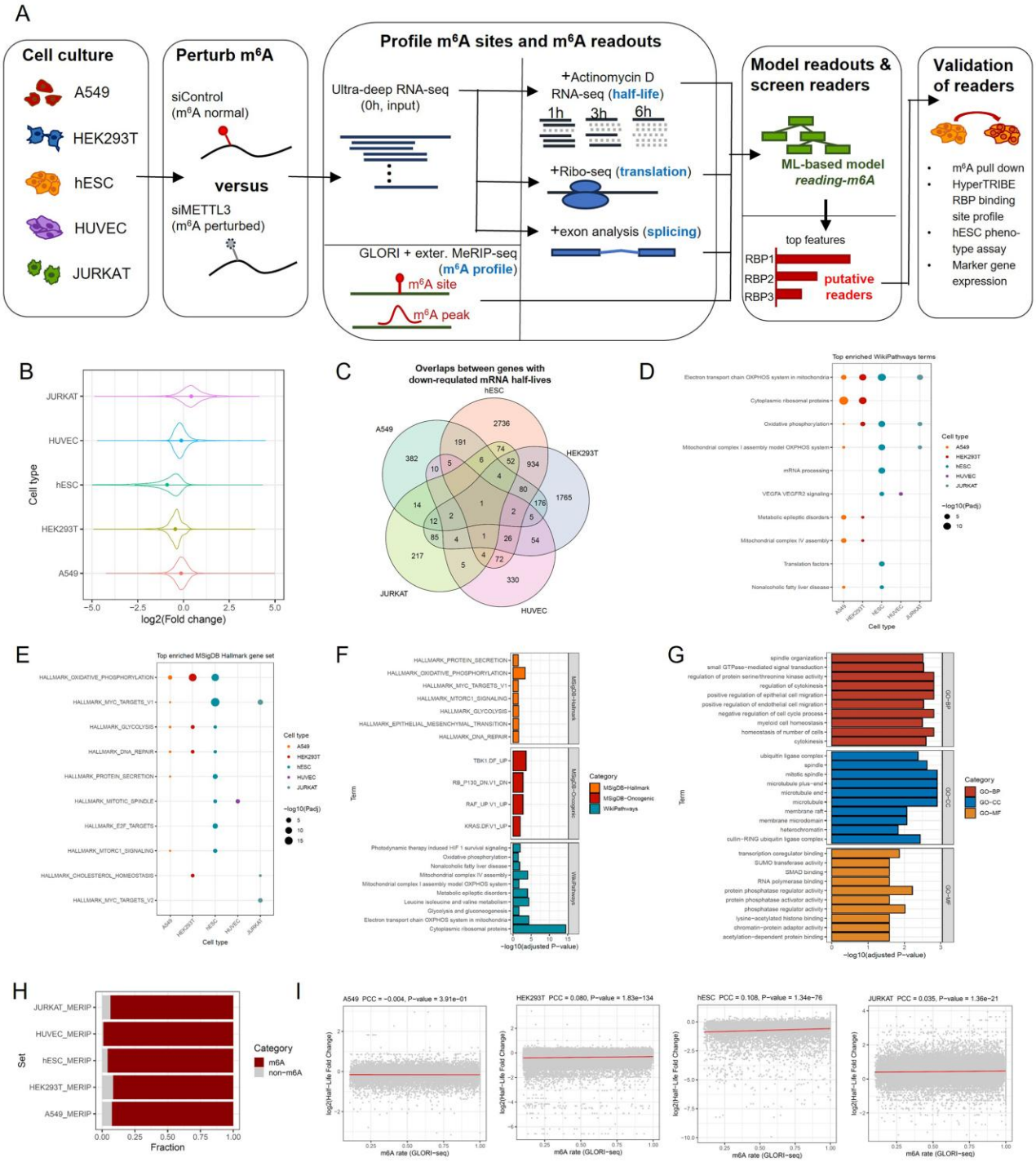


Figure 1. Overview of the experimental design for m^{6}A readout profiling and the half-life readout profiling results.

(A) Overview of the experimental design for m^6A readout profiling. (B) Overview of m^6A-mediated half-life changes. The comparison was performed between m^6A-normal cells (shControl) versus m^6A-disrupted cells (shMETTL3) as the background. (C) Venn diagram of half-life down-regulated genes in different cell types. (D) Top 10 enriched pathways (WikiPathways) of half-life down-regulated genes among different cell types sorted by total -log10(FDR). (E) Top 10 enriched MSigDB Hallmark gene sets of half-life down-regulated genes among different cell types. (E) Top enriched pathways and gene sets for half-life down-regulated genes in A549 cells. (G) Top enriched GO functional terms for half-life down-regulated genes in HUVEC cells. (H) Percentage of half-life down-regulated genes with m^6A methylation peaks identified by MeRIP-seq assay on the corresponding cell types. (I) Correlation between fold change in half-life and m^6A methylation levels assessed by GLORI.

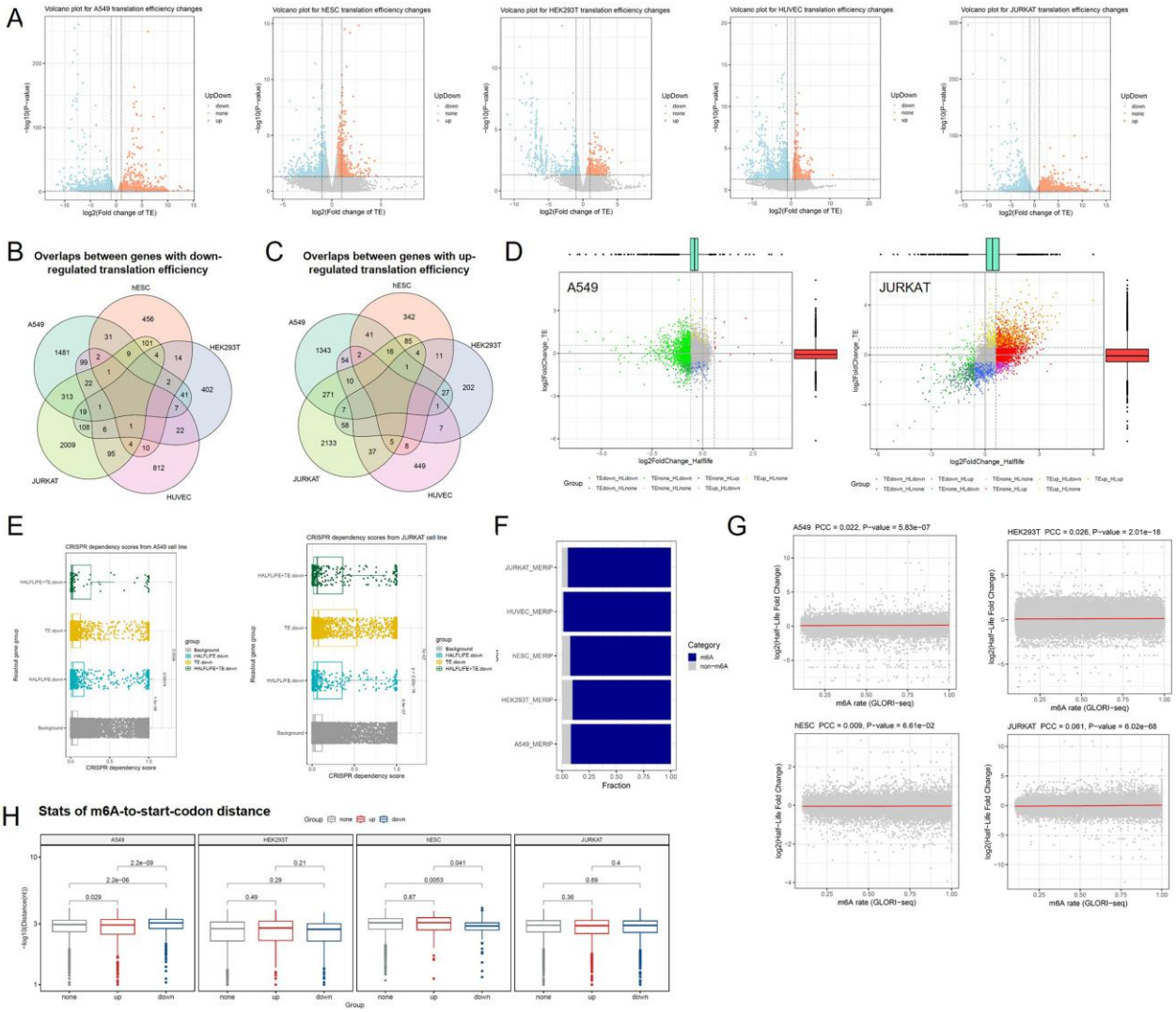


Figure 2. m^{6}A methylation regulates translational efficiency in different cell types.

