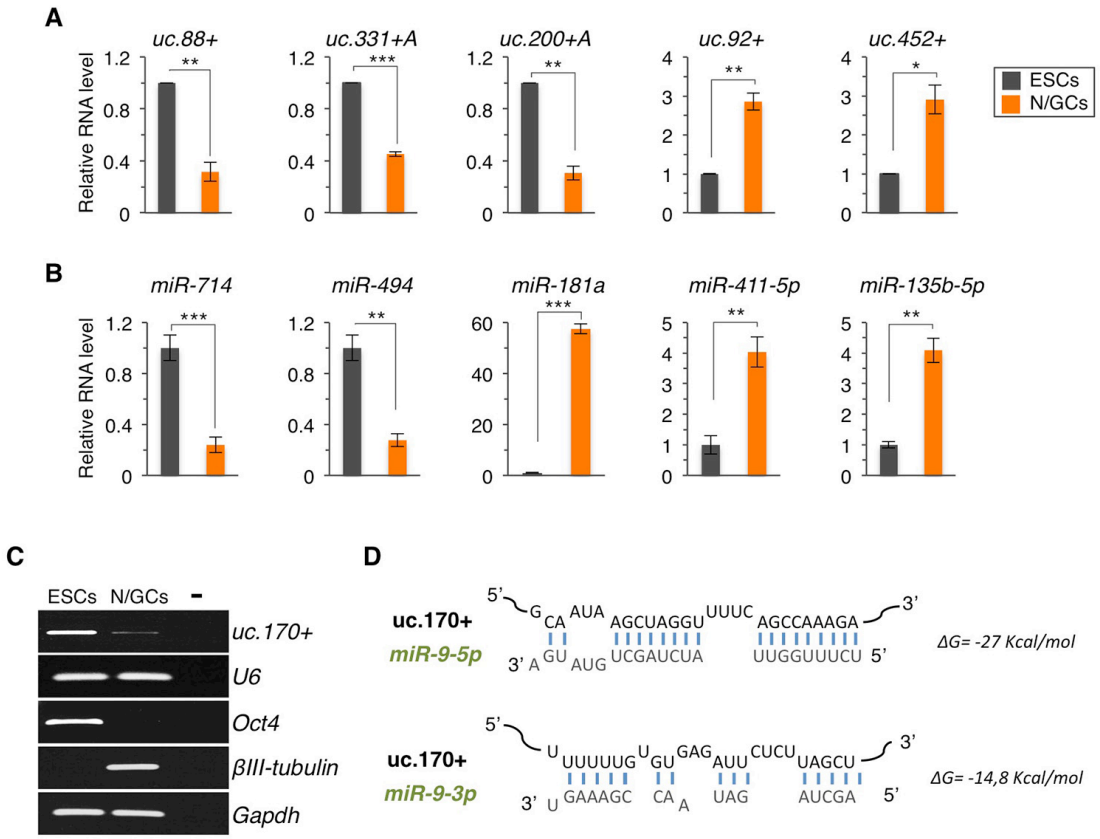
Stem Cell Reports, Volume *10*

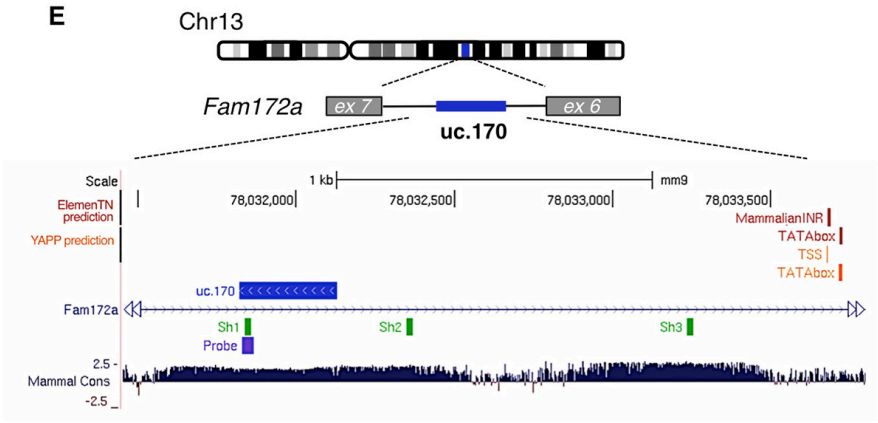
Supplemental Information

An Ultraconserved Element Containing lncRNA Preserves Transcriptional Dynamics and Maintains ESC Self-Renewal

Alessandro Fiorenzano, Emilia Pascale, Miriam Gagliardi, Sara Terreri, Mariarosaria Papa, Gennaro Andolfi, Marco Galasso, Guidantonio Malagoli Tagliazucchi, Cristian Taccioli, Eduardo Jorge Patriarca, Amelia Cimmino, Maria Rosaria Matarazzo, Gabriella Minchiotti, and Annalisa Fico

