12462–12479 *Nucleic Acids Research, 2022, Vol. 50, No. 21* *Published online 23 November 2022*
https://doi.org/10.1093/nar/gkac1084

An RNAi screen of RNA helicases identifies eIF4A3 as a regulator of embryonic stem cell identity

Dan Li[1,2,3], Jihong Yang[1], Vikas Malik[1], Yuting Huang[1], Xin Huang[1], Hongwei Zhou[1] and Jianlong Wang[1,*]

[1]Department of Medicine, Columbia Center for Human Development and Stem Cell Therapies, Columbia Stem Cell Initiative, Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, New York, NY 10032, USA, [2]Department of Cell, Developmental and Regenerative Biology; The Black Family Stem Cell Institute; Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA and [3]The Graduate School of Biomedical Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

Received January 31, 2022; Revised October 21, 2022; Editorial Decision October 24, 2022; Accepted October 26, 2022

ABSTRACT

RNA helicases are involved in multiple steps of RNA metabolism to direct their roles in gene expression, yet their functions in pluripotency control remain largely unexplored. Starting from an RNA interference (RNAi) screen of RNA helicases, we identified that eIF4A3, a DEAD-box (Ddx) helicase component of the exon junction complex (EJC), is essential for the maintenance of embryonic stem cells (ESCs). Mechanistically, we show that eIF4A3 post-transcriptionally controls the pluripotency-related cell cycle regulators and that its depletion causes the loss of pluripotency via cell cycle dysregulation. Specifically, eIF4A3 is required for the efficient nuclear export of *Ccnb1* mRNA, which encodes Cyclin B1, a key component of the pluripotency-promoting pathway during the cell cycle progression of ESCs. Our results reveal a previously unappreciated role for eIF4A3 and its associated EJC in maintaining stem cell pluripotency through post-transcriptional control of the cell cycle.

INTRODUCTION

Embryonic stem cells (ESCs) can self-renew and differentiate into multiple cell lineages, the unique feature collectively known as pluripotency (1). Extensive studies have been devoted to investigating the molecular regulatory mechanisms underlying the pluripotency of ESCs, which led to the discovery of transcription factors (TFs), epigenetic cofactors and signal pathways dedicated to pluripotency (2). In addition, post-transcriptional and translational regulatory layers, although relatively understudied, are also being increasingly scrutinized for their roles in controlling pluripotency (3–6).

DEAD-box (Ddx) and DEAH-box (Dhx) proteins form two major RNA helicase families, named by their characteristic conserved Asp–Glu–Ala–Asp (DEAD)/Asp–Glu–Ala–His (DEAH) motifs. They participate in all aspects of RNA metabolism, from ribosome biogenesis to transcription, RNA maturation, microRNA (miRNA) processing, mRNA export, translation and RNA degradation, with increasingly recognized roles in stem cells and development (3,7).

The cell cycle machinery in pluripotent ESCs has a unique organization with rapid cell division and a truncated G_1 (8,9), controlled by special expression and activity patterns of key cell cycle regulators, including cyclins and CDKs (cyclin-dependent kinases). This unique cell cycle organization changes upon the dissolution of pluripotency and during early embryonic development (8,9). Recent studies have unraveled the molecular links between cell cycle and pluripotency, and established that the unusual cell cycle properties of ESCs actively regulate pluripotency (8,10). During the cell cycle, the regulatory networks impacting the differentiation of ESCs could be activated in the G_1 phase (8,9). Therefore, the G_1 phase represents a window of opportunity for ESCs to exit the self-renewing pluripotent state. In contrast, during the G_2 phase, Cyclin B1 represents a key component of the pluripotency-promoting pathway that activates the expression of many pluripotency factors (11), which in turn influence the cell cycle via the regulation of cell cycle gene expression (8). While the functional contribution of RNA helicases to cell cycle control is well appreciated in various somatic and cancer cells (12,13), how they may contribute to the unique cell cycle organization in pluripotent stem cells (PSCs) remains to be addressed.

To identify Ddx/Dhx factors critical in ESCs and dissect their mechanistic actions on stem cell pluripotency, we performed an RNA interference (RNAi) screen of all Ddx/Dhx factors for their roles in the maintenance of mouse ESCs. Using short hairpin RNAs (shRNAs) target-

*To whom correspondence should be addressed. Tel: +1 212 342 3180; Email: jw3925@cumc.columbia.edu

ing all Ddx/Dhx members, we identified several Ddx/Dhx factors essential for ESC maintenance, among which is Ddx48 (aka eIF4A3), a core component of the exon junction complex (EJC) (14). The EJC is assembled and binds to mRNAs during splicing, and it regulates multiple steps of gene expression encompassing splicing, mRNA nuclear export, translation and nonsense-mediated mRNA decay (NMD) (14). Our study functionally connects eIF4A3 with the pluripotency-specific cell cycle regulators, particularly Cyclin B1, the only cyclin that displays cell cycle-dependent fluctuations in ESCs and promotes pluripotency during ESC cell cycle progression (8). Mechanistically, we found that eIF4A3 directly binds to *Ccnb1* mRNA encoding Cyclin B1 and transports it to the cytoplasm, ensuring its protein synthesis to safeguard ESC identity. The RNAi screen of RNA helicases thus identifies eIF4A3 as a previously unappreciated regulator of ESC identity linking EJC to the pluripotency network.

MATERIALS AND METHODS

Embryonic stem cell culture

Mouse ESCs J1 and CCE were grown in culture plates coated with 0.1% gelatin in an ESC medium containing high-glucose Dulbecco's modified Eagle's medium (DMEM), 15% fetal bovine serum (FBS), 100 μM non-essential amino acids (NEAAs), 1% nucleoside mix, 2 mM L-glutamine, 50 U/ml penicillin/streptomycin, 0.1 mM 2-mercaptoethanol and homemade recombinant leukemia inhibitory factor (LIF) tested for efficient ESC self-renewal, at 37°C and 5% CO_2. The medium was replaced daily, and the cells were passaged every 2–3 days. The mycoplasma test was routinely performed in the lab.

Transfection and lentiviral infection

The cell transfection was performed using Lipofectamine 3000 according to the manufacturer's manual. The lentivirus production and viral infection were performed as described (15). All the knockdown experiments using shRNAs were performed at the same time points as the RNAi screen, which ended on post-infection Day 4.5 with the drug selection (1 μg/ml puromycin) for 2.5 days.

Ddx/Dhx RNA helicase RNAi screen

To perform the RNAi screen of Ddx/Dhx RNA helicases, we selected the constitutive shRNAs (three independent shRNAs for each Ddx/Dhx family member) with the validated knockdown (KD) efficiencies based on the RNAi Consortium (TRC) library database (https://www.broadinstitute.org/rnai-consortium/rnai-consortium-shrna-library), followed by experimental validation in ESCs for the screen hits. The lentiviruses were prepared as previously described (15). On Day 0, ESCs were seeded on the gelatin-coated plates for viral infection as described (15). On Day 1, the virus/medium was changed with the fresh serum/LIF ESC medium for the second viral infection as described (15). From Day 2, the medium was changed daily with the fresh serum/LIF ESC medium containing puromycin (1 μg/ml) to select the infected cells. The uninfected cells, the negative control for drug selection, would all be killed by the drug treatment after 2 days. On Day 4.5, alkaline phosphatase (AP) staining was performed to record the phenotype following the manufacturer's manual (Sigma, 86R-1KT).

eIF4A3 cDNA rescue assay

The rescue cDNA is human eIF4A3. Peptide sequences of eIF4A3 in humans and mice are highly conserved: 408/411 identities. Our results (Figures 2H–J and 6D–F) showed that wild-type (WT) human eIF4A3 can rescue the eIF4A3 KD-induced phenotype in mouse ESCs.

Whole-cell extract preparation, co-immunoprecipitation and western blotting analysis

Whole-cell lysates were prepared from ESCs in Lysis Buffer [50 mM HEPES (pH 7.6), 250 mM NaCl, 0.1% NP-40, 0.2 mM EDTA, 1.4 mM β-mercaptoethanol, 0.2 mM phenylmethylsulfonyl fluoride (PMSF), 1× protease inhibitor cocktail]. Before the immunoprecipitation, the NaCl concentration of the lysates was diluted to 179 mM with Dilution Buffer [20 mM Tris (pH 7.6), 20% glycerol, 0.05% NP-40, 0.2 mM EDTA, 1.4 mM β-mercaptoethanol, 0.2 mM PMSF, 1× protease inhibitor cocktail]. The lysates were incubated with the anti-eIF4A3 antibody (ProteinTech, 17504-1-AP) or IgG control (a normal rabbit IgG polyclonal antibody, PP64 from Millipore Sigma) by rotating overnight at 4°C. On the second day, Protein G–agarose beads were equilibrated with Lysis Buffer diluted with Dilution Buffer (179 mM NaCl). The lysate–antibody mixtures were then added to the equilibrated beads and rotated for 3 h at 4°C. After lysate–antibody–beads incubation, the bound beads were washed five times with Wash Buffer [50 mM HEPES (pH 7.6), 179 mM NaCl, 0.1% NP-40, 0.2 mM EDTA, 1.4 mM β-mercaptoethanol, 0.2 mM PMSF] and eluted with the sodium dodecylsulfate–dithiothreitol (SDS–DTT) loading buffer by boiling for 5 min at 95°C followed by centrifugation at 5000 *g* for 1 min to remove the beads. The eluted proteins were separated on an SDS–polyacrylamide gel electrophoresis (PAGE) gel, and the proteins of interest were visualized by western blotting analysis. Western blotting analyses were carried out with the following antibodies: the primary antibodies eIF4A3 (ProteinTech, 17504-1-AP), Vinculin (Abcam, ab129002), Gapdh (ProteinTech, 10494-1-AP), Hdac2 (Bethyl, A300-705A), Magoh (ProteinTech, 12347-1-AP), Casc3 (ThermoFisher, PA5-115580), Ccnb1 (Sigma, 05-373), Rbm8a (ProteinTech, 67541-1-Ig), β-Actin (Sigma, A5441), Chek (ProteinTech, 13954-1-AP), Haus1 (ProteinTech, 11094-2-AP), Alix/Pdcd6ip (Santa Cruz, sc-53539), Claspin/Clspn (Santa Cruz, sc-376773), Psme3 (ThermoFisher, PA5-21789), Usp39 (ThermoFisher, PA5-97229), Cyclin F (ThermoFisher, PA5-36049), Nsun2 (ProteinTech, 20854-1-AP) and Cenpt (ThermoFisher, PA5-83540); and the horseradish peroxidase (HRP)-conjugated secondary antibodies anti-rabbit IgG HRP-conjugated antibody (BioTechne, HAF008), anti-mouse IgG HRP-linked antibody (Cell Signaling, #7076). For co-immunoprecipitation samples, the true-blot

secondary antibody anti-rabbit IgG HRP TrueBlot (Rockland, 18-8816-31) was used to reduce the IgG detection.

Immunoprecipitation of eIF4A3 protein complexes in ESCs coupled with LC-MS/MS analysis

To identify eIF4A3-interacting partners in mouse ESCs, we used the anti-eIF4A3 antibody (ProteinTech, 17504-1-AP) in three different mouse ESC lines (J1, CCE and CJ7) to independently isolate eIF4A3-associated protein complexes for mass spectrometry (MS) identification, with IgG pull-down as controls (a normal rabbit IgG polyclonal antibody, PP64 from Millipore Sigma). Whole-cell lysates were prepared in Lysis Buffer [50 mM HEPES (pH 7.6), 250 mM NaCl, 0.1% NP-40, 0.2 mM EDTA, 1.4 mM β-mercaptoethanol, 0.2 mM PMSF, 1× protease inhibitor cocktail] from twenty 15 cm dishes of WT ESCs per cell line. The lysates were transferred to a dialyzer with Dialysis Buffer [20 mM HEPES (pH 7.9), 20% glycerol (v/v), 100 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT, 0.2 mM PMSF, 1× protease inhibitor cocktail] at 4°C for 3 h to decrease the salt concentration to 100 mM. The precipitates were removed by centrifugation at 14 000 rpm for 30 min at 4°C. The total lysate protein mass was determined by protein concentration measurement (Bradford). For each immunoprecipitation (IP), 10 mg of proteins were used, and 10% of lysates were saved as Input. IP lysates were diluted with IP-DNP buffer (Dialysis Buffer + 0.02% NP-40) to 12 ml, and benzonase (Pierce, 20 μl 15 U/μl) was added to remove DNA and RNA. The lysates were pre-cleared with the IP-DNP buffer-equilibrated Protein G–agarose beads (100 μl of beads per 10 mg of total protein) for 1 h at 4°C, followed by incubation with 20 μg of anti-eIF4A3 antibody (or IgG) by rotating overnight at 4°C.

The next day, Protein G–agarose beads were equilibrated with IP-DNP buffer. The lysate–antibody mixtures were added to the equilibrated beads and rotated for 3 h at 4°C. The bound beads were washed five times with IP-DNP buffer, and the protein complexes were eluted with the SDS loading buffer by boiling for 5 min at 95°C followed by centrifugation at 5000 *g* for 1 min to remove the beads. The eluted proteins were separated by SDS–PAGE. Whole lanes were excised and subjected to liquid chromatography-tandem MS (LC-MS/MS) analysis. MS data were processed by Thermo Proteome Discoverer software with the SEQUEST engine against the Swiss-Prot mouse protein sequence database. Proteins were filtered by the minimal number of identified unique peptides (≥2). Common contaminating proteins (e.g. keratins) were removed, and the spectral count (PSM, the number of peptide spectrum matches) ratio of eIF4A3 IP/IgG ≥10 was applied. The list was further cleaned with CRAPome by removing common background contaminants (16). The proteins present in all the three lists from the three cell lines were considered the final eIF4A3 interactome in ESCs.

Proliferation assay

One thousand ESCs stably expressing sh*EV* or sh*eIF4A3* were plated on 0.1% gelatin-coated 12-well plates in the ESC medium with 1 μg/ml puromycin. The cells were incubated at 37°C and 5% CO_2, and the medium was changed daily. The cell numbers were counted on days 2, 4 and 6.

