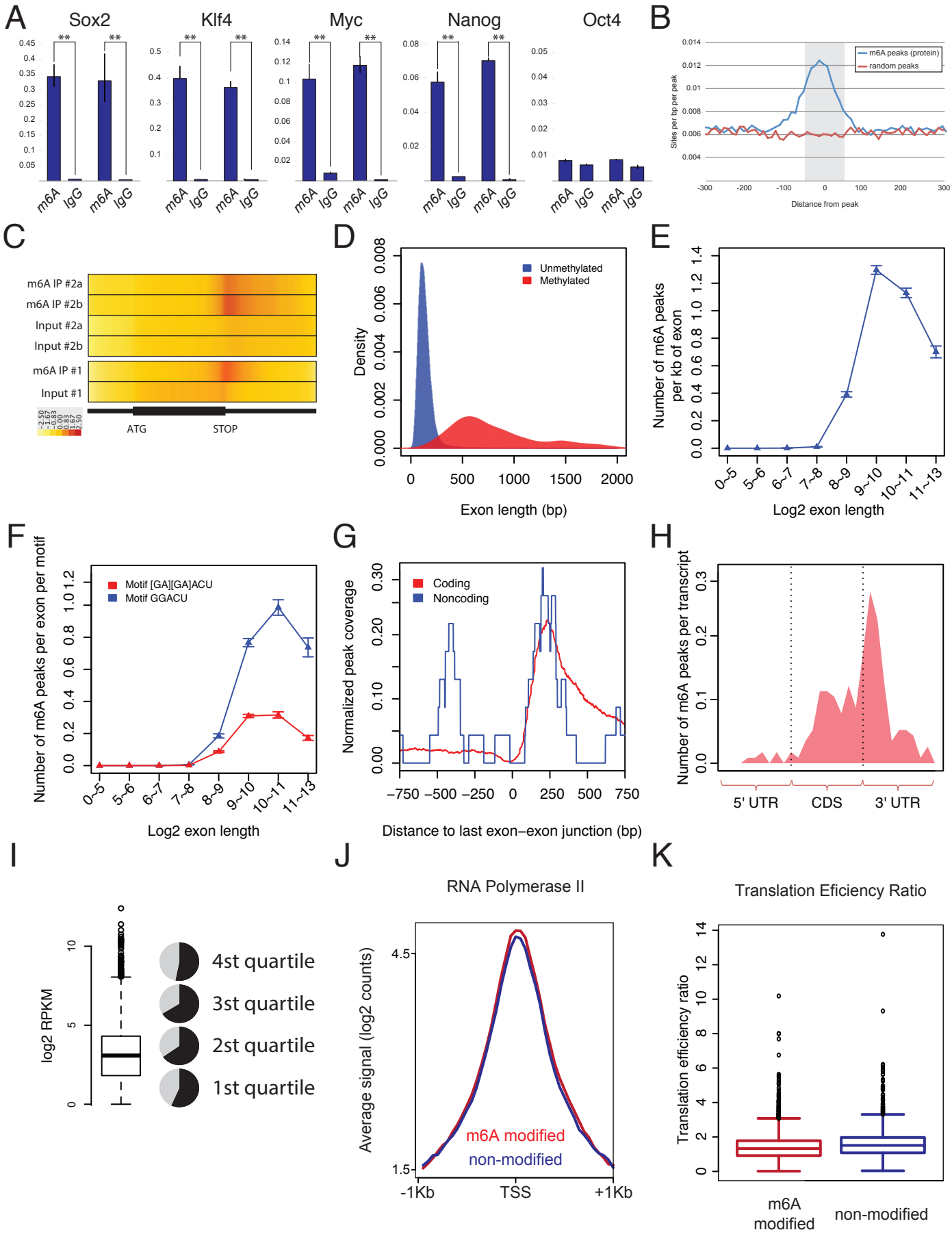
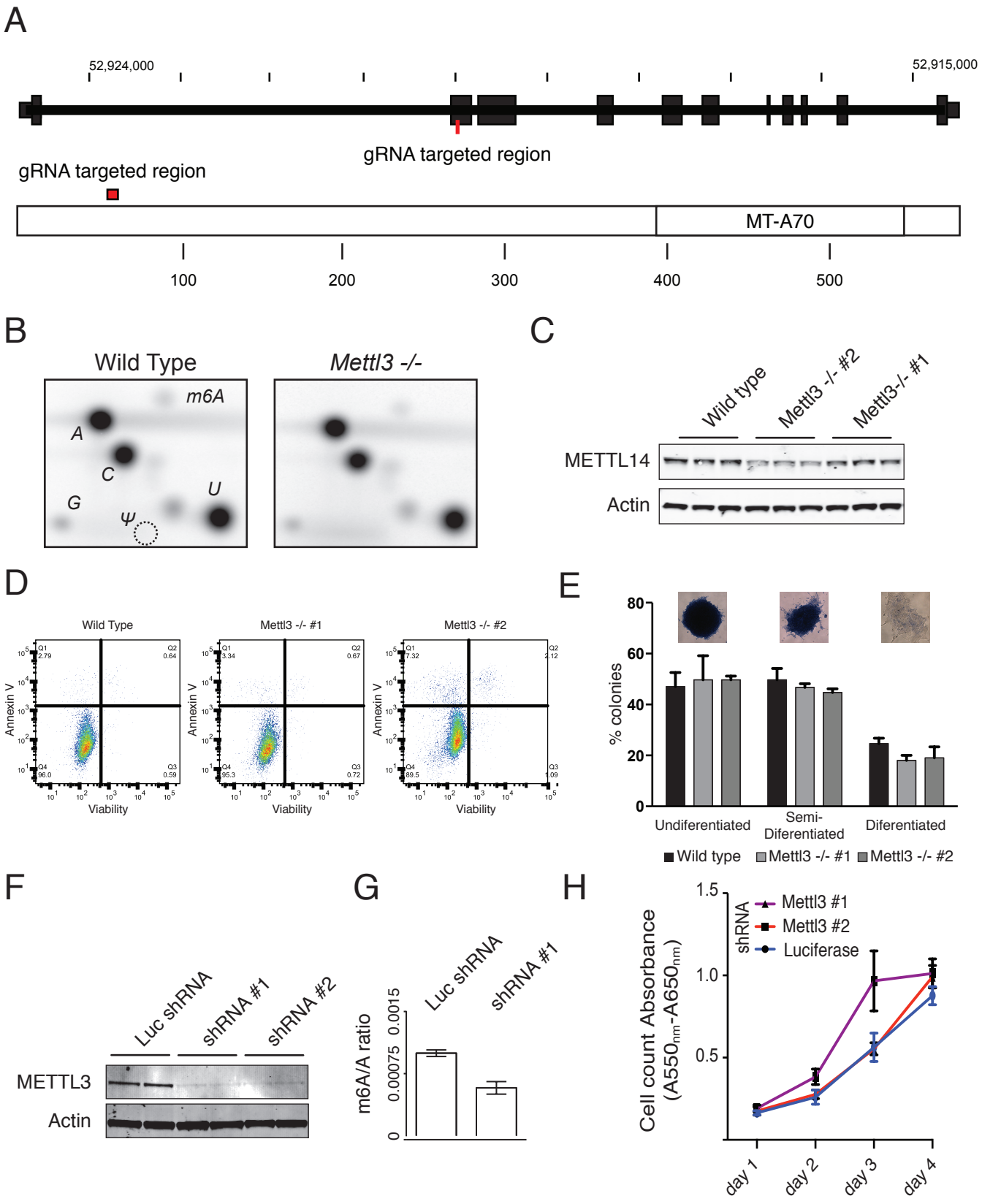
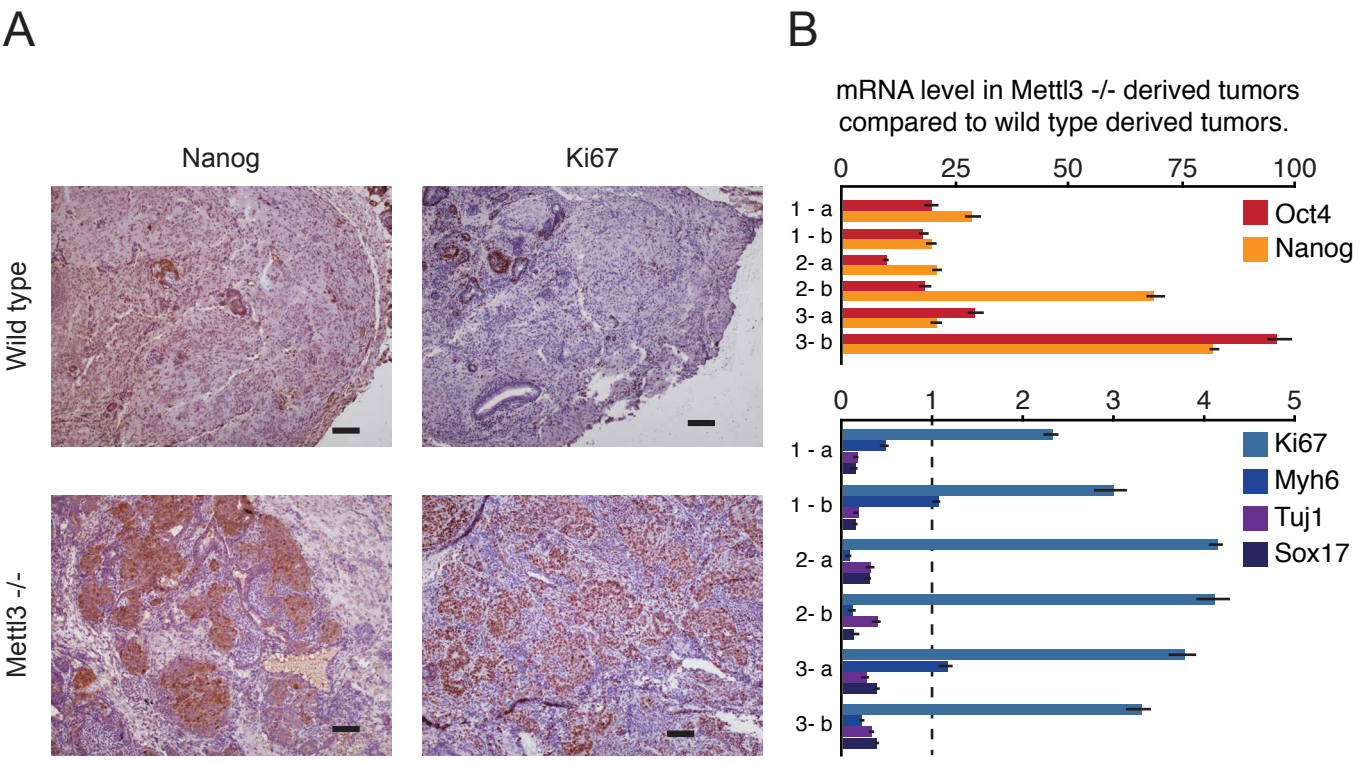
Supplemental Figure 1



