MANUSCRIPT TITLE

KDM3A and KDM3B Maintain Naïve Pluripotency Through the Regulation of Alternative Splicing

AUTHORS

Caleb Dillingham[1,2,3], Harshini Cormaty[1,2,4], Ellen Morgan[1,2], Dakarai Esgdaille[6], Paul Boutz[5], Rupa Sridharan[1,2]

[1]Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, WI, 53715, USA
[2]Department of Cell and Regenerative Biology, University of Wisconsin-Madison, Madison, WI, 53792, USA
[3]Pathology and Laboratory Medicine, University of Wisconsin-Madison, Madison, WI, 53705, USA
[4]Laboratory of Genetics, University of Wisconsin-Madison, Madison, WI, 53706, USA
[5]Department of Biochemistry and Biophysics, Center for RNA Biology, Wilmot Cancer Center, University of Rochester School of Medicine and Dentistry
[6]Department of Biomedical Genetics, University of Rochester School of Medicine and Dentistry

ABSTRACT

Histone modifying enzymes play a central role in maintaining cell identity by establishing a conducive chromatin environment for lineage specific transcription factor activity. Pluripotent embryonic stem cells (ESCs) identity is characterized by lower abundance of gene repression associated histone modifications that enables rapid response to differentiation cues. The KDM3 histone demethylase family removes the repressive histone H3 lysine 9 dimethylation (H3K9me2). Here we uncover a surprising role for the KDM3 proteins in the maintenance of the pluripotent state through post-transcriptional regulation. We find through immunoaffinity purification of the KDM3A or KDM3B interactome and proximity ligation assays that KDM3A and KDM3B interact with RNA processing factors such as EFTUD2 and PRMT5. By generating double “degron” ESCs to degrade KDM3A and KDM3B in the rapid timescale of splicing, we find altered splicing, independent of H3K9me2 status. These splicing changes partially resemble the splicing pattern of the more blastocyst-like ground state of pluripotency and occurred in important chromatin and transcription factors such as *Dnmt3b, Tbx3* and *Tcf12*. Our findings reveal non-canonical roles of histone modifying enzymes in splicing to regulate cell identity.

INTRODUCTION

Gene expression patterns are established by the coordination of transcription factor binding to chromatin (1–3). The chromatin environment is controlled in part by chemical modifications on histones and DNA (4, 5). In general, histone acetylation decompacts chromatin to facilitate transcription factor binding and promotes gene expression (6, 7). The effect of histone methylation is more nuanced and depends on the specific amino acid residue that is modified (8, 9). The abundance of histone modifications at a particular location is controlled by the balance of the activities of enzymes that add or remove each modification that is modulated by the protein complexes they associate with (10, 11). As part of such protein complexes, histone modifying enzymes may serve scaffolding and targeting functions independent of their catalytic activity. Hence histone modifying enzymes with the same catalytic targets can have both redundant and distinct functions.

Histone 3 Lysine 9 dimethylation (H3K9me2) is correlated with gene repression and enriched in tightly compacted heterochromatin (12, 13). H3K9me2 is found on the majority of histone tails in somatic cells and enriched in broad megabase domains that encompass both repetitive elements and protein coding genes (14–17). The enzymes KDM3A and KDM3B demethylate H3K9me1/me2 and have conserved catalytic domains (18–21). In mice, the combined knockout of both KDM3A and KDM3B leads to embryonic lethality at day 4.5 when the epiblast is being formed, insinuating the redundant function of the two enzymes(22). In

contrast, while the KDM3A mouse knockout is obese but viable (23–25), the KDM3B mouse knockout is perinatal lethal, with one-third of the offspring dying within 48 hours of birth (26, 27).

KDM3A and KDM3B also have distinct functions in gaining pluripotency from somatic cells but overlapping functions in the maintenance of pluripotency. Somatic cells can be reprogrammed into induced pluripotent stem cells by overexpression of the Yamanaka transcription factors (28). In this transition, the knockdown of *Kdm3b* severely compromises reprogramming efficiency although KDM3A levels remain unaltered (29, 30). In contrast, the self-renewal of pluripotent embryonic stem cells ESCs is compromised only upon the deletion of both KDM3A and KDM3B (22). Interestingly the KDM3A/KDM3B double knockout resulted in few gene expression changes which is unexpected from a broadly repressive modification (22).

To investigate how KDM3A and KDM3B functions are modulated by their protein interactions, we performed immunoprecipitation of KDM3A and KDM3B in ESCs. We found both enzymes interact with RNA splicing machinery. By performing acute degradation of both proteins using a degron system for targeted ubiquitination and proteolysis (31, 32), we found deficits in splicing that occurred independent of H3K9me2 status. The splicing aberrations occurred in several genes that are important regulators of cell state mimicking the splicing patterns of a more blastocyst like state of pluripotency. Taken together we uncover a role for KDM3A and KDM3B in splicing regulation that controls cell identity during development.

MATERIALS AND METHODS

Cell culture

Tet-inducible V6.5 ESC lines (33) and derivatives were routinely maintained on 0.1% gelatin coated plates in ESC media (KnockOut DMEM, 15% fetal bovine serum (FBS), Glutamax, Penicillin/Streptomycin, non-essential amino acids (NEAA), 2-mercaptoethanol (4 µL/500 mL), and leukemia inhibitory factor (LIF) with irradiated DR4 feeder mouse embryonic fibroblasts (MEFs). E14 ESCs and derivatives grown in ESC media in feeder-free conditions on 0.1% gelatin coated plates; all ESCs confirmed to be mycoplasma free. For growth in 2i/LIF conditions, ESCs were maintained on 0.1% gelatin coated plates in modified 2i/LIF media: 50% DMEM/F12 (Gibco), N-2 Supplement (17502048), Insulin (12.5 mg/L0), progesterone (0.01 mg/L), 50% Neurobasal medium (Gibco), B-27 minus Vitamin A (Thermo Fisher 12587010), 1000 U/mL LIF, b-mercaptoethanol (6.3 ul/L), Penicillin-Streptomycin (Gibco), 3 µM CHIR-99021 (Reprocell 04-0004) , 0.4 PD-0325901 (Reprocell 04-0006), Bovine Serum Albumin (0.0025%), and 5% Fetal Bovine Serum. Passaged every 2-3 days with TrypleE (Gibco).

Cell line derivation

3xFLAG-KDM3A and 3xFLAG-KDM3B Tet-inducible V6.5 ESC lines were generated by FLP-mediated recombination of N-terminal 3X-FLAG-tagged KDM3A or KDM3B at the Col1A1 locus (33). Cells were selected with hygromycin at a final concentration of 100 µg/ml and plated at low density to isolate single clones and screened via Western Blot for endogenous KDM3A or KDM3B and FLAG (M2 FLAG Sigma). FLAG-tagged protein expression was induced with 1ug/ml doxycycline.

Degron cell lines were generated following (31, 32). Briefly, *Kdm3a* microhomology arms were added to the pCRIS-PITChv2-Puro-dTAG (Addgene #91793) donor vector and *Kdm3b* microhomology arms were cloned into the pCRIS-PITChv2-BSD-dTAG (Addgene #91792) donor vector to allow for dual selection with puromycin and blasticidin S hydrochloride. Cas9/sgRNA targeting vectors were constructed according to Sakuma et al. 2016. Briefly, pX330A-1x2 (Addgene #58766) and pX330S-2-PITCh (Addgene #63670) were combined with guide RNAs (gRNAs) targeting *Kdm3a* (5'-GAAACCATGGTGCTCACGCT-3') or *Kdm3b* (5'-

CAACAGCCGCTTGCCCACCG-3') to generate a vector PITCh sgRNA, target-specific sgRNA, and Cas9. Two million E14 ESCs were nucleofected with paired *Kdm3b* targeting and donor plasmids (5 µg each) using the Amaxa 4D-Nucleofector protocol (Lonza) and selected with blasticidin S hydrochloride (Research Products International, B12200) at a final concentration of 10 ug/ml. After selection, ESCs were plated at low density to isolate single clones and screened via Western Blot for KDM3B or HA. A clone expressing degradable KDM3B was selected and nucleofected through the same procedure with 5 µg of the *Kdm3a* targeting and donor plasmids, selected with puromycin at 5 ug/ml, plated at low density to isolate single clones and screened by western blot for endogenous KDM3A and HA. Clones were then expanded and karyotyped by WiCell. All degron degradations were induced using dTAG-13 (Sigma-Aldrich SML2601) at a final concentration of 500nM.

The *Kdm3b*-KO ESC line was generated in the E14 background by using two gRNAs targeting exon 19 (5'- ACGAGATGGCAGGCTCAATC -3') and exon 22 (5'- CTTCAAAGAGTCGCTTTCGG -3'), which were inserted into dual gRNA targeting vector pX333 (Addgene #64073). E14 ESCs were then transfected with 0.5µg of the targeting vector using Lipofectamine 3000 (Thermo Fisher Scientific-L3000001). Transfected ESCs were then plated at low density to isolate single clones and screened via immunofluorescence and Western Blot for endogenous KDM3B.

Feeder MEFs were isolated and maintained in MEF media (DMEM, 10% fetal bovine serum (FBS), Glutamax, Penicillin/Streptomycin, non-essential amino acids (NEAA), 2-mercaptoethanol (4 µL/500 mL) from DR4 mice genetically resistant to geneticin (G418), puromycin, hygromycin, and 6-thioguanine. Feeder cells were irradiated with 9,000 rad after three passages. Mice were maintained according to the UW-Madison institutional animal care and use committee (IACUC)-approved protocol.

RT-qPCR Analysis and semi-quantitative rMATs validation

RNA from cell pellets treated with dTAG-13 for four hours was isolated using the Isolate II RNA Mini Kit (Bioline BIO-52072), and 1 µg was converted to cDNA using the qScript cDNA Synthesis Kit (VWR 101414). Technical replicates of 20 ng (based on RNA concentration) were used to measure Ct on a BioRad CFX96 thermocycler with iTaq UniverSYBR Green SMX (BioRad 1725125) in 20 µl reactions. Relative expression was normalized to the housekeeping gene *Gapdh*. rMATS validation primers were designed using PrimerSeq (34) to distinguish differential splicing events and semi-quantitative PCRs were performed for X cycles dependent on target (27 cycles *Dnmt3b,* 35 cycles *Tbx3,* 40 cycles Tcf12) and analyzed by gel electrophoresis. Gel intensity from each experiment was analyzed by Licor software and reported as arbitrary fluorescence units. Primers used in this study:

Gene	**Forward Primer (5'-3')**	**Reverse Primer (5'-3')**
Dnmt3b	CCCAGTGATGATCGATGCCA	ATGACCGGTACACTCCAGGA
Tbx3	CTACGGGGGAGCAATGGATG	TGGCTCTGACGATGTGGAAC
Tcf12	GGCAAGCTCCTTCATCTCCA	ATGGTCCCAGCAAACTGTGT
c-Myc	AGTGCTGCATGAGGAGACAC	GGTTTGCCTCTTCTCCACAG
Dazl	CAACTGTTAACTACCACTGCAG	CAAGAGACCACTGTCTGTATGC
Gapdh	TTCACCACCATGGAGAAGGC	CCCTTTTGGCTCCACCCT

Immunoblot

Whole cells were lysed in RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% Na deoxycholate, 0.1% SDS, 25 mM Tris pH 7.4) with 1x protease inhibitor (Roche 04693116001), sonicated for 5 secs at 20% amplitude, and quantified with the DC Protein Assay (BioRad 5000112) according to the

manufacturer's instructions. Equal amounts of protein were loaded onto an SDS-Page gel and transferred to a nitrocellulose membrane. Membranes were blocked in blocking buffer (5% milk, 0.1% Tween-20, 1x PBS) followed by incubation with primary antibody in blocking buffer. Membranes were washed 0.1% Tween-20-PBS and incubated with secondary antibody in blocking buffer. Membranes were washed and visualized with ECL reagent. Images were quantified using Image Studio Lite software. Primary antibodies: anti-H3K9me2 (Abcam ab1220, 1:1000, lot#GR33089025), anti-H3K9me3 (Active Motif 39161,1:1000, lot#01022004), anti-KDM3A (Proteintech 12835-1-AP,1:1000, lot#00065897), anti-KDM3B (U.S. Biological 1:1000, lot#L20080658), anti-EFTUD2 (LS-BIO-C346198, 1:1000, lot#207907), anti-RNAPII (Cell Signaling 14958, 1:1000, lot#5), anti-α-TUBULIN (Cell Signaling 3873, 1:2500, lot#16), anti-FLAG M2 (Sigma-Aldrich, F3165, 1:1000).