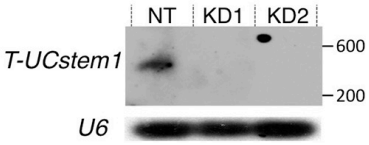
SUPPLEMENTAL INFORMATION





Supplementary Figure 1

A

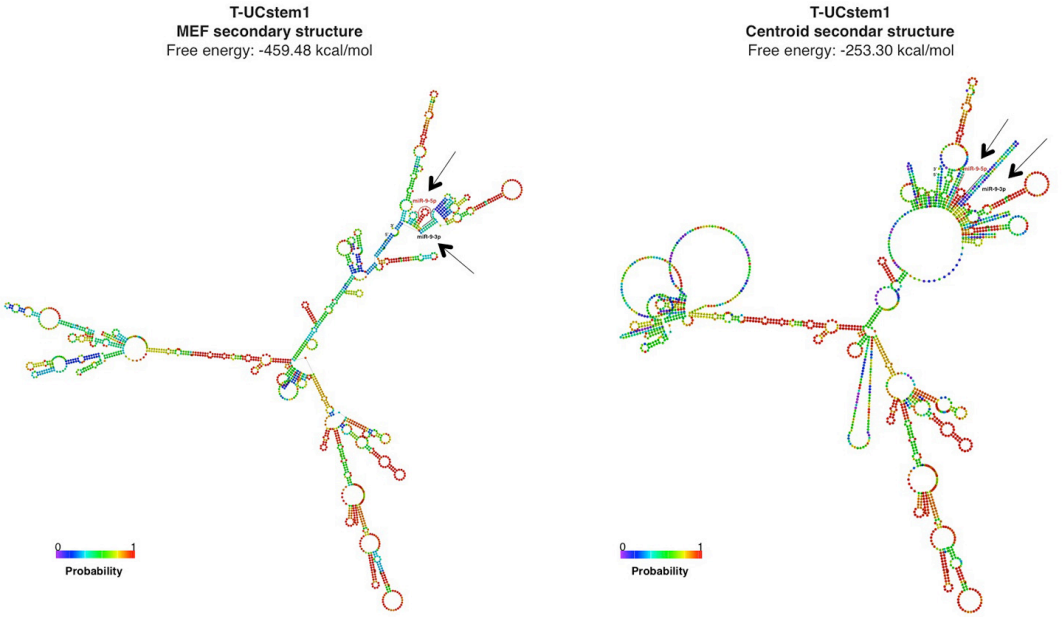


B

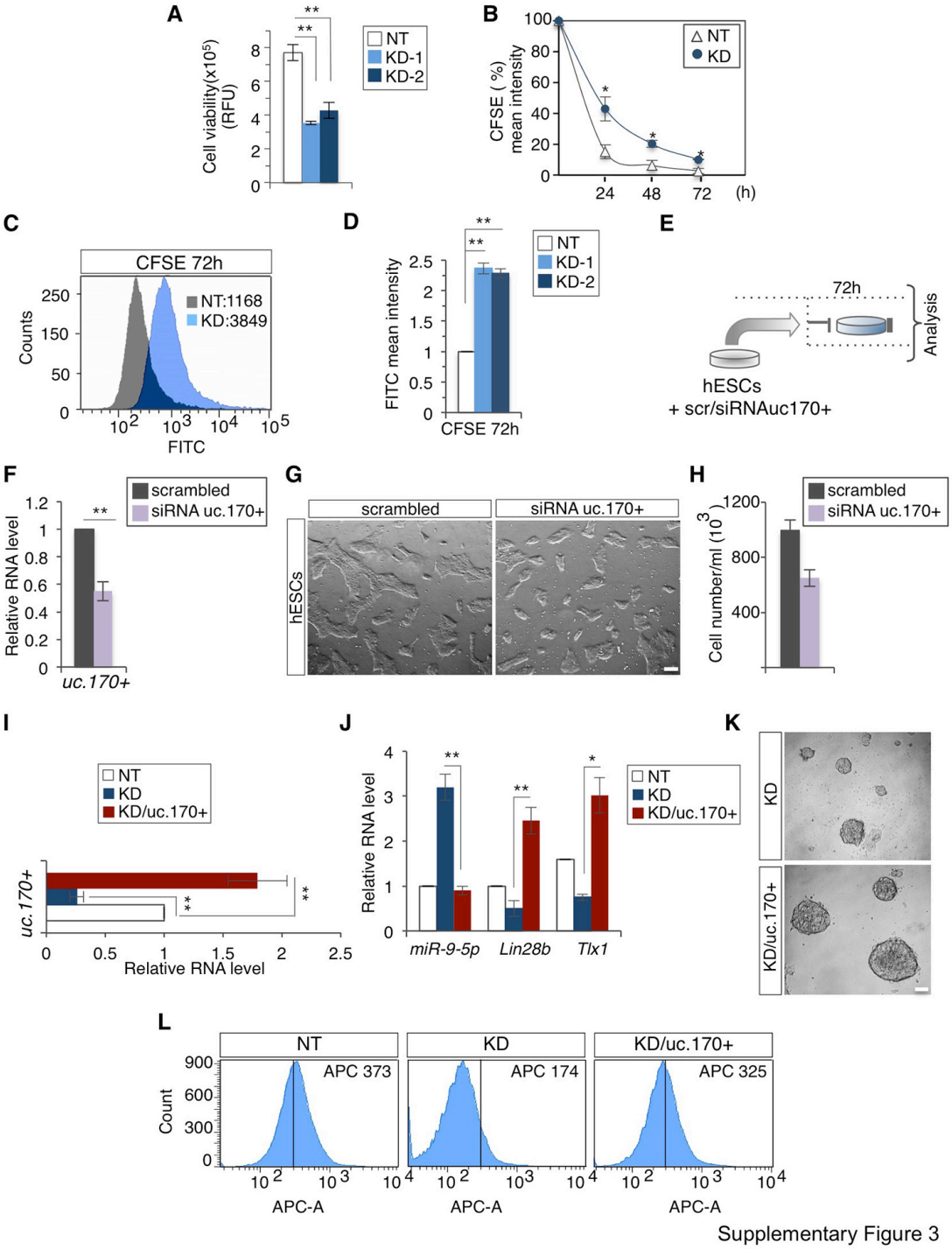
Chr13: 78031716-78033528, strand – (1813bp)