Cell cycle analysis

ESCs (2×10^5 cells) were collected and washed with Dulbecco's phosphate-buffered saline (DPBS), permeabilized with 0.1% Triton X-100 in DPBS and stained with 3 μM 4′,6-diamidino-2-phenylindole (DAPI) at room temperature for 10 min. The samples were then analyzed by flow cytometry using an LSR-II Flow Cytometer (BD Biosciences). Finally, the data were analyzed by FlowJo software using the Dean-Jett-Fox cell cycle model.

Apoptosis detection and flow cytometry

Apoptotic analysis was determined using the fluorescein isothiocyanate (FITC)–Annexin V apoptosis detection kit with PI (propidium iodide) (Biolegend 640914) and performed according to the manufacturer's manual. First, a single-cell suspension was achieved using cell strainers to remove large cell clumps. Then the cells were stained with PI and Annexin V (FITC) to distinguish healthy cells (PI–/FITC–) from apoptotic cells (PI–/FITC+) and dead cells (PI+/FITC+). Both Annexin V- and PI-stained cells were analyzed by flow cytometry. The flow cytometry assays were performed with an LSR-II Flow Cytometer (BD Biosciences), and the data were analyzed by FlowJo software.

RNA extraction and quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was extracted from the mouse ESCs or NIH/3T3 cells as indicated with an RNeasy kit (Qiagen, 74136) following the manufacturer's manual. The RNA purity was validated with A_{260}/A_{280} ratios >1.9. A 1 μg aliquot of total RNA was converted to cDNA using 4 μl of qScript in a 20 μl reaction volume (Quanta, 95048) with the incubation steps following the manufacturer's manual (5 min at 25°C, 30 min at 42°C, 5 min at 85°C and hold at 4°C). Relative gene expression levels were analyzed with Lightcycler 480 SYBR green master mix (Roche, 4729749001) with three replicates on the LightCycler480 real-time PCR system (Roche) or QuantStudio 5 Real-Time PCR System (Applied Biosystems). Gene expression levels were calculated by the $\Delta\Delta C_q$ method with normalization to the *β-actin* expression level. The sequences of primer pairs were from PrimerBank (https://pga.mgh.harvard.edu/primerbank/) and are provided in Supplementary Table S1. More information about the primer pairs can be found in PrimerBank, including amplicon lengths, primer locations, gene ID and validation results (amplification plots, dissociation curves, agarose gel analysis, sequencing and BLAST data). For each primer pair, no-template controls were included as negative controls. The assays were carried out by Jianlong Wang's laboratory and the listed information followed the MIQE guidelines (17).

Cytoplasmic and nuclear RNA isolation

The cytoplasmic and nuclear RNA fractionation was performed as described (18). Briefly, 1×10^6 cells were harvested and washed with ice-cold PBS. After centrifugation

at 300 *g* for 4 min at 4°C, the cell pellets were suspended by 200 μl of lysis buffer (10 mM Tris pH 8.0, 140 mM NaCl, 1.5 mM $MgCl_2$, 0.5% Igepal, 200 U/ml RNaseIn), followed by a 5 min incubation on ice. One-fifth of the lysate was taken as the total RNA sample and added to 1 ml of Trizol for RNA extraction. The rest was centrifuged at 1000 *g* for 3 min at 4°C to pellet the nuclei, and the supernatant was the crude cytoplasmic fraction. For the pure cytoplasmic fraction, the supernatant was centrifuged at 14 000 rpm for 10 min at 4°C and transferred to a new tube for RNA extraction with Trizol. For the pure nuclear fraction, the pellets were washed first with 200 μl of lysis buffer and then with 200 μl of lysis buffer supplemented with 0.5% deoxycholic acid. The purified nuclei were used for RNA extraction with Trizol. Fractionated RNAs from the same number of cells were used for qRT-PCR analysis.

Subcellular protein fractionation

The fractionation of subcellular proteins of ESCs was performed using the Subcellular Protein Fractionation Kit for Cultured Cells (Thermo Scientific, 78840), following the manufacturer's manual.

Immunofluorescence staining

ESCs were fixed with 4% paraformaldehyde (w/v) for 15 min at room temperature, washed with PBS, permeabilized with 0.25% Triton X-100 solution for 5 min at room temperature and blocked with 5% FBS. Fixed cells were incubated with the primary antibody and 5% FBS in PBS overnight at 4°C. The next day, after washing with PBS, cells were incubated with secondary antibodies and 3 μM DAPI with 5% FBS in PBS for 1 h at room temperature in the dark. After washing, cells were imaged with a Leica DMi8 microscope. The primary antibodies are eIF4A3 (ProteinTech, 17504-1-AP), Vinculin (Abcam, ab129002), Oct4 (Santa Cruz, sc5279) and Gata6 (R&D Systems, AF1700).

eIF4A3 inhibitor treatment

eIF4A3-IN-2 (CAS no. 2095677-20-4) was purchased from MedChemExpress (HY-101785). J1 ESCs were treated with 0.33, 1, 3.33 or 10 μM eIF4A3-In-2 for 48 h. Dimethylsulfoxide (DMSO) was the vehicle control.

MTT assay

The cells were seeded into microplates (tissue culture grade, 96 wells, flat bottom) at a concentration of 2×10^4 cells/well in 100 μl of ESC medium with the indicated treatment (DMSO or eIF4A3-IN-2 at different concentrations) and then incubated for 48 h at 37°C and 5% CO_2. After incubation, the MTT assay was performed using the CellTiter 96® Non-Radioactive Cell Proliferation Assay (MTT) Kit (Promega, G4100), following the manufacturer's instructions.

RNAscope assay

The RNAscope assay was performed using RNAscope Multiplex Fluorescent Reagent Kit v2 (ACD Bio, 323100) with RNAscope® Probe-Mm-ccnb1-E5-E7 (ACD Bio, 316241), following the manufacturer's manual.

RNA-seq and data analysis

For the RNA-seq samples, biological replicates were prepared. Total RNA from each sample was extracted using the RNeasy kit (Qiagen, 74136). Then, samples were prepared, indexed and sequenced on the Illumina HiSeq system according to a poly(A) selection protocol as per the manufacturer's instructions (Tecan, 0520).

RNA-seq reads were aligned to the mouse mm9 genome using Bowtie2 (v2.3.4.3) (19), and aligned bam files were sorted by name using the parameter -n. Next, we used the HTSeq software (v0.11.2) and mm9 annotation file from GENCODE (https://www.gencodegenes.org/mouse/release_M1.html) to count reads for each gene using parameters -r name -f bam and -BioMart (20) to retrieve corresponding gene names. The annotation was not isoform specific and did not distinguish the isoforms of the same gene. Thus, for the genes with multiple isoforms, minor changes in splicing may be masked and a particular isoform may be changing more than others, although *Ccnb1* has only one isoform in mice (https://www.ncbi.nlm.nih.gov/gene/268697). Finally, read counts were normalized with the trimmed mean of M-values (TMM) method (21) for differential expression analysis using edgeR (v3.26.8) (22).

For the heatmap, a *Z*-score normalization was performed on the normalized gene expression across samples for each gene. *Z*-scores were calculated on a gene-by-gene (row-by-row) basis by subtracting the mean and then dividing the difference by the standard deviation. For RNA-seq, genes were considered up-/down-regulated if they had a *P*-value <0.05 and a $\log_2$ fold change (eIF4A3 KD/control KD) >2 (up-regulated) or < -2 (down-regulated).

Gene Ontology (GO) analysis and Gene Set Enrichment Analysis (GSEA)

GO analysis was carried out by the DAVID (The Database for Annotation, Visualization, and Integrated Discovery) functional annotation program (https://david.ncifcrf.gov/home.jsp). The terms were ranked according to the *P*-value that the program provided with default parameters.

GSEA v4.1.0 (available at www.gsea-msigdb.org/gsea/index.jsp) was performed to determine whether the indicated set was statistically enriched in eIF4A3 KD versus control KD samples. The gene sets were from the GSEA MSigDB database (the Molecular Signatures Database): formation of primary germ layer (systematic name: M10670), epidermis development (systematic name: M14065), mesoderm development (systematic name: M15421) and endoderm differentiation (systematic name: M34153). The normalized enrichment score (NES) and false discovery rate (FDR) were calculated by GSEA and indicated for each gene set enrichment test.

eCLIP-seq and data analysis

Enhanced UV cross-linking immunoprecipitation (eCLIP)-seq libraries were prepared in biological replicates according to the published eCLIP-seq protocol (23). Briefly,

2×10^7 ESCs were UV cross-linked at 400 mJ/cm^2 with 254 nm radiation. Cells were lysed in iCLIP lysis buffer and sonicated with Bioruptor®. The cell lysates were treated with diluted RNase I to fragment the RNA. The eIF4A3 antibody (ProteinTech, 17504-1-AP) was pre-coupled to Protein G Dynabeads and then added to the cell lysate, followed by overnight incubation at 4°C. Two percent of the lysate was taken as the input sample, and the remaining lysate was magnetically separated and washed. During washing, RNA was dephosphorylated with FastAP and T4 PNK (T4 PNK has the 3′-phosphatase activity) (24), followed by a 3′ RNA adapter ligation with T4 RNA ligase. The protein–RNA complexes were then separated by SDS–PAGE and transferred to nitrocellulose membranes. eCLIP was performed by excising the membrane covering the area from the molecular weight of eIF4A3 (47 kDa) to the molecular weight 75 kDa larger than that of eIF4A3 (i.e. 122 kDa). The SMInput (size-matched input) was used for each biological replicate, and the excised area was the same as its corresponding IP sample. The details regarding the eCLIP workflows are described in the eCLIP-seq protocol (23).

eCLIP-seq reads were processed following the published protocol (23). Adapters were trimmed (cutadapt v1.18), and reads <18 bp were discarded using parameters -m 18 -a NNNNNNNNNNAGATCGGAAGAGCACACG TCTGAACTCCAGTCAC -g ACGCTCTTCCGATCT -A AGATCGGAAGAGCGT -A GATCGGAAGAGCGT C -A ATCGGAAGAGCGTCG -A TCGGAAGAGCGT CGT -A CGGAAGAGCGTCGTG -A GGAAGAGCGT CGTGT for round 1, and parameters -m 18 -A AGAT CGGAAGAGCGT -A GATCGGAAGAGCGTC -A AT CGGAAGAGCGTCG -A TCGGAAGAGCGTCGT -A CGGAAGAGCGTCGTG -A GGAAGAGCGTCGTGT for round 2. Mapping reads were then performed against mouse elements in RepBase (25) with STAR (v2.6.1b) (26). Repeat mapping reads were segregated, and all others were mapped against the mouse mm9 genome with STAR (v2.6.1b). PCR duplicates were removed from uniquely mapped reads. Multiple in-line barcodes were merged for filtered reads, followed by peak identification with the clipper software (v0.2.0, https://pypi.org/project/clipper/) using parameters -s mm9 –Bonferroni –superlocal –threshold-method binomial –save-pickle on read 2 only (Bonferroni correction was employed on our peaks to reduce false positives; a semi-experimental option, '–superlocal', was used to pick up peaks that may be missed with genome-wide or gene-wide thresholds). In addition, based on these candidate peaks, we used the eCLIP size-matched input as control and then compared eCLIP with the input data to obtain peaks enriched in eCLIP samples with an intensity fold change of >2. We used the same code '-s mm9 –Bonferroni –superlocal –threshold-method binomial –save-pickle' for peak calling and 'Peak_input_normalization_wrapper.pl' for peak normalization in the paper mentioned above (23), and then filtered with significance thresholds of $P \leq 10^{-5}$ and fold enrichment ≥4 to obtain our peaks. The filtered peaks with the eCLIP signal compared with the input were then annotated. Peak annotation was performed using HOMER annotatePeaks.pl script to associate peaks with nearby genes.

For the replicate analysis of eCLIP-seq in Supplementary Figure S6B, both reads were counted for input and eCLIP data on the 3′-untranslated region (3′ UTR), intron, coding sequence (CDS) and 5′ UTR, and then normalized by the TMM method (21). Fold changes were calculated for both replicates using the TMM value. The correlation coefficient was calculated based on fold change values to indicate the data repeatability. For peak distribution analysis, we used Guitar (v1.20.1) (27) to visualize the binding frequency on the 5′ UTR, CDS and 3′ UTR.

IGV displays

For the Integrative Genomics Viewer (IGV) displays, we selected 'Normalize Coverage Data' in the preferences of 'Track'. This option applies to coverage tracks that are computed with igvtools. If selected, coverage values would be scaled by (1 000 000/totalCount), where totalCount is the total number of features or alignments.

Ribosome profiling and data analysis

Ribosome profiling was performed on cellular extracts from ESCs infected with control shRNAs (sh*NT* and sh*GFP*) or eIF4A3 shRNAs. Total RNA and ribosome-protected fragment (RPF) libraries were prepared using the TruSeq Ribo Profile (Mammalian) Kit (Illumina, RPHMR12126) according to the manufacturer's reference guide (document #15066016 v01) and the Ribo-Zero Gold Kit (H/M/R) (Illumina, MRZG126). The prepared libraries were sequenced on a HiSeq 2500 system (Illumina).

Ribosome profiling data reads were adapter trimmed using the FASTX-Toolkit (v0.0.14; http://hannonlab.cshl.edu/fastx_toolkit) with the parameter -a AGATCGGAAGAGCACACGTCT. Then, rRNA and tRNA reads were removed using Bowtie (v1.2.3). The remaining reads are aligned to the mouse mm9 genome using TopHat (v2.1.1). The matched total RNA-seq data were also processed with the same procedure as the ribosome profiling data. These aligned bam files were sorted by name with the parameter '-n' and counted by HTSeq (v0.11.2). For genes, we counted using parameters '-r name -f bam'. For transcripts, we used parameters '-r name -f bam –nonunique' to count reads. All read counts were normalized using the TMM method (21).