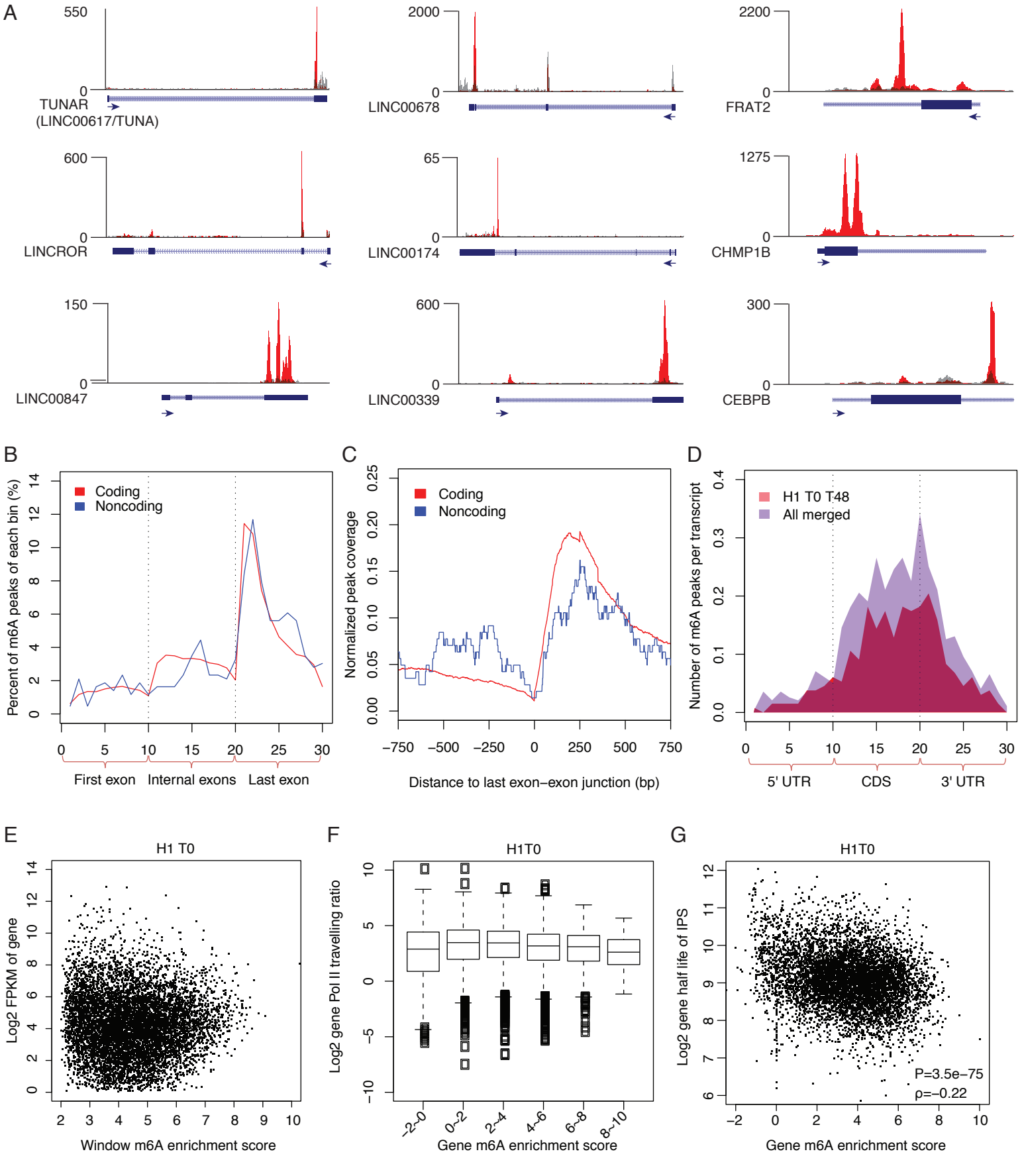
Supplemental Figure 2



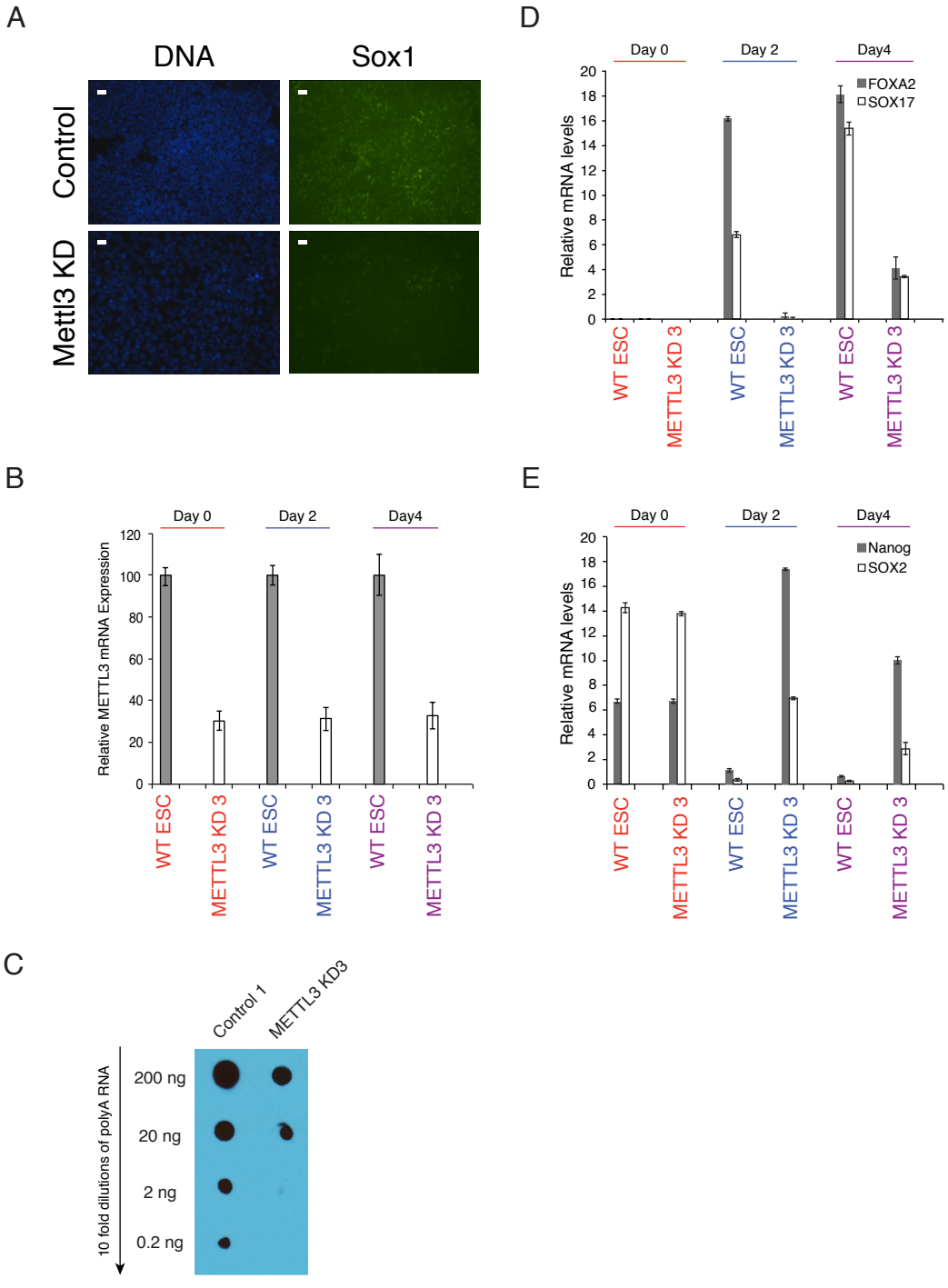
Supplemental Figure 3



Supplemental Figure 4



Supplemental Figure 5



Supplemental Figures

Figure S1. Topology and characterization of m^6A target genes. Related to Figure 1.