(A) Volcano plot of m^{6}A-mediated changes in translational efficiency. The comparison was performed between m^{6}A-normal cells (shControl) versus m^{6}A-disrupted cells (shMETTL3) as the background. (B-C) Venn diagram of translational efficiency down-regulated (B) and up-regulated (C) genes in different cell types. (D) Scatter plot showing translational efficiency versus half-life changes in A549 and JURKAT cells. (E) CRISPR-based gene importance scores of genes with half-life and/or translation efficiency changes in cancer cell lines A549 and JURKAT compared to genome background. (F) Percentage of translation efficiency change genes with m^{6}A methylation peaks identified by MeRIP-seq assay on the corresponding cell types. (G) Correlation between fold change in translation efficiency (in absolute values) and m^{6}A methylation levels assessed by GLORI. (H) Box plots comparing the minimum m^{6}A-site-to-strart-codon distance between genes showing different translational efficiency changes.

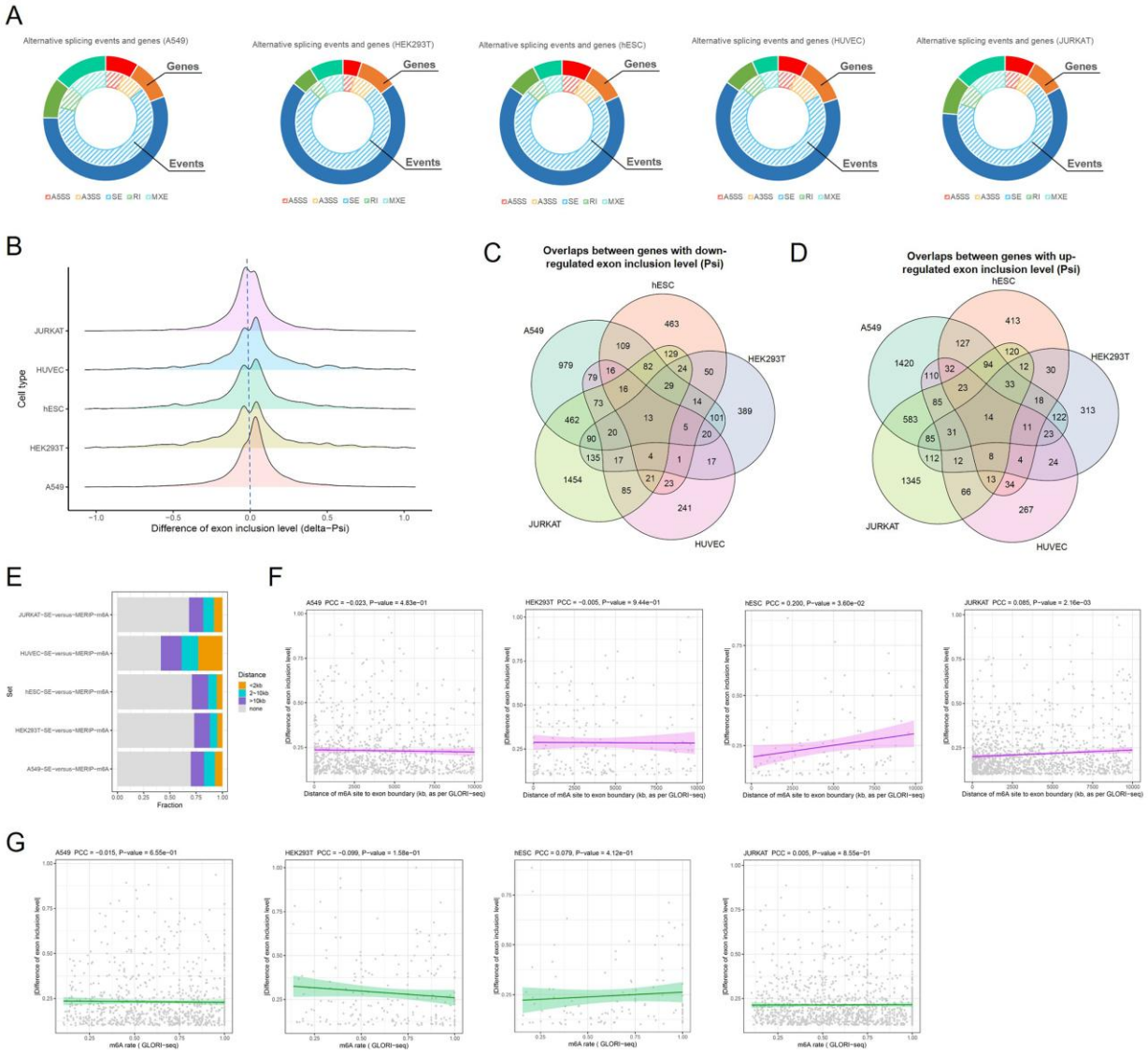


Figure 3. m^{6}A methylation regulates alternative splicing in different cell types.

(A) Circular plot summarizing different types of alternative splicing changes mediated by m^{6}A. The inner and outer loops representing the percentages of alternative splicing events and genes, respectively. (B) Ridge plot summarizing the m^{6}A-mediated exon inclusion ratio (Psi) changes. The comparison was performed between m^{6}A-normal cells (shControl) versus m^{6}A-disrupted cells (shMETTL3) as the background. (C-D) Venn diagram of exon inclusion ratio down-regulated (C) and up-regulated (D) genes in different cell types. (E) Percentage of exons showing exon inclusion ratio changes with m^{6}A methylation peaks (identified by MeRIP-seq assay on the corresponding cell types) near their exon boundaries. (F) Correlation between changes in exon inclusion ratio (in absolute values) and m^{6}A (within 10 kb to exon boundary) methylation levels assessed by GLORI. (G) Correlation between changes in exon inclusion ratio (in absolute values) and the distances from m^{6}A methylation sites to exon boundaries.

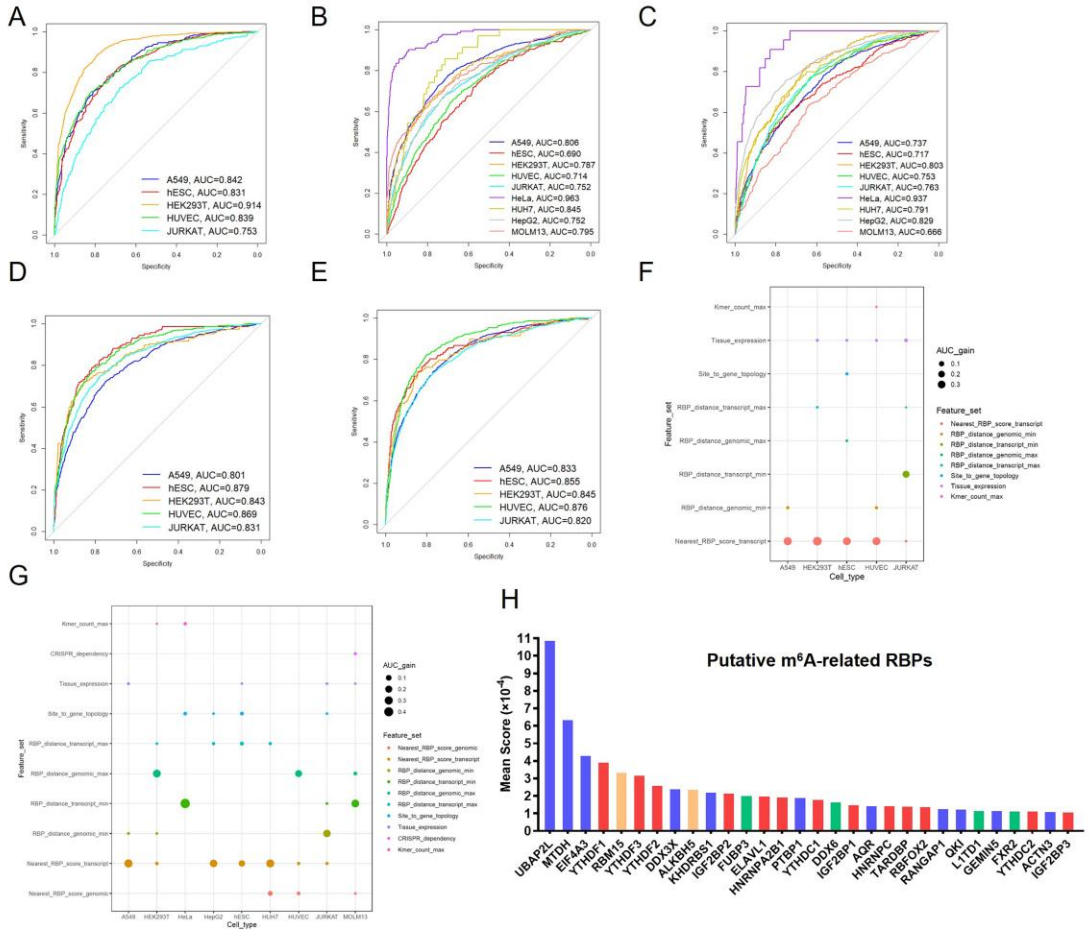


Figure 4. Machine learning models of m^{6}A readouts and their top predictive features.