5'GAGTACGGGAGGGGGCTGTATGCCCATCCAAGTTCTGATGTGACATAATATTAAAGATAGT
TGTTCTATGGAAGGGAAGGGTTAATGGGGATTTTAAACTAATGAACCCAGTGACACAGGCTC
TGAAGTAAGCCCTGCATAGTTCTAAGCCAACCTTAGAATTTCAACCCTTCCAGCCCTGGCTTAG
ATCATACAGGCTCTACTAGAAATGTCAAGAATGTGAAGAGGCGTCTTTAGCCAAAAGGACAC
AACTTTAAGTACATAATGCTAGCTTACACTGGTTCGTTTATTCCCCTCTTCATTAGTGAGAAAAT
GGACTCTCGTAAATGCTTGTTAAAATCCCTACCTACACTCCTCTGGTTAATGTGAGTGATCCAT
TTGTGTTCATTTCTGCCTGACAACGCTTACTGGAAGATTAATTGCCCCATGTCAAACTGTAGCA
ACCCTCGTCCCCTTCTCATTTCCTTTTCTCTCTGCTGTTGAAGTAATAATGGACCTTGTCTTTTGT
TTTAAGACTGGGATTAGGAGAGAGTGAGTTTTTCTTCGGACGACTTTACGGAGTAGAGTTTGG
TTGAATACTTATAAGATTCTCCATGGAAAGGAAGCACATAGTCCTTTTGCCTACTGTTTTACTTC
ACACATCTTGCCATCTTTGTTTAGCCTTCATCCCTATATGATTCTGGTCTAAGTAAATCTACATG
CACCCTATACATACATGTCATTTTCTGTTCTCATACCTGGGGACTTCAAAATTGGCACTGATAAT
TGTCCCCTTCTACAAAGCACATATCCCTTGCAAAACATATTATTTCAGTTCTGTGCATAATGGA
GTCTTAAGTAGAATCCAAAACTCATAGAAAGACCATCTTTCAGGTTCTAAAATTTTATAAAACA
AAATAAAACAAAACAAAACTCAACAATAGCATAAAAATCATAACGGCAGCAACGATATTTAAA
AAAAAATACCCTCGTTGCTTGATTGAAATAAGCAAAAAGATATCACCAAATATAAATTCCAATT
CTCATTTTGCATTTTTTTCAGGAACATTAAAGAGTTTTATGTTATCACAAAACCTCAGAAACTAC
CAAGCCGAAAAGGAGGCTTATTATGAGTAGCCGGCAGTATCTCATTAAATCTCTCAATTGTCT
GATCGTCCACAGCAGACCTCCAAGAGACTTGTATATTATACTCATTGCAATTAACCAAACAAAA
GATTTTAACCGCCAGAGCATTCGGGAAATGTTTGGTTGGGGCTGAAGAAGTGAAATGATATT
CCAGGGCTGGCCAGATGTCACAAGGGGTGATATGCATGTGCTCACTTCATCTGCAGCTTTGTG
CAGGACCTGTCTATTTACAGCACTACAGCTAAGCACTCTGAAGGCCTATTCACTGAGGAGTCC
TTTCAGAAAGTGCTGAAGCACCCCTTAAGCCCACTTAACTACCATTTTCACACACTCTCCCAG
CTCTCCTTTTTGTCCTTGCTTACATTACATCAAACACAGGAACAAAGCAAGAGAACAGAAAC
TCAGAGGCAGAGAATAGACCCATCACAAATATTAATTTGAAAAGGTGTTGAAGTGCAGAAT
CTGCTTTTATGCACAAGGACAACTTGCATTTTTTGTGTGAGATTCTCTTAGCTGCAATAAGCT
AGGTTTTCAGCCAAAGAGAGGCAAAGACTCAAAGTGCAATTATACACAGGGAACTGCTTTA
AATCAAACAATGCTCCGAACTGCTTTAGATCTATAGTGATAAAGACTTGGCAAGCACTGTTAA
ATAGAAGCCCTATATGAGATGCAGAGTTCAGTCTATGGA 3'