(A) m6A enrichment determined by qRT-PCR. Vertical axis represents percentage of recovery. Error bars represent standard deviation of the ΔΔCT value. ** represents p-value < 0.01. Related to Figure 1A. **(B)** Histogram representing motif density in m6A peaks (Blue) and a random control group of windows (Red). Related to Figure 1D. **(C)** Metagene representation of read density obtained in input and after m6A enrichment for genes with at least one modification. Black thick box represents the open reading frame while the black line represents the untranslated regions. The CDS and 3' UTR are divided in 100 bins, while the 5' UTR is divided in 50 bins. Related to Figure 1D. **(D)** The exon length distribution of methylated vs unmethylated internal exons of coding genes is shown. Related to Figure 1G. **(E)** The number of peaks per exon normalized by exon length is shown for different bins of exon length. The error bars represent standard deviations from 1000 times of bootstrapping. Related to Figure 1G. **(F)** The number of peaks per exon normalized by the number of motifs (on sense strand) in the exon is shown. The error bars represent standard deviations from 1000 times of bootstrapping. Related to Figure 1G. **(G)** The density of m6A-seq read coverage increases sharply downstream of the last exon-exon splice junction in both coding and non-coding RNAs. Related to Figure 1G. **(H)** Percentage of m6A peaks that fall into normalized bins across the 5'UTR, CDS, and 3'UTR of single-exon genes is shown. Related to Figure 1G. **(I)** Pie charts representing the fraction of genes with m6A modification for each quartile of expression. Black area represents modified genes. Related to Figure 1H. **(J)** Average coverage of Pol2 signal at the transcriptional start site of modified and unmodified genes. Related to Figure 1I. **(K)** Box plot representing translation efficiency as measured by ribosome profile. Related to Figure 1I.

Figure S2. Characterization of Mettl3 knock out cells. Related to Figure 2.

(A) Representation of the Mettl3 locus, and METTL3 protein, with the CRISPR targeted region marked in red. Related to Figure 2A. **(B)** Representative examples of 2D-TLC plates for mESC wild type and Mettl3 -/- mutant. Nucleotide positions are indicated in the leftmost panel. Related to Figure 2B. **(C)** Western blot for Mettl14 in wild type and two cell lines with Mettl3 KO cell lines. Actin is used as loading control. Related to Figure 2B. **(D)** FACS plots of Annexin V and Aqua Live/Dead fixable Viability dye for Wild type and two Mettl3 KO cell lines. Related to Figure 2C. **(E)** Quantification of colony morphologies for Wild type and two Mettl3 KO cell lines. Experiment performed in triplicate, with at least 50 colonies counted per replicate. Error bars represent standard deviation. Related to Figure 2C. **(F)** Western blot for Mettl3 in wild type and two independent Mettl3 shRNAs. Actin is used as loading control. Related to Figure 2F. **(G)** m6A ratio determined by 2D-TLC in wild type and Mettl3 shRNA line. Related to Figure 2F. **(H)** Cell proliferation assay of wild type and two independent Mettl3 shRNA lines. Related to Figure 2F.

Figure S3. Mettl3 loss of function impairs ESC ability to differentiate. Related to Figure 3.

(A) Representative sections of teratomas stained with hematoxylin and eosin (left), and immunohistochemistry with antibody against Nanog (center) and Ki67 (right). The bar represents 100 μm. Related to Figure 3D. **(B)** Relative mRNA levels between Mettl3 -/- derived tumors and wild type derived tumors for Oct4, Nanog, Ki67, Myh6, Tuj1 and Sox17. Error bars represent standard deviation of the ΔΔCT value. Related to Figure 2E and 2F.

Figure S4. m^{6}A-seq profiling of hESC during endoderm differentiation. Related to Figure 5.

(A) Examples of m6A location in multi-exon non-coding RNAs and single-exon mRNAs. UCSC Genome browser plots of m6A-seq reads (red) along indicated RNAs in undifferentiated hESCs (i.e. T0). The grey reads are from non-immunoprecipitated control input libraries. The read density is calculated from the average of the two replicate T0 samples. Arrow indicates the direction of transcription. Related to Figure 5D. **(B)** Multi-exon coding and non-coding RNAs exhibit enrichment of m6A sites near the last exon-exon splice junction. The distribution of m6A peaks across the length of the mRNAs (n= 9489) and noncoding RNAs (n= 207) is shown. The 5' most (first) exon, all internal exons, and the 3' most (last) exon are divided into 10 bins and the percentage of m6A peaks that fall within each bin are shown. Related to Figure 5I. **(C)** The density of m6A-seq read coverage increases sharply downstream of the last exon-exon splice junction in both coding (n= 5231) and non-coding RNAs (n= 68). Related to Figure 5I. **(D)** Single-exon genes tend to have more m6A sites at their 3' end. The percentage of m6A peaks that fall into normalized bins across the 5'UTR, CDS, and 3'UTR of single-exon genes is shown for hESC cells (T0 and T48 combined, n=137) as well as in merged data ("All merged"; n=200) from hESCs, 293T (Meyer et al., 2012) and HepG2 (Dominissini et al., 2012). Related to Figure 5I. **(E)** Scatter plot representation of m6A enrichment score (on the X axis) and gene expression level in FPKM (on the Y axis) for each m6A peak. Related to Figure 5J. **(F)** m6A peak intensity is not correlated with nascent RNA transcription based on pausing index. The m6A enrichment scores vs GRO-seq deterimined the Pol II traveling ratio is plotted. The pausing index equal GRO-seq density at promoter defined as -300 and +300 of TSS divided by GRO-seq density in the gene body defines as +300 to end of the gene. Related to Figure 5J. **(G)** mRNA half-life is anti-correlated with m6A enrichment in genes. Related to Figure 5J.

Figure S5. METTL3 is required for normal human ESC endoderm differentiation. Model of METTL3 function(s). Related to Figure 7.

(A) Staining for Sox1 and DNA of neural stem cells in Mettl3 knock down (KD) and control cells. **(B)** Knockdown of METTL3 leads to a reduction in METTL3 mRNA levels. qRT-PCR for METTL3 mRNA was performed from RNA extracted from control WT hESC cells versus hESCs with anti-METTL3 shRNA (KD) clone #3 across the three indicated time points during endoderm differentiation. Error bars represent standard deviation across 3 replicates per time point. Related to Figure 7B. **(C)** Knockdown of METTL3 leads to a functional reduction in m6A levels. An anti-m6A dot blot was performed on 10x fold dilution of polyA selected RNA from wildtype (WT) hESC cells versus anti-METTL3 knockdown (KD) clone #3. Related to Figure 7C. **(D)** Knockdown of METTL3 leads to a delayed and reduced induction of endodermal marker genes. qRT-PCR was performed on indicated genes and time points. Error bars represent standard deviation across 3 replicates per time point. Related to Figure 7D and 7E. **(E)** Knockdown of METTL3 leads prevents the normal reduction of stem maintenance/marker genes. qRT-PCR was performed for indicated genes and time points. Error bars represent standard deviation across 3 replicates per time point Related to Figure 7D and 7E.

Supplemental Table

Supplemental Table 1: All mouse high confidence peaks. Related to Figure 1 and Figure 4. Coordinates of m^6A peaks in mouse genome (mm9), position of the m^6A peak in the transcript, type of transcript and Gene symbol. For the Difference in Mettl3 the ratio between the IP and the Input is represented.

Supplemental Table 2: Nanostring counts after m^6A immunoprecipitation. Related to Figure 1. Gene symbol with counts for Input, m^6A immunoprecipitation and IgG. The ratio to the input and fold enrichment over the gene body of Actb are represented.

Supplemental Table 3: All human high confidence peaks. Related to Figure 5. Coordinates of m^6A peaks in human genome (mm9), type of transcript and Gene symbol.

Supplemental Table 4: DPMI between T0 and T48. Related to Figure 5. Coordinates of m^6A peaks in human genome (mm9), type of transcript and Gene symbol. For each row is indicated if DPMI is over 1.5 or 2 fold.

Supplemental Table 5: Human and mouse methylated gene comparasion. Related to Figure 6. Gene ID in human and mouse and type of homology.

Supplemental Table 6: hESC and mESC common peaks. Related to figure 6. Gene ID in human and mouse and chromosome coordinates of common peaks.

Supplemental Movies

Video of Cardiomyocite colonies differentiated from wild type (Movie S1) or Mettl3 KO (Movie S2) mouse embryonic stem cells. Videos where capture after 12 days of differentiation. Related to Figure 3A.

Supplemental Methods

mESC cell Culture and differentiation

J-1 murine embryonic stem cells were grown under typical feeder free ES cell culture conditions. Cells were grown in gelatinized (0.2% Gelatin) tissue culture plates in mESC media (KnockOut DMEM (Gibco, Life Technologies; 10829-018) supplemented with 1000 U/ml leukemia inhibitory factor (Millipore; ESG1107), 1x non-essential amino acids (Gibco, Life Technologies; 11140-050), 1x Glutamax (Gibco, Life Technologies; 35050-061), 10% Pen Strep (Gibco, Life Technologies; 151140-122) and 15% Fetal Bovine Serum (HyClone, SH30071.03)).

For cardiomyocite differentiation, mESCs were plated at a density of 2 x 10^5 cells/mL in ultra-low attachment plates in cardiomyocyte differentiation media (CMD) (DMEM [GIBCO], 15% FBS [Hyclone], 1% penicillin/streptomycin, 1% GlutaMax and 1mM Ascorbic Acid [Sigma]) to induce EB formation. Media was changed on day 3 and on day 6, EBs were re-suspended in fresh CMD media and replated on 0.2% gelatin coated dishes. Media was changed on day 9 and on day 12 the number of contracting patches of cells was quantified in triplicate for each cell line.