(A-E) ROC curve assessing the performances of machine learning models (termed as Reading-m^{6}A) for predicting the m^{6}A readout effects in terms of half-life down-regulated genes (A), translational efficiency down-regulated genes (B), translational efficiency up-regulated genes (C), exon inclusion ratio down-regulated exons (D), and exon inclusion ratio up-regulated exons (E) in different cell types. (F) Bubble plot summarizing the performance contributions (as evaluated by the gain of AUC) of different feature sets for predicting half-life down-regulated genes. (G) Bubble plot summarizing the performance contributions of different feature sets for predicting translational efficiency down-regulated genes. (H) Bar plot summarizing the RNA binding proteins (RBPs) with high feature importance score in the machine learning models, where red color indicates known m^{6}A readers, orange color indicates known m^{6}A writers/erasers, green color indicates the newly identified m^{6}A reader in this study (FUBP3, DDX6, L1TD1 and FXR2), and blue color indicates the RBPs showing no significant m^{6}A binding ability in our RNA pull-down assay.

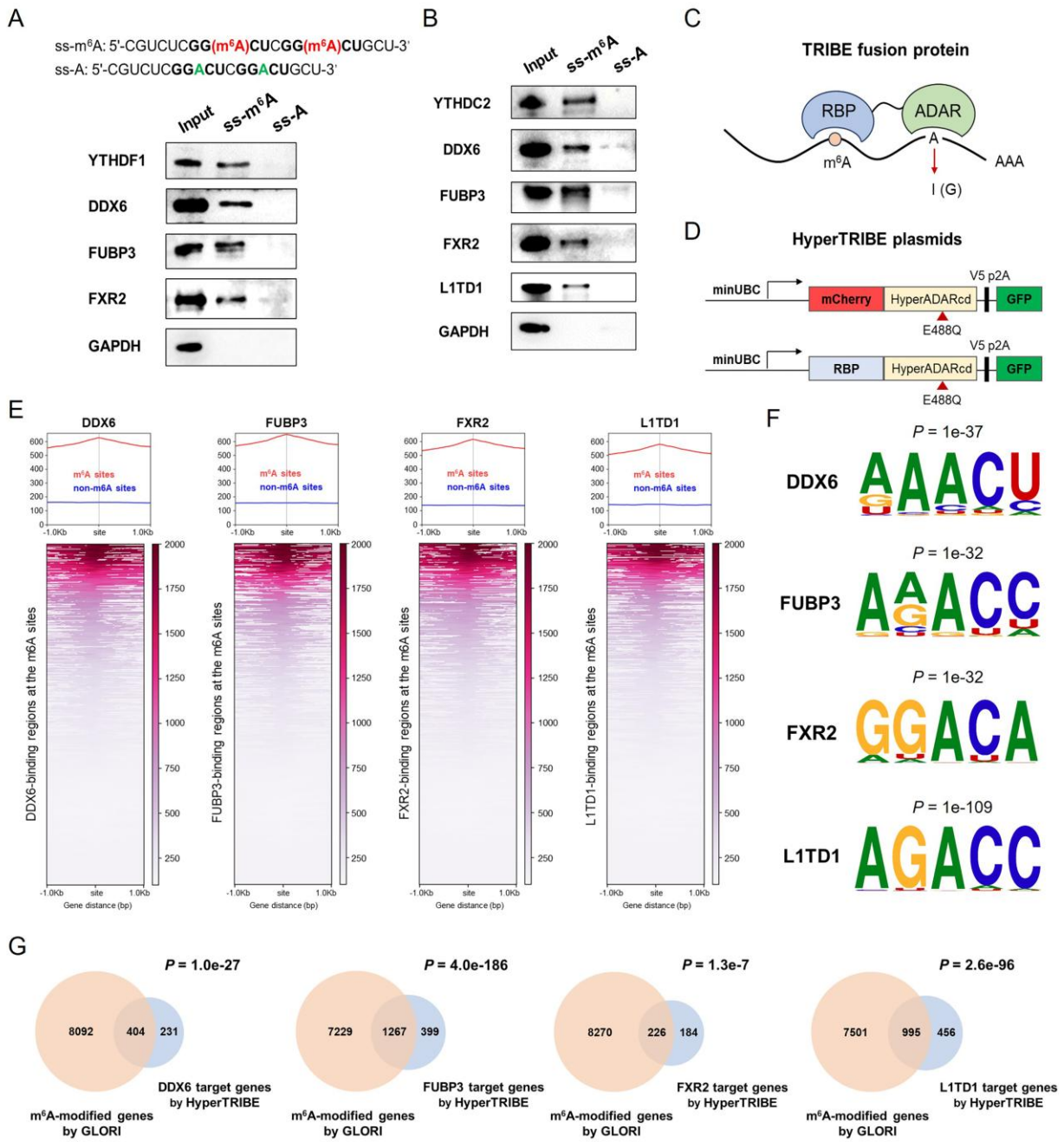


Figure 5. Identification of m^{6}A readers in hESCs.

(A) RNA pull-down assays that verify the binding of RBPs to m^{6}A in A549 cells. (B) RNA pull-down assays that verify the binding of RBPs to m^{6}A in hESCs. (C) Overview of principle of the TRIBE/HyperTRIBE for transcriptome-wide identification of RBP binding sites. (D) Diagram of plasmid constructs for HyperTRIBE. (E) Heatmap showing the distribution of HyperTRIBE-identified RBP binding sites near the m^{6}A methylation sites determined by GLORI in hESCs. (F) DRACH-like motifs in the proximal sequences of RBP binding sites

in hESCs. (G) Venn diagram showing the intersection between RBP target genes identified by HyperTRIBE and m^6A target genes identified by GLORI in hESCs.