RESULTS

An shRNA screen identifies Ddx/Dhx factors required for ESC maintenance

To systematically explore the roles of Ddx/Dhx family members in ESCs, we employed three independent shRNAs for each Ddx/Dhx member in our loss-of-function screen (Supplementary Table S1). We infected J1 mouse ESCs with the prepared shRNA lentiviruses and, after two rounds of infections, selected the infected cells with puromycin until all cells in the control died. Since undifferentiated ESCs express high levels of AP while differentiated cells do not, we performed AP staining at Day 4.5 post-virus infection/shRNA treatment as the readout for ESC maintenance (Figure 1A). Accordingly, KD of the Ddx/Dhx mem-

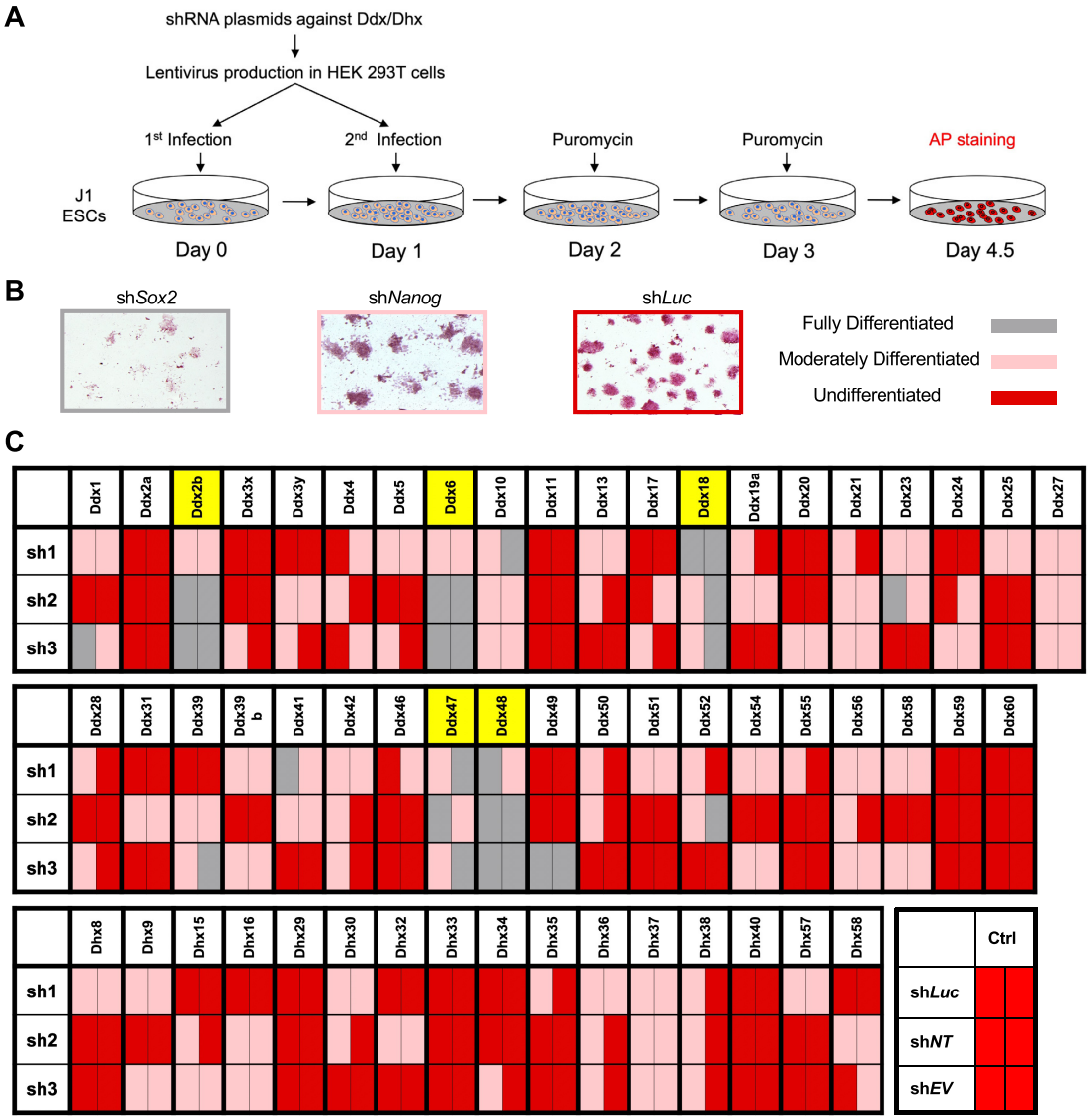


Figure 1. An RNAi screen identifies essential Ddx/Dhx RNA helicases for ESC maintenance. (**A**) Schematic of the RNAi screen to identify Ddx/Dhx RNA helicase dependency in ESCs. AP, alkaline phosphatase. (**B**) Representative images of the AP-stained colonies from the RNAi screen, including the positive controls (sh*Sox2* and sh*Nanog*) and the negative control (sh*Luc*). The image border colors denote ESC states as indicated. (**C**) The RNAi screen results with the keyboard schematic summary. The box colors match the colors used in (**B**). Selected top candidates are highlighted in yellow. The 'sh1–3' rows are three short hairpins for each gene, and the two columns within each gene denote the two biological replicates of the screen.

bers important for ESC maintenance would reduce AP activity with differentiation morphology. In the screen, the positive controls were sh*Sox2* (inducing severe differentiation) and sh*Nanog* (inducing moderate differentiation), and the negative controls (no differentiation) were sh*Luciferase* (sh*Luc*), sh*Non-Target* (sh*NT*) and pLKO.1 empty vector (sh*EV*). To categorize the AP staining result of all shRNAs, we used different colors to indicate various degrees of differentiation: gray for severe differentiation, pink for moderate differentiation and red for no differentiation, compared with controls (Figure 1B). The representative Ddx/Dhx RNAi screening results are shown in Supplementary Figure S1 and summarized in Figure 1C. For *Ddx*/*Dhx* genes whose official gene names are not *Ddx*/*Dhx*, we used their *Ddx*/*Dhx* aliases for consistency, i.e. *Ddx2a* (official gene name: *eIF4A1*), *Ddx2b* (*eIF4A2*), *Ddx48* (*eIF4A3*), *Ddx39b* (*Bat1a*) and *Ddx13* (*Skiv2l*). The results were highly reproducible from two independent screens (represented by two columns within each gene) (Figure 1C).

We identified five Ddx/Dhx members (Ddx2b, Ddx6, Ddx18, Ddx47 and Ddx48) whose depletion induced significant differentiation in ESCs, indicating their critical roles in ESC maintenance. We validated these five candidates by repeating the shRNA experiment in another mouse ESC line, CCE, with the same protocol described (Figure 1A), and confirmed that the pluripotency loss and the differentiation induction were reproducible and consistent with the target KD efficiency (Supplementary Figure S2A, B). Among these five candidates, Ddx2b (28,29), Ddx6 (30–33), Ddx18 (34) and Ddx47 (35) were previously reported to play im-

portant roles in ESCs, not only validating our screening results but also highlighting Ddx48 as a novel candidate. Examination of their expression levels among various tissues and cell lines revealed that, like Ddx18, Ddx48 is also particularly enriched in ESCs (Supplementary Figure S3), indicating its ESC-specific roles. Ddx48 is also called eIF4A3, a component of the EJC known for its functions in various aspects of RNA metabolism, including splicing, mRNA export and RNA stability control in mostly non-pluripotent cellular systems (14,36–40). We thus focused on studying the functional significance of eIF4A3 in ESC maintenance to further understand the pluripotency control mechanism.

eIF4A3 and its helicase activity are required for ESC maintenance

We first examined the expression profiles of *eIF4A3* in the establishment of pluripotency through somatic cell reprogramming, during early development and in the process of ESC differentiation. We found that the mRNA level of *eIF4A3* increases during MEF reprogramming (Supplementary Figure S4A) (41). During mouse embryogenesis, both RNA and RPF (suggesting mRNA translation level) levels of eIF4A3 are up-regulated in the post-ZGA (zygotic genome activation) phase (Supplementary Figure S4B) (42). Finally, the RNA level of eIF4A3 decreases in the differentiation of ESCs to all the three primary lineages (Supplementary Figure S4C) (43). These expression profiles are consistent with the critical role of eIF4A3 in ESC maintenance.

To further characterize the effects of eIF4A3 depletion on ESC maintenance, we first validated the KD efficiency of eIF4A3 shRNAs in ESCs by western blotting analysis (Figure 2A, B). We also confirmed that eIF4A3 KD decreases ESC proliferation (Figure 2C). To characterize the cellular identity of the eIF4A3-depleted ESCs, we performed RNA-seq of ESCs collected at Day 4.5 after the infection with sh*eIF4A3*-2/3 or sh*Control* (sh*Luc* and sh*EV*) viruses (Figure 2D; Supplementary Table S2). We found that 898 genes were up-regulated in their RNA levels upon eIF4A3 KD, with the enrichment of development-related GO terms such as 'multicellular organism development' and 'metanephros development' (Figure 2E). Conversely, there were 383 down-regulated genes upon eIF4A3 KD, with the enrichment of metabolism-related GO terms (Figure 2F). GSEA showed the enrichment of gene sets representing 'formation of primary germ layer' and all the three differentiation lineages in the eIF4A3 KD cells (Figure 2G). These results are consistent with the differentiation phenotype of eIF4A3-depleted ESCs and also with the well-known differential metabolism of undifferentiated and differentiated states of ESCs (44–46).

eIF4A3 is an ATP-dependent Ddx RNA helicase. To investigate whether its helicase activity is required for ESC maintenance, we transfected ESCs with the 3×Flag-tagged *eIF4A3* transgene encoding either the WT or ATPase mutant (MT) harboring the DEAD to NEAD mutation (47). Before eIF4A3 KD, the expression levels of the ectopic *3×Flag-eIF4A3* (WT and MT) transgenes were similar to those of the endogenous eIF4A3 (Figure 2H, I, left). After eIF4A3 KD with the shRNA targeting the 3′ UTR of the endogenous *eIF4A3* mRNAs to deplete the endogenous eIF4A3, the protein expression levels of the ectopic 3×Flag-eIF4A3, both WT and MT, were maintained at the expected close to endogenous levels (Figure 2H, I, right). Importantly, we found that 3×Flag-eIF4A3 WT, but not the MT, can fully rescue the cellular differentiation associated with eIF4A3 KD, indicated by the pluripotency marker Oct4 and the differentiation marker Gata6 (Figure 2J), excluding the off-target effects of shRNAs. These results demonstrate that the helicase activity is required for the functions of eIF4A3 in ESC maintenance.

eIF4A3 is an integral part of the EJC required for ESC maintenance

eIF4A3 has been identified as a component of the EJC core in many cellular systems (14). To investigate whether eIF4A3 also functions through the EJC in ESCs, we performed IP of eIF4A3 and its associated protein partners from ESC lysates using an anti-eIF4A3 antibody followed by MS and western blotting analysis. We confirmed the enrichment of eIF4A3 after IP and identified a total of 248 eIF4A3-interacting proteins encompassing Rbm8a, Magoh and Casc3 (Figure 3A, B; Supplementary Table S3), the core EJC components that were at least 10-fold more enriched by IP with anti-eIF4A3 over IgG in all three ESC lines (Supplementary Table S3). GO analysis identified the enrichment of terms associated with mRNA splicing, translation and nuclear export (Figure 3C), consistent with the broad functions of the EJC. Immunofluorescence using the anti-eIF4A3 antibody (Supplementary Figure S4D, E) and subcellular protein fractionation followed by western blotting analysis (Supplementary Figure S4F) further confirmed that, similar to NIH/3T3 cells, eIF4A3 is highly enriched in the nucleus with relatively low expression in the cytoplasm. These data demonstrate that eIF4A3 is an integral component of the EJC core in mouse ESCs and suggest that eIF4A3/EJC plays a prominent regulatory role in the ESC nucleus besides some accessory functions in the cytoplasm (14).

To investigate whether the EJC is overall important in ESC maintenance, we performed KD of each of the four core components of the EJC, followed by gene expression and cell morphology examinations. We found that the depletion of eIF4A3, Magoh and Rbm8a caused the down-regulation of pluripotency genes, including the triumvirate *Oct4*, *Nanog* and *Sox2* (Figure 3D). Although Casc3 deficiency did not down-regulate pluripotency gene transcripts as much as the other components, the depletion of all the four components individually, including Casc3, led to the loss of pluripotency with a significant or moderate differentiation phenotype in ESCs, indicated by the loss of Oct4 and/or the gain of Gata6 expression (Figure 3E). These results demonstrate that the EJC is required for ESC maintenance. Interestingly, we noted that the depletion of eIF4A3 also compromised the expression levels of *Magoh* and *Rbm8a*, but not *Casc3* (Figure 3D). However, the depletion of Magoh or Rbm8a did not affect expression of the other core components (Figure 3D). These results suggest that eIF4A3, as a core component of the EJC, may play a