For Neuron differentiation, Mouse embryonic stem cells were grown in mESC medium (DMEM (Invitrogen), 12% knockout replacement serum (Invitrogen), 3% cosmic calf serum (Thermo Scientific) supplemented with non-essential amino acids (Invitrogen), penicillin-streptomycin (Invitrogen), sodium pyruvate (Invitrogen), 2-mercaptoethanol (Invitrogen) and LIF). Cells were dissociated in 2.5% trypsin for 5 minutes, pelleted, and resuspended on a gelatinized plate in MEF medium (DMEM, 10% cosmic calf serum, non-essential amino acids, penicillin-streptomycin, sodium pyruvate, 2-mercaptoethanol) for 30 minutes to remove feeders. 5x10^6 mESCs were then replated onto 10cm bacterial plates in MEF medium and cultured for 4 days. On day 4, cells were replated under adherent culture conditions. Medium was replaced with ITSFn

medium (DMEM:F12 (Invitrogen), insulin [5ug/ml], apotransferrin [50ug/ml], sodium selenate [30nM], fibronectin [250ng/ml]) the following day and replaced every other day. Cells were cultured for 10 days in ITSFn before fixation.

For the cell proliferation assay (MTT) 5 thousand cells where cultured in 24 well dish and the assay performed according to the manufacturer's protocol (Roche; 11465007001). For the single colony assays and Nanog staining, 1 thousands cells where cultured, per well, on a six well dish.

For Alkaline Phosphatase Staining, at day 6 cells were fixed (50% Methanol, 50% Acetone) and stained for Alkaline Phosphatase with Vector Blue Alkaline Phosphatase Substracte Kit (Vector; 5300), according to manufacturer's protocol.

For Nanog and Oct4 staining cells where fixed with 4% paraformaldehyde (PFA) (Thermo Scientific, 28909). Cardiomyocites were cultured in chamber slides and fixed on day 12 with 4% PFA and N cells where fixed for 20 minutes in 4% PFA. Cells where washed 3 times with PBS and blocked in PBS with 0.1% Triton and 5% FBS (for N cells, CCS was used instead of FBS) for 20 minutes. Cells where then incubated with primary antibody [Rabbit anti-Nanog Antibody, Bethyl; mouse anti-Oct-3/4, Santa cruz, mMF20, Developmental studies Hybridoma bank; anti-Tuj1, Covance (1:1000), rabbit anti-Nanog, ReproCell (1:200)] for 30 minutes in blocking medium. After 3 PBS washes, cells where incubated with secondary antibody (Alexa 488 Goat anti-mouse, Alexa Goat anti-Rabbit, donkey Alexa-555 anti-mouse, donkey Alexa-488 anti-Rabbit (1:1000; Invitrogen)) in blocking medium. Cells where washed 3 times and Nuclei were counterstained with DAPI. Images where collected on a Zeiss Observer.Z1 using AxioVision software.

hES cell culture, transfection and differentiation

H1 (WA01) cells were cultured in feeder-free condition using mTESR1 media (Stem Cell Technologies Cat.# 05850) on 6-well plates coated with matrigel (BD Biosciences, Cat.# 354603), as described (Sigova et al., 2013).

Transfection of shMETTL3 RNA (DF/HCC DNA Resource Core Cat.#HsSH00253093) and scrambled shRNA (DF/HCC DNA Resource Core, pLKO-scramble, Cat.#EvNO00438085) was performed using Lipofectamine LTX (Life Technologies Cat.#25338100). Two days after transfection, cells were treated with 0.5 microgram per milliliter of puromycin (Life Technologies Cat.# A113802) for 6 days. For each shRNA, two independent puromycin-resistant colonies were picked from independent wells and expanded and Maintained under puromycin for analysis. Before Endodermal differentiation puromycin was withdrawn. Endodermal differentiation was then induced by resting cells in RPMI (Life Technologies Cat.#11875-093) with B27 supplement (Life Technologies Cat.#17504-044) for 24 hours followed by addition of Activin (R&D Systems), as described (Sigova et al., 2013). Day 2 and Day 4 of differentiation were measured from the time that Activin was added. Neuronal induction of human ESCs was performed according to the published protocol (Chambers et al., 2009). Adherent human embryonic stem cells were treated with potent and specific inhibitors of SMAD signaling. Two pools of chemicals were used for this purpose. In one pool, N2 media was supplemented with 5 uM solution of SB431542 and 2uM Dorsomorphin. In another pool, SB431542 and LDN for the final concentration of 0.1 uM were used. Media was replaced everyday the chemical treatment was stooped on day eight, and the cells were passage one day after. A fraction of the cells were fixed and stained at this stage to compare the neuronal marker expression -Pax6 and Sox1 between the two conditions. The remaining cells were cultured as spheres in neuronal media containing N2 and B27. These cells were seeded after 4 days of non-adherent culture condition and stained for neuronal specific markers. Results of staining were similar to the result of staining for adherent cells. We note that METTL3 knockdown had minimal effects on neural differentiation driven by LDN in contrast to Dorsomporhin.

RNA extraction, DNASE I treatment and poly A selection

mESC total RNA was isolated from cells according to manufacturer's instructions using TRIzol reagent (Ambion). The RNA was re-suspended in ultrapure H_2O, treated with DNAse I (Ambion) for 30 min at 37ºC and subjected to RNA clean up reaction with RNeasy Midi Kit (Qiagen), according to manufacture's protocol. RNA was eluted in ultrapure H_2O. PolyA RNA selection was performed using MicroPoly(A) Purist (Life Technologies) according to the manufacturer's protocol. The second polyA RNA selection was performed using the eluate of the first polyA RNA selection as starting material according to the manufacture's instruction.

hESC total RNA was isolated from cells according to manufacturer's instructions using TRIzol LS reagent (Ambion). Total RNA was treated using DNAse I (Promega) for 20 minutes at 37ºC. The treated RNA was then acid phenol/chloroform extracted and chloroform extracted. The RNA was precipitated using 300mM final concentration of $NaCl_2$ spiked with 1μl of 50mg/ml of Ultra Pure Glycogen (Promega) and 2.5 volume of 100% ethanol at -20ºC either for 2 hours or overnight. The precipitated RNA was then centrifuged using a refrigerated table-top at maximum speed (>13,000g) at 4ºC for 20 minutes. The precipitated RNA was then washed with 70ºC ethanol and centrifuged at maximum speed for an additional 10 minutes. The final pellet was then re-suspended in ultra pure H_2O. PolyA RNA selection was performed twice using Dynabeads mRNA Purification Kit (Invitrogen Cat. #610.06) according to the manufacturer's protocol. The second polyA RNA selection was performed using the eluate of the first polyaA RNA selection as starting material according to the manufacture's instruction.

For all RNA samples, the concentration, purity and integrity of the RNA were verified using a NanoDrop and Bioanalyzer.

RNA m6A IP

Mouse ESC protocol 1

PolyA+ RNA was purified with one round of selection with MicroPoly(A)Purist Kit (Ambion; AM1919). The PolyA+ RNA was fragmented to ~100 nucleotide fragments by incubation with Zinc Chloride buffer (10 mM ZnCl2, 10 mM Tris-HCl, pH 7.0). After the RNA was incubated at 94°C for 30 seconds, Zinc Chloride buffer, previously warmed to 94°C, was added and incubated for 2 minutes. The reaction was stopped with 0.2M EDTA, and the RNA precipitated with standard ethanol precipitation. 15 µg of anti-m6A polyclonal antibody (Synaptic Systems) were pretreated with agarose beads coated with ssDNA to reduced background (Xiao et al., 2011). Antibody was conjugated to Dynabeads Protein G (Life Technologies; 10003D) overnight at 4°C. 200 µg of fragmented RNA were incubated with the antibody in 1x DamIP buffer (10 mM sodium phosphate buffer, pH 7.0, 0.3 M NaCl, 0.05% (w/v) Triton X-100) supplemented with 1% SuperRNAse Inhibitor (Ambion), for 3 hours at 4°C. After incubation, the antibody was washed 5 times with DamIP buffer and the RNA eluted with 0.5 mg ml–1 N6-methyladenosine (Sigma-Aldrich) in DamIP buffer (Xiao and Moore, 2011). 1 volume of Ethanol was added to the eluted RNA, and the RNA recovered an RNeasy mini column.

Library construction:

The imunoprecipitated RNA, and an equivalent amount of input RNA where used for library generation with the dUTP protocol, as described (Levin et al., 2010) except libraries were size selected by gel purification after ligation and after PCR amplification. Libraries where sequenced using an Illumina HiSeq at the Stanford Center for Genomics and Personalized Medicine.