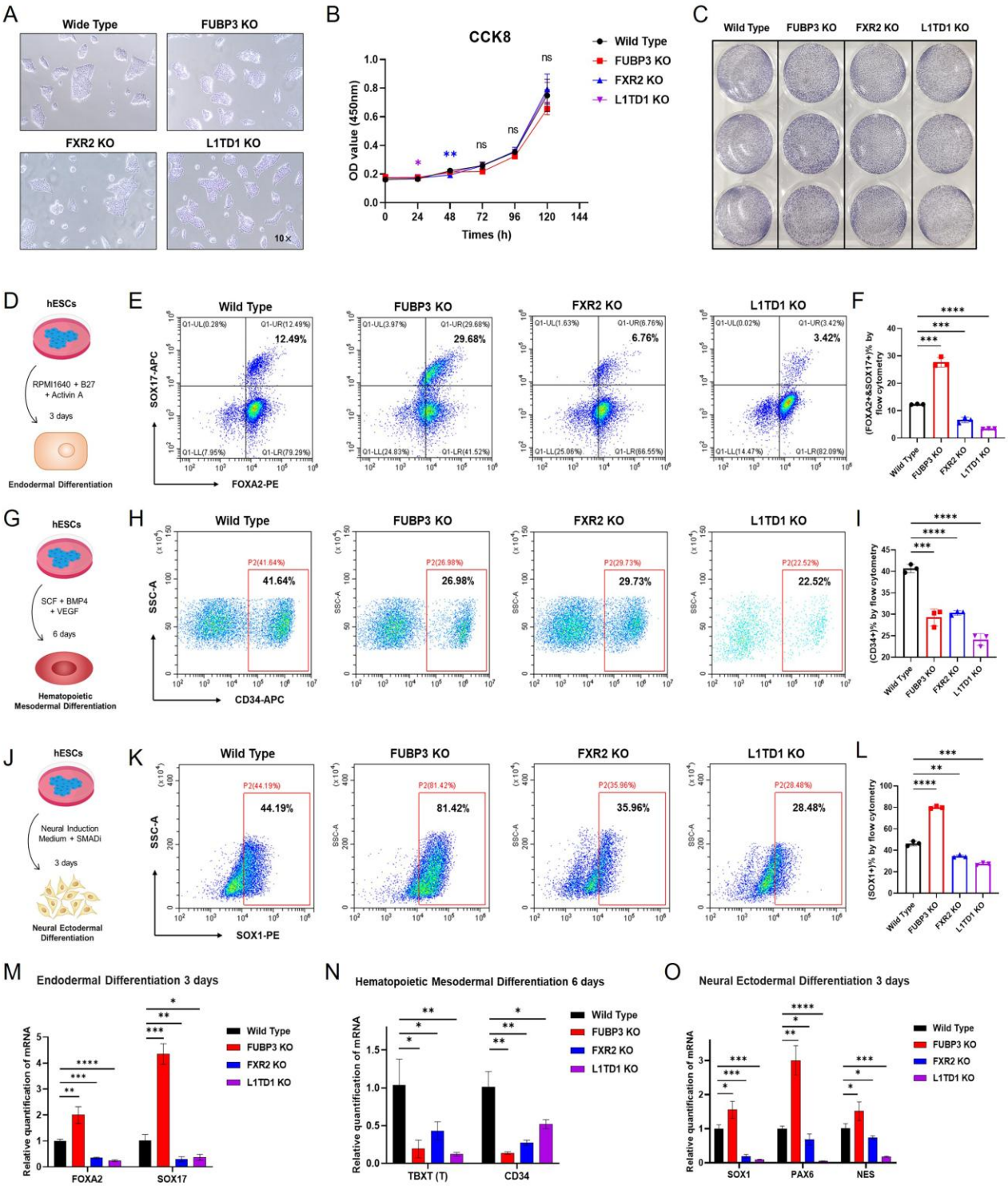
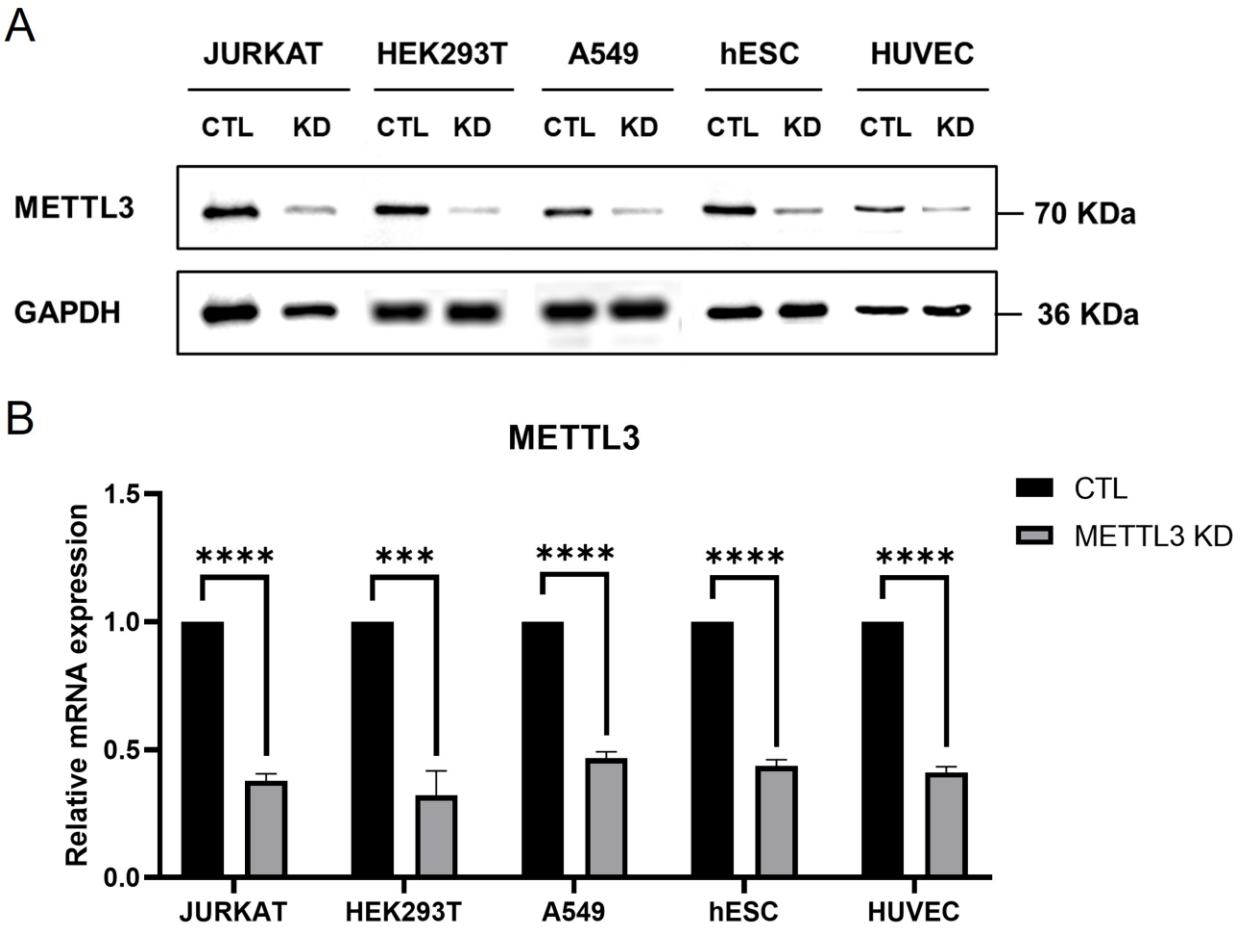


Figure 6. Assessment of m^{6}A readers' regulatory effects on hESC proliferation and differentiation.

(A) Representative images of morphology of WT versus *FUBP3*, *FXR2* and *L1TD1* KO hESCs under bright field. (B) Growth curve assessing cell proliferation kinetics of *FUBP3*, *FXR2* and *L1TD1* KO versus WT hESCs. (C) ALP staining assay of WT versus *FUBP3*, *FXR2* and *L1TD1* KO hESCs. (D) Schematic representation of the induction endodermal differentiation experiment of hESCs. (E-F) Flow cytometry and statistical histogram analysis of endodermal differentiation of WT versus *FUBP3*, *FXR2* and *L1TD1* KO hESCs. (G) Schematic representation of the induction hematopoietic mesodermal differentiation experiment of hESCs. (H-I) Flow cytometry and statistical histogram analysis of hematopoietic mesodermal differentiation of WT versus *FUBP3*, *FXR2* and *L1TD1* KO hESCs. (J) Schematic representation of the induction neural ectodermal differentiation experiment of hESCs. (K-L) Flow cytometry and statistical histogram analysis of neural ectodermal differentiation of WT versus *FUBP3*, *FXR2* and *L1TD1* KO hESCs. (M-O) RT-qPCR analysis of definitive endoderm markers SOX17, FOXA2 (M); mesoderm marker TBXT (also known as T) and CD34 (N); and neural-ectoderm markers SOX1, PAX6, NES in WT versus *FUBP3*, *FXR2* and *L1TD1* KO hESCs. (O). Figures in (D), (G) and (J) were drawn by using pictures from Biovisart (https://biovisart.com.cn) and Bioicons (https://bioicons.com/) with permission. For all histograms, * means $P < 0.05$, ** means $P < 0.01$, *** means $P < 0.001$, **** means $P < 0.0001$, ns means not significant.