Immunofluorescence

Performed as previously described (35). Briefly, coverslips were fixed in 4% paraformaldehyde-PBS, permeabilized in 0.5% Trition-X-PBS, and washed in 0.2% Tween-20-PBS. Coverslips were blocked for 30 min in blocking buffer (1x PBS, 5% donkey serum, 0.2% Tween-20, and 0.2% fish skin gelatin). Cells were stained for 1 hr with primary antibody in blocking buffer, rinsed 2x in wash buffer, and stained for 1 hr with secondary (1:1000) in blocking buffer. Coverslips were rinsed with wash buffer, stained with DAPI (0.1 µg/ml) in wash buffer, and rinsed with wash buffer. Imaging was performed on Nikon Eclipse Ti using NIS Elements software at 20x zoom. Primary antibodies: anti-HA (Santa Cruz sc7392, 1:250, lot#G0622).

Duolink Proximity Ligation Assay

Coverslips were prepared as in the Immunofluorescence section and then stained with the Duolink® In Situ Detection Reagents Red (DUO92008) following the manufacturer's instructions. Briefly, coverslips were blocked for 30 minutes at 37°, stained with primary antibody in Duolink antibody dilutent for 1 hr at room temperature, incubated with Duolink mouse and rabbit probes for 1 hr at 37°, ligated for 30 minutes at 37°, amplified for 100 minutes at 37°, and finally mounted with Aquapolymount (Polysciences, 18606). Confocal images were acquired using a Nikon A1R confocal microscope with Nikon NIS-Elements software and analyzed with NIS-Elements. Z-stacks were taken in 1 µM slices, approximately 5 slices taken per colony, with slice in the middle of colony selected. For the transcriptional inhibition PLA, cells were treated with 5,6-Dichloro-1-beta-Ribo-furanosyl Benzimidazole (DRB) for 3 hrs at a final concentration of 10µM and then processed at the same Duolink™ parameters above and finally imaged with Nikon Eclipse Ti using NIS Elements software. Primary antibodies: anti-HA (Santa Cruz sc7392, 1:250, lot#G0622), anti-PRMT5 (EMD Millipore 07-405, 1:250, lot#3288011) anti-KDM3B (Cell Signaling 5377, 1:250, lot#1), anti-EFTUD2 (LS-BIO-C346198, 1:250, lot#207907), and anti-OCT4 (Cell Signaling 2840, 1:250, lot#16).

Protein Extraction and Immunoprecipitation

Total protein was extracted using a modified nuclear extraction method (36) as previously described (37). Briefly, ESCs were feeder depleted, and extracted in 5x volume buffer A (10 mM HEPES [pH 7.9],1.5 mM MgCl2, 10 mM KCl, 1 mM DTT, Phosstop™- phosphatase inhibitor (Roche 04906837001) and cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche 04693132001) for 15 minutes on ice after gentle douncing. Nuclei were pelleted at 1500xg for 10 minutes, washed in buffer A, and resuspended in 2x volume buffer C (0.42 M KCl, 20 mM HEPES [pH 7.9], 0.2 mM EDTA [pH 8.0], 5% glycerol, 1 mM DTT, and phosphatase and protease inhibitor cocktail and dounced thoroughly. After douncing, $MgCl_2$ was spiked in to a final concentration of 1.5 mM, benzonase added (250 units/ml; Sigma-Aldrich E1014), and

incubated for 10 minutes at room temperature. Extracts were then rotated at 4°C for 2 hours. Soluble nuclear extracts were isolated by centrifugation at 23,000xg for 30 minutes at 4°C.

For FLAG-IP-MS, 10-15 mg of nuclear extract was dialyzed overnight to 0.3M KCl using a 7,000 MWCO Slide-A-Lyzer Dialysis Cassette G2 (Thermo Scientific). Extracts were then incubated with FLAG-M2 agarose (Sigma, 50% slurry) bead slurry at a ratio of 90 ul beads slurry to 400 ug of dialyzed protein for 2 hours rotating at 4°C. Samples were washed thrice with IP buffer (0.3 M KCl, 20 mM HEPES, and phosphatase and protease inhibitor cocktail), collected by spinning at 1000g x minutes between washes. IP complexes were eluted with 200 ng/uL of 5X FLAG peptide (manufactured by Peprotech).

Eluted protein complexes were trichloroacetic acid (TCA) precipitated, in-solution digested with trypsin (38), desalted using Agilent Bond Elut OMIX C18 SPE pipette tips per manufacturer protocol and eluted in 10µl of 60/40/0.1% ACN/H_2O/TFA. Dried to completion in the speed-vac and finally reconstituted in 40µl of 0.1% formic acid. Peptides were analyzed by nanoLC-MS/MS using the Agilent 1100 nanoflow system (Agilent) connected to hybrid linear ion trap-orbitrap mass spectrometer (LTQ-Orbitrap Elite™, Thermo Fisher Scientific) equipped with an EASY-Spray™ electrospray source (held at constant 35°C). Chromatography of peptides prior to mass spectral analysis was accomplished using capillary emitter column (PepMap® C18, 3µM, 100Å, 150x0.075mm, Thermo Fisher Scientific). Protein annotations, significance of identification and spectral based quantification was done with Scaffold software (version 4.11.0, Proteome Software Inc., Portland, OR). Peptide identifications were accepted if they could be established at greater than 7.0% probability to achieve an FDR less than 1.0% by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 99.0% probability to achieve an FDR less than 1.0% and contained at least 2 identified peptides.

Crapome Filtering and String Network Analysis

KDM3A and KDM3B FLAG-IP mediated interactome was first filtered for non-nuclear compartments using DAVID GO analysis (https://david.ncifcrf.gov/), then remaining proteins were further filtered using the “crapome” contaminant repository (https://reprint-apms.org/). Contaminant controls (cc2,cc3,cc4,cc5,cc6,cc28) were selected based on similarity of nuclear extraction method, instrumentation, and sensitivity along with a FLAG-IP from our uninduced ESCs. After analysis with the Crapome pipeline, proteins with a minimum spectral count of 3, saint probability score of 0.7, and fold change (FC) cutoff of 2 were kept for further analysis. String network analysis (https://string-db.org/) was performed on the overlapping or unique interactomes of KDM3A and KDM3B. Experiments and databases were used for active interaction sources at a high confidence, hiding disconnected nodes. The resulting network was exported to Cytoscape (39) for further visual editing based on FC. The largest bubbles have a FC >10, medium FC 10<x>5, and small FC x<5.

RNA isolation and library preparation

Samples for RNA-seq were prepared as previously described (40). ESCs (WT and degron) were grown in three independent replicate samples and treated with dTAG-13 for 4 hours, washed with PBS, harvested with trypsin and stored in Trizol (Sigma 15596018). RNA was isolated by adding one-fifth the volume of chloroform to each sample and phases separated by max centrifugation. The upper aqueous layer was isolated, RNA was precipitated with 0.53 volumes of ethanol, and applied to a RNeasy column (Qiagen 74104). RNA was washed with 500 µl RW1 and DNA was digested on the column with DNase (Qiagen 79254) for 30 min. RNA was washed according to the RNeasy protocol and eluted in 30 ul of H2O. RNA was quantitated with Qubit and 1 µg of RNA was used to . The cDNA library was constructed using TruSeq RNA

Sample Preparation kit V2 (RS-122-2002) according to the manufacturer's instructions. Libraries were assessed with Bioanalyzer3.0.

RNA-Seq computational and statistical analysis

RNA-seq analysis was performed as described in (41). Libraries were sequenced to a depth of ~50 million reads/sample were sequenced PE150 by Novogene on a Illumina Novaseq platform. Sequencing quality was assessed with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Sequenced reads were processed using Trimmomatic (42) with the following parameters: LEADING:3 TRAILING:3 CROP:135 HEADCROP:15 SLIDINGWINDOW:4:15 MINLEN:36. Reads were aligned to the mm9 genome using RSEM-1.2.4 (43) with a mismatch per seed set to 2 and seed length set to 28 (- bowtie-m 200 --bowtie-n 2 --forward-prob 0.5 --seed-length 28 --paired-end). RSEM-1.2.4 alignment yielded transcripts per million (TPM) for each gene. A median TPM value from the three replicates was used for all main figures. Differentially expressed (DE) genes were called with EBSeq (44). Differentially expressed genes were filtered to have a posterior probability DE greater than 0.95 and a 2 or 0.5 greater posterior fold change (PostFC). Expression changes reported are PostFC values determined by EBSeq unless otherwise noted to be TPM. Gene Ontology was performed using DAVID (45, 46). DAVID parameters included: GOTERM_BP_Direct and GOTERM_MF_Direct, viewed by functional annotation clustering. GO Terms within functional annotation clusters with a p-value X<.001 were considered significant.

Differential Splicing Analysis

Sequencing reads were first mapped to the mouse genome, mm9, using STAR aligner (version 2.7.8a) (47). The mm9 genomic index was created using the genomegenerate function from STAR with ENCODE fasta and comprehensive gene annotation (NCBIm37). To identify canonical differential splicing events, mapped data was analyzed with rMATS 4.1.2 (48–50) was used for counting junction reads amongst three replicates of each condition. Filtering for significant splicing events was based on FDR<0.1, minimum inclusion level difference of 5%, and 15 minimum read count in at least one included or skipped isoform. Heatmap generated using pheatmap in Rstudio with kmeans clustering of 5 for inclusion level difference of aberrantly spliced genes.

Transcriptomic annotation was performed as originally described (51)with some updates. Mapped reads were filtered to remove repeat RNAs and exonic reads. Reads attributable to alternative polyadenylation sites were removed based on 3' end sequencing of mouse embryonic stem cells. Intronic reads were partitioned and allocated to each intron proportional to its weight assuming all introns of a gene are present in equal abundance. Differential analysis using DESeq was then used to determine introns enriched in read coverage (observed read counts) compared to the expected counts using an FDR adjusted p-value threshold of 0.01 and fold change threshold of 2. We used custom scripts to derive a transcriptome annotation based on Mm9 gene start and end boundaries, the genomic locations of detained introns as determined above, and the genomic locations of all splice junctionsTo identify intronic polyadenylation (alternative last exon (ALE) choice) events and detained introns, the Dexseq (52) pipeline was used. Events were counted using the dexseq_count.py script of mapped reads against a custom annotation for mouse ESCs previously generated in (51). For the data presented in Figure 4, only events with a padj value<0.1 and from events annotated as Detained Introns or Class I ALEs were used. All scripts are available on request.

ChIP-Seq analysis

ChIP-Seq analysis was performed as in as described (53). Briefly, H3K9me2 ChIP-Seq data for WT ESCs with the accession number GSE153561 (54) and GSE177058 (55) were downloaded

from Gene Expression Omnibus. H3K9me2 ChIP-Seq data for WT and KDM3A/KDM3B DKO ESCs with the accession number DRA006496 was downloaded from the DDBJ database (22). All data was aligned to mm9 using Bowtie2 (56) with the default parameters. Sam files were converted into Bam files and sorted with samtools-1.2 (57) with the default parameters. Peaks were called relative to the input with MACS2 (58) using the following parameters: --broad -p 0.0001. Peak files were annotated using Homer (annotatePeaks.pl). Annotated peaks from both (54)and (55) were overlapped with Venny2.1 (https://bioinfogp.cnb.csic.es/tools/venny/) and used to determine H3K9me2 bound genes in figure 3. ChIP enrichment was visualized using Integrative Genomics Viewer (IGV) (59). H3K9me2 quantification was calculated from all data sets with bedtools multiBamCov and normalized by read counts per million and per kilobase of gene length. Metaplots were generated using Deeptools (60). First, bigwigs were created using bamCompare to calculate the log2 of H3K9me2 IP over input specifically across skipped exons of interest with --binSize 10 --normalizeUsing CPM --ignoreForNormalization chrX --scaleFactorsMethod None. Quantification of the skipped exons was visualized using deeptools with computeMatrix scale-regions -b 500 -a 500 --regionBodyLength 500 and plotted with plotProfile -m regionsLabel Pos Neg --numPlotsPerRow 1