Mouse ESC protocol 2

Second set of libraries was generated as described in (Schwartz et al., 2013). Total RNA was subjected to two rounds of selection with MicroPoly(A)Purist Kit

(Ambion; AM1919). 5ug of RNA were fragmented as described above. After fragmentation RNA was incubated with 30 units of Polynucleotide Kinase in 50mM Tris-HCl pH 7.6, 8 mM EDTA and 2 mM DTT. RNA was purified on a quiagen RNeasy column, and 10% was saved to be used as input. RNA was denatured and incubated with 25ul of protein G beads (previously bound to 3 ug of anti-m6A polyclonal antibody (Synaptic Systems) in 1x IPP buffer (150 mM NaCl, 10 mM TRIS-HCL and 0.1% NP-40). After 3 hours, beads where washed 2 times with IPP buffer, 2 times with low salt buffer, 2 times with high salt buffer and 1 time with IPP buffer. RNA was eluted from the beads with 30 ul of RLT buffer, for 5 minutes. The RNA eluate was added to 20ul of myone Silane beads re-suspended in 30 ul of RLT. 60 ul of Ethanol where added to the beads and incubated for 2 minutes. The beads where then washed 2 times with 70% Ethanol and the RNA eluted in 160 ul of IPP buffer. The eluted RNA was added to 25 ul of Protein A beads previously bound to 3ug of anti-m6A polyclonal antibody (Synaptic Systems). After 3 hour incubation beads where washed and RNA eluted as described above. RNA was eluted in 100 ul of RNAse free water.

Library construction:

After isolating fragmented m6A enriched RNA we constructed deep sequencing libraries as Rouskin et al. with the following modifications. RNA was first ligated to 25 pmol of pre-adenylated L3 (IDT) adaptor overnight at 16°C. The ligated samples were subjected to 8% PAGE separation, stained and imaged with SybrGold (Life Technologies) and ligated material was excised. The resulting gel slices were crushed and the RNA was eluted in 400uL of Crush Soak Buffer (500mM NaCl and 1mM EDTA) and 5uL of SUPERaseIn (Life Technologies) overnight at 4°C. Eluted RNA was purified with SpinX columns (Corning), precipitated, and reverse transcribed (RT) with RT oligos modified from the iCLIP method ((Konig et al., 2010), sequences below). cDNAs size selected on a 6% PAGE and eluted in 400uL of Crush Soak Buffer at 50°C overnight. Eluted cDNA was purified with SpinX columns, precipitated, and circularized using CircLigaseII (Epicentre) for 2 hours at 60°C in a 20uL reaction. Circular cDNAs were purified

with MiniElute columns and Buffer PNI (Qiagen) and eluted in 20uL of EB Buffer. PCR amplification was performed in 50uL reactions with 25uL 2x Phusion High Fidelity Master Mix, 2.5uL of 10uM P3/P5 PCR primers (Ule, NSMB 2009/2010), and 22.5uL of circularized cDNA. Samples required between 15-25 cycles of PCR. PCR reactions were purified using AMPure XP beads (Beckman) and final library DNA was eluted in 20uL of water. Quantification was performed by BioAnalyzer analysis of the DNA, which was then sent for deep sequencing on an Illumina HiSeq2500 machine (Elim Biopharm, Hayward, CA).

Oligo and adapater sequences:

preA_L3 /5rApp/AGA TCG GAA GAG CGG TTC AG/3ddC/; P5 AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC T; P3 CAA GCA GAA GAC GGC ATA CGA GAT CGG TCT CGG CAT TCC TGC TGA ACC GCT CTT CCG ATC T; RToligo1 (Barcode) /5phos/NNN NNA ACC NNN NAG ATC GGA AGA GCG TCG TGA T/iSp18/GGATCC/iSp18/TACTGAACCGC

Human ESC protocol:

Of note for each biological replicate for m^6A-seq, we started with 400mg of total RNA yielding approximately 10μg of double polyA selected RNA which was re-suspended in a final volume of 50μl using UltraPure H_2O (Life Technologies). 250μl of digestion/fragmentation buffer (10nM $ZnCl_2$, 10mM Tris HCl, pH7.0) was added to the 50μl of 2x polyA RNA. The 300μl of PolyA RNA/fragmentation buffer was heated at 94ºC for exactly 5 minutes. 50μl of 0.5M EDTA was added to stop the fragmentation reaction and immediately put on ice.

The 2X polyA fragmented RNA was then heated at 65°C for 5 minutes and immediately put on ice. 50 μl of m^6A-DynaBeads (The m^6A antibody-Synaptic Systems was coupled to Dynabeads using the Life Technologies coupling kit cat# 14311D) were equilibrated by washing twice for 5 minutes in 500μl of m^6A-Binding Buffer (50mM Tris-HCl, 150mM NaCl2, 1% NP-40, 0.05% EDTA). The RNA was then added to the equilibrated m^6A-DynaBeads. The RNA was allowed

to bind to the m^6A-Dynabeads (in 500μl volume of m^6A-Dynabeads/m^6A-Binding Buffer at room temperature while rotating (tail-over-head) at 7 rotations per minutes for 1 hour. The tubes containing the samples were placed on a magnet allowing the beads complexes to cluster for one minute or until the solution become clear. The liquid phase was carefully collected and placed on ice as this 500μl fraction represents the "Supernatant" of the m^6A IP. Following the collection of the supernatant fraction, series of washes were performed using various buffers (see as follow). For all wash steps to the exception of the elution step, the beads were washed 3 minutes then place on a magnet and the wash buffers were discarded. Following the supernatant collection. **Wash step 1**: The reminding fractions bound to the beads were washed twice in 500μl of m^6A-Binding Buffer (Tris-HCl 50mM, $NaCl_2$ 150mM, NP-40 1%, EDTA 0.05%). **Wash Step 2:** The RNA/beads complexes were washed once in 500μl of Low Salt Buffer (SSPE 0.25 X, EDTA 0.001M, Tween-20 0.05%, NaCl 37.5 mM). **Wash Step 3:** The RNA/beads complexes were washed once in 500μl of High Salt Buffer (SSPE 0.25X, EDTA, 0.001M, Tween-20 0.05%, NaCl 137.5 mM). **Wash Step 4:** The RNA/beads complexes were washed twice in 500μl of in TET (T.E. + 0.05% Tween-20). **Elution Step:** The m^6A-RNA was eluted from the beads by repeating four times the following: 125μl of Elution Buffer (DTT 0.02M, NaCl 0.150M, Tris-HCl pH7.5 0.05M, EDTA 0.001M, SDS 0.10%) was added to the beads and incubated at 42ºC for 5 minutes. At the end of the 5 minutes the beads were gently vortexed and placed on the magnet. The liquid phase was collected and transferred to a fresh tube as this will represent the eluate fraction containing the m^6A "enriched RNA". An additional 125μl of elution buffer was then added to the beads and the processed was repeated. The liquid phase obtained at each step was added to the "fresh tube" containing the 125μl of eluate from the previous step so the total final eluate volume was 500μl.

All RNA fractions were extracted as follow. 500μl of acid phenol-chloroform (acid-phenol:chloroform, pH 4.5 (with IAA, 125:24:1) Ambion) were added to the 500μl sample. The sample was centrifuged at 4ºC at 10,000g for 7.5 minutes. The

upper phase was carefully collected making sure not to touch the inter-phase and transfer to a clean 1.5 ml tube. 500ml of chloroform was added to the fresh tube vortexed briefy and centrifuged at 4ºC at 10,000g for 7.5 minutes. The upper phase was transferred to a fresh 1.5 ml tube and $NaCl_2$ ethanol precipitated overnight at -20ºC in presence 1μl of (20mg/ml) Ultra Pure Glycogen. The following day the sample was centrifuged at 4ºC for 20 minutes at 16,000g. The pellet was then washed in 70% ethanol centrifuged and additional 10 minutes at 4ºC at 16,000g. The pellet was then let to dry at room temperature for 10 minutes prior to be re-suspended in the desired volume of Ultra-Pure H_20 (Invitrogen Cat# 10977-015).

Library construction:

100ng (100ng of input and 100ng of post m^6A-IP positive fraction) were used for library construction and RNAseq using TrueSeq Stranded mRNA Sample Preparation Guide, entering the protocol by adding the Fragment, Prime, Finish Mix, skipping the elution step and proceeding immediately to the synthesis of the First Strand cDNA. From that point on, the exact steps of the Illumina TruSeq Stranded mRNA sample Preparation Guide were followed to the end. RNA Sequencing. Each individual library fragment size was verified on Agilent Bioanalyzer 2100 with High Sensitivity chip. Final quantification was done by qPCR on Perkin Elmer 2500Fast with Kapa library quantification kit (#KK4824). Libraries were pooled at equimolar concentrations according to the manufacturer guidelines (TruSeq Stranded mRNA Sample Preparation Guide- September 2012). After clustering on Illumina cBot, samples were run on Illumina HiSeq 2000.

For m6AIP-RT-qPCR, and m6AIP-Nanostring, experiment were performed as described above (protocol 1), except 2ug of fragmented RNA, and 1ug of antibody were used. Rabbit IgG was used as a non-specific antibody control for immunoprecipitation in parallel to the anti-m6A polyclonal antibody (Synaptic Systems).

Real Time PCR

For the mouse experiments, RNA was analyzed on a LightCycler 480 by RT-qPCR with One-Step RT-PCR Master Mix SYBR Green (Stratagene). For gene expression experiments, each PCR reaction was performed in 12 μl with 45ng of total RNA, 0.8 μl of RT block/enzyme mixture, 1.2μl primers at 1.25μM each and 6μl of MasterMix (final volume 12 μl). The PCR was carried on using a standard protocol with melting curve. The amount of target were calculated using the formula: Amount of target = $2^{-\Delta\Delta C(T)}$ (Livak and Schmittgen, 2001). Two tailed T test for unequal, unpaired data sets with heteroscedastic variation was used to compare samples. Primer sequences available upon request.