1005 ## Supplementary Figures



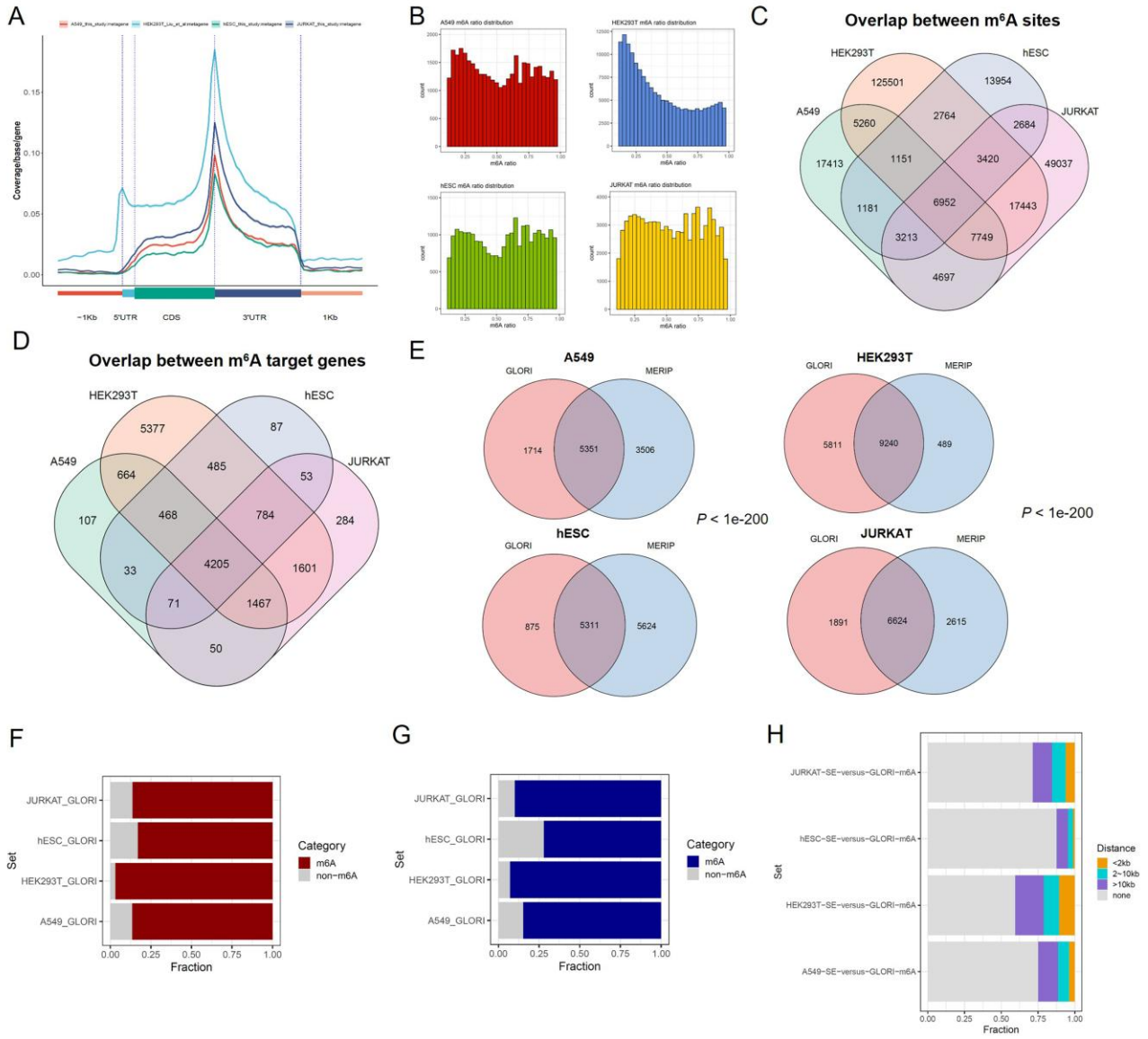
1006

1007 **Supplementary Figure S1. Verification of *METTL3* knockdown effects.**

1008 (A) Western blot analysis for expression of METTL3 in shControl and METTL3 KD cells. (B) RT-qPCR
1009 analysis for the relative expression of *METTL3* mRNA in shControl and METTL3 KD cells. **** means P <
1010 0.0001.

1011

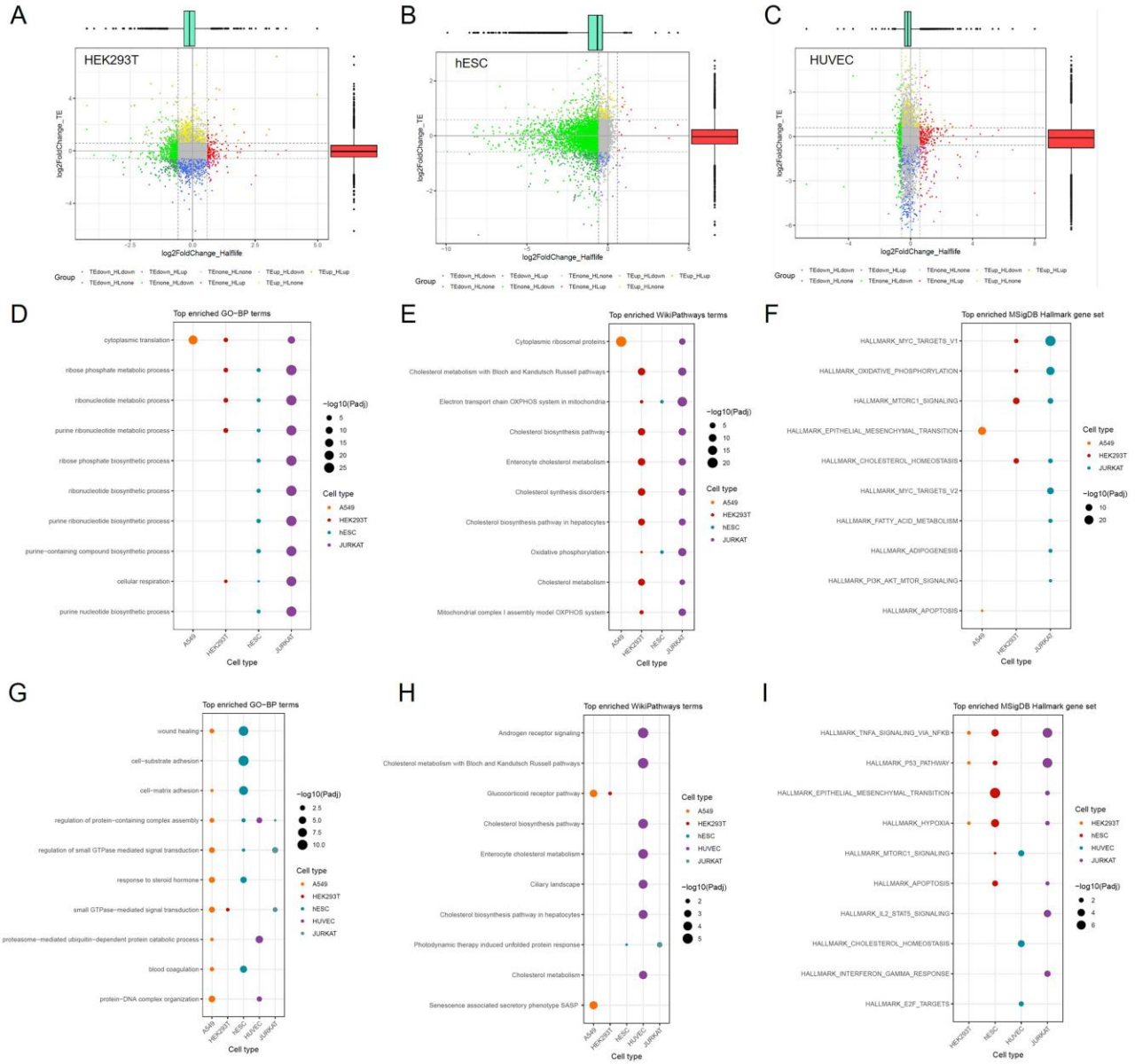
1012



Supplementary Figure S2. Overview of m^{6}A modification sites identified by GLORI.