RESULTS

KDM3A and KDM3B interact with splicing machinery in mouse embryonic stem cells

KDM3A and KDM3B have identical catalytic activity and partially redundant functions in several model systems. In mouse embryonic stem cells (ESCs) the deletion of both *Kdm3a* and *Kdm3b* decreases cell number within 4 days (22). KDM3A and KDM3B are divergent at the protein level outside of the catalytic Jumonji domain and a zinc finger domain that is required for catalytic activity. To gain insight into KDM3A and KDM3B function we performed protein interactome analysis in ESCs. We first generated ESC lines with a FLAG-epitope tagged KDM3A or KDM3B under a doxycycline inducible promoter (Figure 1A and B) and achieved ~1.5-2 fold over endogenous levels after induction (Figure 1A and B). We then performed immunoprecipitation for each protein from nuclear extracts isolated with benzonase digestion followed by mass spectrometry (Figure 1B, Supplementary Figure S1A). After stringent filtering for common contaminants of FLAG pulldowns (Crapome database (61) Saint score 0.7, Methods), we identified 122 proteins that were commonly enriched over background uninduced controls in both the KDM3A and KDM3B interactomes and 37 and 35 unique interactions respectively (Figure 1C). We further classified the interactions using STRING network analysis (Figure 1D, Supplementary Figure S1B and S1C). KDM3A and KDM3B demethylate the repressive H3K9me2 that is enriched in heterochromatin that would ultimately result in an increase in transcription. Therefore, we expected to find interactors that are transcriptional activators. Instead, we found associations with repressor proteins including the Heterochromatin Protein 1 family (CBX1,CBX3, CBX5), DNA methyltransferase complex DNMT3A and DNMT3L, components of the polycomb complex, EED, histone deacetylase HDAC6 and components of the nuclear lamina such as BANF1 (Figure 1D).These associations suggest that KDM3A and KDM3B are localized to a heterochromatin environment within the cell where they may initiate the first step of derepression. A majority of the interacting proteins were associated with the RNA processing machinery (Figure 1C and D) along with the FACT and PAF complexes that associate with actively elongating RNAPII. The RNA processing proteins included members of the spliceosomal machinery, which removes introns from pre-mRNA (Ohi 2017), and RNA helicases, which are important for general RNA metabolism and rearrangement of ribonucleoprotein complexes (62). We also found components of the exon junction complex, which control post-transcriptional export and stability of mRNAs (63) (Figure 1D and E). To validate these interactions under physiological conditions in situ, we performed a proximity ligation assay (Methods, Duolink™) for the interaction of KDM3B and EFTUD2, a GTPase that

functions in the spliceosome. We found a highly specific signal only in the presence of both KDM3B and EFTUD2 antibodies that was absent when cells were stained with either antibody alone (Figure 1F), and in an ESC line that was knocked out for KDM3B (Figure 1F, Supplementary Figure S1D). Further, an abundant ESC specific nuclear protein OCT4 (POU5F1), that was not detected in the KDM3A/KDM3B interactome, did not show an interaction by proximity ligation assay (Figure 1F). Because KDM3B is suggested to associate better with chromatin under ongoing transcription (54), we inhibited cells with 5,6-Dichloro-1-beta-Ribo-furanosyl Benzimidazole (DRB) that pauses RNA polymerase II and repeated the PLA. We found the relative association of KDM3B with EFTUD2 decreased upon DRB treatment suggesting that KDM3B splicing associations are modulated but not dependent on ongoing transcription (Supplementary Figure S1E).

Delayed accumulation of H3K9me2 upon acute degradation of KDM3A and KDM3B proteins

The majority (~80% in human cells) of mRNA processing is thought to occur co-transcriptionally, with some variation between species and cell types (64–67). Most active spliceosomes are associated with chromatin (68). Traditional methods of genomic deletion such as conditional deletion or transcriptional intervention with RNAi knockdown would only remove the KDM3A and KDM3B proteins slowly, confounding direct and indirect effects on splicing. Therefore, we needed a model of rapid and inducible protein degradation. We gene-edited an inducible “degron” tag into the endogenous locus of both KDM3A and KDM3B (Figure 2A). To facilitate CRISPR-Cas9 mediated gene editing we used microhomology mediated end joining (MMEJ) (32) (Supplementary Figure S2A) to insert the degron tag into the 5’ end of each KDM3 locus to avoid the 3’ end that encodes the Jumonji catalytic domain. The degron tag contains a 2xhemagglutinin (HA) tag and is derived from a mutated variant of the FKBP12 protein, $FKBP12^{F36V}$, which responds to the dTAG-13 ligand for targeting the tagged protein to the ubiquitin proteasome pathway (31). We initially obtained clones that were degron tagged at both *Kdm3b* alleles with decreased expression of the KDM3B protein (Figure 2B). Even though depletion of KDM3B has been shown in some cases to lead to genome instability (69), importantly, these clones maintained a stable karyotype throughout the derivation process (Supplementary Figure S2C). The Kdm3b-degron tagged clones were then targeted with the *Kdm3a* degron construct and heterozygous clones were obtained where one allele of *Kdm3a* codes for an in-frame degron tag and the second allele was knocked-out by an insertion at the 3’ repair site (Supplementary Figure S2B and D). Both KDM3A and KDM3B proteins were completely degraded upon exposure to the small molecule dTAG-13 (Figure 2B). From a timecourse of dTAG-13 exposure we determined that KDM3A/KDM3B could be degraded as soon as 1 hour (Figure 2D). By contrast, the target of KDM3A/3B catalytic activity, H3K9me2, took 48 hours to accumulate post-degradation (Figure 2C). This increased H3K9me2 did not lead to a further increase in H3K9me3 within 48 hours (Figure 2C). Due to a lack of endogenous KDM3A antibody for immunofluorescence, we performed immunofluorescence and PLAs with HA and EFTUD2 in the degron and confirmed this interaction responds in a dTAG-13 dependent manner (Supplementary Figure S2E and F). This established a system in which we could study the direct effects of KDM3A and KDM3B loss in a rapid manner.

Complete degradation of KDM3A and KDM3B leads to differential gene expression agnostic of H3K9me2 status

We performed RNA-seq of the degron clones at 4 hours after exposure to dTAG-13 to capture immediate changes in both gene expression and splicing. The target of KDM3A and KDM3B, H3K9me2, is enriched in broad megabase domains that encompass tissue-specific genes

during development (17). After degradation, we found 420 genes with differential expression (DEG) upon complete loss of all KDM3A/KDM3B protein compared to wildtype ESCs (Figure 3A) (Supplementary Table 2). The proteins in the shared KDM3A and KDM3B interactome are not direct transcriptional targets since they were not differentially expressed after KDM3A/KDM3B degradation (data not shown). The upregulated and downregulated genes were enriched for functional categories of cell migration and signaling pathways, and interferon response respectively (Figure 3A). Despite this functional distinction both the up and down-regulated DEG were lower expressed as compared to the global median of genes in ESCs (Figure 3B).

To determine if the low expression of DEGs was because of prevailing H3K9me2, we overlapped the DEGs with consensus peaks from two published H3K9me2 chromatin immunoprecipitation sequencing (ChIP-seq) data sets (54, 55). A significant proportion of the DEGs (179 genes) were enriched for H3K9me2, indicating that H3K9me2 may participate in localizing KDM3A and KDM3B to affected genes (Supplementary Figure S3A). Irrespective of whether they were up or downregulated upon KDM3A/KDM3B degradation, the DEGs that had a H3K9me2 peak were lower expressed than those that lacked H3K9me2 (Figure 3C). Among the genes that contained a H3K9me2 peak, the overall H3K9me2 enrichment was not significantly different between up and down DEG (Figure 3D).

Finally, to determine the genes that may be directly regulated by the loss of KDM3A and KDM3B, we analyzed a third set of H3K9me2 ChIP-seq containing KDM3A/KDM3B DKO ESCs and WT ESCs (22)). Assuming loss of the two demethylases would lead to the spread of repressive H3K9me2, we reasoned genes that were downregulated and bound by H3K9me2 in the DKO ESCs could be direct targets of the enzymes and found 25 genes that fit this criterion (Figure 3E). We did not find a significant difference between the H3K9me2 quantification of DEG genes in the WT or DKO ESCs (Supplementary Figure S3B). Taken together with the lack of change in global H3K9me2 levels (Figure 2C) and the small number of DEG that are enriched for H3K9me2 at 4 hours (Figure 3E), we conclude that these gene expression changes are not dependent on H3K9me2 status but may instead correlate with the presence of KDM3A or KDM3B proteins within these domains. Upon the degradation of KDM3A or KDM3B, essential protein complexes that are scaffolded or nucleated by KDM3A or KDM3b may be disrupted, leading to changes in gene expression or alternative splicing.

Alternative splicing defects upon acute loss of KDM3A/KDM3B occur in signaling and chromatin related genes

Due to the interaction of KDM3A and KDM3B with splicing machinery, we analyzed the RNA sequencing for splicing aberrations using rMATs. rMATS detects differential splicing events by modeling exon inclusion versus exon skipping levels in multiple replicates to calculate a ‘percent spliced in (PSI)’ for each exon considered (48, 70). rMATs reports canonical splicing events such as skipped exons, alternative 5’ and 3’ splice site selection, and mutually exclusive exons (Figure 4B). We performed rMATS with cutoffs of FDR < 0.1, PSI difference of 5%, and a 15 read minimum for each junction and identified 901 significant events within 729 genes (Figure 4B).

We also performed DEXseq (52) analysis to identify non-canonical events such as detained introns (DIs), where mRNA containing one or more introns is trapped in the nucleus and ultimately degraded, or post-transcriptionally spliced and exported into the cytoplasm in response to an additional cue. DEXseq calculates the ratio between the read counts of a particular exon and the counts at the exons at the rest of the gene to calculate relative exon usage (52). By combining DEXseq with a custom annotation including detained introns (51) we assessed the splicing effects upon the acute loss of KDM3A and KDM3B. At a DEXseq

threshold of padj < 0.1 we identified several genes that had detained introns (34 events), as well as some with alternative polyadenylation sites (32 events) (Figure 4B).

The assessment of DIs is relevant because we found an interaction of KDM3A/KDM3B with PRMT5, a known regulator of detained introns in several cell types (71); and of splicing in ESCs (72). While PRMT5 was strongly enriched in the KDM3A interactome (Supplementary Table 1), it is also a common contaminant in FLAG-mediated immunoprecipitation mass spectrometry (Crapome(61)). Therefore, we first validated the interaction independently using an HA-epitope that is introduced along with the degron tag in the KDM3A/KDM3B degron ESC line. We found a HA-KDM3A/KDM3B interaction with PRMT5 using PLA that was lost upon the addition of dTAG-13. (Supplementary Figure S4A). An endogenous KDM3B-PRMT5 interaction was also observed in a second, feeder-dependent ESC line (Supplementary Figure S4B).

Combining the significant events from the rMATS and DEXseq analyses, we found the splicing of 781 genes to be impacted by loss of KDM3A/KDM3B (Supplementary Table 3) but only 21 genes were both differentially expressed and changed in splicing patterns (Figure 4A) suggesting that the changes in isoforms do not immediately lead to expression changes (73). We found a greater number of exon skipping/inclusion events (446) with a negative PSI (WT>DDeg+) than positive PSI (WT<DDeg+, 249), insinuating the loss of KDM3A and KDM3B leads to more exon skipping. Other than a 3-fold increase in detained introns in the DDeg+ condition, we found the directionality within other event classes to be evenly distributed between the WT and DDeg+ conditions (Figure 4B).

We then evaluated whether the genes affected in the DDeg+ condition, were enriched for specific functions. While KDM3A/3B responsive genes had a significant enrichment for gene ontology terms such as Transferase Activity and Phosphorylation (Figure 4C), genes with a positive PSI value were associated with Negative Regulation of Telomere Maintenance, mRNA processing, Chromatin Remodeling, and Transferase Activity (Figure 4C, 4D). By contrast, genes with a negative PSI value were associated with Kinase Activity and Transferase Activity (Figure 4C, 4D). The largest class of splicing event, skipped exons (SEs), included several proteins that had sequence-specific DNA binding such as transcription factors *Tbx3, Meis1,* and *Zfp263.* To gain further insight into the splicing events that were most susceptible to the acute loss of the KDM3 proteins, we analyzed the 901 significant canonical splicing events found in the WT/DDeg+ analysis and compared this to the PSI values between WT and the DDeg ESCs prior to the addition of the dTAG-13 ligand (WT vs DDeg-), which have slightly reduced expression of KDM3A and KDM3B and performed k-means clustering into 5 clusters (Figure 4E).

In general, most of the splicing events were further exacerbated upon the addition of dTAG-13. The cluster agnostic to the dTAG-13 addition included genes that function in RNA processing such as *Ythdf3* and *Prp40b*. Chromatin-associated proteins such as *Atrx* and *Med12* and chromatin-modifying enzymes *Dnmt3b*, *Setd6* and *Suv39h2* had splice events included more in the WT cells (WT>DDeg+) (Figure 4E). *Rbfox2,* RNA-binding protein that regulates alternative splicing events, was also in the WT>DDeg+ group (Supplementary Figure S4C). The largest clusters with splice events included more in DDeg+ cells (WT<DDeg+) included genes with functions in ubiquitinase activity such as *Rnf13, Rnf214 and Trim3*. The clusters that were most sensitive to the amount of KDM3 proteins included kinases in the MAP pathway *Mapk5* and transcription factors such as *Tbx3* and Tcf12 that are important for pluripotency.