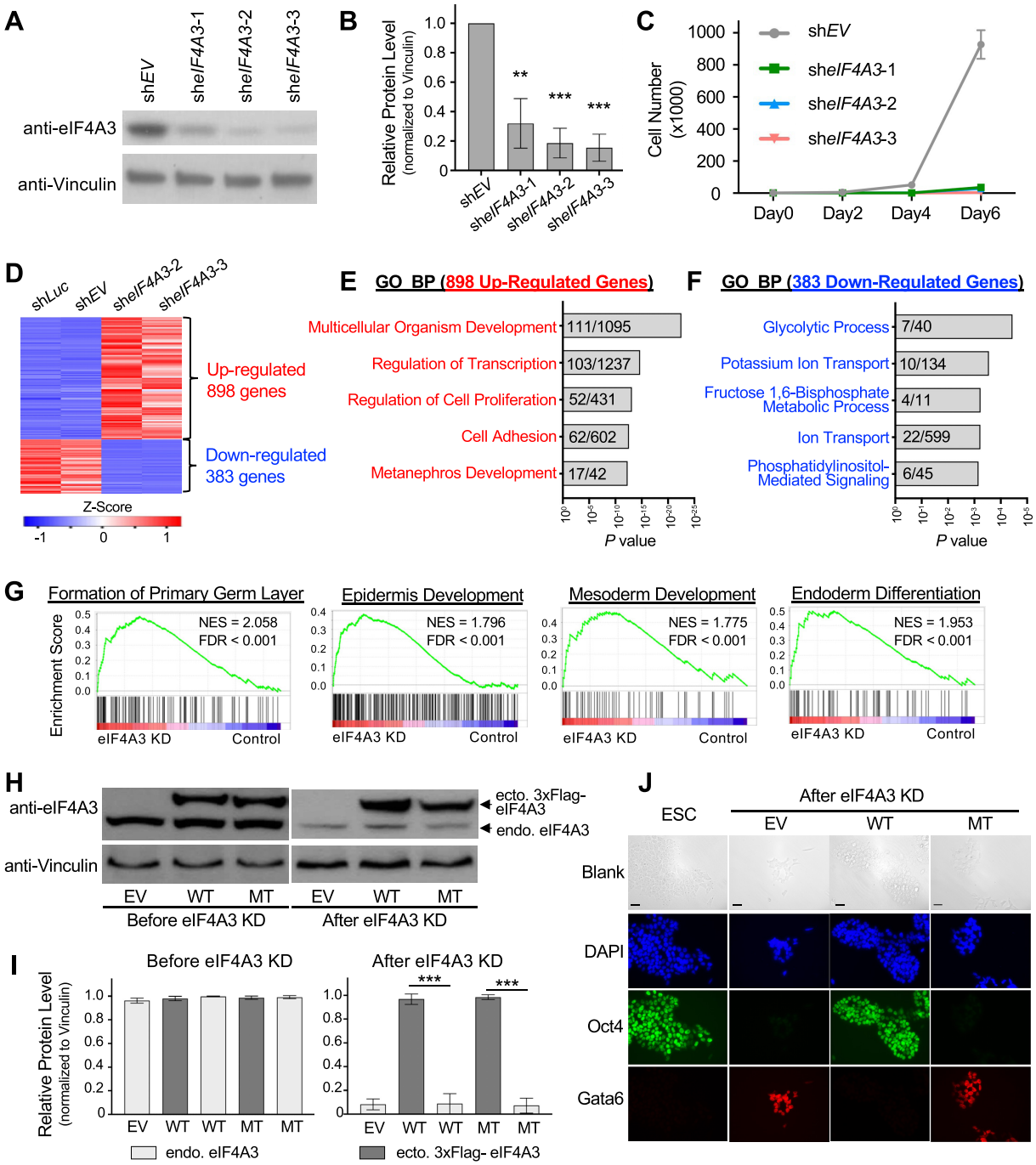


Figure 2. eIF4A3 is required for ESC maintenance. (**A** and **B**) Western blot analysis (A) and the protein quantification (B) validating efficient eIF4A3 KD (she*IF4A3*) compared with control KD (sh*EV*). EV, empty vector. Vinculin is the loading control. (**C**) A time-course growth assay indicating the proliferation defects of eIF4A3 KD (she*IF4A3-1/2/3*) relative to control KD (sh*EV*) ESCs. (**D–F**) RNA-seq data analysis comparing the transcriptomes in eIF4A3 KD (she*IF4A3-2/3*) and control KD (sh*Luc* and sh*EV*) ESCs, depicted by heatmap (D) and GO analysis of the up-regulated (E) and down-regulated (F) transcripts upon eIF4A3 KD. sh*Luc*, sh*Luciferase*. For each term, the numerator represents the hit number in our datasets, and the denominator represents the total gene number in each term. (**G**) GSEA results of primary germ layer gene sets (formation of the primary germ layer, epidermis development, mesoderm development and endoderm differentiation) by comparing eIF4A3 KD with control KD cells from RNA-seq data. NES, normalized enrichment score. (**H**) Western blots of eIF4A3 and Vinculin in ESCs ectopically (ecto.) expressing EV, epitope-tagged WT eIF4A3 (3×Flag-eIF4A3 WT) or ATPase mutant eIF4A3 (3×Flag-eIF4A3 MT) before (left) and after (right) eIF4A3 KD. Vinculin is the loading control. The she*IF4A3* targets the 3′ UTR of endogenous (endo.) *eIF4A3*, so it does not target the transgenic *eIF4A3* (3×Flag-eIF4A3 WT and MT). (**I**) Protein quantification of eIF4A3 normalized to Vinculin in (H). Error bars represent the standard deviation (SD) in experimental duplicates. (**J**) Immunofluorescence of Oct4 and Gata6 of eIF4A3 KD ESCs rescued with the transgene of EV, epitope-tagged WT eIF4A3 (3×Flag-eIF4A3 WT) or ATPase mutant eIF4A3 (3×Flag-eIF4A3 MT), ESC as the control. Scale bar, 20 μm. All *P*-values were calculated using two-tailed unpaired Student's *t*-test, **P*-value <0.05; ***P*-value <0.01; ****P*-value <0.001; 'ns' means not significant.

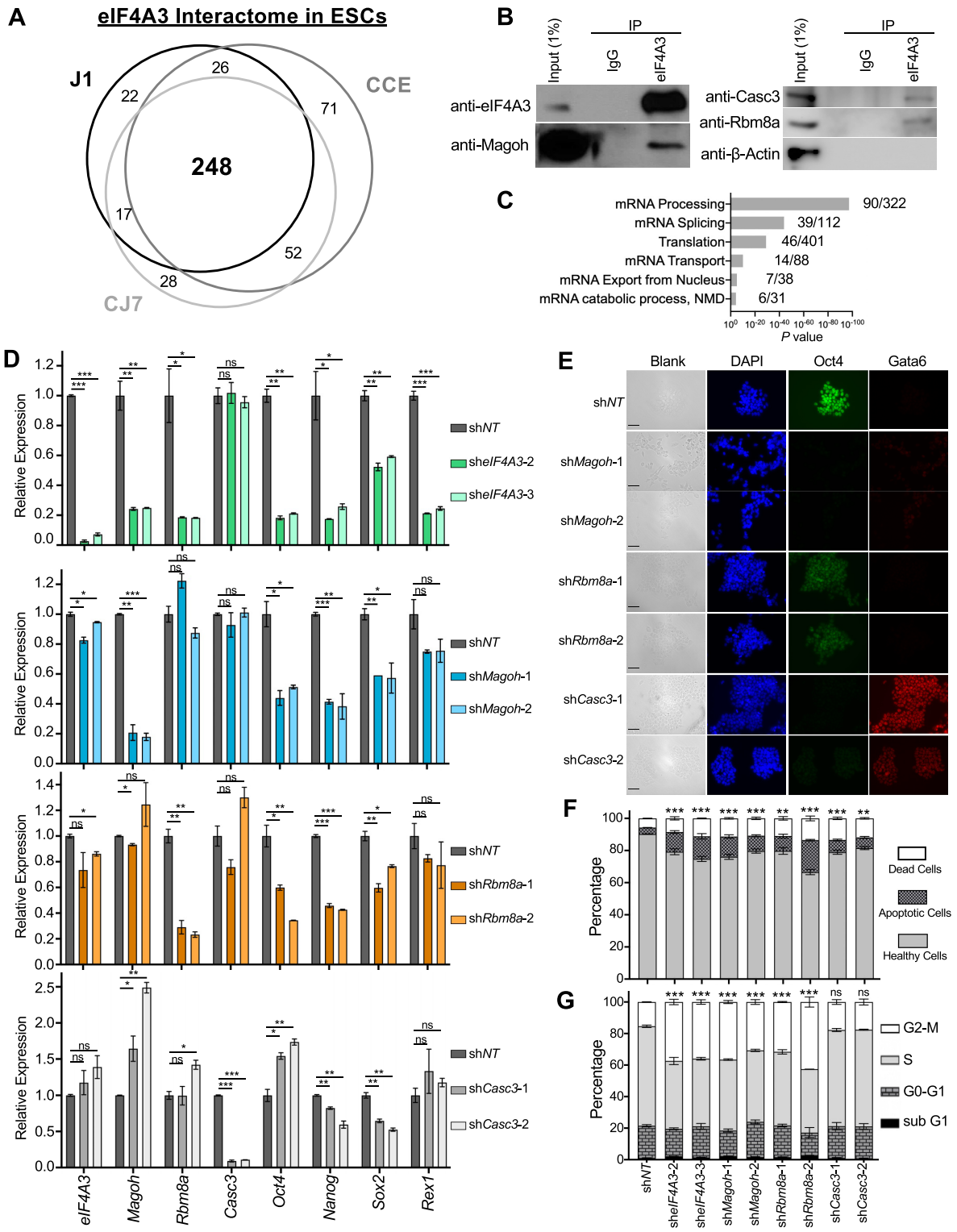


Figure 3. The eIF4A3 interactome and effects of eIF4A3/EJC loss on cell cycle regulation in ESCs. (**A**) Venn diagrams showing the eIF4A3 interactome in ESCs. eIF4A3 IP-MS was performed in three ESC lines (J1, CCE and CJ7). (**B**) Validation of the physical interactions of eIF4A3 with Magoh, Casc3, Rbm8a and β-Actin (the negative control prey) by IP of eIF4A3 complexes in ESCs followed by western blotting analysis with the indicated antibodies. IgG serves as the negative IP control. (**C**) GO analysis of the eIF4A3 interactome in ESCs (the 248 proteins in A). For each term, the numerator represents the hit number in our datasets, and the denominator represents the total gene number in each term. (**D** and **E**) qRT-PCR analysis (D) and immunofluorescence of Oct4 and Gata6 (E) of ESCs with control KD (sh*NT*) or KD of each EJC core component indicated. The qRT-PCR (D) is normalized to *β-actin* expression relative to sh*NT*. Scale bar, 20 μm. (**F** and **G**) The apoptosis (F) and cell cycle (G) analyses of ESCs upon KD with sh*NT* or shRNAs targeting *eIF4A3*, *Magoh*, *Rbm8a* or *Casc3*. Error bars represent the SD in experimental triplicates.

critical role in maintaining the complex integrity of the EJC in ESCs.

We also noticed that depletion of the four core components of the EJC in ESCs resulted in increased cell death (Figure 3F). We then performed cell cycle analysis and found that Casc3 deficiency did not change the ESC cell cycle pattern. However, the depletion of eIF4A3, Magoh or Rbm8a caused the G_2–M arrest in ESCs (Figure 3G). eIF4A3, Magoh and Rbm8a are highly expressed in ESCs, and Casc3 also has a relatively high expression level in ESCs (Supplementary Figures S3 and S5A). To examine whether such effects of the EJC deficiency are specific to pluripotent ESCs, we knocked down these components in mouse embryonic fibroblasts (MEFs; NIH/3T3) using the same shRNAs with validated KD efficiency (Supplementary Figure S5B, C). We found that the EJC core deficiency did not cause obvious morphological changes of NIH/3T3 cells (Supplementary Figure S5D) except for some subtle changes of apoptotic/dead cell populations (Supplementary Figure S5E) and cell cycle pattern (Supplementary Figure S5F), which is in stark contrast to what was observed in ESCs (Figure 3D–G; Supplementary Figure S2A). In addition, we also found that the depletion of eIF4A3 increased the RNA expression levels of the other core components in NIH/3T3 cells, which is opposite to what was observed in ESCs (compare the top panels in Supplementary Figure S5B and Figure 3D). However, such an increase in RNA levels did not influence the protein levels of the other core components in NIH/3T3 cells (Supplementary Figure S5C), implying that the eIF4A3-dependent EJC integrity is present in ESCs but not NIH/3T3 cells.

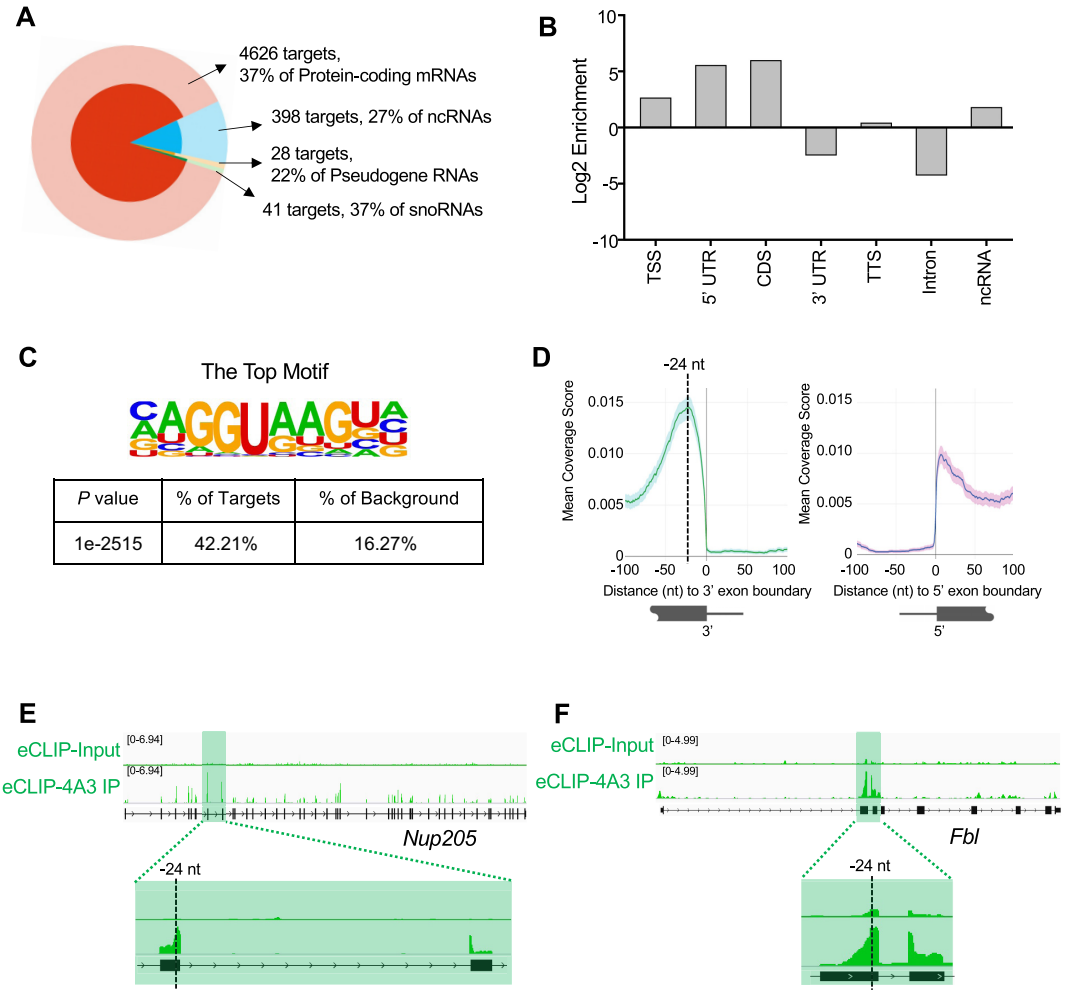
These results demonstrate that eIF4A3 is an integral part of the EJC in ESCs and that the EJC has PSC-specific functions in maintaining both expression levels of pluripotency transcripts and cell cycle patterns in ESCs.