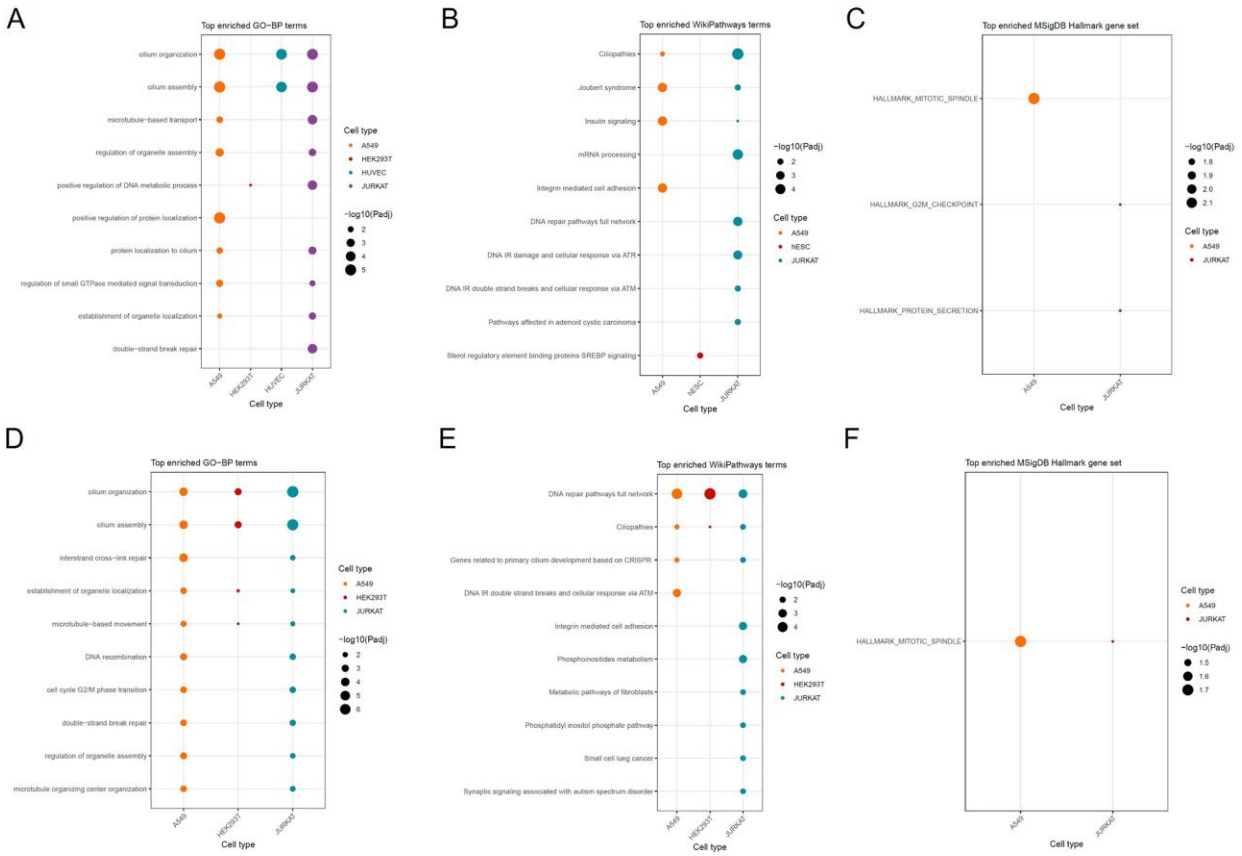
(A) Metagene plot showing the distribution of GLORI-identified m^{6}A modification sites in the target genes. Western blot analysis for expression of METTL3 in shControl and shMETTL3 cells. (B) Histogram illustration of m^{6}A modification ratio estimated by GLORI. (C) Venn diagram showing the overlap of m^{6}A modification sites between different cell types. (D) Venn diagram showing the overlap of m^{6}A target genes between different cell types. (E) Venn diagram showing the overlap between m^{6}A target genes identified by GLORI and those identified by MeRIP-seq in the corresponding cell types. (F) Percentage of half-life down-regulated genes with m^{6}A methylation peaks identified by GLORI on the corresponding cell types. (G) Percentage of genes showing translational efficiency changes with m^{6}A methylation peaks identified by GLORI on the

corresponding cell types. (H) Percentage of exons showing exon inclusion ratio changes with m^6A methylation sited identified by GLORI on the corresponding cell types) near their exon boundaries.



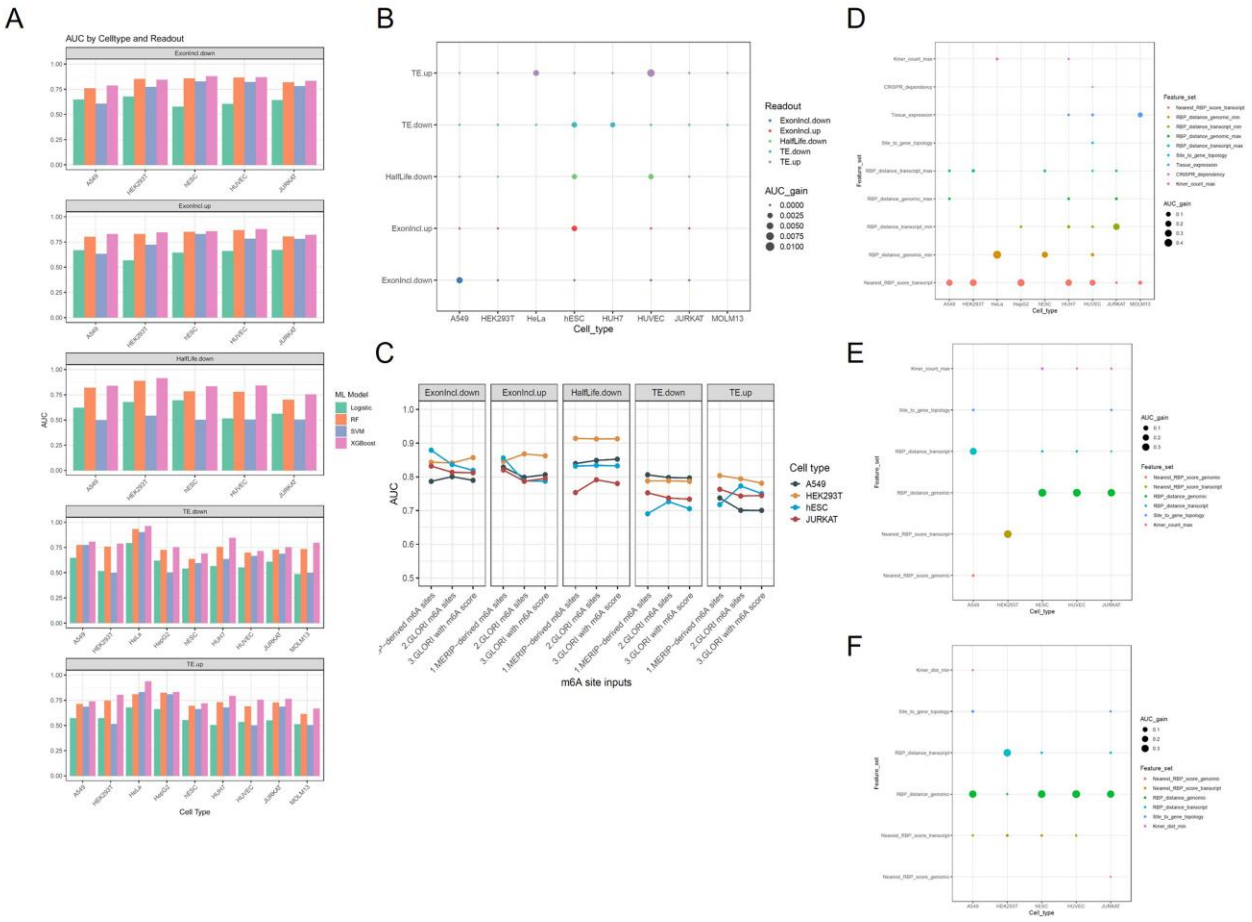
Supplementary Figure S3. Distribution and functional association of genes showing m^{6}A-mediated translational efficiency changes.

(A-C) Scatter plot showing translational efficiency versus half-life changes in HEK293T (A), hESC (B) and HUVEC cells (D) Top 10 enriched Gene Ontology-Biological Process (GO-BP) functional terms for translational efficiency down-regulated genes. (E) Top 10 enriched WikiPathways pathways for translational efficiency down-regulated genes. (F) Top 10 enriched HsigDB Hallmark gene set for translational efficiency down-regulated genes. (G-I) The corresponding results for translational efficiency up-regulated genes.