KDM3A/KDM3B depletion leads to the missplicing of regulators of the naïve state of pluripotency

Mouse pluripotent states exist in a continuum of pluripotency ranging from the “ground state” that resembles the early blastocyst to a “naïve state” resembling early epiblast and “primed state” obtained from the late epiblast. The ground state ESCs can be obtained from embryos by

deriving ESCs in 2i/LIF conditions without serum while the “naïve state” can be captured in serum/LIF conditions. ESCs can be converted from the naïve state (in which all our experiments were performed) to ground state by passaging in 2i/LIF conditions (Figure 5A). We were struck by the differential splicing of several regulators of the pluripotency continuum in our dataset. To investigate this idea further we evaluated splicing in the ground state of pluripotency.

We performed rMATS analysis to compare canonical splicing events between our WT ESCs which were cultured in serum/LIF and ESCs grown in 2i/LIF conditions from two sets of published data (Methods). There were ~4000-5000 splicing events detected between serum and 2i/LIF conditions indicating that cell fate is influenced by the state of splicing in the cell. (Supplementary Figure S5A and B). We then compared the splicing events that changed upon degradation of all KDM3 alleles (Double degron conditions- 901 events from Figure 4E) to those that changed in the 2i/LIF conditions (Supplementary Figure S5A and B). There was a larger overlap in splicing events with a negative PSI value in both the DDeg+ (WT<DDeg+) and the 2i/LIF conditions (198 events from Yang et al. (74) among 304 shared events and 156 events from Kinoshita et. al. (75) from 249 shared events), Supplementary Figure S5A and B). By contrast there was a smaller overlap in splicing with a positive PSI value in WT ESCs than in DDeg+ (WT>DDeg+) and 2i/LIF conditions (64 events from Yang et al. (74) of 304 shared events and 36 events from Kinoshita et al. (75) from 249 shared events, Supplementary Figure S5A and B). When comparing the magnitude of changes, we found a large positive correlation in the inclusion levels of splice forms between the DDeg+ and the 2i/LIF conditions (Pearson = 0.73 with Kinoshita et al. (75) and 0.51 with Yang et al. (74) (Figure 5B and Supplementary Figure S5C). Therefore, upon degradation of the KDM3 proteins the alternative splicing of genes trends toward a more ground pluripotent state pattern.

Among the splicing events we focused on candidates that have been implicated in cell fate switches in pluripotency and were similarly spliced in the WT-DDeg+ and, WT-2i/LIF comparison. *Dnmt3b*, a de novo methyltransferase that is lowly expressed in ground state cells and gains in expression toward primed pluripotency. The different isoforms of Dnmt3b that are detected in our data lack two exons that code for conserved motifs within the catalytic domain and are thought to be catalytically inactive (76–78). *Tbx3* is a transcription factor with two isoforms that regulate the expression of pluripotency gene *Nanog* (79). *Tcf12* is a transcription factor whose deletion compromises mesodermal specification and reduces the expression of *Nanog* and other transcription factors in ESCs and can destabilize self-renewal capacity (80, 81).

To validate our results, we performed semi-quantitative RT-PCR analysis for three candidate genes in two independent clones of KDM3A/KDM3B with degron tags and in WT ESCs grown in serum/LIF and 2i/LIF conditions (Figure 5C). In every case we found that the splicing pattern in KDM3A/KDM3B degraded clones trended towards the ESCs grown in 2i/LIF conditions, indicating a shift towards ground state.

The transcriptomes of serum and 2i/LIF ESCs are highly similar (82). Among the few genes that are significantly different between the two conditions, we found that genes higher in 2i/LIF, such as *Dazl*, were slightly increased in the double degron cell lines, while genes higher in serum/LIF, such as *c-Myc*, were not downregulated (Figure 5D). Notably, the exposure of double degron cell lines to 2i/LIF conditions increased *Dazl* levels compared to WT cells and decreased c-Myc levels further than in WT cells (Figure 5D). Taken together we conclude that altering KDM3 levels impacts important regulators of cell fate by splicing leading to a compromised naïve state of pluripotency.

DISCUSSION

In this study, we have found the histone demethylases, KDM3A and KDM3B, to be important regulators of RNA splicing in mouse ESCs (mESCs) independent of the H3K9 methylation status of target genes. Histone demethylases have a profound impact on processes such as development, gene regulation, and disease manifestation (83). The primary function of histone demethylases is in modulating transcriptional activity by altering local chromatin structure and providing a platform for the recruitment of histone modification "reader" proteins that transduce downstream effects. Histone demethylases can also have non-histone targets, expanding their range of action (84, 85). Beyond these catalytic activities, histone demethylases are members of larger protein complexes that modulate their activity, raising the possibility of non-catalytic functions such as scaffolding. In pluripotent cells KDM3A and KDM3B are implicated to have overlapping function, since only the deletion of all four alleles induces lethality in mESCs and developing epiblasts (22). Supporting this notion by performing a discovery-based protein interaction analysis we have found that the majority of KDM3A and KDM3B partners are the same. That most of these overlapping binding partners were related to RNA splicing machinery insinuates that KDM3A and KDM3B mediate post-transcriptional control in addition to their traditionally acknowledged transcriptional roles in pluripotency.

Splicing regulation is essential in ESCs as improper splicing can lead to loss of self-renewal, differentiation defects and cell death(86–88). In human ESCs, only specific spliced isoforms of the master pluripotency transcription factor OCT4 can support self-renewal (89). During somatic cell reprogramming the entire pattern of splicing is switched from that of a somatic to a pluripotent cell (90). A specific isoform of FOXP1 generated through alternative splicing has been shown to promote the maintenance of ESC pluripotency and contribute to efficient reprogramming of somatic cells (91). Beyond these examples of transcription factors, the deletion of splicing factor *Rbfox2* in human ESCs, where it binds to almost 7% of transcripts and autoregulates its own splicing, leads to cells death (92). These studies underscore the importance of alternative splicing in cell fate determination. From our results, the loss of KDM3A and KDM3B leads to a disruption of the pluripotency continuum and leaves the cells skewed towards the ground state (2i/LIF). If ESCs are unable to exit the pluripotent state properly, this would lead to differentiation defects and may explain the early developmental lethality observed in KDM3A/KDM3B double-knockout mice. The loss of KDM3A/KDM3B leads to the improper splicing of *Rbfox2, Tbx3, Dnmt3b,* and *Tcf12*, all of which have been implicated in the proper maintenance of pluripotency (92–95). The variant of RBFOX2 affected by the KDM3 proteins prevents higher order assembly that affects splicing regulation. As ESCs are switched from serum to 2i/LIF media, an environment of DNA hypomethylation is rapidly induced, thought to be driven by *Prdm14* and downregulation of *Dnmt3a/3b* expression (96, 97). Our results implicate another layer of DNA hypomethylation from an increase in the catalytically inactive spliced isoform of DNMT3B detected in the loss of KDM3A/KDM3B. *Tbx3* is of particular interest because the expression level has been linked to the cell's lineage commitment (98) and a specific alternative splicing event at exon 2 determines the protein's binding ability to the core pluripotency regulator Nanog, controlling transcriptional output (79).

How are the splicing targets of KDM3A and KDM3B selected? One avenue could be by RNA modifications (99). N6-methyladenosine (m6A) regulates the fate of mRNA molecules by altering stability, localization, and splicing (100, 101). Misregulation of m6A has developmental consequences as depletion of m6A from epiblast blocks naïve pluripotency exit through the retention of methylated transcripts and results in early embryonic lethality (102). In mESC, m6A on newly transcribed RNA leads to the recruitment of KDM3B and subsequent gene activation by the removal of H3K9me2 (54). These observations imply a connection where RNA methylation precedes histone demethylation. However, we did not find a reciprocal relationship, where KDM3B splicing targets were preferentially m6a methylated (data not shown). Given the

necessity of fine-tuning post-transcriptional regulation through alternative splicing for development, further study of the mechanism behind KDM3A and KDM3B splicing regulation is warranted.

Beyond pluripotency, there are examples of KDM3A affecting the splicing of single genes in cancer cells. In prostate cancer cells, KDM3A serves as a co-activator for the alterative splicing of androgen receptor variant 7, a major driver of therapy resistance in prostate cancer progression (103). In response to DNA damage in breast cancer cells, KDM3A is phosphorylated, and regulates SAT1 alternative splicing, independent of demethylating activity (104). Taken together our results emphasize this additional layer of post transcriptional gene regulation in the non-transformed developmental context and suggest that known phenotypes of the individual KDM3A and KDM3B knockouts in metabolism and germline specification will have to be re-examined.

DATA AVAILABILITY

The datasets supporting the conclusions of this article are available in the National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO) repository with the accession code GSE230341. Reviewers can access this with token otebqwiebbkjhij at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE230341
Proteomics Mass Spectrometry Data has been submitted to PRIDE Database reference#1-20230523-165433 (under processing for assession number)
Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

We would like to thank Dr. Coral Wille for ChIP-seq analysis and Hailey Thurston for technical assistance; members of the Sridharan lab for discussion and critical reading of the manuscript; and the Mass Spectrometry Core facility at UW-Madison Biotechnology Center.

FUNDING

This was supported by the National Institutes of Health [1R01HD105151 to R.S.]; The UW-Madison Stem Cell and Regenerative Medicine Center fellowship to C.M.D.; National Institutes of Health [5T32GM007133 to H.C.]; National Institutes of Health [R01GM141544 to P.L.B.]; National Institutes of Health [T32 GM135134 D.E.E.]; Funding for open access charge: NIH [1R01HD105151].

Conflict of interest statement
None declared.

REFERENCES

1. Holmberg,J. and Perlmann,T. (2012) Maintaining differentiated cellular identity. *Nat Rev Genet*, 13, 429–439.

2. Iwafuchi-Doi,M. and Zaret,K.S. (2016) Cell fate control by pioneer transcription factors. *Development*, 143, 1833–1837.

3. Shchuka,V.M., Malek-Gilani,N., Singh,G., Langroudi,L., Dhaliwal,N.K., Moorthy,S.D., Davidson,S., Macpherson,N.N. and Mitchell,J.A. (2015) Chromatin Dynamics in Lineage Commitment and Cellular Reprogramming. *Genes-basel*, 6, 641–661.

4. Gaspar-Maia,A., Alajem,A., Meshorer,E. and Ramalho-Santos,M. (2011) Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Bio*, 12, 36.

5. Black,J.C. and Whetstine,J.R. (2011) Chromatin landscape. *Epigenetics*, 6, 9–15.

6. Talbert,P.B., Meers,M.P. and Henikoff,S. (2019) Old cogs, new tricks: the evolution of gene expression in a chromatin context. *Nat Rev Genet*, 20, 283–297.

7. Shvedunova,M. and Akhtar,A. (2022) Modulation of cellular processes by histone and non-histone protein acetylation. *Nat Rev Mol Cell Bio*, 23, 329–349.

8. Berger,S.L. (2007) The complex language of chromatin regulation during transcription. *Nature*, 447, 407–412.

9. Li,B., Carey,M. and Workman,J.L. (2007) The Role of Chromatin during Transcription. *Cell*, 128, 707–719.

10. Kooistra,S.M. and Helin,K. (2012) Molecular mechanisms and potential functions of histone demethylases. *Nat Rev Mol Cell Bio*, 13, 297–311.

11. Black,J.C., Van Rechem,C. and Whetstine,J.R. (2012) Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact. *Mol Cell*, 48, 491–507.

12. Barski,A., Cuddapah,S., Cui,K., Roh,T.-Y., Schones,D.E., Wang,Z., Wei,G., Chepelev,I. and Zhao,K. (2007) High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell*, 129, 823–837.

13. Hublitz,P., Albert,M., Hfmpeters,A. and Peters,A.H.F.M. (2009) Mechanisms of transcriptional repression by histone lysine methylation. *Int J Dev Biol*, 53, 335–354.

14. Lienert,F., Mohn,F., Tiwari,V.K., Baubec,T., Roloff,T.C., Gaidatzis,D., Stadler,M.B. and Schübeler,D. (2011) Genomic Prevalence of Heterochromatic H3K9me2 and Transcription Do Not Discriminate Pluripotent from Terminally Differentiated Cells. *Plos Genet*, 7, e1002090.

15. Wen,B., Wu,H., Shinkai,Y., Irizarry,R.A. and Feinberg,A.P. (2009) Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet*, 41, 246–250.

16. Poleshko,A., Shah,P.P., Gupta,M., Babu,A., Morley,M.P., Manderfield,L.J., Ifkovits,J.L., Calderon,D., Aghajanian,H., Sierra-Pagán,J.E., *et al.* (2017) Genome-Nuclear Lamina Interactions Regulate Cardiac Stem Cell Lineage Restriction. *Cell*, 171, 573-587.e14.

17. Padeken,J., Methot,S.P. and Gasser,S.M. (2022) Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance. *Nat Rev Mol Cell Bio*, 23, 623–640.