eIF4A3 post-transcriptionally controls pluripotency-related cell cycle regulators in ESCs

To understand the molecular mechanism by which eIF4A3 and its associated EJC control ESC maintenance, we aimed to identify the direct RNA targets of eIF4A3 in ESCs. We performed anti-eIF4A3 eCLIP to identify genome-wide RNA targets of eIF4A3 in mouse ESCs (Supplementary Figure S6A, B) (23), and identified 30 763 significant sequence peaks corresponding to 5093 annotated transcripts enriched in anti-eIF4A3 eCLIP relative to the Input control (Figure 4A; Supplementary Table S4). The targets are mostly protein-coding mRNAs, with some non-coding RNAs (RNAs), pseudogene RNAs and small nucleolar RNAs (snoRNAs) (Figure 4A). Most binding peaks are located within 5′ UTRs and CDS regions (Figure 4B; Supplementary Figure S6C, D) with enrichment at the exon–exon junctions, particularly around 24 nt upstream of exon junctions (Figure 4C–F; the motif in Figure 4C is around the 3′ exon boundary in Figure 4D), consistent with the eIF4A3/EJC binding position identified in other cellular systems (48,49). Thus, these analyses confirmed the expected and specific binding pattern of eIF4A3 as a component of the EJC on the target mRNAs.

eIF4A3/EJC is involved in multiple regulatory layers, including RNA splicing, nuclear export, mRNA translation and RNA degradation (14,48). While not excluding the possible roles of eIF4A3 in other aspects of RNA metabolism including RNA splicing (see the Discussion), we reasoned that the proteins are the final converging point of all those multiple steps of RNA metabolism and thus decided to focus on the translation status of these mRNAs upon eIF4A3 KD relative to control KD. For eIF4A3 and associated EJCs located within the coding regions of mRNAs that are shuttled from the nucleus to the cytoplasm, translating ribosomes can remove and disassemble them during the initial (or pioneer) round of mRNA translation (48). We performed ribosome profiling (50) in ESCs upon KD with either control shRNA or sh*eIF4A3* (Supplementary Figure S7A) to identify RPFs, which represent the mRNA fragments bound by translating ribosomes. The total RNA samples were purified and subjected to deep sequencing in parallel (Supplementary Figure S7A). We identified 13 611 transcripts (Supplementary Table S5), among which 1978 transcripts underwent significant ribosome binding changes, reflected by RPF changes, upon efficient eIF4A3 KD (Figure 5A; Supplementary Figure S7B). To identify the critical pathway underlying the post-transcriptional control by eIF4A3 in the context of EJCs being the node of post-transcriptional networks (14), we focused on the 1421 transcripts out of those 1978 transcripts that were not significantly altered in RNA levels (Supplementary Table S5). By integrating with the anti-eIF4A3 eCLIP-seq dataset, we further narrowed down the list to 503 transcripts that were also direct eIF4A3-binding targets (Figure 5A). On these 503 transcripts, eIF4A3 occupancy also represents an EJC-like enrichment ∼24 nt upstream of the 3′ exon boundary (Supplementary Figure S7C). Interestingly, GO analysis of these 503 transcripts revealed that the top three terms on the biological process are all related to cell cycle regulation (Figure 5B), consisting of 40 genes (Figure 5C). These results suggest that eIF4A3 (and probably the EJC) may regulate pluripotency through post-transcriptional, and subsequently or consequently, translational control of the ESC-specific cell cycle.

We then focused on these 40 transcripts (Figure 5C) and asked which ones were specific for PSCs. First, we examined their expression levels in reprogramming to pluripotency (51) and embryoid body (EB) differentiation (52) (Supplementary Figure S7D). We identified nine genes showing the most enriched expression in the pluripotent states in both processes, i.e. the expression levels increase during MEF reprogramming to induced PSCs and decrease during EB differentiation (bold brown text in Supplementary Figure S7D). In contrast, only one common factor shows the opposite expression trend in the two processes (bold green text in Supplementary Figure S7D). We then examined the expression levels of these 10 candidates in different mouse tissues and cell lines, which confirmed relatively high expression of those nine candidates and low expression of that single candidate in ESC lines (Supplementary Figure S8A). Finally, consistent with their expression patterns in pluripotent states and the post-transcriptional/translational regulation mediated by eIF4A3, the expected decreased/increased ribosome binding patterns as well as their protein level



P value	% of Targets	% of Background
1e-2515	42.21%	16.27%

Figure 4. eIF4A3 binds at the exon junction sites of protein-coding mRNAs in ESCs. (**A**) Pie chart depiction of the categories of eIF4A3 target RNAs in ESCs, identified by the eIF4A3 eCLIP-seq data. The slices represent the ratio of all the RNA transcripts falling into the four indicated types in the Input control of the eCLIP-seq. In each slice, the inside layer with a darker color represents the ratio of the eIF4A3 targets (enriched in the eIF4A3 eCLIP-seq compared with the Input control) in all the RNA transcripts of each type. (**B**) The enrichment of peaks in characteristic RNA regions relative to the occupancy of each characteristic region in the genome from the eIF4A3 eCLIP-seq data. (**C**) The top enriched motif in eIF4A3-bound peaks from the eIF4A3 eCLIP-seq data. (**D**) Metagene profile of eIF4A3-binding sites on exons throughout the transcriptome in ESCs relative to 5′ and 3′ splice sites. The dark gray boxes represent exonic regions. To avoid the inherent interference from the other side of the short exons, only the exons >150 nt were profiled here. (**E** and **F**) IGV snapshots of eIF4A3 target examples from eIF4A3 eCLIP-seq data (green). The peak enrichment around the –24 nt site is zoomed in and shown as the bottom panels.

changes upon eIF4A3 KD were observed on these candidate transcripts without changes at the RNA levels (Figures 5D and 6G; Supplementary S8B, C).

Taken together, these data demonstrate that eIF4A3 maintains ESC pluripotency through a post-transcriptional mechanism with translational (protein synthesis) consequence of pluripotency-related cell cycle regulators.

eIF4A3 directs the nuclear export of Ccnb1 mRNAs to maintain ESC pluripotency

To explore whether we could pinpoint the specific post-transcriptional steps when eIF4A3 may act in controlling the expression of ESC-specific cell cycle regulators, we focused on the candidate gene *Ccnb1* encoding Cyclin B1 for the following reasons. First, Cyclin B1 is the only cyclin with a strong cell cycle-dependent behavior in mouse ESCs (8), and it is also the only cyclin required for mouse early embryogenesis (53). Second, the pluripotency transcripts are sensitive to the changes in Ccnb1 expression such that the Ccnb1 depletion down-regulates the RNA levels of pluripotency mRNAs with the G_2–M arrest in ESCs (11). Third, as shown before, the expression of Ccnb1 is highly enriched in the pluripotent states (Supplementary Figures S7D and S8A). Consistent with the post-transcriptional control, we found that neither the splicing pattern nor the abundance of *Ccnb1* mRNAs was affected by eIF4A3 depletion (Figure 5D–F). However, eIF4A3 loss led to a significant decrease in ribosome binding on *Ccnb1* mRNAs (Figure 5D). Such a requirement of eIF4A3 for ribosome binding to mRNAs is ostensibly contradictory to the well-established model whereby ribosomes disassemble eIF4A3/EJCs from mRNAs during the pioneer round of translation immediately upon nucleocytoplasmic export (48). To reconcile this

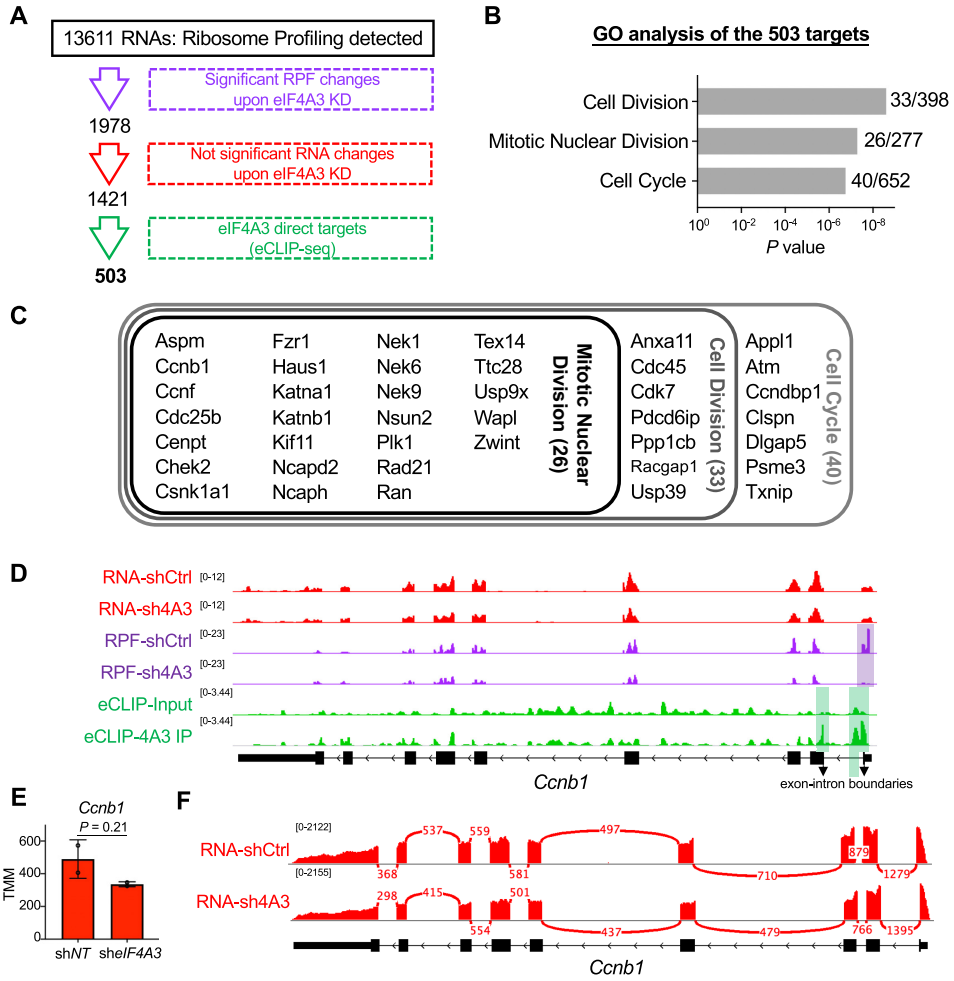


Figure 5. eIF4A3 directly controls the mRNAs of cell cycle regulators in a post-transcriptional manner. (**A**) Flow chart showing the filtering process to identify the genes under eIF4A3-mediated post-transcriptional control. The genes with 'significant RPF changes upon eIF4A3 KD' were defined as the genes with a *P*-value of RPF changes <0.05. The genes with 'Not significant RNA changes upon eIF4A3 KD' were defined as the genes with a *P*-value of RNA changes ≥0.05. 'eIF4A3 direct targets (eCLIP-seq)' were from the eIF4A3 eCLIP-seq data in ESCs with enriched CLIP signal compared with the input control. (**B**) GO analysis of the final 503 targets in (A). For each term, the numerator represents the hit number in our datasets, and the denominator represents the total gene number in each term. (**C**) Venn diagrams showing overlaps in the genes associated with the cell cycle-related GO terms in (B). (**D**) IGV snapshots on *Ccnb1* showing the in-parallel RNA-seq (red) and RPF profiling (purple) datasets from the ribosome profiling between control KD (sh*Ctrl*) and eIF4A3 KD (sh*4A3*). The major down-regulation of RPF on *Ccnb1* upon eIF4A3 KD happened on the translation initiation region (5′ UTR and the beginning of the CDS region), shaded in purple. The eIF4A3 binding (eCLIP-seq) profiling data (green) with binding of eIF4A3 around the exon junction site (the first and the third peaks from the right) and the intron (the second peak from the right) is shaded in green. (**E**) Comparison of transcript levels of *Ccnb1* in sh*NT* and she*IF4A3*. The data are from the in-parallel RNA-seq data of the ribosome profiling shown as the red tracks in (D). (**F**) Comparison of Sashimi plots of *Ccnb1* in sh*NT* and she*IF4A3* KD ESCs. *Ccnb1* only has one isoform in mice.

contradiction, a parsimonious hypothesis is that eIF4A3 may be necessary for the nuclear export of *Ccnb1* mRNAs to the cytoplasm to be accessible to ribosome binding (Figure 6A).

To test this hypothesis, we first isolated cytoplasmic and nuclear RNAs and used the fractionated RNAs from the same numbers of cells for qRT-PCR analysis to examine the subcellular localization patterns of *Ccnb1* mRNAs in J1 and CCE mouse ESCs. We found that *Ccnb1* mRNAs are highly cytoplasmic in both cell lines (Figure 6B). However, upon eIF4A3 KD, the subcellular localization of *Ccnb1* was altered with markedly increased nuclear percentiles (Figure 6C), which could be rescued by the ectopic expression of the WT but not the ATPase mutant (MT) eIF4A3 (Figure 6D–F). These findings support a model whereby eIF4A3 targets *Ccnb1* transcripts, with the involvement of its RNA helicase activity, for transporting to the cytoplasm.