For human experiments, a first mixed made of 10pg to 5μg of RNA in 5μl volume, 2μl of random hexamers (Roche), 1μl of dNTPmix (10mM each) and 5μl of ultrapure H_2O was first generated, heated at 65ºC for 5 minutes and immediately put on ice. 4μl of 5X First Strand Buffer was added along with 1μl of 0.1M DTT, 1μl RNAse inhibitor and 1μl of Superscript III reverse transcriptase (Invitrogen). The 20μl reverse transcription reaction was then incubated 5 minutes at room temperature, then 60 minutes at 50ºC then 15 minutes at 70ºC. The freshly synthesized cDNA was treated with 1μl of RNAse H at 37ºC for 20 minutes. For Sybergreen quantitative real time PCR assays, each PCR reaction was done in a 20μl volume made of 10μl of master mix (SYBR GreenER qPCR SuperMix for iCycler -Invitrogen), 5μl of primer mix at 1.2μM (each) and 5μl of cDNA template at 20ng/μl. The PCR was carried on using a standard protocol with melting curve. The amount of target were calculated using the formula: Amount of target = $2^{-\Delta\Delta C(T)}$ (Livak and Schmittgen, 2001). The qPCR using Taqman reagents was done in a 10μl volume made of 5μl of Universal PCR Master Mix (Applied Bosystems Cat.#4304437), 0.5μl of TaqMan probe mix (each), 2μl of cDNA template at 50ng/μl and 2.5μl of H_2O. The PCR was carried on using a standard protocol with melting curve. The amount of target were calculated as above. The TaqMan probes were purchased from Applied Biosystems; 18s (AB Hs99999901_s1), FOXA2 (AB Hs00232764_m1), SOX17

(AB Hs 00751752_s1), NANOG (AB Hs 02387400_g1), and SOX2 (AB 010533049_s1). Two tailed T test for unequal, unpaired data sets with heteroscedastic variation was used to compare samples.

RNA stability assay

Wild type and Mettl3 KO cells were treated with 0.8 μM Flavopiridol for 3 hours. RNA extraction and qRT_PCR as described above.

shRNAs targeting shRNAs

Short Hairpin RNAs targeting the mouse Mettl3 sequences GCACACTGATGAATCTTTA and GCACTTCCTTACAAAGCT were generated in the pSicoR plasmid backbone (Addgene 12084, (Ventura et al., 2004)). The plasmid pSicoR shluc (Addgene 14782, (Konig et al., 2010)) was used as a negative control. The plasmids were co-transfected into 293T cells with pMd2G and psPAX2 with Fugene HD (Promega, E2311) according to manufacturer's instructions. Virus where collected after 48 hours. The collected media was filtered through a 0.45 μm membrane and the virus concentrated with Lenti-X concentrator (Clontech; 631231). J-1 mESC cells were infected in the presence of 2 μg per ml polybrene. After 24 hours, cells where selected with puromycin. After selection, cells where replated at low density and single clones where collected. Real time PCR was used to choose determine efficiency of the Knock Down.

The shRNA hairpins targeting human Mettl3 were purchased from DF/HCC DNA Resource Core. Multiple sh clones were purchased against METTL3 (HsSH00253093, HsSH00253439, HsSH00253446, HsSH00253487, HsSH00253494). After testing of their individual knockdown efficiency both by qRT-PCR and anti-METTL3 western blot in 293T, we identified number HsSH00253093 (insertSequence: CCG GGC TGC ACT TCA GAC GAA TTA TCT CGA GAT AAT TCG TCT GAA GTG CAG CTT TTT; Target Sequence:

GCTGCACTTCAGACGAATTAT) as giving optimal knockdown and this was used to generate H1-ESCs knockdown cell lines. The scrambled shRNA control pLKO-Scramble (Cat# EvO00438085) was also obtained from the DF/HCC DNA Resource Core.

CRISPR mutants

gRNA sequences where chosen and designed from CRISPR design tool (Hsu et al., 2013). DNA blocks containing all of the components necessary for gRNA expression (Mali et al., 2013) were synthesized by IDT and cloned in Topo-Blunt plasmid (Invitrogen). Plasmids for guide RNA were co-nucleofected (Lonza; VPH-1001), according to manufacturer's instructions, with a human codon optimized Cas9 expression plasmid and a plasmid with a puromicine resistance cassette. Cells were plated at low density for single colony isolation. The remaining cells were cultured for surveyor assay. After 24 hours, cells were selected with puromicine for 48 hours. DNA extraction and surveyor assay as described in (Cong et al., 2013). Single colonies where selected and tested by western blot for loss of Protein. DNA sequencing of the targeted locus was used to confirm presence of mutations that abrogate protein production.

Annexin V analysis

One million cells were collected and washed twice with PBS. The cells were incubated with 1 µl of Live/Dead Fixable Aqua (Life Technologies) for 30 minutes, protected from light. The cells were then washed twice with FACS buffer and re-suspended in 1X Binding buffer followed by an incubation with 5 µl of fluorochrome conjugated Annexin V for 15 min. The cells were washed once with FACS buffer and resuspended in 500 µl of Binding buffer. Samples were analyzed on a special order FACS Aria II (BD Biosciences).

Western Blot

Cells were collected and lysed in RIPA buffer (400 mM NaCl, 1% Igepal, 0,5 % Sodium Deoxycholate, 0.1% SDS and 10mM Tris-Cl pH 8.0) for 30 min on ice. The lysate was centrifuged for 10 minute and the supernatant collected. Protein was quantified with BCA Protein Assay Kit (Pierce). Proteins where resolved on a NuPAGE 4-12 % Bis-Tris Midi Gel and transferred to Immobilon-FL membrane. Primary antibodies used are: (Rabbit anti-METTL3/MT-A70, Bethyl A301-568; Mouse anti-beta actin, mAbcam 8224 and Rabbit anti-PARP, Cell Signaling, 9542). Secondary antibodies used: IRDye 680RD Goat anti-Mouse IgG (H + L) (Licor) and IRDye 800CW Goat anti-Rabbit IgG (H + L) (Licor). Images where collected on a Licor Odyssey imaging system.

2D-TLC measurement of m6A

2D-TLC was performed as described by (Jia et al., 2011). 100 to 200 ng of polyA+ RNA, selected for two rounds, was digested with 2000 units of RNAse T1 (Ambion) in a final volume of 25 µl, with 1x PNK buffer and incubated at 37°C for 1 hour. The RNA was labeled with 10 units of PNK (NEB) and 1 µl [ϒ-32P]ATP (6000 Ci/mmol; Perkin-Elmer). The reaction was cleaned with a G25 column and precipitated with Standard Ethanol precipitation. The RNA was re-suspended in 10 µl of 50mM sodium acetate (pH 5.5) and digested with 1 Unit of nuclease P1 (USBiological; N7000). 1 µl was loaded on a Cellulose TLC glass plate (EMD chemicals; 5716-7). The first dimension was resolved in isobutyric acid:0,5 M NH4OH (5:3, v/v) and the second dimension resolved in isopropanol:HCl:water. The plates were exposed on a phosphor screen and scanned on a GE typhoon TRIO at the Stanford Functional Genomics Facility.

m^6A level dot-blots

Amersham Hybond-XL (Cat.# RPN303s) membrane was rehydrated in H_2O for 3 minutes. The membrane was then “sandwiched” in Bio-Dot Microfiltration Apparatus (BioRad, cat. #170-6545). Each well was then filled with

H_2O and flushed by gentle suction vacuum until it appeared dry. 5μl of H_2O alone was then applied to the membrane in each well followed by addition of indicated amount of RNA and this was allowed to bind to the membrane by gravity. The apparatus was disassembled and the membrane was cross-linked in a UV STRATALINKER 1800 using the automatic function and then the membrane was placed back into the apparatus. The membrane was then blocked 10 minutes using sterile RNAse DNase free TBST + 5% milk. The m^6A primary antibody (Anti-m^6A, Synaptic Systems, Cat. #202 003) was then added at a concentration of 1:500 at room temperature for 1 hour in TBST + 5% milk. The membrane was then washed four times in PBST. The membrane was then incubated with the secondary anti rabbit antibody (1:5000 dilution) for 30 minutes in TBST + 5% milk. The membrane was washed 4 times 5 minutes in TBST and expose on an auto radiographic film using Pierce ECL Western Blotting Substrate.

Mass spectrometric quantification of m^6A

Enzymatic hydrolysis of RNA to ribonucleosides was carried out as described previously,(Taghizadeh et al., 2008) with modifications. Following addition of 100 nM [^{15}N]- ethenocytidine and 10 μM [^{15}N]-guanosine as internal standards for m^6A and adenosine respectively (due to similar masses and retention times), RNA (200 ng) was digested with 2 U nuclease P1 (Sigma Aldrich, St. Louis, MO) at 37 ºC for 3 h in 55 µl in buffer containing 16 mM sodium acetate (pH 6.8), 1.8 mM zinc chloride, 9 µg/mL coformycin, 45 µg/mL tetrahydrouridine, 2.3 mM desferroxamine, 0.45 mM butylated hydroxytoluene, followed by addition of 45 µl of 27 mM of sodium acetate (pH 7.8), 17 U calf thymus alkaline phosphatase (New England Biolabs, Ipswich, MA) and 0.1 U snake venom phosphodiesterase (Sigma Aldrich) with incubation overnight at 37 ºC. The digestion mixture was later deproteinized by centrifugal filtration (Nanosep 10K; Pall Corporation, Port Washington, NY), and 10 µl of the mixture was analyzed by a liquid chromatography-coupled triple quadrupole mass spectrometry (LC-QQQ). HPLC was performed on an Agilent series 1200

instrument (Agilent Technologies, Santa Clara, CA) consisting of a binary pump, a solvent degasser, a thermostatted column compartment and an autosampler. The nucleosides were resolved on a Dionex Acclaim PolarAdvantage C16 column (3 μm particles, 120 Å pores, 2.1 x 150 mm; 30 °C) at 300 μL/min using a solvent system consisting of 0.1% acetic acid in H_2O (A) and 0.1% acetic acid in acetonitrile (B), with the elution performed isocratically at 0% B for 29 min, followed by a column washing at 70% B and column equilibration. Mass spectrometry detection was achieved using an Agilent 6410 QQQ mass spectrometer in positive electrospray ionization mode with the following parameters: ESI capillary voltage, 3000 V; gas temperature, 340 °C; drying gas flow, 10 L/min; nebulizer pressure, 20 psi; fragmentor voltage, 150 V. The nucleosides were quantified using the nucleoside → base ion mass transitions of 282.1 → 150.1 (m^6A), and 268.1 → 136.1 (A). Absolute quantities of m^6A and A were determined from calibration curves prepared daily.