18. Cloos,P.A.C., Christensen,J., Agger,K. and Helin,K. (2008) Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. *Gene Dev*, 22, 1115–1140.

19. Kim,J.-Y., Kim,K.-B., Eom,G.H., Choe,N., Kee,H.J., Son,H.-J., Oh,S.-T., Kim,D.-W., Pak,J.H., Baek,H.J., *et al.* (2012) KDM3B Is the H3K9 Demethylase Involved in Transcriptional Activation of lmo2 in Leukemia. *Mol Cell Biol*, 32, 2917–2933.

20. Klose,R.J., Kallin,E.M. and Zhang,Y. (2006) JmjC-domain-containing proteins and histone demethylation. *Nat Rev Genet*, 7, 715–727.

21. Yamane,K., Toumazou,C., Tsukada,Y., Erdjument-Bromage,H., Tempst,P., Wong,J. and Zhang,Y. (2006) JHDM2A, a JmjC-Containing H3K9 Demethylase, Facilitates Transcription Activation by Androgen Receptor. *Cell*, 125, 483–495.

22. Kuroki,S., Nakai,Y., Maeda,R., Okashita,N., Akiyoshi,M., Yamaguchi,Y., Kitano,S., Miyachi,H., Nakato,R., Ichiyanagi,K., *et al.* (2018) Combined Loss of JMJD1A and JMJD1B Reveals Critical Roles for H3K9 Demethylation in the Maintenance of Embryonic Stem Cells and Early Embryogenesis. *Stem Cell Rep*, 10, 1340–1354.

23. Tateishi,K., Okada,Y., Kallin,E.M. and Zhang,Y. (2009) Role of Jhdm2a in regulating metabolic gene expression and obesity resistance. *Nature*, 458, 757–61.

24. Okada,Y., Scott,G., Ray,M.K., Mishina,Y. and Zhang,Y. (2007) Histone demethylase JHDM2A is critical for Tnp1 and Prm1 transcription and spermatogenesis. *Nature*, 450, 119–123.

25. Qin,L., Xu,Y., Yu,X., Toneff,M.J., Li,D., Liao,L., Martinez,J.D., Li,Y. and Xu,J. (2017) The histone demethylase Kdm3a is required for normal epithelial proliferation, ductal elongation and tumor growth in the mouse mammary gland. *Oncotarget*, 8, 84761–84775.

26. Liu,Z., Oyola,M.G., Zhou,S., Chen,X., Liao,L., Tien,J.C.-Y., Mani,S.K. and Xu,J. (2015) Knockout of the Histone Demethylase Kdm3b Decreases Spermatogenesis and Impairs Male Sexual Behaviors. *Int J Biol Sci*, 11, 1447–57.

27. Liu,Z., Chen,X., Zhou,S., Liao,L., Jiang,R. and Xu,J. (2015) The Histone H3K9 Demethylase Kdm3b Is Required for Somatic Growth and Female Reproductive Function. *Int J Biol Sci*, 11, 494–507.

28. Takahashi,K. and Yamanaka,S. (2006) Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, 126, 663–676.

29. Tran,K.A., Jackson,S.A., Olufs,Z.P.G., Zaidan,N.Z., Leng,N., Kendziorski,C., Roy,S. and Sridharan,R. (2015) Collaborative rewiring of the pluripotency network by chromatin and signalling modulating pathways. *Nat Commun*, 6, 6188.

30. Chen,J., Liu,H., Liu,J., Qi,J., Wei,B., Yang,J., Liang,H., Chen,Y., Chen,J., Wu,Y., *et al.* (2013) H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat Genet*, 45, 34–42.

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

32. Sakuma,T., Nakade,S., Sakane,Y., Suzuki,K.-I.T. and Yamamoto,T. (2016) MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat Protoc*, 11, 118–133.

33. Beard,C., Hochedlinger,K., Plath,K., Wutz,A. and Jaenisch,R. (2006) Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis*, 44, 23–28.

34. Tokheim,C., Park,J.W. and Xing,Y. (2014) PrimerSeq: Design and Visualization of RT-PCR Primers for Alternative Splicing Using RNA-seq Data. *Genom Proteom Bioinform*, 12, 105–109.

35. Tran,K.A., Dillingham,C.M. and Sridharan,R. (2019) Coordinated removal of repressive epigenetic modifications during induced reversal of cell identity. *Embo J*, 38, e101681.

36. Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res*, 11, 1475–1489.

37. Zaidan,N.Z., Walker,K.J., Brown,J.E., Schaffer,L.V., Scalf,M., Shortreed,M.R., Iyer,G., Smith,L.M. and Sridharan,R. (2018) Compartmentalization of HP1 Proteins in Pluripotency Acquisition and Maintenance. *Stem Cell Rep*, 10, 627–641.

38. Wiśniewski,J.R., Zougman,A., Nagaraj,N. and Mann,M. (2009) Universal sample preparation method for proteome analysis. *Nat Methods*, 6, 359–362.

39. Shannon,P., Markiel,A., Ozier,O., Baliga,N.S., Wang,J.T., Ramage,D., Amin,N., Schwikowski,B. and Ideker,T. (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res*, 13, 2498–2504.

40. Zaidan,N.Z. and Sridharan,R. (2020) HP1γ regulates H3K36 methylation and pluripotency in embryonic stem cells. *Nucleic Acids Res*, 48, gkaa1091-.

41. Wille,C.K. and Sridharan,R. (2022) DOT1L inhibition enhances pluripotency beyond acquisition of epithelial identity and without immediate suppression of the somatic transcriptome. *Stem Cell Rep*, 17, 384–396.

42. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120.

43. Li,B. and Dewey,C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *Bmc Bioinformatics*, 12, 323.

44. Leng,N., Dawson,J.A., Thomson,J.A., Ruotti,V., Rissman,A.I., Smits,B.M.G., Haag,J.D., Gould,M.N., Stewart,R.M. and Kendziorski,C. (2013) EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics*, 29, 1035–1043.

45. Huang,D.W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, 4, 44–57.

46. Sherman,B.T., Hao,M., Qiu,J., Jiao,X., Baseler,M.W., Lane,H.C., Imamichi,T. and Chang,W. (2022) DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res*, 50, W216–W221.

47. Dobin,A. and Gingeras,T.R. (2015) Mapping RNA-seq Reads with STAR. *Curr Protoc Bioinform*, 51, 11.14.1-11.14.19.

48. Shen,S., Park,J.W., Lu,Z., Lin,L., Henry,M.D., Wu,Y.N., Zhou,Q. and Xing,Y. (2014) rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc National Acad Sci*, 111, E5593–E5601.

49. Shen,S., Park,J.W., Huang,J., Dittmar,K.A., Lu,Z., Zhou,Q., Carstens,R.P. and Xing,Y. (2012) MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data. *Nucleic Acids Res*, 40, e61–e61.

50. Park,J.W., Tokheim,C., Shen,S. and Xing,Y. (2013) Deep Sequencing Data Analysis. *Methods Mol Biology*, 1038, 171–179.

51. Boutz,P.L., Bhutkar,A. and Sharp,P.A. (2015) Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Gene Dev*, 29, 63–80.

52. Anders,S., Reyes,A. and Huber,W. (2012) Detecting differential usage of exons from RNA-seq data. *Genome Res*, 22, 2008–2017.

53. Wille,C.K., Zhang,X., Haws,S.A., Denu,J.M. and Sridharan,R. (2022) Antagonistic H3K79me-H3K9ac crosstalk determines elongation at housekeeping genes to promote pluripotency. *Biorxiv*, 10.1101/2022.09.26.509534.

54. Li,Y., Xia,L., Tan,K., Ye,X., Zuo,Z., Li,M., Xiao,R., Wang,Z., Liu,X., Deng,M., *et al.* (2020) N6-Methyladenosine co-transcriptionally directs the demethylation of histone H3K9me2. *Nat Genet*, 10.1038/s41588-020-0677-3.

55. Yelagandula,R., Stecher,K., Novatchkova,M., Michetti,L., Michlits,G., Wang,J., Hofbauer,P., Vainorius,G., Pribitzer,C., Isbel,L., *et al.* (2023) ZFP462 safeguards neural lineage specification by targeting G9A/GLP-mediated heterochromatin to silence enhancers. *Nat Cell Biol*, 25, 42–55.

56. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods*, 9, 357–359.

57. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and Subgroup,1000 Genome Project Data Processing (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25, 2078–2079.

58. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W., *et al.* (2008) Model-based Analysis of ChIP-Seq (MACS). *Genome Biol*, 9, R137.

59. Robinson,J.T., Thorvaldsdóttir,H., Winckler,W., Guttman,M., Lander,E.S., Getz,G. and Mesirov,J.P. (2011) Integrative genomics viewer. *Nat Biotechnol*, 29, 24–26.

60. 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–W191.

61. Mellacheruvu,D., Wright,Z., Couzens,A.L., Lambert,J.-P., St-Denis,N.A., Li,T., Miteva,Y.V., Hauri,S., Sardiu,M.E., Low,T.Y., *et al.* (2013) The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. *Nat Methods*, 10, 730–736.

62. Owttrim,G.W. (2013) RNA helicases. *Rna Biol*, 10, 96–110.

63. Wang,Z., Ballut,L., Barbosa,I. and Hir,H.L. (2018) Exon Junction Complexes can have distinct functional flavours to regulate specific splicing events. *Sci Rep-uk*, 8, 9509.

64. Ameur,A., Zaghlool,A., Halvardson,J., Wetterbom,A., Gyllensten,U., Cavelier,L. and Feuk,L. (2011) Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat Struct Mol Biol*, 18, 1435–1440.

65. Tilgner,H., Knowles,D.G., Johnson,R., Davis,C.A., Chakrabortty,S., Djebali,S., Curado,J., Snyder,M., Gingeras,T.R. and Guigó,R. (2012) Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res*, 22, 1616–1625.

66. Neugebauer,K.M. (2019) Nascent RNA and the Coordination of Splicing with Transcription. *Csh Perspect Biol*, 11, a032227.

67. Brugiolo,M., Herzel,L. and Neugebauer,K.M. (2013) Counting on co-transcriptional splicing. *F1000prime Reports*, 5, 9.

68. Girard,C., Will,C.L., Peng,J., Makarov,E.M., Kastner,B., Lemm,I., Urlaub,H., Hartmuth,K. and Lührmann,R. (2012) Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nat Commun*, 3, 994.

69. Saavedra,F., Gurard-Levin,Z.A., Rojas-Villalobos,C., Vassias,I., Quatrini,R., Almouzni,G. and Loyola,A. (2020) JMJD1B, a novel player in histone H3 and H4 processing to ensure genome stability. *Epigenet Chromatin*, 13, 6.

70. Muller,I.B., Meijers,S., Kampstra,P., Dijk,S. van, Elswijk,M. van, Lin,M., Wojtuszkiewicz,A.M., Jansen,G., Jonge,R. de and Cloos,J. (2021) Computational comparison of common event-based differential splicing tools: practical considerations for laboratory researchers. *Bmc Bioinformatics*, 22, 347.

71. Braun,C.J., Stanciu,M., Boutz,P.L., Patterson,J.C., Calligaris,D., Higuchi,F., Neupane,R., Fenoglio,S., Cahill,D.P., Wakimoto,H., *et al.* (2017) Coordinated Splicing of Regulatory Detained Introns within Oncogenic Transcripts Creates an Exploitable Vulnerability in Malignant Glioma. *Cancer Cell*, 32, 411-426.e11.

72. Li,Z., Yu,J., Hosohama,L., Nee,K., Gkountela,S., Chaudhari,S., Cass,A.A., Xiao,X. and Clark,A.T. (2015) The Sm protein methyltransferase PRMT5 is not required for primordial germ cell specification in mice. *Embo J*, 34, 748–758.

73. Pan,Q., Shai,O., Misquitta,C., Zhang,W., Saltzman,A.L., Mohammad,N., Babak,T., Siu,H., Hughes,T.R., Morris,Q.D., *et al.* (2004) Revealing Global Regulatory Features of Mammalian Alternative Splicing Using a Quantitative Microarray Platform. *Mol Cell*, 16, 929–941.

74. Yang,P., Humphrey,S.J., Cinghu,S., Pathania,R., Oldfield,A.J., Kumar,D., Perera,D., Yang,J.Y.H., James,D.E., Mann,M., *et al.* (2019) Multi-omic Profiling Reveals Dynamics of the Phased Progression of Pluripotency. *Cell Syst*, 8, 427-445.e10.

75. Kinoshita,M., Barber,M., Mansfield,W., Cui,Y., Spindlow,D., Stirparo,G.G., Dietmann,S., Nichols,J. and Smith,A. (2021) Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency. *Cell Stem Cell*, 28, 453-471.e8.