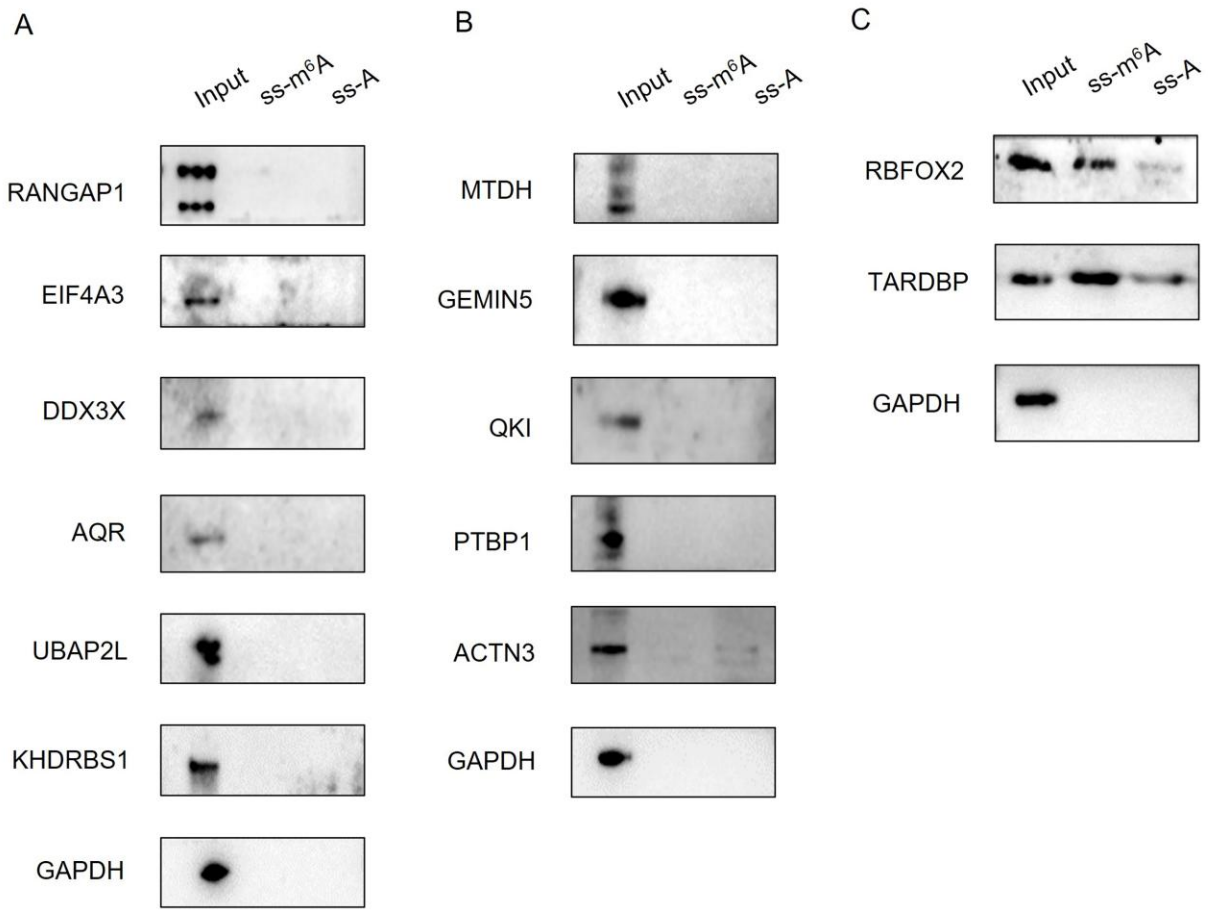
Supplementary Figure S4. Functional association of genes with exon skipping events.

(A) Top 10 enriched GO-BP functional terms for exon inclusion ratio down-regulated genes. (B) Top 10 enriched WikiPathways pathways for exon inclusion ratio down-regulated genes. (C) Top 10 enriched HsigDB Hallmark gene set for exon inclusion ratio down-regulated genes. (D-F) The corresponding results for exon inclusion ratio up-regulated genes.



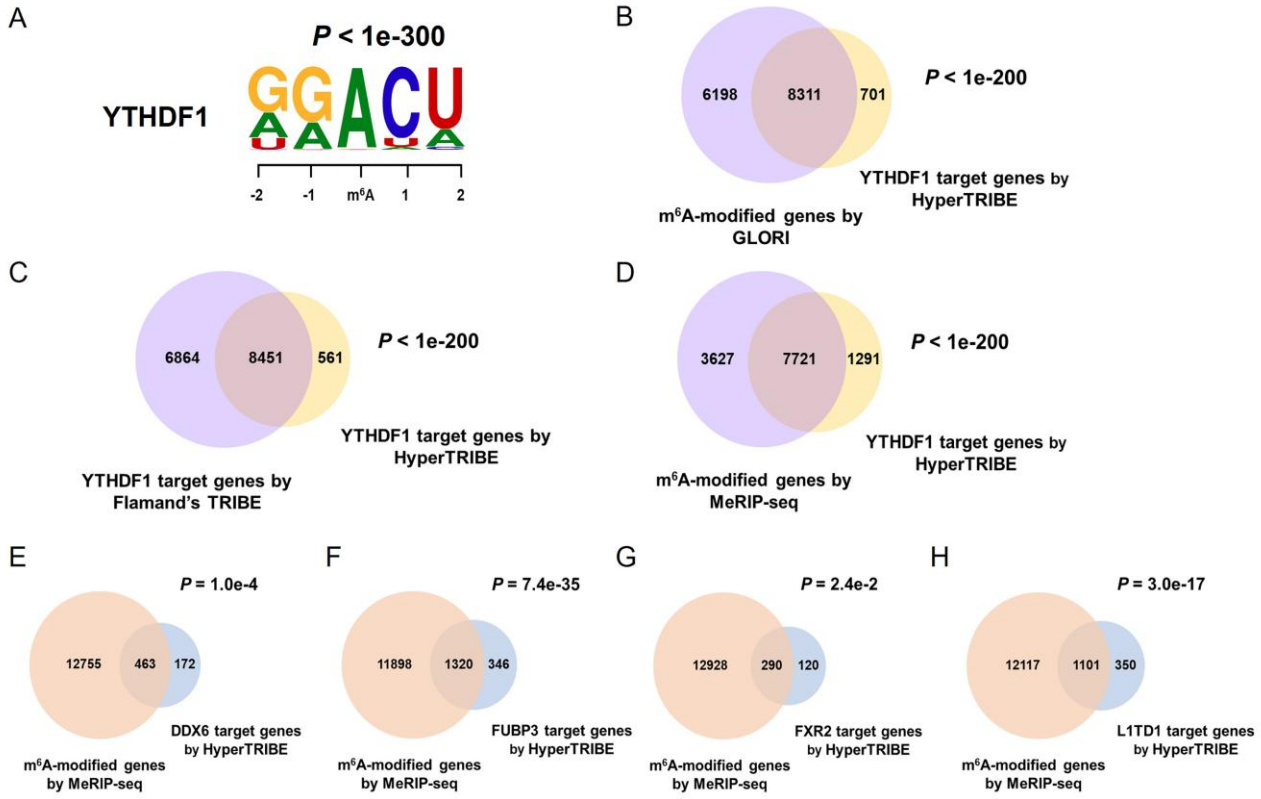
Supplementary Figure S5. Ablation experiments and feature set contribution of the machine learning predication models of m^{6}A readouts.

(A) Comparison between XGBoost and other classical machine learning models in predicting different categories of m^{6}A readouts. The classical machine learning models here include random forest (RF), logistic regression (Logistic) and support vector machine (SVM). ExonIncl means exon inclusion ratio; TE means translational efficiency. (B) Bubble plot summarizing the gain of AUC after parameter optimization of the XGBoost model. (C) Line plot comparing the performance of XGBoost models with MeRIP-seq-derived m^{6}A sites, GLORI-derived m^{6}A sites or GLORI-derived m^{6}A sites plus the corresponding m^{6}A methylation rate. (D) Bubble plot summarizing the performance contributions of different feature sets for predicting translational efficiency up-regulated genes. Bubble plot summarizing the performance contributions of different feature sets for predicting exon inclusion ratio down-regulated exons. (F) Bubble plot summarizing the performance contributions of different feature sets for predicting exon inclusion ratio up-regulated exons.