Microarray Data Acquisition and Data Analysis.

RNA was extracted as described above and submitted for Hybridization on GeneChip Mouse Exon 1.0 ST Array at the Protein and Nucleic Acid Facility of the Stanford School of Medicine. For gene expression analysis, arrays were RMA normalized using justRMA package in R. After normalization, probes with average expression of all arrays less than 100 were filtered out as not expressed probes. For each expressed probe, its expressions were log2ed, and the gene expression was defined as the average expression of all the expressed probes that attached to this gene. Student T-test comparing wide-type versus knockout signals in the arrays were used to calculate the significance of the expression changes, and false discovery rate (FDR) was estimated using p.adjust package in R. Differential expression was defined using the following filters: significance analysis of microarrays 3.0 (Tusher et al., 2001) with a false discovery rate less than 5%, an average fold change ≥2 in any group, and an average raw expression intensity ≥100in any group.

m6A methylation IP analysis

Libraries generated with iCLIP adaptors (mouse, protocol 2) where separated by barcode, and perfectly matching reads were collapsed and barcodes removed. For all libraries, single-end RNA-Seq reads were mapped to the mouse (mm9 assembly) of human genome (hg19 assembly) using TopHat (version 1.1.3) (Trapnell et al., 2009). Only uniquely mapped reads were subjected to downstream analyses.

The mouse RNA-seq reads, recorded in BAM/SAM format were transformed to bedGraph format, indicating the number of reads on each genomic position. A non-redundant mm9 transcriptome was assembled from UCSC RefSeq genes, UCSC genes, and predictions from (Ulitsky et al., 2011) and (Guttman et al., 2011). Gene expression in the form of RPKM was calculated using a self-developed script.

For human RNA-seq reads, FPKMs of Ensembl genes (release 64) were calculated using Cufflinks (vesion 2.0.2) (Trapnell et al., 2010) and differentially expressed genes between input RNAs of T0 and T48 were determined by Cuffdiff (version v2.0.2) (Trapnell et al., 2013).

To make UCSC read coverage tracks, the read coverage at each single nucleotide was normalized to library size for input and eluate (m^6A RIP) respectively. For human samples, we normalized the read densities by adjusting the library sizes (total uniquely mapped reads) to be the same (average total uniquely mapped reads of initial sequencing runs of 4 samples) for input and eluate (m^6A RIP) respectively. The average normalized read densities of replicates A and B were shown in the Figures.

Peak calling and analysis

Search for enriched peaks was performed by scanning each gene using 100-nucleotide sliding windows, and calculate an enrichment score for each sliding window (Dominissini et al., 2012). Windows with RPKM ≥ 5 in the eluate,

enrichment score ≥ 2 in genes with RPKM in the input sample ≥1 were defined as enriched in m6A pull down. Enriched windows with score greater than neighboring windows where selected as m6A peaks. To determine “high-confidence”, we first intersected the peaks in biological replicates, requiring at least 0.5 overlap using the BedTools package (Quinlan and Hall, 2010). Peaks that did not intersect where merged, and peaks that merged end to end where also kept for downstream analysis. The peaks where re-defined as 100 nt windows centered at the middle of the intersected/merged peaks. For Human m6A peak detection, eluate window RPKM ≥ 10 instead of 5 were used. Common peaks were determined in the same way as described in mouse. For each time point, the common peaks of the two replicates were referred to as “high-confidence” peaks.

To study the peak distributions on transcripts, we assigned each “high-confidence” peak (using middle point) to the collapsed transcript (mouse) or to the longest isoform of each Ensembl gene. 100 bins of equal length were made for 5’UTR, CDS and 3’UTR respectively and the average number of peaks for each bin was calculated. The peak intensity was calculated as the ratio of window RPKM between eluate and input for each peak. To compare the peak intensities between two samples, we used sample specific peaks as well as common peaks and required input window RPKM ≥ 20 to obtain reliable peak intensity values.

A gene’s enrichment score was defined as the maximum enriched window in this gene. HOMER software package (Heinz et al., 2010) was used for de novo discovery of the methylation motif, using the high confidence peaks. Random windows for control where obtained using the BedTools package (Quinlan and Hall, 2010).

GO (Gene Ontology) analyses for methylated genes were conducted using DAVID (Huang da et al., 2009) with genes with RPKM ≥ 1 (mouse) or FPKM ≥ 1 (human) as background.

Dataset comparasion

Pol II occupancy, obtained from (Rahl et al., 2010), at transcriptional start sites was determined using an in-house developed script based on annotations downloaded from the UCSC table browser. Mouse mRNA half life and Protein translation efficiency was extracted from (Ingolia et al., 2011; Sharova et al., 2009) for genes with RPKM >= 1 in the input. Plotting and statistical test performed in R. For genes with multiple Half life values reported, the average value was used. We obtained human mRNA half-life of induced pluripotent stem (IPS) cells from published thesis (Neff et al., 2012). The m6A enrichment score was calculated as the maximum window scores of all windows of each gene including unmethylated genes, the windows with input window RPKM <1 were removed from the calculation.

Gene Set Enrichment Analysis

Genes were ranked by their enrichment score, and equally divided into 10 groups. For each group, a multi-dimensional gene set enrichment analysis over DAVID Gene Ontology terms and stem cell gene sets (Wong et al., 2008) was performed using Genomica (Segal et al., 2005; Segal et al., 2004; Segal et al., 2003). A P-value of < 0.01 from hyper geometric test between a gene group and gene set was defined as significant.

Determination of differentially methylated peaks

To determine effects of Mettl3 loss of function on m6A peaks, we calculated the peak intensity for the high confidence peaks identified in wild type cells. Peaks with significant changes in peak intensity (p.value < 0.05) where considered for further analysis. To determine the effect of differentiation in hESC, the union of m^6A peaks of T0 and T48 (initial sequencing run, with comparable sequencing depth for both time points) were analyzed to determine the differentially methylated peaks between T0 and T48 that meet the following criteria: 1) Input gene FPKM ≥ 1 in all 4 samples; 2) Input window RPKM ≥ 10 in

all 4 samples; 3) At least 1.5 fold (or 2 fold) change of peak intensities in both replicates in the same direction; 4) The maximum peak intensity of all samples ≥ 2; 5) In each replicate, the sample with higher peak intensity must be called as having peak. To determine the union of m6A peaks of T0 and T48, we pooled all the peaks of 4 samples and merged the same peaks and peaks with 50bp overlapped, the unmerged peaks were then merged if they were end-to-end peaks spanning 200 bp. We took the center 100 bp of merged peaks as union peaks if they meet the following criteria in either T0 or T48: 1) both replicates had the peaks; 2) The center 100bp had window score ≥ 2 in both replicates.

Heatmap and clustering analysis

Heatmaps of all 4 samples were made based on Z score scaled log2 values for peak intensities or gene expression levels (FPKMs) respectively. For analysis of the differentially expressed genes, the genes and samples were clustered by average linkage hierarchical clustering using 1-Pearson correlation coefficient of log2(FPKM) as the distance metric. For peak intensity analysis, the peaks and samples were clustered in the same way using 1-Pearson correlation coefficient of log2(peak intensity) as the distance metric.

Analysis of m6A sites in non-coding RNAs

The longest isoforms of Ensembl genes were used to study the distribution of m6A peaks on coding and noncoding transcripts. Noncoding transcripts overlapping with any isoforms of coding genes were removed, and transcripts with less than 3 exons were also removed. The analysis used the peaks found wild type mESC cells or the union of H1 T0 (all data), H1 T48, 293T, HepG2 (including stimulated samples) and human brain (Dominissini et al., 2012; Meyer et al., 2012). To study the m6A peak distributions on transcripts, in each transcript we made 10 bins of equal length for the first exon, internal exons and the last exon respectively, and the percentage of peaks in each bin was calculated for coding and noncoding transcripts. Additionally, the peak coverage

around the last exon-exon splice junction was also analyzed for coding and noncoding transcripts. The peaks used in this analysis included the wild type mESC or H1 T0 (all data), H1 T48, 293T, HepG2 (including stimulated samples) and human brain (Dominissini et al., 2012; Meyer et al., 2012). The peak coverage (number of peaks covering the site) normalized by the total number of overlapped peaks was calculated for the 750 bp regions flanking the last splice junction. Therefore, the transcripts with less than 750 bp on either side were also removed from the analysis.