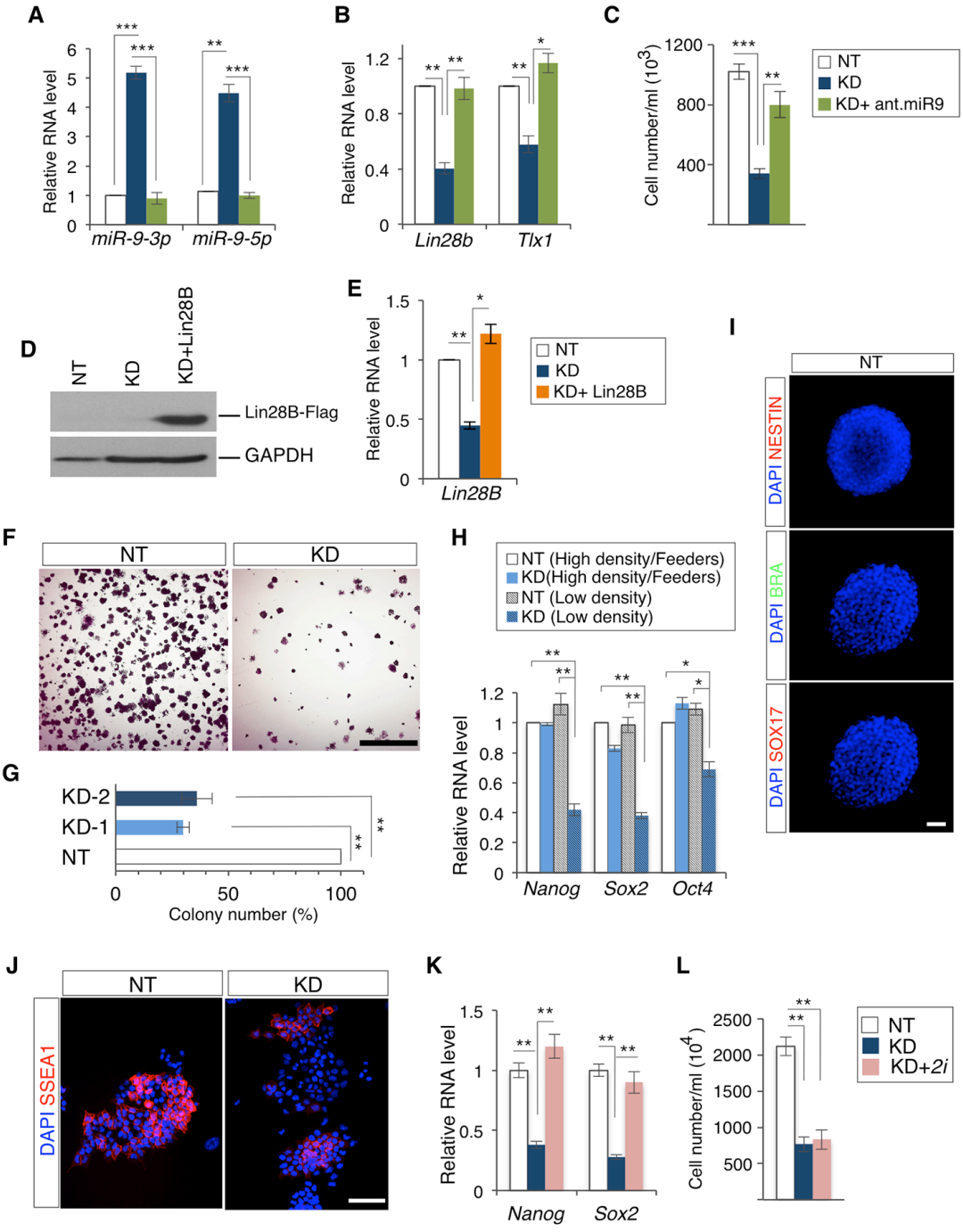
C



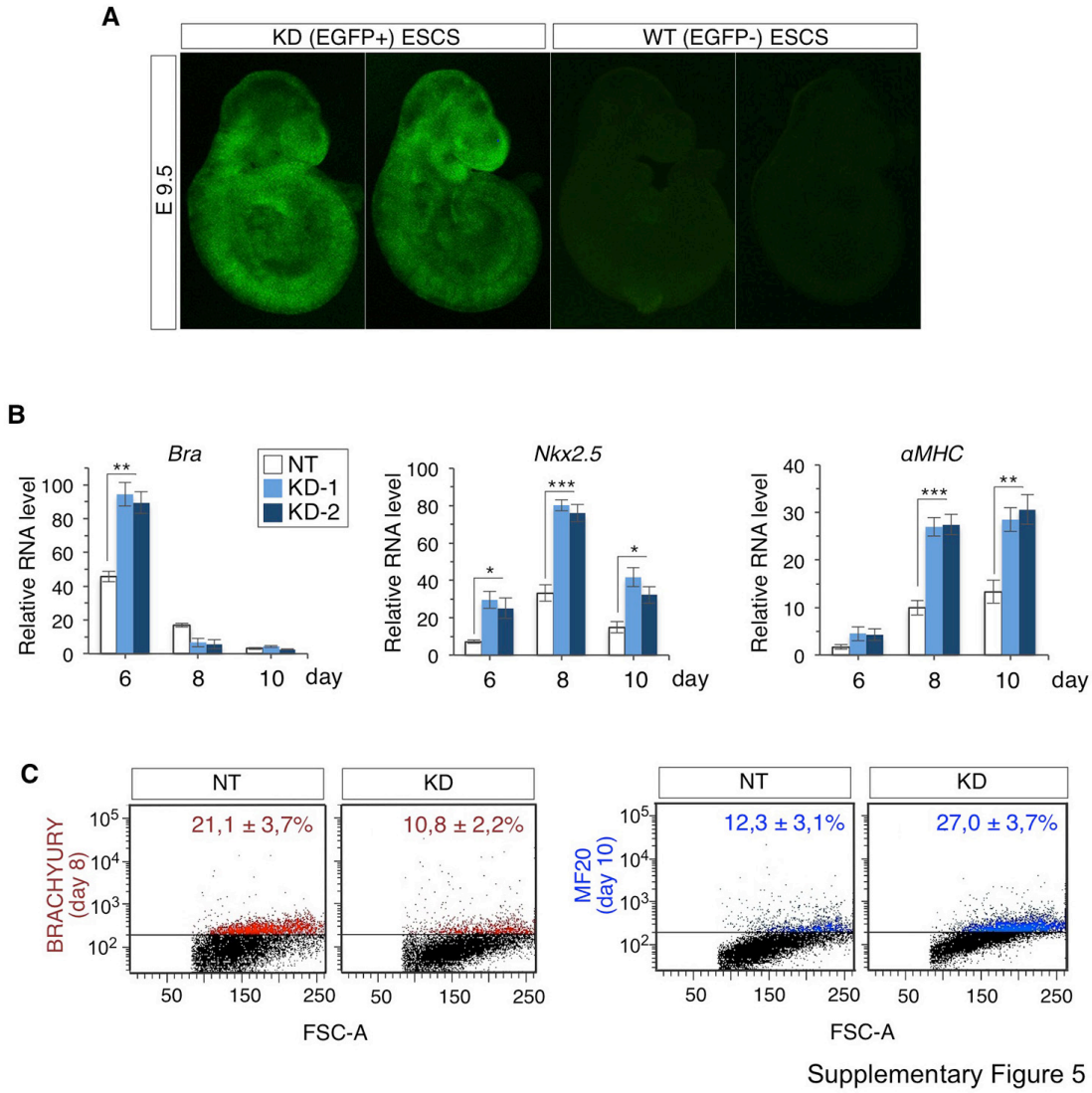
Supplementary Figure 2



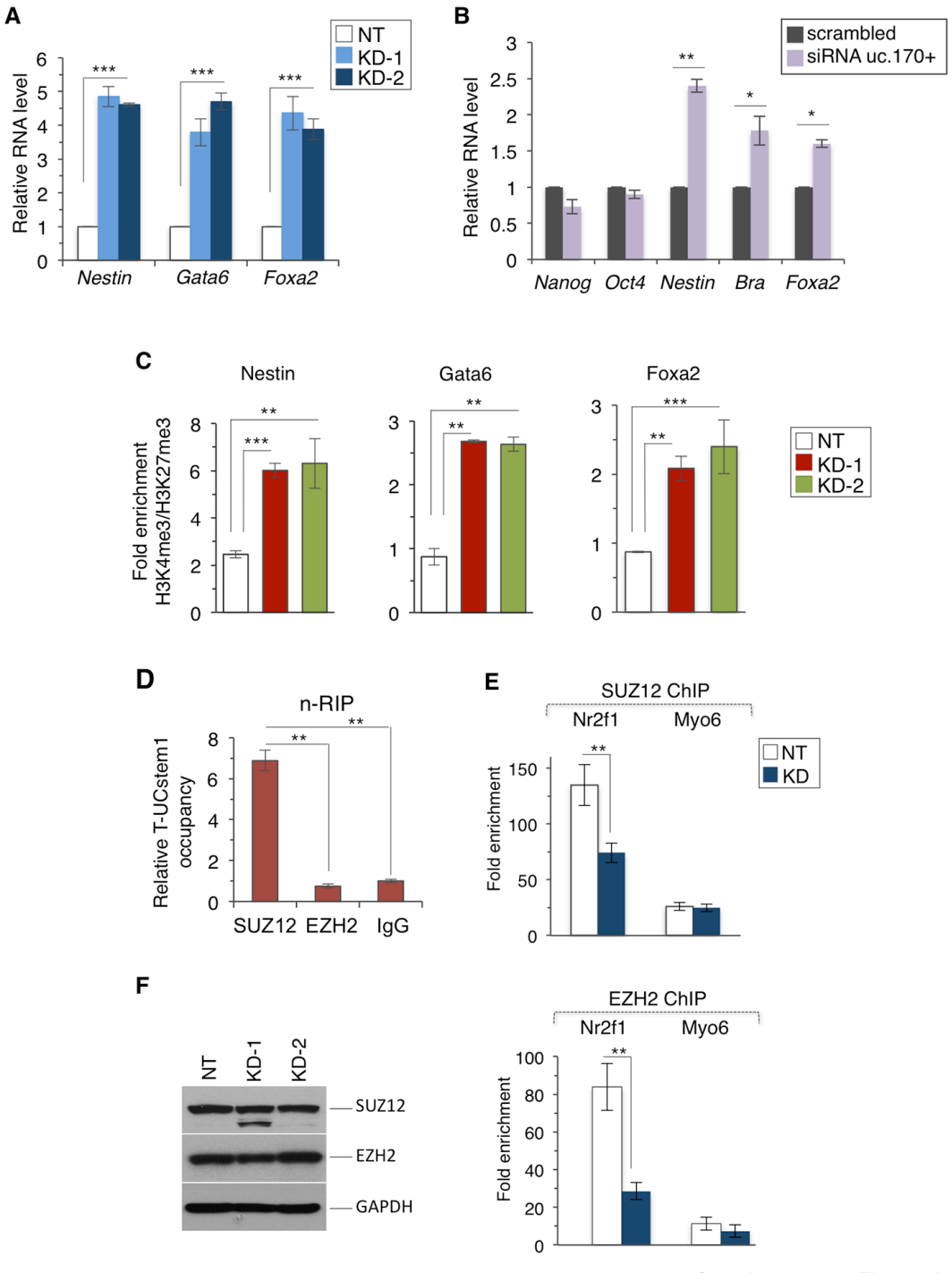
Supplementary Figure 3



Supplementary Figure 4



Supplementary Figure 5



Supplementary Figure 6

Table S3. Primers sets used in this study.

GENE NAME	FORWARD PRIMER	REVERSE PRIMER
T-UCstem1	TGAGTCTTTGCCTCTCTTTGG	AAGTGCTGAAGCACCCCTTA
Gapdh	TGGGGGAACTTAAAGTGCAG	GATGTAGGCAGCTGTCATTC
Fam172a	TGAAAAAGGACGAACCACCT	TCGCCCAGAGCTTCATATCT
Nanog	AAGTACCTCAGCCTCCAGCA	GTGCTGAGCCCTTCTGAATC
Nr2f1	ATCCGCATCTTTCAGGAACA	TGATTTCTCCTGCAGGCTTT
Oct3/4	TCAGCTTGGGCTAGAGAAGG	TGACGGGAACAGAGGGAAAG
Sox1	GCAGCTATCAACCAGATCC	GATGTAGGCAGCTGTCATTC
Brachyury	GAACCTCGGATTCACATCGT	TTCTTTGGCATCAAGGAAGG
Hes1	TGAAGGATTCCAAAAATAAAA TTCTCTGGG	CGCCTCTTCTCCATGATAGGC TTTGATGAC
Tlx1	GGTCACCCCTATCAGAACCG	TTTTACTTGCGCATCGGTCA
Sox17	AGCTAAGCAAGATGCTAGGC AAG	TCTCTGCCAAGGTCAACGC
Foxa2	ACCTGAGTCCGAGTCTGAGC	TGTAGCTGCGTCGGTATGTC
Gata6	GCCAACTGTCACACCACAAC	GGTTTTCGTTTCCTGGTTTG
Nkx2.5	CAGTGGAGCTGGACAAAGCC	TAGCGACGGTTGTGGAACCA