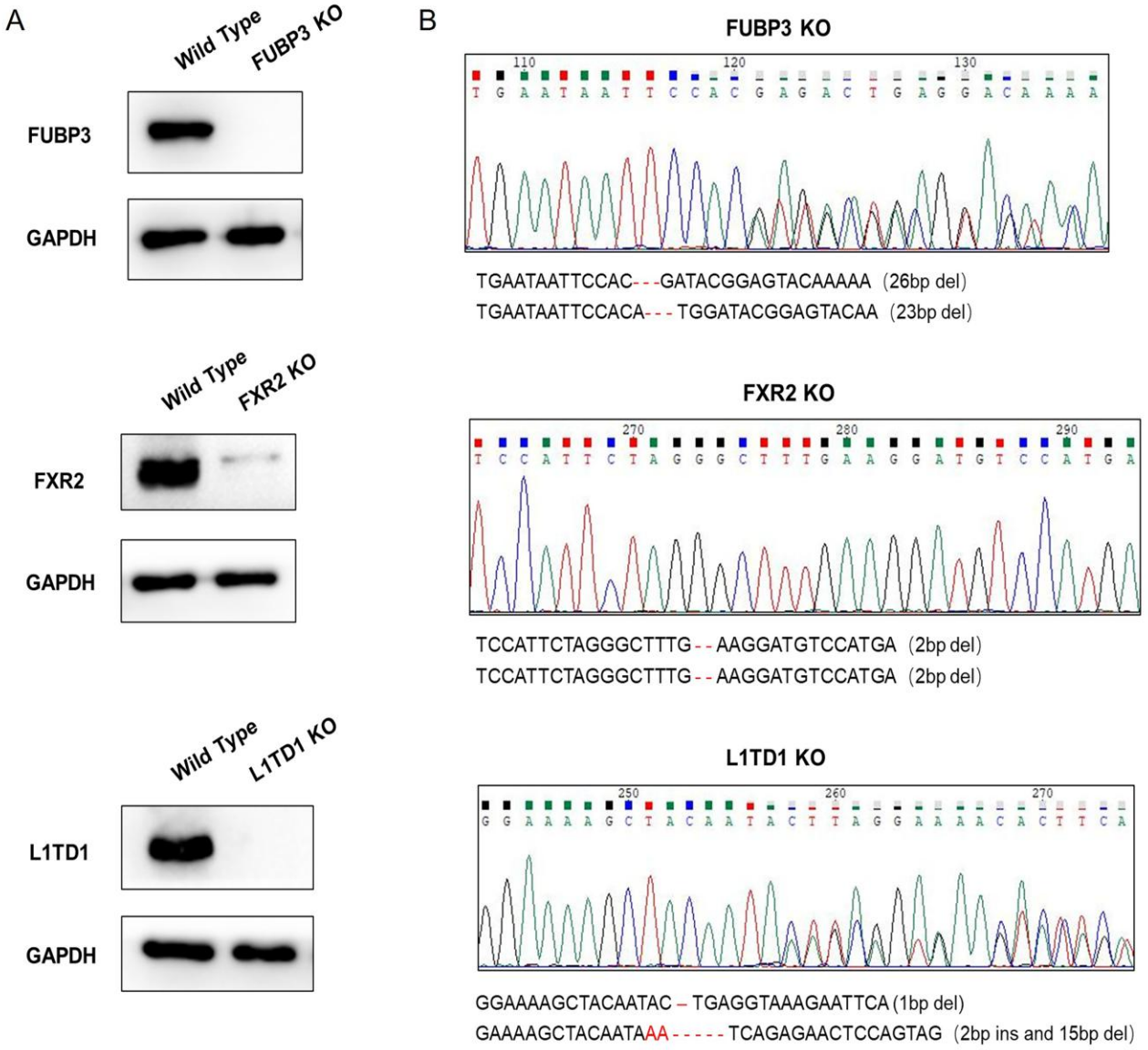
Supplementary Figure S6. Screening of m^{6}A readers by RNA pull-down assay.

Three batches of RNA pull-down assay were performed. (A) RNA pull-down assay batch 1. (B) RNA pull-down assay batch 2. (C) RNA pull-down assay batch 3. RBFOX2 and TARDBP (also known as TDP43) are known m^{6}A readers that serve as positive controls here.



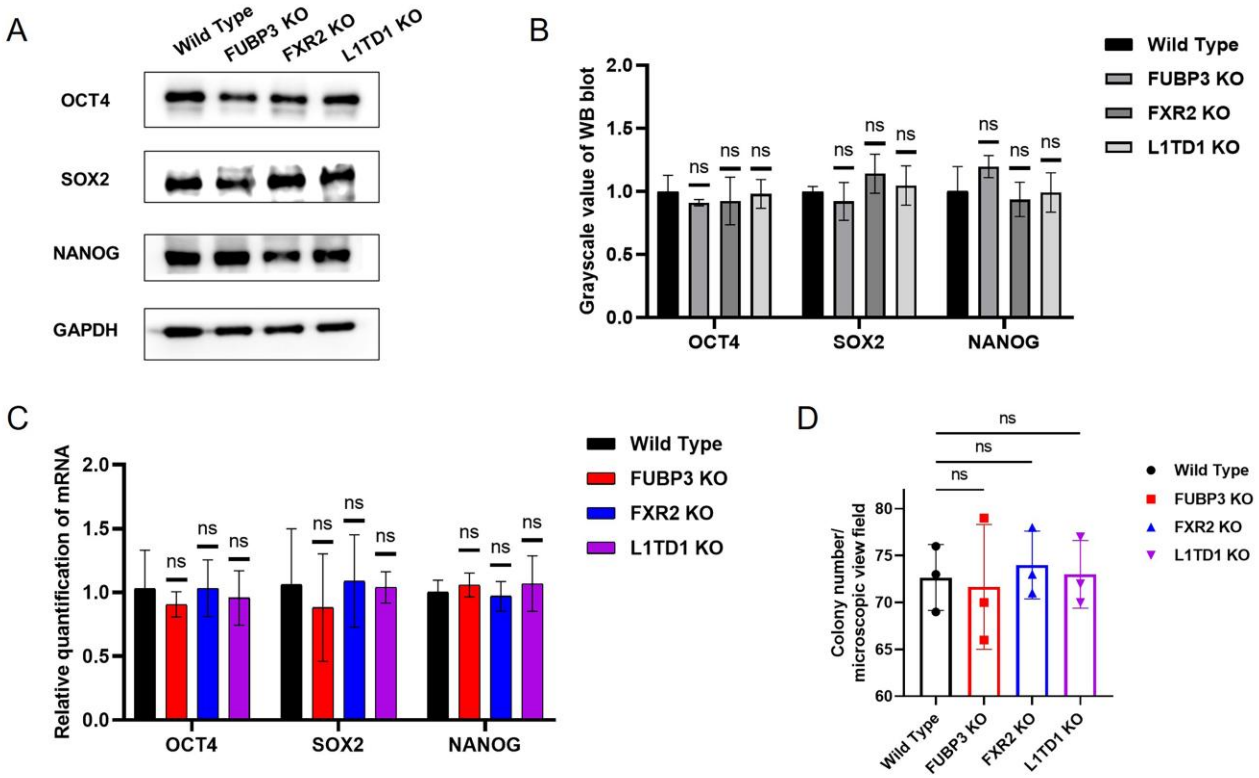
Supplementary Figure S7. Overlaps between the binding target genes of m^{6}A readers and m^{6}A modification.

HyperTRIBE assay of YTHDF1 in HEK293T was performed as the positive control assay. Then the HyperTRIBE assay was performed in hESCs for the four putative m^{6}A readers (DDX6, FUBP3, FXR2 and L1TD1) found in this study. (A) DRACH-like motifs in the proximal sequences of YTHDF1 binding sites in HEK293T. (B) Venn diagram showing the intersection between YTHDF1 target genes identified by HyperTRIBE and m^{6}A target genes identified by GLORI in HEK293T. (C) Venn diagram showing the intersection between YTHDF1 target genes identified by HyperTRIBE and those identified by Flamand et al. (D) Venn diagram showing the intersection between YTHDF1 target genes identified by HyperTRIBE and m^{6}A target genes identified by MeRIP-seq in HEK293T. (E-H) Venn diagrams showing the intersection between DDX6 (E), FUBP3 (F), FXR2 (G) and L1TD1 (H) target genes identified by HyperTRIBE and m^{6}A target genes identified by MeRIP-seq in hESC.



Supplementary Figure S8. Validation of *FUBP3*, *FXR2* and *L1TD1* KO cell lines by Western blot and sanger sequencing.

(A) The validation of *FUBP3*, *FXR2* and *L1TD1* KO cell lines by Western blot. (B) The validation of *FUBP3*, *FXR2* and *L1TD1* KO cell lines by sanger sequencing.



Supplementary Figure S9. Regulatory effects of m^{6}A readers on self-renewal of hESCs.

(A) Representative western blot analysis of expression of OCT4, SOX2 and NANOG in WT versus *FUBP3*, *FXR2* and *L1TD1* KO hESCs. (B) The statistics of grayscale value of western blot, ns means not significant. (C) RT-qPCR analysis of OCT4, SOX2, NANOG mRNA expression in WT versus *FUBP3*, *FXR2* and *L1TD1* KO hESCs. ns means not significant. (D) Statistics of the ALP-positive colony numbers from WT versus *FUBP3*, *FXR2* and *L1TD1* KO hESCs. ns means not significant.