Cyclin B1 is important for the cell cycle and the expression of pluripotency transcripts in ESCs (11). To substantiate our model, we overexpressed *Ccnb1* cDNA (the CDS without introns) in eIF4A3-depleted ESCs (Figure 6G, H) and asked whether Cyclin B1 overexpression could rescue the eIF4A3 KD phenotype. Of note, the intron-less *Ccnb1* cDNA transcript is suboptimal for the EJC targeting (54), and thus it is minimally affected by eIF4A3 KD, allowing for the overexpression (Figure 6G, H). In eIF4A3 KD ESCs, the pluripotency (indicated by Oct4 immunofluorescence) and the G_2–M arrest could be partially rescued by ectopic expression of Ccnb1, but the differentiation state (indicated by Gata6 immunofluorescence) was not rescued

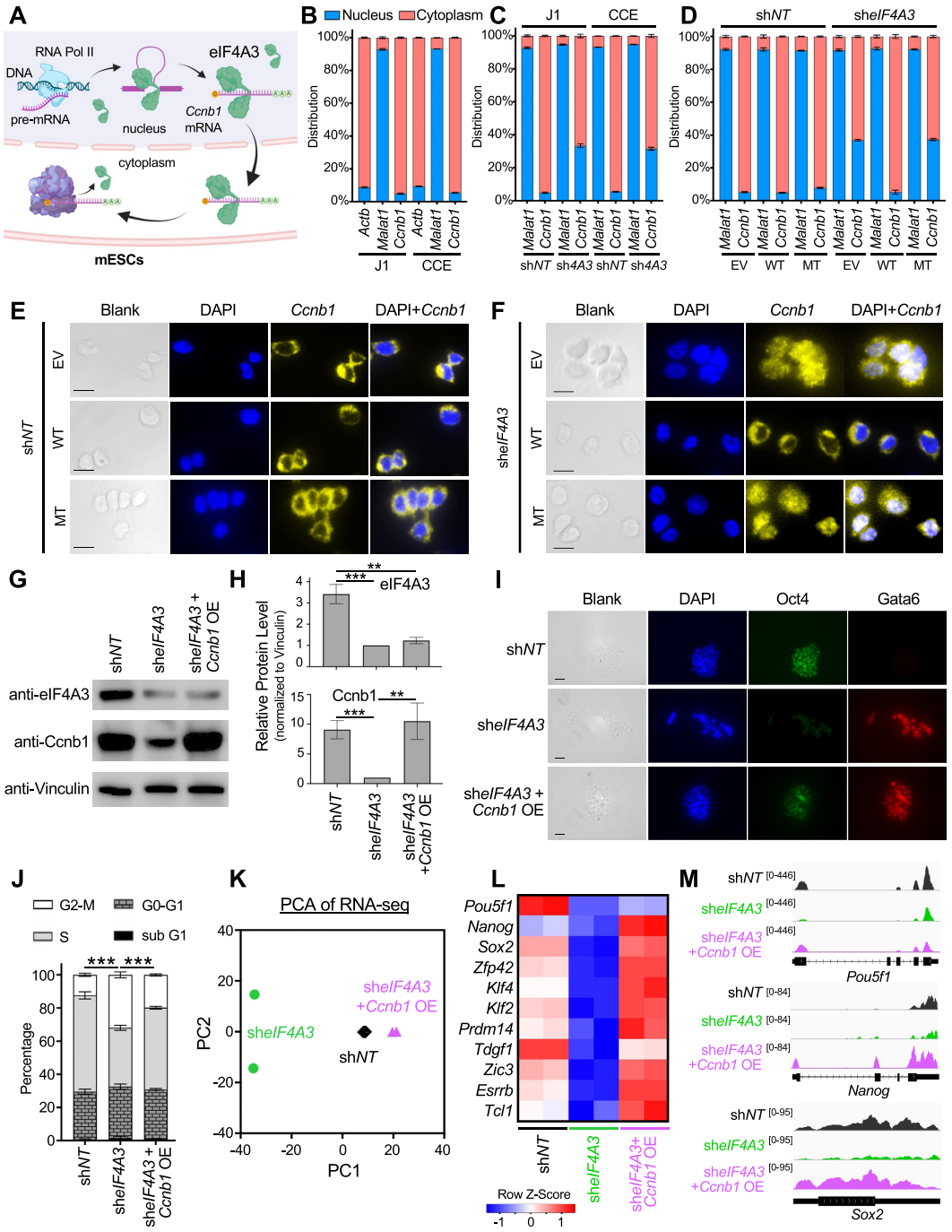


Figure 6. eIF4A3 controls the nuclear export of *Ccnb1* mRNAs to maintain the ESC cell cycle profile. (**A**) Schematic of the eIF4A3-mediated control of *Ccnb1* mRNAs in the ESC nucleus and cytoplasm. (**B**) *Ccnb1* mRNAs are mainly localized in the cytoplasm in J1 and CCE ESC lines, as revealed by qRT-PCR. *Actb* and *Malat1* are cytoplasmic and nuclear markers, respectively. (**C**) The nuclear and cytoplasmic distribution of *Malat1* and *Ccnb1* transcripts in ESCs upon control (sh*NT*) or eIF4A3 (sh*4A3*) KD, as revealed by qRT-PCR analysis. (**D**) The nuclear and cytoplasmic distribution of *Malat1* and *Ccnb1* transcripts in ESCs ectopically expressing EV, epitope-tagged WT eIF4A3 (3×Flag-eIF4A3 WT) or ATPase mutant eIF4A3 (3×Flag-eIF4A3 MT) upon control (sh*NT*, left) or eIF4A3 (she*IF4A3*, right) KD, as revealed by qRT-PCR. (**E** and **F**) Validation of subcellular localization of *Ccnb1* mRNAs under the indicated conditions by RNAscope (*in situ* hybridization for RNA detection). Scale bar, 10 μm. (**G** and **H**) Western blot (G) and the protein quantification (H) of eIF4A3 and Ccnb1 in ESCs upon KD with sh*NT*, she*IF4A3* and she*IF4A3* plus *Ccnb1* OE (OE: overexpression). Vinculin is the loading control (G), and the quantification is normalized to the level of Vinculin (H). Error bars represent the SD in experimental duplicates. (**I**) Immunofluorescence of Oct4 and Gata6 in ESCs with sh*NT*, she*IF4A3* or she*IF4A3* plus *Ccnb1* OE. Scale bar, 20 μm. (**J**) The cell cycle analysis of ESCs upon KD with sh*NT*, she*IF4A3* and she*IF4A3* plus *Ccnb1* OE. Error bars represent the SD in experimental triplicates. (**K**) Principal component analysis of the RNA-seq data from ESCs upon KD with sh*NT*, she*IF4A3* or she*IF4A3* plus *Ccnb1* OE. (**L**) Heatmap showing the expression of pluripotency genes from the RNA-seq data of ESCs upon KD with sh*NT*, she*IF4A3* or she*IF4A3* plus *Ccnb1* OE. (**M**) The IGV snapshots of the RNA-seq data of ESCs upon KD with sh*NT*, she*IF4A3* or she*IF4A3* plus *Ccnb1* OE at *Pou5f1*, *Nanog* and *Sox2* genomic regions. All *P*-values were calculated using two-tailed unpaired Student's *t*-test, **P*-value <0.05; ***P*-value <0.01; ****P*-value <0.001; 'ns' means not significant.

(Figure 6I, J). We also collected RNAs with biological replicates of control KD, eIF4A3 KD and eIF4A3 KD plus Ccnb1 overexpression ESCs for RNA-seq analyses (Supplementary Table S6). Principal component analysis (PCA) showed that the transcriptome profiles of eIF4A3 KD plus Ccnb1 overexpression ESCs are more similar to those of the control KD ESCs at both PC1 and PC2 dimensions than those of eIF4A3 KD alone (Figure 6K). Importantly, the expression of pluripotency transcripts was rescued, although to different extents, by ectopic overexpression of Ccnb1 in eIF4A3 KD ESCs (Figure 6L). Of note, upon Ccnb1 overexpression, *Pou5f1* (the gene coding for Oct4) is slightly rescued while *Nanog* has a higher expression level than the control, and *Sox2* is fully rescued to a similar level to that in the control (Figure 6L, M), indicating different responses of pluripotency factors to Ccnb1 overexpression. These data demonstrate that eIF4A3 regulates the nuclear export of *Ccnb1* mRNAs, indirectly controlling its translation in maintaining ESC pluripotency.

The specific eIF4A3 inhibitor mimics eIF4A3 depletion effects in ESCs

A highly selective inhibitor of eIF4A3, eIF4A3-IN-2 (CAS no. 2095677-20-4), was recently discovered with specific suppression of eIF4A3 but not other eIF4A family members (55). To complement our eIF4A3 KD experiments, we treated mouse ESCs with this inhibitor in a series of concentrations for 48 h (Supplementary Figure S9A). We found that a higher concentration of this inhibitor caused a greater loss of the ESC pluripotency accompanied by an induction of differentiation, indicated by the loss of Oct4 and the gain of Gata6 expression (Supplementary Figure S9A), consistent with the ESC phenotype associated with eIF4A3 KD described above (Supplementary Figure S2A). To determine the toxicity of this inhibitor to ESCs, we performed an MTT assay and found that low concentrations (0.33 and 1 μM) of eIF4A3-IN-2 did not cause cell death while high concentrations (3.33 and 10 μM) are toxic to the cells (Supplementary Figure S9B).

We also tested the subcellular localization of *Ccnb1* mRNAs in the control and eIF4A3-IN-2-treated mouse ESCs. While *Ccnb1* mRNAs are preferentially located in the cytoplasm of the control (DMSO) ESCs (Supplementary Figure S9C, D, left), eIF4A3-IN-2 treatment markedly increased the nuclear retention of *Ccnb1* transcripts (Supplementary Figure S9C, D, right). These results indicate that this specific eIF4A3 inhibitor blocks the nuclear export of *Ccnb1* mRNAs, mimicking eIF4A3 KD effects in ESCs (Figure 6D–F).

Together, these results complement and reinforce our findings using eIF4A3 shRNAs and suggest that this inhibitor could be useful for studying eIF4A3 in various cell types, e.g. treating cancers with elevated eIF4A3 and Cyclin B1 expression (40,56).

DISCUSSION

Starting with a specific RNAi screen of Ddx/Dhx RNA helicases, we identified and validated the critical roles of eIF4A3 in pluripotency maintenance. Though eIF4A3 was also identified as essential for ESCs in previous screens (57,58), its specific roles in pluripotency were not explored, and whether eIF4A3 functions through the EJC in ESCs remains undefined. Many studies of eIF4A3/EJC-related pathophysiological defects (39,40,59,60) focused on the transcriptome changes associated with the RNA splicing and NMD effects upon eIF4A3/EJC alterations (61–64). In contrast, our study of eIF4A3/EJCs in ESCs revealed a post-transcriptional role for eIF4A3 in targeting and regulating a pluripotency-specific cell cycle network beyond RNA splicing and NMD functions. In wild-type ESCs (equivalent to control KD; Figure 7, left part), eIF4A3 targets the pluripotency-associated *Ccnb1* mRNA encoding Cyclin B1 for nuclear export. The abundance of Cyclin B1 protein, required for ESC self-renewal and pluripotency (11), depends on the efficient mRNA nuclear export mediated by eIF4A3. However, upon eIF4A3 KD, *Ccnb1* mRNAs are retained in the nucleus and thus cannot be translated, leading to the down-regulation of Cyclin B1. Consequently, the loss of Cyclin B1 disrupts the cell cycle progression by arresting the G_2–M transition, resulting in the loss of ESC maintenance and/or differentiation (Figure 7, right part).

eIF4A3 is commonly known as a core component of the EJC. Previous studies show that the EJC is a multifunctional protein complex that affects diverse aspects of RNA metabolism, including splicing, mRNA nuclear export, mRNA translation and mRNA stability (14). Our study suggests that, in maintaining pluripotency, eIF4A3 mainly exports the mature spliced mRNAs encoding cell cycle regulators with minimal roles in their splicing or RNA stability. Interestingly, we noticed an intronic signal in *Ccnb1* (Figure 5D, the shaded long green box, eCLIP-seq data). EJC forms on a pre-mRNA and becomes more stable on a spliced exon–exon junction (14). The intronic peak in *Ccnb1* indicates that eIF4A3 can bind to that intronic part, although eIF4A3 KD does not influence the splicing of *Ccnb1* in ESCs (Figure 5F). We also noticed that, upon eIF4A3 KD, the reduction of Cyclin B1 protein was more pronounced than that of eIF4A3 (Figure 6G, H) and that the RPF reduction on *Ccnb1* mRNA mainly happens at the translation initiation region (5′ UTR and the beginning of the coding region) (Figure 5D, shaded in purple, the RPF tracks). In ribosome profiling, it is well known that for the proteins bound at translation initiation regions, their expression changes can lead to RPF alteration mainly at these translation initiation regions, with much less or even no alteration on the rest of mRNA bodies, and such a reduction can dramatically decrease the protein levels of these mRNAs (29). This could apply to eIF4A3 KD-induced Ccnb1 change in our study, considering that eIF4A3 is mainly bound around the first exon–exon junction of *Ccnb1* mRNA with the enrichment at the translation initiation region (Figure 5D). Therefore, eIF4A3 may regulate Ccnb1 protein synthesis through indirect and direct translational control via nuclear export of and binding to *Ccnb1* mRNA, respectively.