Exon length analysis

Middle points of all high-confidence peaks in the two time points were assigned to exons of the longest isoforms of Ensembl coding genes. Only internal exons were used in the subsequent analysis. Exon length and number of m6A motifs were used to normalize the number of peaks in each exon. Error bar indicates variations estimated via 1000 times of bootstrapping for each bin of exon length.

Single exon gene analysis

Ensembl genes without any multi-exon isoforms were considered as single exon genes. The peak distribution of the longest isoform of single exon protein-coding genes was analyzed in the same way as for multi-exon protein-coding genes, except that 10 bins were made for each 5'UTR, CDS and 3'UTR.

Comparison of m6A peaks between mouse and human ESCs

We used common peaks of 3 mESC and common peaks of 2 hESC for mouse and human ESC m6A comparison. To compare the methylated genes between mESC and hESC at gene level, only Ensembl genes with the annotated one to one ortholog between human and mouse were considered in the comparison, and the genes must have gene expression value (RPKM or FPKM) greater than 1 in all samples of both hESC and mESC. To compare the m6A

peak intensities between human and mouse ESCs, we aligned all the mESC peaks to human genome based on the UCSC pairwise genome alignment (http://hgdownload.soe.ucsc.edu/), the orthologous mouse-human regions of merged peaks (at least 1 bp overlap) and species specific peaks were used for the comparison. For merged peaks, we took the center 100 bp regions and only used those had window scores ≥ 2 in all samples of both species. Only Ensembl genes with the annotated one to one orthologs between human and mouse were considered. To obtain reliable peak intensity values, we required gene RPKM or FPKM ≥ 1 and input window RPKM ≥ 5 in all samples of both species.

GRO-seq analyses and RNA polymerase II traveling ratio calculation

GRO-seq data for hESCs (replicate 1-3) and GRO-seq data for 48 hours of endodermal differentiation (replicate 1) (Sigova et al., 2013) (GSE 41009) were analyzed. FASTQ files were mapped to hg19 using Bowtie2 with the parameters -k2 -L24 -N1 --local. Calculation of the traveling ratio was adapted from (Rahl et al., 2010). Briefly, each gene was divided into the proximal promoter and gene body. The proximal promoter was defined as the region from 30 bp upstream to 300 bp downstream of the transcription start site. The gene body was defined as 300 bp downstream of the TSS to the end of the annotated gene. The number of GRO-seq reads that mapped to the promoter proximal region and gene body was determined for each gene in each experimental condition. The total number of reads mapped to each region was divided by the length of the region to determine the read density. The RNA polymerase II traveling ratio (TR) was calculated for each gene by dividing the density of the promoter proximal region by the density of the gene body region.

Analysis of the relationship between m^6A and RNA polymerase II travelling ratio

To compare the m^6A peak intensity and RNA polymerase II travelling ratio, the m6A enrichment score was calculated as the maximum window scores of all

windows of each gene including unmethylated genes, the windows with input window RPKM <1 were removed from the calculation.

Teratoma generation and histopathology

Mettl3 wild type and mutant cells were trypsinized and 2.5 X 10^6 cells were subcutaneously injected into 8-week-old female SCID/Beige mice (Charles River). Teratoma progression was monitored by volume measurement every other day after a visible tumor mass formed. In the fourth week after injection, the mice were euthanized and the tumors were harvested, weighed, measured and then were processed for histological analysis. All the animal studies were approved by Stanford University IACUC guidelines.

For histological analysis, teratomas were fixed with 4% paraformaldehyde, processed for routine histopathology, embedded in paraffin and 4 micron sections were stained with hematoxylin and eosin (H&E); or stained by immunohistochemistry (IHC) with VECTASTAIN ABC Kit (PK-4000, Vector laboratories) and DAB Peroxidase Substrate Kit (SK-4100 , Vector laboratories) following the manufacturer's instructions. Antibodies used for IHC were: anti-Nanog (1:500; A300-397A, Bethyl) and anti-Ki67 (1:100; RM-9106, Thermo). Tumors were evaluated and images where captured using a Zeiss Axioskop 2 microscope with a DS-Ri1 camera and NIS-Elements D image software.

Antibodies used in this study:

Rabbit polyclonal anti-m^6A (Synaptic Systems, 202 003); Rabbit polyclonal anti-METTL3 (Proteintech, 15073-1-AP); Rabbit polyclonal anti-METTL3 (Bethyl, A301-568); Rabbit pre-immune serum (Sigma, R9133); Mouse monoclonal anti-beta actin (mAbcam, 8224); Rabbit polyclonal anti-PARP (Cell Signaling, 9542); Rabbit polyclonal anti-Nanog (Bethyl, A300-397A); Rabbit polyclonal anti-Nanog (ReproCell); Mouse monoclonal anti-Oct-3/4 (Santa cruz, sc-5279); Mouse monoclonal anti-Tuj1 (MMS-435P); mMF20 (Developmental studies Hybridoma bank); Rabbit monoclonal anti -Ki67 (Thermo, RM-9106); Donkey anti-Rabbit

antibody (Amersham, NA934); Goat anti-Mouse IgG (H + L) IRDye 680RD (Licor) ; Goat anti-Rabbit IgG (H + L) IRDye 800CW (Licor); Goat anti-mouse Alexa-488; Goat anti-Rabbit Alexa-555; Donkey anti-mouse Alexa-555; Donkey anti-rabbit Alexa-488.

m^{6}A antibody titration

We generated an m^{6}A antibody titration curve to identify the point of saturation of the anti-m^{6}A antibody in the context of performing m^{6}A RIPs (Figure S1). To do so, we utilized an *in vitro* generated transcript from a plasmid containing full length GAPDH transcript. The plasmid was first linearized by restriction digest using SalI just downstream of the GAPDH cDNA cloning site. The linearized plasmid was gel purified and *in vitro* T7 mediated transcription was performed using the Ambion MEGAscript Kit (AM1334) as described in the user manual. The incorporation of m^{6}A to the m^{6}A transcripts was done by adding TriLink N^6-Methyladenosine-5'-Triphosphate (cat# N1013) at the indicated concentration to unmodified ATP of the kit (ex a 2% m^{6}A transcript was made by mixing 98% ATP with 2% m^{6}A nucleotide) according to the manufacturer instructions. The anti-m^{6}A RIP was performed as described in the m^{6}A-seq section, with the exception that intact full length GAPDH transcript was utilized as input for the RIP step.

References

Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sedalain, M., Studer, L. (2009) Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat Biotechnol 27, 275-280.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A.*, et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819-823.

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.

Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L.*, et al.* (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature *477*, 295-300.

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.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O.*, et al.* (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol *31*, 827-832.

Huang da, W., Sherman, B.T., Zheng, X., Yang, J., Imamichi, T., Stephens, R., and Lempicki, R.A. (2009). Extracting biological meaning from large gene lists with DAVID. Curr Protoc Bioinformatics *Chapter 13*, Unit 13 11.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789-802.

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.G.*, et al.* (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol *7*, 885-887.

Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol *17*, 909-915.

Levin, J.Z., Yassour, M., Adiconis, X., Nusbaum, C., Thompson, D.A., Friedman, N., Gnirke, A., and Regev, A. (2010). Comprehensive comparative analysis of strand-specific RNA sequencing methods. Nat Methods *7*, 709-715.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823-826.

Neff, A.T., Lee, J.Y., Wilusz, J., Tian, B., and Wilusz, C.J. (2012). Global analysis reveals multiple pathways for unique regulation of mRNA decay in induced pluripotent stem cells. Genome research *22*, 1457-1467.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432-445.

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

Segal, E., Friedman, N., Kaminski, N., Regev, A., and Koller, D. (2005). From signatures to models: understanding cancer using microarrays. Nat Genet *37 Suppl*, S38-45.

Segal, E., Friedman, N., Koller, D., and Regev, A. (2004). A module map showing conditional activity of expression modules in cancer. Nat Genet *36*, 1090-1098.

Segal, E., Shapira, M., Regev, A., Pe'er, D., Botstein, D., Koller, D., and Friedman, N. (2003). Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat Genet *34*, 166-176.

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

Sigova, A.A., Mullen, A.C., Molinie, B., Gupta, S., Orlando, D.A., Guenther, M.G., Almada, A.E., Lin, C., Sharp, P.A., Giallourakis, C.C*., et al.* (2013). Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. Proc Natl Acad Sci U S A *110*, 2876-2881.

Taghizadeh, K., McFaline, J.L., Pang, B., Sullivan, M., Dong, M., Plummer, E., and Dedon, P.C. (2008). Quantification of DNA damage products resulting from deamination, oxidation and reaction with products of lipid peroxidation by liquid chromatography isotope dilution tandem mass spectrometry. Nat Protoc *3*, 1287-1298.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol *31*, 46-53.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol *28*, 511-515.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A *98*, 5116-5121.

Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell *147*, 1537-1550.

Ventura, A., Meissner, A., Dillon, C.P., McManus, M., Sharp, P.A., Van Parijs, L., Jaenisch, R., and Jacks, T. (2004). Cre-lox-regulated conditional RNA interference from transgenes. Proc Natl Acad Sci U S A *101*, 10380-10385.

Wong, D.J., Liu, H., Ridky, T.W., Cassarino, D., Segal, E., and Chang, H.Y. (2008). Module map of stem cell genes guides creation of epithelial cancer stem cells. Cell Stem Cell *2*, 333-344.

Xiao, R., and Moore, D.D. (2011). DamIP: using mutant DNA adenine methyltransferase to study DNA-protein interactions in vivo. Curr Protoc Mol Biol *Chapter 21*, Unit21 21.