76. Aoki,A., Suetake,I., Miyagawa,J., Fujio,T., Chijiwa,T., Sasaki,H. and Tajima,S. (2001) Enzymatic properties of de novo-type mouse DNA (cytosine-5) methyltransferases. *Nucleic Acids Res*, 29, 3506–3512.

77. Weisenberger,D.J., Velicescu,M., Cheng,J.C., Gonzales,F.A., Liang,G. and Jones,P.A. (2004) Role of the DNA Methyltransferase Variant DNMT3b3 in DNA Methylation. *Mol Cancer Res*, 2, 62–72.

78. Chen,T., Ueda,Y., Dodge,J.E., Wang,Z. and Li,E. (2003) Establishment and Maintenance of Genomic Methylation Patterns in Mouse Embryonic Stem Cells by Dnmt3a and Dnmt3b. *Mol Cell Biol*, 23, 5594–5605.

79. Zhao,D., Wu,Y. and Chen,K. (2014) Tbx3 isoforms are involved in pluripotency maintaining through distinct regulation of Nanog transcriptional activity. *Biochem Bioph Res Co*, 444, 411–414.

80. Wang,D., Claus,C.L., Vaccarelli,G., Braunstein,M., Schmitt,T.M., Zúñiga-Pflücker,J.C., Rothenberg,E.V. and Anderson,M.K. (2006) The Basic Helix-Loop-Helix Transcription Factor HEBAlt Is Expressed in Pro-T Cells and Enhances the Generation of T Cell Precursors. *J Immunol*, 177, 109–119.

81. Hough,S.R., Thornton,M., Mason,E., Mar,J.C., Wells,C.A. and Pera,M.F. (2014) Single-Cell Gene Expression Profiles Define Self-Renewing, Pluripotent, and Lineage Primed States of Human Pluripotent Stem Cells. *Stem Cell Rep*, 2, 881–895.

82. Marks,H., Kalkan,T., Menafra,R., Denissov,S., Jones,K., Hofemeister,H., Nichols,J., Kranz,A., Stewart,A.F., Smith,A., *et al.* (2012) The Transcriptional and Epigenomic Foundations of Ground State Pluripotency. *Cell*, 149, 590–604.

83. Tran,K.A., Dillingham,C.M. and Sridharan,R. (2019) The role of α-ketoglutarate–dependent proteins in pluripotency acquisition and maintenance. *J Biol Chem*, 294, 5408–5419.

84. Ramadoss,S., Guo,G. and Wang,C.-Y. (2016) Lysine demethylase KDM3A regulates breast cancer cell invasion and apoptosis by targeting histone and the non-histone protein p53. *Oncogene*, 36, 47–59.

85. Majello,B., Gorini,F., Saccà,C.D. and Amente,S. (2019) Expanding the Role of the Histone Lysine-Specific Demethylase LSD1 in Cancer. *Cancers*, 11, 324.

86. Kalsotra,A. and Cooper,T.A. (2011) Functional consequences of developmentally regulated alternative splicing. *Nat Rev Genet*, 12, 715–729.

87. Salomonis,N., Schlieve,C.R., Pereira,L., Wahlquist,C., Colas,A., Zambon,A.C., Vranizan,K., Spindler,M.J., Pico,A.R., Cline,M.S., *et al.* (2010) Alternative splicing regulates mouse embryonic stem cell pluripotency and differentiation. *Proc National Acad Sci*, 107, 10514–10519.

88. Das,S., Jena,S. and Levasseur,D.N. (2011) Alternative Splicing Produces Nanog Protein Variants with Different Capacities for Self-renewal and Pluripotency in Embryonic Stem Cells*. *J Biol Chem*, 286, 42690–42703.

89. Wang,X. and Dai,J. (2010) Concise Review: Isoforms of OCT4 Contribute to the Confusing Diversity in Stem Cell Biology. *Stem Cells*, 28, 885–893.

90. Ohta,S., Nishida,E., Yamanaka,S. and Yamamoto,T. (2013) Global Splicing Pattern Reversion during Somatic Cell Reprogramming. *Cell Reports*, 5, 357–366.

91. Gabut,M., Samavarchi-Tehrani,P., Wang,X., Slobodeniuc,V., O'Hanlon,D., Sung,H.-K., Alvarez,M., Talukder,S., Pan,Q., Mazzoni,E.O., *et al.* (2011) An Alternative Splicing Switch Regulates Embryonic Stem Cell Pluripotency and Reprogramming. *Cell*, 147, 132–146.

92. Yeo,G.W., Coufal,N.G., Liang,T.Y., Peng,G.E., Fu,X.-D. and Gage,F.H. (2009) An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. *Nat Struct Mol Biol*, 16, 130–137.

93. Esmailpour,T. and Huang,T. (2012) TBX3 Promotes Human Embryonic Stem Cell Proliferation and Neuroepithelial Differentiation in a Differentiation Stage-dependent Manner. *Stem Cells*, 30, 2152–2163.

94. Wu,B., Li,Y., Li,B., Zhang,B., Wang,Y., Li,L., Gao,J., Fu,Y., Li,S., Chen,C., *et al.* (2021) DNMTs Play an Important Role in Maintaining the Pluripotency of Leukemia Inhibitory Factor-Dependent Embryonic Stem Cells. *Stem Cell Rep*, 16, 582–596.

95. Li,Y., Brauer,P.M., Singh,J., Xhiku,S., Yoganathan,K., Zúñiga-Pflücker,J.C. and Anderson,M.K. (2017) Targeted Disruption of TCF12 Reveals HEB as Essential in Human Mesodermal Specification and Hematopoiesis. *Stem Cell Rep*, 9, 779–795.

96. Leitch,H.G., McEwen,K.R., Turp,A., Encheva,V., Carroll,T., Grabole,N., Mansfield,W., Nashun,B., Knezovich,J.G., Smith,A., *et al.* (2013) Naive pluripotency is associated with global DNA hypomethylation. *Nat Struct Mol Biol*, 20, 311–316.

97. Hackett,J.A., Sengupta,R., Zylicz,J.J., Murakami,K., Lee,C., Down,T.A. and Surani,M.A. (2013) Germline DNA Demethylation Dynamics and Imprint Erasure Through 5-Hydroxymethylcytosine. *Science*, 339, 448–452.

98. Russell,R., Ilg,M., Lin,Q., Wu,G., Lechel,A., Bergmann,W., Eiseler,T., Linta,L., Kumar P.,P., Klingenstein,M., *et al.* (2015) A Dynamic Role of TBX3 in the Pluripotency Circuitry. *Stem Cell Rep*, 5, 1155–1170.

99. Li,S. and Mason,C.E. (2014) The Pivotal Regulatory Landscape of RNA Modifications. *Annu Rev Genom Hum G*, 15, 1–23.

100. 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 m6A Reader YTHDC1 Regulates mRNA Splicing. *Mol Cell*, 61, 507–519.

101. 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.

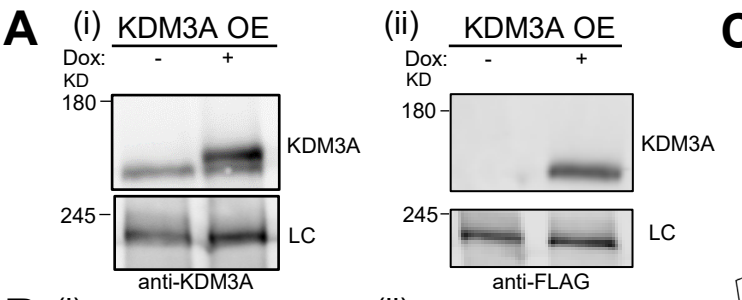
102. Geula,S., Moshitch-Moshkovitz,S., Dominissini,D., Mansour,A.A., Kol,N., Salmon-Divon,M., Hershkovitz,V., Peer,E., Mor,N., Manor,Y.S., *et al.* (2015) m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science*, 347, 1002–1006.

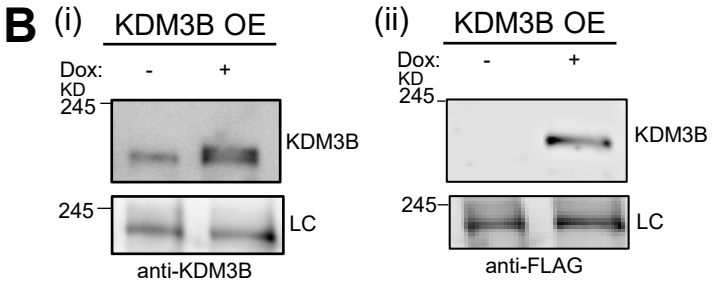
103. Fan,L., Zhang,F., Xu,S., Cui,X., Hussain,A., Fazli,L., Gleave,M., Dong,X. and Qi,J. (2018) Histone demethylase JMJD1A promotes alternative splicing of AR variant 7 (AR-V7) in prostate cancer cells. *Proc National Acad Sci*, 115, 201802415.

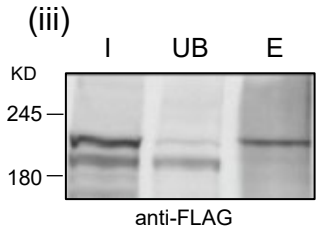
104. Baker,M., Petasny,M., Taqatqa,N., Bentata,M., Kay,G., Engal,E., Nevo,Y., Siam,A., Preiser,S. and Salton,M. (2021) KDM3A regulates alternative splicing of cell-cycle genes following DNA damage. *Rna*, 10.1261/rna.078796.121.

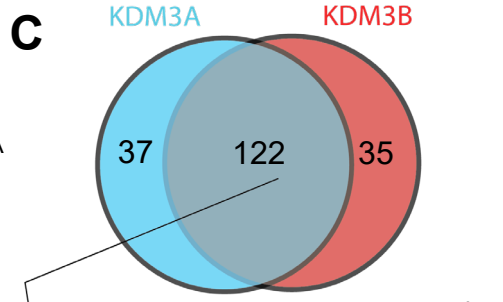
105. Singh,G., Pratt,G., Yeo,G.W. and Moore,M.J. (2015) The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. *Annu Rev Biochem*, 84, 325–54.

Figure 1:

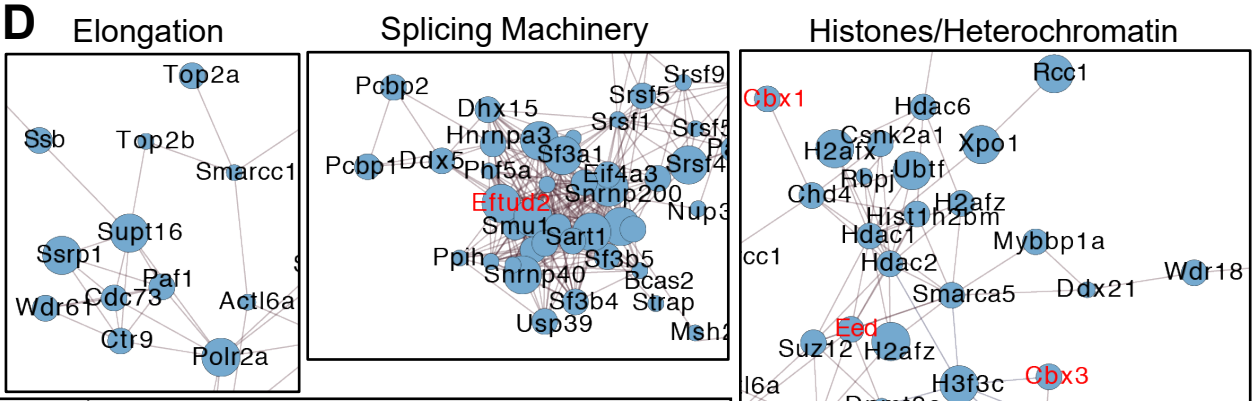








Gene Ontology	p-value
mRNA Processing	2.3e-44
mRNA splicing via spliceosome	3.5e-38
Positive Regulation of Transcription	8.8e-6
Transcription Corepressor Binding	1.3e-5



E

mRNA Interactome Proteins	KDM3A/KDM3B Interactors
EJC Interacting Proteins (18)	ACIN1, PNN, SRRM1, SRRM2, TRA2A, TRA2B
SR Proteins (11)	SSRF1, SSRF2, SSRF4, SSRF5, SSRF6, SSRF9
hnRNP Proteins (30)	HNRNPA3, PCBP1, PCBP2, TARDBP
mRNA processing proteins (23)	EFTUD2, PABPC1, PRPF8, SNRNP40, SNRNPD2, SF3B4