αMHC	TGAAAACGGAAAGACGGTGA	TCCTTGAGGTTGTACAGCACA
Nestin	AGGAGAGAACCACGACCCAC	GCTGCTGGGTCTCTTGTTCG
βIII Tubulin	CATGGACAGTGTTCGGTCTG	TGCAGGCAGTCACAATTCTC
Sox2	CACAACTCGGAGATCAGCAA	CTCCGGGAAGCGTGTACTTA
Lin28	TGGGGGAACTTAAAGTGCAG	AAGATGGCTCAAACCACACC
Pou5f2	TGGGAGCTATGTTTGGGAAG	CTGCATATGCCCAGAAGGTT
Uc.88+	GGAAGCAGAAGTCGGGAAG A	GAGGGCTGATTAGCATGCAG

Uc.331+A	CACTACAGCTCTCTGTGCTTTTAC	CTTACGTTCAGGATCACTGG
Uc.200+A	CTGGGTTAAATGCTTGTTGCC	ACAGCTCTGTGAAGGCAGTC
Uc.92+	GAGTGGAGAGACAGCTCCTA	GGGAAATGACTGCTAGACTA
Uc.452+	CCAGAGCAAGTACTTGCAAG	CCATCCATCTTGGGGGCTCA
Gata6 promoter	ACTTTTTCTGGAGCTCGCGT	GTTCCGCACGTGGAAATAGC
Nestin promoter	GGTGCGTGACTACCAGGAG	TGCACCTCTAAGCGACTCTC
Foxa2 promoter	CCTGGAGAGACCCGTTTAGC	CCACCTACTGCCCTGTTTGT
Myo6 promoter	GCTCCGTAGCAGTGACGTG	GAGCACCGGAGACGACAG
Nr2f1 promoter	TGGGAGAGTCGAGCAGGATC	AGCGCTGCCTTCCTGAATG
T-UCstem1 genomic locus	ACCCAGTGACATCATGTTTTG	CTTCCCACAATGACCTATGTCA
T-UCstem1 (RIP)	AATCGTCCACAGCAGACCTC	AGGAGAGCTGGGAGAGTGTG

Table S4. Antibodies used in this study, related to Figure 3, 4, 5 and 6.