The helicase activity of eIF4A3 is required for both the maintenance of ESCs (Figure 2H–J) and the efficient nuclear export of *Ccnb1* mRNAs (Figure 6D–F). Such a demand for the helicase activity indicates that eIF4A3 may

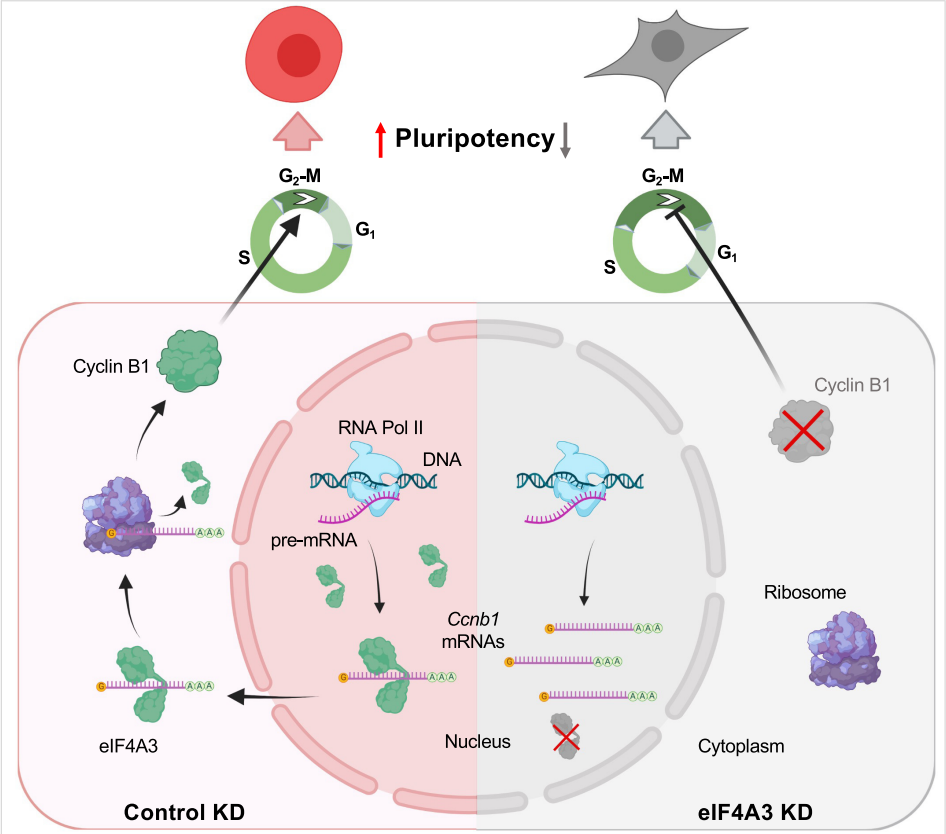


Figure 7. A summary model depicting how eIF4A3 may control the cell cycle to safeguard ESC identity. Briefly, in control KD ESCs (left, representing the normal ESC state), eIF4A3 directly targets *Ccnb1* mRNAs and transports them to the cytoplasm for mRNA translation. *Ccnb1* encodes Cyclin B1, which is required for the cell cycle progression and promotes pluripotency of ESCs. However, in eIF4A3 KD ESCs (right), *Ccnb1* mRNAs cannot be transported to the cytoplasm, resulting in the translational inhibition of Cyclin B1. The lack of Cyclin B1 causes the G_2–M arrest and disrupts the pluripotency of ESCs.

play a role in the remodeling of *Ccnb1* transcripts to form specific RNA secondary structures conducive to binding nuclear export proteins. Such remodeling may be important to ensure that only fully matured transcripts, including *Ccnb1* mRNAs, are exported for translation in the cytoplasm (7,65). eIF4A3 is deposited on exon–exon junctions and can serve as a docking platform for other mRNA export factors (14). Accordingly, it may also recruit other RNA helicases for RBP–RNA remodeling during RNA export, such as Ddx3x, Ddx5, Ddx21 and Mov10, all identified in the eIF4A3 ESC interactome (Figure 3A; Supplementary Table S3). We also noted the presence of other nuclear export factors in the eIF4A3 interactome, including the Thoc (THO transcription factor/nuclear export complex) family proteins such as Thoc2/Thoc5 that can direct the nuclear export of pluripotency transcripts (66,67) and Zc3h14 that is located on the nucleoplasmic side of the nuclear pore complex (14), compatible with the finding that eIF4A3 is more enriched in the ESC nuclei (Supplementary Figure S4D, F). Additional interacting partners such as ribosomal proteins and spliceosome components are considered essential proteins, which may function together with eIF4A3 for ESC survival. Our studies, however, highlight their specific roles in pluripotency, as previously validated for Thoc2/Thoc5 (66,67) and Cyclin B1 (11), as well as for ribosomes (4) and splicing factors (68–70) (Supplementary Table S3). Our eIF4A3 interactome encompassing the EJC in ESCs will provide a rich resource to further explore the functional relationships between those factors and eIF4A3/EJC-mediated RNA metabolism in controlling pluripotency. For example, the dramatic post-transcriptional effect of Casc3 KD on Oct4 expression (compare the RNA in Figure 3D and the protein in Figure 3E) in ESC pluripotency warrants further investigation.

Moreover, from our eIF4A3 eCLIP-seq in ESCs, apart from the deposition on the canonical EJC regions (the 3′ exon boundary motif present in 42.21% of the targets, Figure 4C), eIF4A3 also binds to non-canonical EJC regions in ESCs, such as 5′ exon boundaries and, to a much lesser extent, introns (Figure 4B, D), as in other non-pluripotent cells (71). Ccnb1 exemplifies this, with the second and third binding peaks representing an intronic region and a 5′ exon boundary, respectively (shaded in green from right, Figure 5D). It would be interesting to investigate whether the other EJC components or interacting partners identified in the interactome bind to these non-canonical regions together with eIF4A3 and how these non-canonical binding events may contribute to eIF4A3's functions in ESCs.

In addition to *Ccnb1* mRNAs, many other mRNAs encoding cell cycle regulators were also subject to direct post-transcriptional regulation by eIF4A3 (Figure 5C; Supple-

mentary Figure S7D). Therefore, it is plausible that the dysregulation of these cell cycle regulators upon eIF4A3 depletion could also disrupt the cell cycle pattern in ESCs. eIF4A3 is known to be phosphorylated by CDK1 and CDK2 in a cell cycle-dependent manner, and such modification affects the RNA binding capacity of eIF4A3 and triggers EJC remodeling (72). While our eCLIP-seq and ribosome profiling data conclusively establish the post-transcriptional effects of eIF4A3 loss on the expression of the key components of the cell cycle machinery (Figure 5A–C) and our RNA-seq data also imply the effects of eIF4A3 loss on the transcript levels of some cell cycle regulators (Figure 2D, E), it remains to be determined how eIF4A3/EJC and the cell cycle machinery may cross-talk in shaping the unique cell cycle pattern in ESCs. Moreover, the immunofluorescence results of Ccnb1 overexpression in eIF4A3 KD ESCs (Figure 6I) suggest that, apart from the pluripotency induction via activation of Ccnb1, eIF4A3 may also repress differentiation through other pathways that cannot be rescued by Ccnb1 overexpression in eIF4A3 KD ESCs. Such repression of differentiation may be through other cell cycle factors whose transcripts are bound by eIF4A3. Interestingly, like pluripotent ESCs, neural stem cells have rapid cell division cycles, and the KD of *eIF4A3*, *Magoh* and *Rbm8a*, but not *Casc3*, leads to an increase in the G_2–M phase of neural stem cells (73). Also the *Magoh* mutant can prolong mitosis of neural progenitors (74). These results suggest a potentially common theme by which the EJC regulates cell cycles in embryonic and neural stem cells. Apart from the unique cell cycle pattern, ESCs also have special metabolic requirements to power rapid cell division (44,75) and their metabolism is quickly remodeled during the loss and acquisition of pluripotency (44–46). eIF4A3 KD down-regulates many metabolism-related genes, such as the genes involved in the glycolytic process (Figure 2F). Since cell cycle regulators may also have roles in the coordination of growth and metabolism (76,77), future studies are warranted to resolve how eIF4A3 may coordinate the cell cycle reorganization and metabolic remodeling in maintaining ESC pluripotency.

Finally, we restricted our screen to the maintenance of mouse ESCs in the conventional serum/LIF culture, representing a metastable pluripotent state. Therefore, it is conceivable that we may not identify Ddx/Dhx factors that play important roles in other pluripotent states such as naïve (2i/LIF) or primed (Fgf2/Activin) pluripotency or during the early differentiation process of ESCs. For example, Ddx1 and DDX5/DDX17 were reported to be critical for the maintenance of mouse ESCs cultured in serum/2i/LIF (78) and for neural differentiation of human PSCs (79), but they were missed in our screen. Since most Ddx/Dhx helicases have relatively high expression levels in ESCs (Supplementary Figure S3), their additional functions in stem cells and early development will probably be uncovered in future studies.

DATA AVAILABILITY

Sequencing data generated or used in this study can be found in the following NCBI GEO (https://www.ncbi.nlm.nih.gov/geo/) series accession numbers. RNA-seq for eIF4A3 KD: eIF4A3 KD data, GSE194012; control KD data, GSM4556193 (rnaseq_Con_rep1) and GSM4556194 (rnaseq_Con_rep2) in GSE150673. Note: the RNA-seq for these two studies were performed together in the same experimental setting, sharing the same control data. RNA-seq for Ccnb1 overexpression in eIF4A3 KD, GSE194012. eCLIP-seq data, GSE194012. Ribosome profiling: eIF4A3 KD data, GSE194012; control KD data, GSM4556259(totalRNA_Con_rep1), GSM4556260 (totalRNA_Con_rep2), GSM4556255 (RPF_Con_rep1) and GSM4556256 (RPF_Con_rep2) in GSE150676. Note: the ribosome profiling for these two studies were performed together in the same experimental setting, sharing the same control data. The eIF4A3 immunoprecipitation MS data have been deposited in the ProteomeXchange Consortium (http://www.proteomexchange.org/) via the PRIDE partner repository with the dataset identifier PXD031362.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank G. Yeo and E.V. Nostrand for eCLIP-seq technical advice, M.J. Moore for providing the eIF4A3 wild-type construct, the ISMMS genomics core facility, NYU Langone's genome technology center and Illumina Tech Support for technical assistance, and a shared instrumentation grant for the LSR II Flow Cytometer (S10RR027050) to the Columbia Center for Translational Immunology (CCTI) Flow Cytometry Core. The figures (Figures 6A, 7; Supplementary Figures S6A, S7A) were created with BioRender.com.

Author contributions: D.L. designed and conducted the experiments, and wrote the manuscript. D.L., J.Y., V.M., and X.H. performed the data analysis. Y.H. and H.Z. provided reagents and experimental support. J.W. conceived, designed and supervised the studies, and wrote and approved the final manuscript.

FUNDING

National Institute of General Medical Sciences [GM129157]; New York State Stem Cell Science [C32583GG and C32569GG]; National Institute of Child Health and Human Development [HD095938 and HD097268]. Funding for open access charge: NIH.

Conflict of interest statement. None declared.

REFERENCES

1. Young,R.A. (2011) Control of the embryonic stem cell state. *Cell*, **144**, 940–954.
2. Theunissen,T.W. and Jaenisch,R. (2017) Mechanisms of gene regulation in human embryos and pluripotent stem cells. *Development*, **144**, 4496–4509.
3. Li,D., Kishta,M.S. and Wang,J. (2020) Regulation of pluripotency and reprogramming by RNA binding proteins. *Curr. Top. Dev. Biol.*, **138**, 113–138.
4. Li,D. and Wang,J. (2020) Ribosome heterogeneity in stem cells and development. *J. Cell Biol.*, **219**, e202001108.