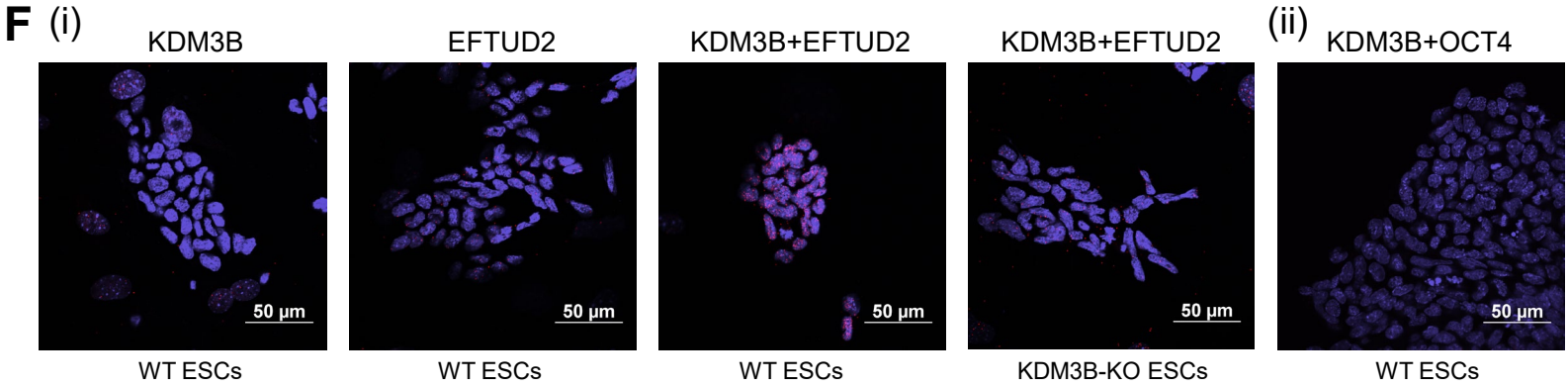


Figure 1: KDM3A and KDM3B interact with splicing machinery in ESCs

(A) Immunoblot of (i) KDM3A, (ii) FLAG in ESCs with doxycycline (dox)-inducible expression of FLAG-KDM3A. LC= loading control = RNAPII. OE=Overexpression
(B) Same as above except in dox-inducible FLAG-KDM3B ESCs (i) KDM3B (ii) FLAG (iii) FLAG following immunoprecipitation with FLAG antibody. I=Input, UB=Unbound, E=Elute
(C) Venn diagram of KDM3A and KDM3B interactome. Filtered for nuclear compartment using DAVID GO (https://david.ncifcrf.gov/) analysis and frequent FLAG pulldown contaminants using "Crapome" (https://reprint-apms.org/) repository at SAINT score of 0.7 (61). Gene Ontology (GO) enrichment of shared interactome are listed at enrichment score>2.5 and p-value<.00001.
(D) String network analysis (https://string-db.org/) of KDM3A and KDM3B shared interactome with high confidence interaction scores (0.7 minimum) from database and experimental sources. Node size represents fold change (FC) of KDM3A over uninduced spectral counts. FC>10 = large node, 10>FC>5 = medium node, FC<5 = small node.
(E) KDM3A and KDM3B interactors overlapped with functional categories of mRNP components from Singh et al. (105) or categorized as other mRNA processing proteins. Shared= black, KDM3A unique=blue, and KDM3B unique=red.
(F) Single slice of 1 uM confocal images of *in situ* Duolink™ Proximity ligation assay (PLA) of (i) KDM3B single antibody, EFTUD2 single antibody, and KDM3B+EFTUD2 antibodies in WT ESCs or KDM3B+EFTUD2 antibodies in KDM3B KO ESCs. (ii) PLA of KDM3B and OCT4 in WT ESCs. Scale bar = 50μM with 60x objective.

Figure 2:

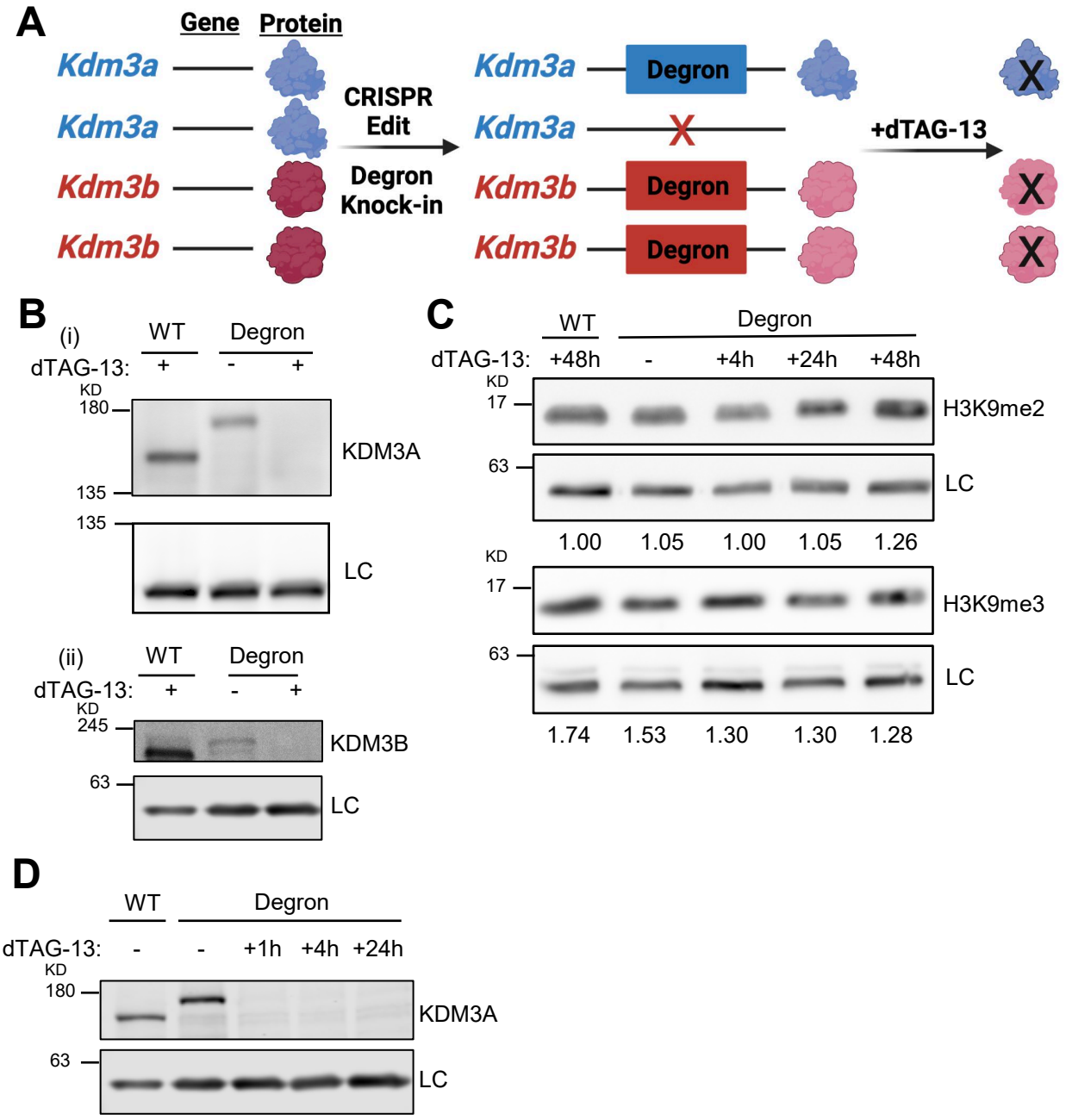

A
Gene
Protein
Kdm3a
Kdm3a
Kdm3b
Kdm3b
CRISPR Edit
Degron Knock-in
Degron
+dTAG-13
B
(i)
WT
Degron
dTAG-13:
+
-
+
KD
180
135
KDM3A
135
LC
(ii)
WT
Degron
dTAG-13:
+
-
+
KD
245
KDM3B
63
LC
C
WT
Degron
dTAG-13:
+48h
-
+4h
+24h
+48h
KD
17
H3K9me2
63
LC
1.00
1.05
1.00
1.05
1.26
KD
17
H3K9me3
63
LC
1.74
1.53
1.30
1.30
1.28
D
WT
Degron
dTAG-13:
-
-
+1h
+4h
+24h
KD
180
KDM3A
63
LC

Figure 2: Inducible double knockout of KDM3A and KDM3B results in delayed global accumulation of H3K9me2

(A) Schematic of gene editing strategy to generate degron inducible KDM3A and KDM3B proteins.

(B) Immunoblot for (i) KDM3A or (ii) KDM3B in WT ESCs or KDM3A/KDM3B degron ESCs. Note increased size of degron tagged protein. dTAG-13 ligand induced samples (+) treated for 24 hours. (i) Loading control= LC = EFTUD and (ii) Loading control= LC = TUBULIN

(C) Immunoblot of H3K9me2, H3K9me3, and TUBULIN (LC) at indicated timepoints of degradation. Ratio of H3K9me2 or H3K9me3 over TUBULIN, quantified in Licor software.

(D) Immunoblot for KDM3A, KDM3B, and TUBULIN (LC) in 24 hours timecourse of dTAG-13 induced degradation.

Figure 3:

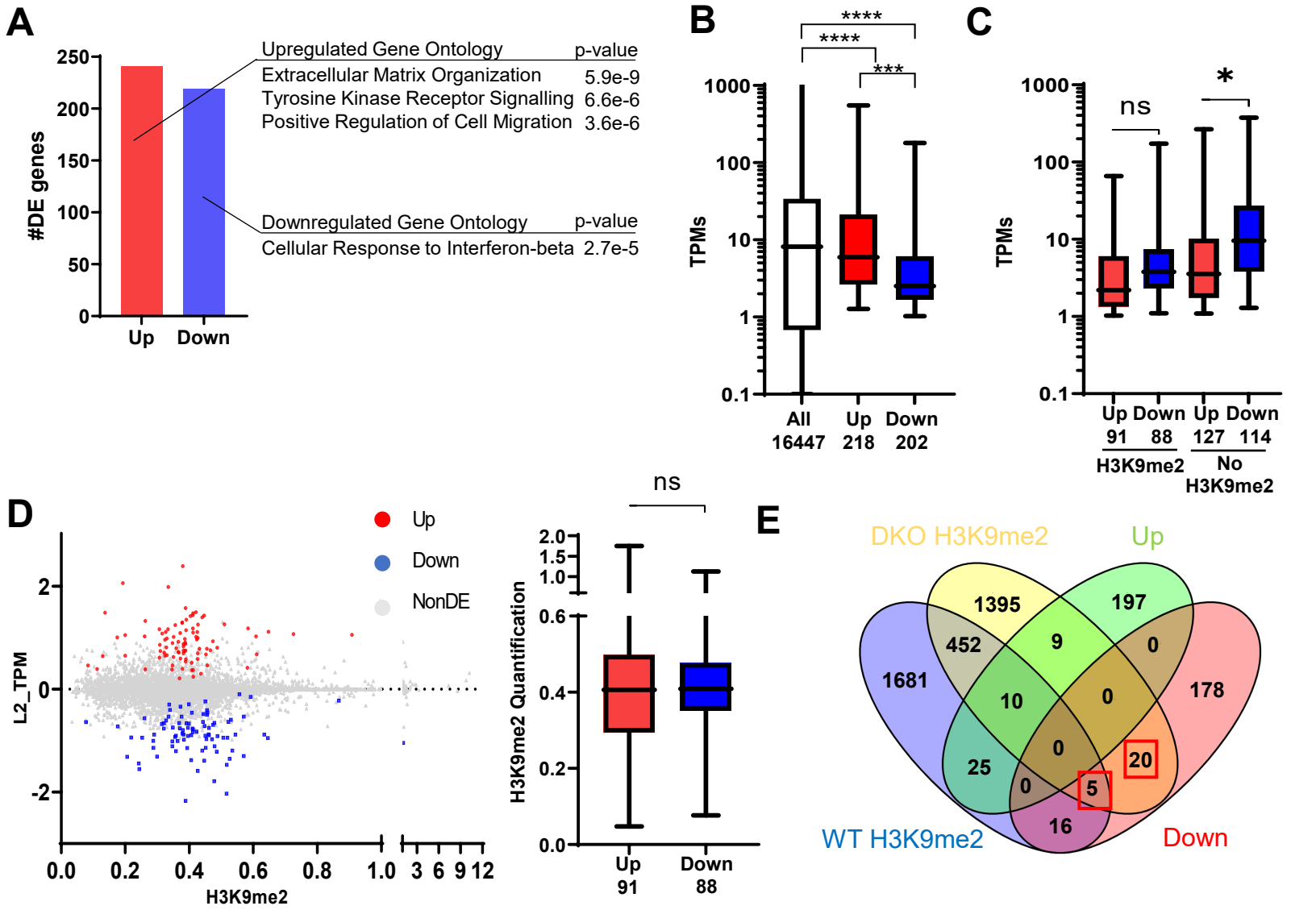

A
#DE genes
250
200
150
100
50
0
Up
Down
Upregulated Gene Ontology p-value
Extracellular Matrix Organization 5.9e-9
Tyrosine Kinase Receptor Signalling 6.6e-6
Positive Regulation of Cell Migration 3.6e-6
Downregulated Gene Ontology p-value
Cellular Response to Interferon-beta 2.7e-5
B
TPMs
1000
100
10
1
0.1
****
****
***
All 16447
Up 218
Down 202
C
TPMs
ns
*
Up 91
Down 88
Up 127
Down 114
H3K9me2
No H3K9me2
D
Up
Down
NonDE
L2_TPM
H3K9me2
0.0
0.2
0.4
0.6
0.8
1.0
3
6
9
12
ns
H3K9me2 Quantification
2.0
1.5
1.0
0.6
0.4
0.2
0.0
Up 91
Down 88
E
DKO H3K9me2
Up
WT H3K9me2
Down
1395
197
452
9
0
1681
10
0
178
0
25
20
0
5
16

Figure 3: Acute degradation of KDM3A and KDM3B leads to differential gene expression agnostic of H3K9me2 status

(A) Number of differentially expressed genes (DEG) between Double Degron (DDeg+) exposed to dTAG-13 ESCs and wildtype ESCs. Gene ontology categories for DEGs shown at minimum p-value < 0.00001.