ANTIBODY	SOURCE	CAT. NO.	APPLICATION
Nanog	Cell Signaling	8822	IF (1:400) WB (1:1000)
Oct4	Santa Cruz B.T.	sc-8628	IF (1:400)
Brachyury	Santa Cruz B.T.	sc-17745	IF (1:200) WB (1:1000)
Sox17	Santa Cruz B.T.	sc-17318	IF (1:200)
PARP	Cell Signaling	9542	WB (1:1000)
Nestin	Santa Cruz B.T.	sc-33677	IHC (1:400) IF (1:500) FACS (1:500)
MF20	DSHB	2147781	FACS (1:50)
Sox2	Cell Signaling	14962	IF (1:200) WB (1:500)
βIII-tubulin	Sigma-Aldrich	T4026	FACS (1:500) IF (1:1000)
GFAP	Dako	IS524	FACS (1:500) IF (1:1000)
SSEA1	Cell Signaling	4744	IF (1:1000)
Caspase9	Cell Signaling	9508	WB (1:1000)
p27	Santa Cruz B.T.	(C-19)	WB (1:400)
p21	Santa Cruz B.T.	sc - 397	WB (1:400)
Ezh2	Active Motif	39875	ChIP & RIP (5 µg)
Ezh2	BD Bioscences	612666	WB (1:2000)
Suz12	Active Motif	39357	ChIP & RIP (5 µg) WB (1:300)
H3K4me3	Abcam	Ab8580	ChIP (5 µg)
H3K27me3	Abcam	mab ab6002	ChIP (5 µg)
IgG	Millipore	12-370	ChIP & RIP (5 µg)
α-Flag Lin28	homemade		WB (1:2000)

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

ESC culture and differentiation

NT and T-UCstem1 KD ESCs were cultured in serum/LIF/Feeders, as previously described (Bedzhov et al., 2014). 2i/LIF comprises the Mek inhibitor PD0325901 (Selleckchem.com, 1 μM), the Gsk3 inhibitor CHIR99021 (Selleckchem.com, 3 μM), and leukemia inhibitory factor (Lif, Millipore100 U/ml) in F12/Neurobasal medium (Guo et al., 2009). The serum-free mono-step neural differentiation protocol was previously described (Fico et al., 2008). Briefly, cells were seeded at 1.5×10^3 cells/cm^2 in knockout D-MEM (Gibco) supplemented with 15% knockout serum replacement (Gibco), 0.1 mM β-mercaptoethanol (Sigma-Aldrich), 2 mM glutamine, and 100 U/ml penicillin/streptomycin (Gibco) and grown for up to 12 days. At the indicated time points, cells were either fixed for immunofluorescence analysis or collected for RNA extraction. For *in vitro* differentiation towards cardiomyocytes, ESCs were cultivated in embryoid bodies (EBs) as previously described (D'Aniello et al., 2013). ESCs were allowed to differentiate through EBs in hanging drop (300 cells/drop) placed on the lids of tissue culture dishes for 2 days without the addition of growth factors. After further 3 days of culture in suspension, 5-day-old EBs were plated on gelatin-coated plates for further analysis.

T-UCstem1 shRNA sequences

In particular, to silence T-UCstem1, we used three different shRNA reported below:

Sh1 5'-caccGCAAAGACTCAAAGTGCAATTcgaaAATTGCACTTTGAGTCTTTGC-3'

Sh2 5'-caccGCAGACCTCCAAGAGACTTGTcgaaACAAGTCTCTTGGAGGTCTGC-3'

Sh3 5'-caccGCTTACACTGGTTCGTTTATTcgaaAATAAACGAACCAGTGTAAGC-3'

Among the KD ESC clones obtained, we used KD-1 and KD-2 clones, derived by using Sh1 and Sh2 respectively.

***In silico* prediction alignment in ultraconserved RNA sequences base pairing with miRNAs**

The target prediction tools *miRBase* (Kozomara and Griffiths-Jones, 2014) and *RNAhybrid* (version 2.1, Bielefeld University) were used to identify putative miRNA target sites in T-UCE sequences described by Bejerano *et al*. (Bejerano et al., 2004) T-UCE::miRNA duplex formation was evaluated under highly stringent conditions using pvalue< 0.05 and ΔG< -12 Kcal/mol.

***In silico* prediction of T-UCstem1 promoter and secondary structure**

We used two independent promoter prediction tools: 1) ElemeNT (Sloutskin et al., 2015), reporting a TATA box at 1595bp (TATAAAAA PWM score 1.00) and an Inr (score 0.0279) at 41bp downstream the TATA box (1,554bp upstream uc.170+); and 2) Eukaryotic Core Promoter Predictor (YAPP; http://www.bioinformatics.org/yapp/cgi-bin/yapp.cgi), reporting a TATA box at 1595bp (ATTATAAAAATG score 0.96) and a TSS at 35bp downstream to TATA box (1,554bp upstream uc.170+). Default parameters were used for the above mentioned prediction tools.

For the T-UCstem1 secondary structure prediction Minimun Energy Free (MEF) and Centroid algorithms (available in the RNAfold suite) were used with the default parameters (Gruber et al., 2008).

Colony Formation Assay

Colony-formation assay was performed as previously described (Chambers et al., 2007). Briefly, ESCs were trypsinized to obtain a single cell suspension and plated at low density (100 cells/cm^2). After 6 days, colonies were fixed in 4% paraformaldehyde (PFA) and stained with crystal violet and alkaline phosphatase as previously described (D'Aniello et al., 2015). Images were collected on a DMI6000B

microscope (Leica Microsystems). The morphological classification (domed/flat) was performed blinded by two investigators.