5. Chua,B.A., Van Der Werf,I., Jamieson,C. and Signer,R.A.J. (2020) Post-transcriptional regulation of homeostatic, stressed, and malignant stem cells. *Cell Stem Cell*, **26**, 138–159.
6. Saba,J.A., Liakath-Ali,K., Green,R. and Watt,F.M. (2021) Translational control of stem cell function. *Nat. Rev. Mol. Cell Biol.*, **22**, 671–690.
7. Bourgeois,C.F., Mortreux,F. and Auboeuf,D. (2016) The multiple functions of RNA helicases as drivers and regulators of gene expression. *Nat. Rev. Mol. Cell Biol.*, **17**, 426–438.
8. Liu,L., Michowski,W., Kolodziejczyk,A. and Sicinski,P. (2019) The cell cycle in stem cell proliferation, pluripotency and differentiation. *Nat. Cell Biol.*, **21**, 1060–1067.
9. Boward,B., Wu,T. and Dalton,S. (2016) Concise review: control of cell fate through cell cycle and pluripotency networks. *Stem Cells*, **34**, 1427–1436.
10. Soufi,A. and Dalton,S. (2016) Cycling through developmental decisions: how cell cycle dynamics control pluripotency, differentiation and reprogramming. *Development*, **143**, 4301–4311.
11. Gonzales,K.A., Liang,H., Lim,Y.S., Chan,Y.S., Yeo,J.C., Tan,C.P., Gao,B., Le,B., Tan,Z.Y., Low,K.Y. *et al.* (2015) Deterministic restriction on pluripotent state dissolution by cell-cycle pathways. *Cell*, **162**, 564–579.
12. Sergeeva,O. and Zatsepin,T. (2021) RNA helicases as shadow modulators of cell cycle progression. *Int. J. Mol. Sci.*, **22**, 2984.
13. Zhang,L. and Li,X. (2021) DEAD-Box RNA helicases in cell cycle control and clinical therapy. *Cells*, **10**, 1540.
14. Le Hir,H., Sauliere,J. and Wang,Z. (2016) The exon junction complex as a node of post-transcriptional networks. *Nat. Rev. Mol. Cell Biol.*, **17**, 41–54.
15. Ivanova,N., Dobrin,R., Lu,R., Kotenko,I., Levorse,J., DeCoste,C., Schafer,X., Lun,Y. and Lemischka,I.R. (2006) Dissecting self-renewal in stem cells with RNA interference. *Nature*, **442**, 533–538.
16. 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.
17. Bustin,S.A., Benes,V., Garson,J.A., Hellemans,J., Huggett,J., Kubista,M., Mueller,R., Nolan,T., Pfaffl,M.W., Shipley,G.L. *et al.* (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.*, **55**, 611–622.
18. Guo,C.J., Ma,X.K., Xing,Y.H., Zheng,C.C., Xu,Y.F., Shan,L., Zhang,J., Wang,S., Wang,Y., Carmichael,G.G. *et al.* (2020) Distinct processing of lncRNAs contributes to non-conserved functions in stem cells. *Cell*, **181**, 621–636.
19. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with bowtie 2. *Nat. Methods*, **9**, 357–359.
20. Durinck,S., Moreau,Y., Kasprzyk,A., Davis,S., De Moor,B., Brazma,A. and Huber,W. (2005) BioMart and bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*, **21**, 3439–3440.
21. Robinson,M.D. and Oshlack,A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.*, **11**, R25–R25.
22. McCarthy,D.J., Chen,Y. and Smyth,G.K. (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.*, **40**, 4288–4297.
23. Van Nostrand,E.L., Pratt,G.A., Shishkin,A.A., Gelboin-Burkhart,C., Fang,M.Y., Sundararaman,B., Blue,S.M., Nguyen,T.B., Surka,C., Elkins,K. *et al.* (2016) Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods*, **13**, 508–514.
24. Cameron,V. and Uhlenbeck,O.C. (1977) 3′-Phosphatase activity in T4 polynucleotide kinase. *Biochemistry*, **16**, 5120–5126.
25. Bao,W., Kojima,K.K. and Kohany,O. (2015) Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA*, **6**, 11.
26. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
27. Cui,X., Wei,Z., Zhang,L., Liu,H., Sun,L., Zhang,S.W., Huang,Y. and Meng,J. (2016) Guitar: an R/Bioconductor package for gene annotation guided transcriptomic analysis of RNA-related genomic features. *Biomed. Res. Int.*, **2016**, 8367534.
28. Chen,C., Zhang,X., Wang,Y., Chen,X., Chen,W., Dan,S., She,S., Hu,W., Dai,J., Hu,J. *et al.* (2021) Translational and post-translational control of human naïve versus primed pluripotency. *Iscience*, **25**, 103645.
29. Li,D., Yang,J., Huang,X., Zhou,H. and Wang,J. (2022) eIF4A2 targets developmental potency and histone H3.3 transcripts for translational control of stem cell pluripotency. *Sci. Adv.*, **8**, eabm0478.
30. Nicklas,S., Okawa,S., Hillje,A.L., Gonzalez-Cano,L., Del Sol,A. and Schwamborn,J.C. (2015) The RNA helicase DDX6 regulates cell-fate specification in neural stem cells via miRNAs. *Nucleic Acids Res.*, **43**, 2638–2654.
31. Wang,Y., Arribas-Layton,M., Chen,Y., Lykke-Andersen,J. and Sen,GL. (2015) DDX6 orchestrates mammalian progenitor function through the mRNA degradation and translation pathways. *Mol. Cell*, **60**, 118–130.
32. Freimer,J.W., Hu,T.J. and Blelloch,R. (2018) Decoupling the impact of microRNAs on translational repression versus RNA degradation in embryonic stem cells. *Elife*, **7**, e38014.
33. Di Stefano,B., Luo,E.C., Haggerty,C., Aigner,S., Charlton,J., Brumbaugh,J., Ji,F., Rabano Jiménez,I., Clowers,K.J., Huebner,A.J. *et al.* (2019) The RNA helicase DDX6 controls cellular plasticity by modulating P-body homeostasis. *Cell Stem Cell*, **25**, 622–638.
34. Zhang,H., Wu,Z., Lu,J.Y., Huang,B., Zhou,H., Xie,W., Wang,J. and Shen,X. (2020) DEAD-box helicase 18 counteracts PRC2 to safeguard ribosomal DNA in pluripotency regulation. *Cell Rep.*, **30**, 81–97.
35. You,K.T., Park,J. and Kim,V.N. (2015) Role of the small subunit processome in the maintenance of pluripotent stem cells. *Genes Dev.*, **29**, 2004–2009.
36. Kervestin,S. and Jacobson,A. (2012) NMD: a multifaceted response to premature translational termination. *Nat. Rev. Mol. Cell Biol.*, **13**, 700–712.
37. Boehm,V. and Gehring,N.H. (2016) Exon junction complexes: supervising the gene expression assembly line. *Trends Genet.*, **32**, 724–735.
38. Woodward,L.A., Mabin,J.W., Gangras,P. and Singh,G. (2017) The exon junction complex: a lifelong guardian of mRNA fate. *Wiley Interdiscip Rev RNA*, **8**, https://doi.org/10.1002/wrna.1411.
39. Kanellis,D.C., Espinoza,J.A., Zisi,A., Sakkas,E., Bartkova,J., Katsori,A.M., Bostrom,J., Dyrskjot,L., Broholm,H., Altun,M. *et al.* (2021) The exon-junction complex helicase eIF4A3 controls cell fate via coordinated regulation of ribosome biogenesis and translational output. *Sci. Adv.*, **7**, eabf7561.
40. Zhu,Y., Ren,C. and Yang,L. (2021) Effect of eukaryotic translation initiation factor 4A3 in malignant tumors. *Oncol. Lett.*, **21**, 358.
41. Malik,V., Glaser,L.V., Zimmer,D., Velychko,S., Weng,M., Holzner,M., Arend,M., Chen,Y., Srivastava,Y., Veerapandian,V. *et al.* (2019) Pluripotency reprogramming by competent and incompetent POU factors uncovers temporal dependency for oct4 and sox2. *Nat. Commun.*, **10**, 3477.
42. Xiong,Z., Xu,K., Lin,Z., Kong,F., Wang,Q., Quan,Y., Sha,Q.Q., Li,F., Zou,Z., Liu,L. *et al.* (2022) Ultrasensitive Ribo-seq reveals translational landscapes during mammalian oocyte-to-embryo transition and pre-implantation development. *Nat. Cell Biol.*, **24**, 968–980.
43. Gifford,C.A., Ziller,M.J., Gu,H., Trapnell,C., Donaghey,J., Tsankov,A., Shalek,A.K., Kelley,D.R., Shishkin,A.A., Issner,R. *et al.* (2013) Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell*, **153**, 1149–1163.
44. Somasundaram,L., Levy,S., Hussein,A.M., Ehnes,D.D., Mathieu,J. and Ruohola-Baker,H. (2020) Epigenetic metabolites license stem cell states. *Curr. Top. Dev. Biol.*, **138**, 209–240.
45. Mathieu,J. and Ruohola-Baker,H. (2017) Metabolic remodeling during the loss and acquisition of pluripotency. *Development*, **144**, 541–551.
46. D'Aniello,C., Cermola,F., Patriarca,E.J. and Minchiotti,G. (2019) Metabolic–epigenetic axis in pluripotent state transitions. *Epigenomes*, **3**, 13.
47. Pause,A. and Sonenberg,N. (1992) Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *EMBO J.*, **11**, 2643–2654.
48. Schlautmann,L.P. and Gehring,N.H. (2020) A day in the life of the exon junction complex. *Biomolecules*, **10**, 866.

49. Hauer,C., Sieber,J., Schwarzl,T., Hollerer,I., Curk,T., Alleaume,A.M., Hentze,M.W. and Kulozik,A.E. (2016) Exon junction complexes show a distributional bias toward alternatively spliced mRNAs and against mRNAs coding for ribosomal proteins. *Cell Rep.*, **16**, 1588–1603.
50. Ingolia,N.T., Ghaemmaghami,S., Newman,J.R.S. and Weissman,J.S. (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, **324**, 218–223.
51. Dominic Heng,J.-C., Feng,B., Han,J., Jiang,J., Kraus,P., Ng,J.-H., Orlov,Y.L., Huss,M., Yang,L., Lufkin,T. *et al.*(2010)The nuclear receptor Nr5a2 can replace oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell*, **6**, 167–174.
52. Genomics,B.M.C., Hailesellasse Sene,K., Porter,C.J., Palidwor,G., Perez-Iratxeta,C., Muro,E.M., Campbell,P.A., Rudnicki,M.A. and Andrade-Navarro,M.A.(2008)Gene function in early mouse embryonic stem cell differentiation.*BMC Genomics*, **8**, 85.
53. Strauss,B., Harrison,A., Coelho,P.A., Yata,K., Zernicka-Goetz,M. and Pines,J. (2018) Cyclin B1 is essential for mitosis in mouse embryos, and its nuclear export sets the time for mitosis. *J. Cell Biol.*, **217**, 179–193.
54. Ideue,T., Sasaki,Y.T., Hagiwara,M. and Hirose,T. (2007) Introns play an essential role in splicing-dependent formation of the exon junction complex. *Genes Dev.*, **21**, 1993–1998.
55. Iwatani-Yoshihara,M., Ito,M., Ishibashi,Y., Oki,H., Tanaka,T., Morishita,D., Ito,T., Kimura,H., Imaeda,Y., Aparicio,S. *et al.* (2017) Discovery and characterization of a eukaryotic initiation factor 4A-3-selective inhibitor that suppresses nonsense-mediated mRNA decay. *ACS Chem. Biol.*, **12**, 1760–1768.
56. Ye,C., Wang,J., Wu,P., Li,X. and Chai,Y. (2017) Prognostic role of cyclin B1 in solid tumors: a meta-analysis. *Oncotarget*, **8**, 2224–2232.
57. Fazzio,T.G., Huff,J.T. and Panning,B. (2008) An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell*, **134**, 162–174.
58. Hackett,J.A., Huang,Y., Günesdogan,U., Gretarsson,K.A., Kobayashi,T. and Surani,M.A. (2018) Tracing the transitions from pluripotency to germ cell fate with CRISPR screening. *Nat. Commun.*, **9**, 4292.
59. McMahon,J.J., Miller,E.E. and Silver,D.L. (2016) The exon junction complex in neural development and neurodevelopmental disease. *Int. J. Dev. Neurosci.*, **55**, 117–123.
60. Miller,E.E., Kobayashi,G.S., Musso,C.M., Allen,M., Ishiy,F.A.A., de Caires,L.C. Jr, Goulart,E., Griesi-Oliveira,K., Zechi-Ceide,R.M., Richieri-Costa,A *et al.* (2017) EIF4A3 deficient human iPSCs and mouse models demonstrate neural crest defects that underlie Richieri–Costa–Pereira syndrome. *Hum. Mol. Genet.*, **26**, 2177–2191.
61. Akhtar,J., Kreim,N., Marini,F., Mohana,G., Brune,D., Binder,H. and Roignant,J.Y. (2019) Promoter-proximal pausing mediated by the exon junction complex regulates splicing. *Nat. Commun.*, **10**, 521.
62. Leung,C.S. and Johnson,T.L. (2018) The exon junction complex: a multitasking guardian of the transcriptome. *Mol. Cell*, **72**, 799–801.
63. Gangras,P., Gallagher,T.L., Parthun,M.A., Yi,Z., Patton,R.D., Tietz,K.T., Deans,N.C., Bundschuh,R., Amacher,S.L. and Singh,G. (2020) Zebrafish rbm8a and magoh mutants reveal EJC developmental functions and new 3′UTR intron-containing NMD targets. *PLoS Genet.*, **16**, e1008830.
64. Zhang,Y. and Sachs,M.S. (2015) Control of mRNA stability in fungi by NMD, EJC and CBC factors through 3′UTR introns. *Genetics*, **200**, 1133–1148.
65. Stewart,M. (2019) Polyadenylation and nuclear export of mRNAs. *J. Biol. Chem.*, **294**, 2977–2987.
66. Wang,L., Miao,Y.L., Zheng,X., Lackford,B., Zhou,B., Han,L., Yao,C., Ward,J.M., Burkholder,A., Lipchina,I. *et al.* (2013) The THO complex regulates pluripotency gene mRNA export and controls embryonic stem cell self-renewal and somatic cell reprogramming. *Cell Stem Cell*, **13**, 676–690.
67. Saunders,A. and Wang,J. (2014) Export and expression: mRNAs deliver new messages for controlling pluripotency. *Cell Stem Cell*, **14**, 549–550.
68. Li,Y., Wang,D., Wang,H., Huang,X., Wen,Y., Wang,B., Xu,C., Gao,J., Liu,J., Tong,J. *et al.* (2021) A splicing factor switch controls hematopoietic lineage specification of pluripotent stem cells. *EMBO Rep.*, **22**, e50535.
69. Ratnadiwakara,M., Archer,S.K., Dent,C.I., Ruiz De Los Mozos,I., Beilharz,T.H., Knaupp,A.S., Nefzger,C.M., Polo,J.M. and Anko,M.L. (2018) SRSF3 promotes pluripotency through nanog mRNA export and coordination of the pluripotency gene expression program. *Elife*, **7**, e37419.
70. Kim,Y.D., Lee,J., Kim,H.S., Lee,M.O., Son,M.Y., Yoo,C.H., Choi,J.K., Lee,S.C. and Cho,Y.S. (2017) The unique spliceosome signature of human pluripotent stem cells is mediated by SNRPA1, SNRPD1, and PNN. *Stem Cell Res.*, **22**, 43–53.
71. Saulière,J., Murigneux,V., Wang,Z., Marquenet,E., Barbosa,I., Le Tonquèze,O., Audic,Y., Paillard,L., Roest Crollius,H. and Le Hir,H. (2012) CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nat. Struct. Mol. Biol.*, **19**, 1124–1131.
72. Incheol Ryu,A., Won,Y.-S., Ha,H., Kyu Song,H., Jung,H., Ki Kim Correspondence,Y., Ryu,I., Kim,E., Park,Y., Kyung Kim,M. *et al.* (2019) eIF4A3 phosphorylation by CDKs affects NMD during the cell cycle. *Cell Rep.*, **26**, 2126–2139.
73. Silver,D.L., Watkins-Chow,D.E., Schreck,K.C., Pierfelice,T.J., Larson,D.M., Burnetti,A.J., Liaw,H.J., Myung,K., Walsh,C.A., Gaiano,N. *et al.* (2010) The exon junction complex component magoh controls brain size by regulating neural stem cell division. *Nat. Neurosci.*, **13**, 551–558.
74. Pilaz,L.J., McMahon,J.J., Miller,E.E., Lennox,A.L., Suzuki,A., Salmon,E. and Silver,D.L. (2016) Prolonged mitosis of neural progenitors alters cell fate in the developing brain. *Neuron*, **89**, 83–99.
75. Dahan,P., Lu,V., Nguyen,R.M.T., Kennedy,S.A.L. and Teitell,M.A. (2019) Metabolism in pluripotency: both driver and passenger?*J. Biol. Chem.*, **294**, 5420–5429.
76. Burgess,R.J., Agathocleous,M. and Morrison,S.J. (2014) Metabolic regulation of stem cell function. *J. Intern. Med.*, **276**, 12–24.
77. Huber,K., Mestres-Arenas,A., Fajas,L. and Leal-Esteban,L.C. (2021) The multifaceted role of cell cycle regulators in the coordination of growth and metabolism. *FEBS J.*, **288**, 3813–3833.
78. Suzuki,T., Katada,E., Mizuoka,Y., Takagi,S., Kazuki,Y., Oshimura,M., Shindo,M. and Hara,T. (2021) A novel all-in-one conditional knockout system uncovered an essential role of DDX1 in ribosomal RNA processing. *Nucleic Acids Res.*, **49**, e40.
79. Suthapot,P., Xiao,T., Felsenfeld,G., Hongeng,S. and Wongtrakoongate,P. (2022) The RNA helicases DDX5 and DDX17 facilitate neural differentiation of human pluripotent stem cells NTERA2. *Life Sci.*, **291**, 120298.