(B) Boxplot of absolute expression (TPMs) in ESCs for all annotated genes from RSEM quantification with a non-zero median TPM (All expressed genes), TPMs from up or down DEGs. Numbers indicate number of genes per condition. ****p-value < 0.0001 assessed by unpaired Welch's t-test.

(C) Boxplot of average TPMs+1 from DEGs with (H3K9me2) or without (No H3K9me2) a called consensus H3K9me2 peak (54, 55). *p-value < 0.05 assessed by unpaired Welch's t-test

(D) Left: Scatterplot of log2(DDeg+/WT) fold change of TPMs of Non-DE and DEG against H3K9me2 enrichment (normalized read counts per million and by 1kb per gene) at those genes. Right: Boxplot of total H3K9me2 enrichment at up or down DEG.

(E) Four-way Venn overlap of DEG and H3K9me2 peaks in WT ESCs and KDM3A and KDM3B DKO ESCs (22).

Figure 4:Alternative splicing defects upon acute loss of KDM3A/KDM3B occur in signaling and chromatin related genes

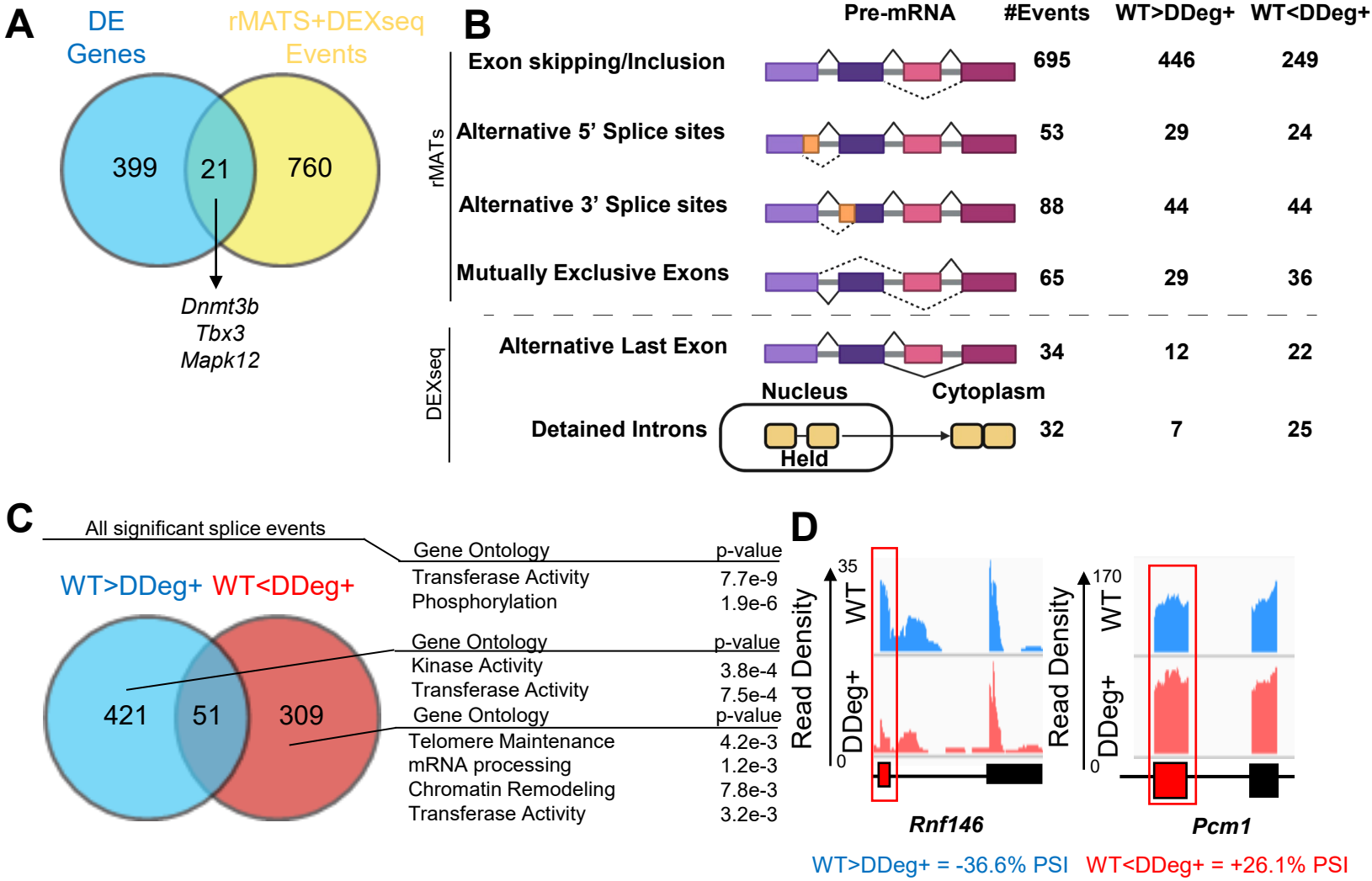


Figure 4: Alternative splicing defects upon acute loss of KDM3A/KDM3B occur in signaling and chromatin related genes

(A) Venn overlap of DE Genes (blue) and aberrantly spliced genes (yellow) called from rMATs and Dexseq analysis.

(B) Left panel: schematic of alternative splicing pathways, results above the dotted line are from rMATs and below are from Dexseq. Right panel: Number of significant events called in each class of alternative splicing pathway, FDR<0.1, inclusion level difference x>+/-0.5, and 15 read minimum.

(C) Venn overlap of aberrantly spliced genes sorted by higher inclusion in WT (blue) or in DDeg+ (red). Gene ontology categories for aberrantly spliced genes shown, minimum p-value < 0.001. IGV tracks for example genes of exons with higher percent spliced in (PSI) for WT or DKO conditions. PSI value calculated in rMATs from 3 replicates in each condition. Negative value indicates event is included more in the WT condition and vice versa. The affected exon is marked in red.

(D) (i) Bar graph shows the number of genes within each category, binned by gene length normalized by intron number from longest to shortest. Categories include genes with significant spliced events sorted by negative inclusion (blue) or positive inclusion (red) and plotted on the left y-axis, and the distribution of any gene detected (black) by rMATs is included as comparison on the right y-axis. (ii) Same as above except splice events are binned by decreasing expression (TPMs).

(E) Clustering (k-means=5) of differentially spliced events by inclusion level difference score generated in rMATs. WT vs uninduced degron (DDeg-) and WT vs induced degron (DDeg+) lines are shown. A positive number indicates greater inclusion in WT cells. Example genes from clusters and significant gene ontology terms (minimum p-value < 0.001) are in right panel.

Figure 5: Loss of KDM3A and KDM3B leads to differences in stable naïve state.

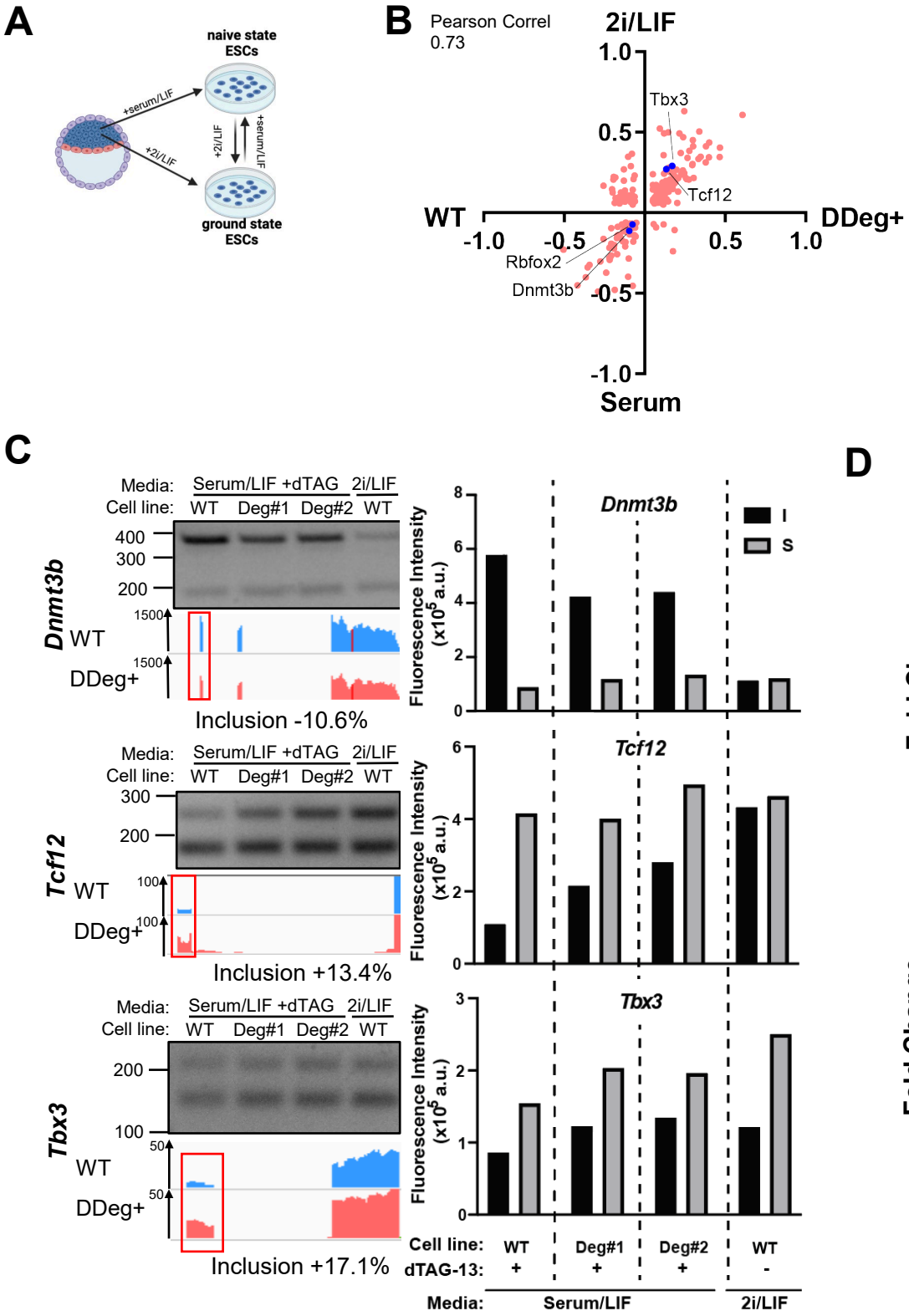


Figure 5: Loss of KDM3A and KDM3B leads to differences in stable naïve state.

(A) Schematic detailing derivation of naïve ESCs (serum/LIF) or ground ESCs (2i/LIF).

(B) Scatterplot of splicing events called by rMATs in both WT vs DDeg+ ESCs as well as serum vs 2i/LIF ESCs (74) measured by inclusion level difference score. Pearson correlation between the inclusion level difference of specific events called in both comparisons. Blue dots represent events chosen for further confirmation.

(C) Semi-quantitative PCR of splicing targets in WT ESCs, DDeg+ clones at 4 hours of degradation, and 2i/LIF ESCs. Representative IGV tracks depicting affected exon (red box) are below. Electrophoresis image was quantified using Licor software, fluorescence intensity reported as arbitrary units (a.u.) and represented in right panel. I=included isoform and S=Skipped isoform.

(D) Relative expression of c-Myc or Dazl in WT or DDeg+ ESCs in serum/LIF or 2i/LIF conditions. WT cells in serum/LIF set to 1.