hESC Culture and Differentiation

The human Embryonic Stem Cell (hESC) line H9 (WiCell, cat. no. hPSCreg WAe009-A) and RC17 (Roslin Cells, cat. no. hPSCreg RCe021-A) were used in this study and were kindly provided by Prof. Malin Parmar (Wallenberg Neuroscience Center, Lund University, Sweden). Undifferentiated hESCs were maintained in culture on Lam-521 (100 µg/ml; Biolamina, cat. no. LN-521) in iPS-Brew XF (StemMACS; Miltenyi, cat. no. 130-104-368) and passaged with EDTA (0.5 M, pH 8.0; Thermo Fisher Scientific, cat.no 15575020) ((Kirkeby et al., 2017)). H9 and RC17 cell lines were differentiated toward human ventral midbrain dopaminergic neurons progenitors according to Nolbrant et al ((Nolbrant et al., 2017). Briefly, the hESCs were seeded to a concentration 10,000 cells per cm^2, in DMEM/F12:Neurobasal (1:1), N2 supplement (1:100; Thermo Fisher Scientific, cat. no. A1370701) onto plastic- ware coated with Lam111 (100 ug/ml; Biolamina, cat. no. LN-111). Patterning factors SB431542 (10 uM; Miltenyi, cat. no. 130-106-543), Noggin (100 ng/ml; Miltenyi, cat. no. 130-103-456), Shh-C24II (300 ng/ml; Miltenyi, cat. no. 130-095-727) and CHIR99021 (Miltenyi, cat. no. 130-106-539) were also present in the medium from day 0 to day 8. From day 9 of differentiation, FGF8b (100ng/ml; Miltenyi, cat. no. 130-095-740) was added to the medium and on day 11, cells were re-plated at 800.000 cells per cm^2 in DMEM/Neurobasal, B27 supplement (1:100; Thermo Fisher Scientific, cat. no. 12587010). Patterning factors BDNF (20ng/ml, Miltenyi, cat. no. 130-096-286), AA (0,2mM, Sigma-Aldrich, cat. no. A4403-100MG) FGF8 (100ng/ml) were added. For terminal differentiation, the cells were re-plated at day 16 to a concentration 155,000 cells per cm^2 in DMEM/Neurobasal with B27 supplement with BDNF (20ng/ml), AA (0,2mM) GDNF (10ng/ml, R&D Systems, cat. no. 212-GD-010) db-cAMP (500uM Sigma-Aldrich, cat. no. D0627-1G) and DAPT (1uM, *N*-[(3,5-diuorophenyl)acetyl]-l-alanyl-2-phenyl]glycine-1, 1-dimethylethyl

ester; R&D Systems, cat. no. 2634). The experiments on hESCs were performed in the laboratory of Prof Malin Parmer at the Department of Experimental Medical Science and Lund Stem Cell Center BMC, Lund University, 22632 Lund, Sweden.

Cloning and Site-Directed Mutagenesis

After WT cloning, QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA) was used to generate the deletion of 4bp from the site of complementarity between uc.170+ and both miR9-5p and miR9-3p, by PCR using the WT psiCHECK+uc.170+ construct as the template. The following primers containing the deletion were designed and used for site-directed mutagenesis according to the manufacturer's protocol:

miR9-5p Fw 5'-GCTGCAATAAGCTAGGTTTTCAAAGAGAGGCAAAGACTC-3,

miR9-5p Rv 5'-GAGTCTTTGCCTCTCTTTGAAAACCTAGCTTATTGCAGC-3;

miR9-3p Fw 5'-GAGATTCTCTTGCAATAAGCTAGGTTTTCAGCCAAAGAGAGGC-3',

miR9-3p Rv 5'-GCCTCTCTTTGGCTGAAAACCTAGCTTATTGCAAGAGAATCTC-3'.

WT and mutant inserts were confirmed by sequencing.

Rapid Amplification of cDNA Ends (RACE)-PCR

To identify the 5′ and 3′ ends of the uc.170+ transcript, total RNA from ESCs was extracted and treated with DNase I (RNase-free) endonuclease, and the SMARTer RACE 5'/3' kit (Clontech) was used to generate RACE-ready cDNA according to the manufacturer's instructions. The cDNA ends were amplified with SeqAmp DNA Polymerase (Takara), and gene-specific primers (GSP1: 5′-AGGGGTGATATGCATGTGCT-3′; GSP2: 5′-TGAGAAGGGGACGAGGGTTGCTACA -3′) were used. Furthermore, nested PCR analysis was performed with the nested universal primer provided with the kit

(SMARTer RACE cDNA Amplification kit) and two nested gene-specific primers (NESTED-GSP1: 5′- TGCTGAAGCACCCCTTAAGCCCACT-3′; NESTED-GSP2: 5′-GGGCATACAGCCCCCTCCCCGTACTC-3′). Mouse heart RNA and transferrin receptor-specific primers provided with the kit were used as reaction controls. The PCR fragments were then run on a 1.5% agarose gel, and DNA was extracted from the gel using a QIAquick Gel Extraction kit (QIAGEN) according to the manufacturer's instructions. The RACE products were then cloned into a TOPO TA pCR2.1 cloning vector (Thermo Fisher Scientific) according to the manufacturer's instructions. The sequences were obtained by using the T7 and GSPs primers and verified by using the UCSC Genome Browser (University of California Santa Cruz).

Flow cytometry

Single cell suspensions of ESC-derived neurons, were obtained using either trypsin-EDTA or TrypleSelect 1x (Gibco), fixed, stained with the appropriated primary and secondary antibodies according to the manufacturer's protocols, and were analysed with a BD FACS CantoIITM cytofluorimeter (BD Biosciences). Details and list of antibodies are in Table S4.

Cell Cycle and Proliferation assays

For Cell Cycle analysis, the cells were dissociated to single cell suspension, fixed with cold 70% ethanol before propidium iodide (PI) staining (20 µg/ml) and were analysed by flow cytometry using a BD FACS CantoIITM cytofluorimeter (BD Biosciences). Cell viability was measured using the colorimetric CyQuant® cell proliferation assay (Invitrogen), following the manufacturer's instructions. Absorbance was analysed at 480–520 nm, using the Fluoroskan Ascent FL Microplate Fluorometer and Luminometer (Thermo Fisher Scientific, Waltham, MA, USA). For the proliferation assay, the Click-iT EdU Flow Cytometry Assay (Invitrogen) was used. Briefly, cells were incubated with 5-ethynyl-2'-deoxyuridine (EdU) (10 µM;

overnight at 37 °C), dissociated, fixed and permeabilized, following the manufacturer's instructions. Samples were analysed at FACS-Canto using the DivaTM software (BD Biosciences, San Jose, CA, USA). For cell division analysis CellTrace™ CFSE Cell Proliferation Kit was used according to the manufacturer's instructions (Thermo Fisher Scientific). CFSE [5-(and-6)-carboxyfluorescein diacetate succinimidyl ester] was used to trace multiple generations using dye dilution by flow cytometry.

RNA Extraction, Northern Blot analysis, quantitative RT-PCR and copy number determination

Total RNAs were isolated using TRIzol Reagent (Thermo Fisher Scientific) following the manufacturer's instructions. The concentration of RNA was determined by 260/280 nm absorbance using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific), and the integrity of RNA was checked using gel electrophoresis.

Agarose gel–based Northern blotting was performed and transferred onto Hybond-*n* + membrane (Amersham Pharmacia Biotech). Biotinylated probes used were complementary to the sequence of T-UCstem1 (5'-CCTGTGTATAATTGCACTTTGAGTCTTTGCCTCTCTTTG -3') or the U6. Detection Chemiluminescent Nucleic Acid Detection Module (Thermo Fisher Scientific) were used according to the manufacturer's instructions. The size of the detected RNA was determined by using a size marker run on the same gel. Total RNA (1µg) was reverse-transcribed using, QuantiTect Reverse Transcription kit (Qiagen). RT-qPCR was performed using strand-specific primers for T-UCE analysis and random primers for coding-gene expression. A miRCURY LNA Universal RT miR PCR kit (Exiqon) was used for miRNA analysis according to the manufacturer's instructions. Small nuclear RNA U6 was used as a reference for T-UCEs and miRNAs. The sequences of primers are reported in Table S3.

The relative amount of specific transcripts was measured by RT-PCR analysis. Briefly, it was performed using an iQ SYBR Green Supermix (BioRad) protocol with a CFX96Deep Well system RealTime PCR Detection System (Bio-Rad), according to the manufacturer's instructions. The copy number of transcripts per cell was calculated by the comparative cycle threshold method presented by Livak and Schmittgen (Livak and Schmittgen, 2001).

RNA fractionation

The RNA fractionation was performed as described by Cabianca DS *et al.* (Cabianca et al., 2012). Briefly, ESC were detached by treating with 1X Trypsin, counted and centrifuged at RT 168 g for 5 min. The pellet was lysed with 175 µl/10^6 cells of cold RLN1 solution (50 mM Tris HCl pH 8.0; 140 mM NaCl; 1.5 mM $MgCl_2$; 0,5% NP-40; 2mM Vanadyl Ribonucleoside Complex; Sigma) and incubated 5 min on ice. Next, the suspension was centrifuged at 4°C and 300 g for 2 min and the supernatant, corresponding to the *cytoplasmic fraction*, was transferred into a new tube and stored in ice. The pellet containing nuclei was extracted with 175 µl/10^6 cells of cold RLN2 solution (50 mM Tris HCl pH 8.0; 500 mM NaCl; 1.5 mM $MgCl_2$; 0,5% NP-40; 2mM Vanadyl Ribonucleoside Complex) and 5 min incubated in ice. The suspension was centrifuged at 4°C and 16360 g for 2 min and the supernatant, corresponding to the *nuclear-soluble fraction*, was transferred into a new tube and stored in ice. The remaining pellet corresponds to the *chromatin-associated fraction*.

Total RNA was extracted by using Trizol reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions; in particular, the extraction from aqueous solutions was followed for the cytoplasmic and nuclear-soluble fractions, whereas the chromatin-associated fraction was considered as a pellet. *Gapdh* and *Xist* were used as quality control of the RNA fractionation.

Western Blotting

Whole cell lysates were prepared with ice-cold immunoprecipitation assay (RIPA) lysis buffer. Detection was performed with ECL reagents (Amersham Biosciences). Details and list of antibodies are in Table S4.

Luciferase Reporter Assay and miRNA mimic and Lin28 transfection

Uc.170+ was cloned into the NotI and XhoI sites in psiCheck™ vector (Promega) immediately downstream Renilla luciferase reporter gene. This plasmid contains a firefly luciferase expression cassette that acts as an internal normalization of luciferase activity. Dual-Luciferase Reporter Assay System (DLR assay system, Promega, Madison, WI) was used to measure luciferase activity of 293FT cells co-transfected with uc.170+ cloned in psiCHECK2 together with miR-9-5p or miR9-3p using lipofectamine 2000 (Invitrogen). Luciferase assays were analyzed based on ratio of Renilla/Firefly to normalize over the cell number and transfection efficiency.

For miRNA mimics and Inhibitors (AntagomiR) transfection, ESCs (200,000 cells/well) were plated in six-well plates and transfected with mimics miR-9-5p/3p or antagomiR-9-5p/3p (Exiqon) and AllStars Negative Control (scrambled) (Exiqon) using RNAiMAX reagent (Invitrogen, cat. 13778150).

NT and T-UCstem1 KD ESCs (200,000cells/well) were also transfected with the empty vector or plasmid expressing Flag-tagged Lin28 (kindly provided by Dr. Silvia Parisi), using lipofectamine 2000 (Invitrogen).

EGFP-labelled ESCs and Chimera generation

GFP was inserted in both NT and T-UCstem1 KD ESCs at the Rosa26 locus by using the R26P-SA-EGFPpuro plasmid (Addgene). Ten days after transfection, puromycin-selected clones were verified for correct self-renewal and differentiation properties. Chimeras were obtained by injecting NT and T-UCstem1 KD GFP-

labelled ESCs (13–16) into 4- to 8-cell-stage embryos using standard techniques. Chimeric mouse generation was performed by morula injection of NT and TUCstem1 KD GFP-labelled ESCs. Resultant embryos were cultured for 48 h *in vitro* and implanted by uterus transfer into pseudopregnant foster mothers using standard methods. Pregnant mice were killed at day E9.5 and whole embryos were photographed with fluorescence microscope. Experiments were done in accordance to the law on animal experimentation (article 7; D.L. 116/92) under the Animal Protocol approved by the Italian Ministry of Health.

Teratoma Assay

ESCs were trypsinized into single-cell suspension and resuspended in phosphatase buffered saline (PBS). ESCs ($3x10^6$) were injected subcutaneously into hind limbs of severe combined immunodeficiency mice (SCID). Teratomas were collected, fixed in 4% PFA, sectioned and stained with hematoxylin/eosin or subjected to immunohistochemistry for the histological analysis.

Immunohistochemistry

Samples were processed with the standard streptavidin–biotin-immunoperoxidase method (DAKO Universal Kit, DAKO Corp., Carpinteria, CA, USA). Diaminobenzidine was used as the final chromogen, and hematoxylin as the nuclear counter stain.

Immunofluorescence

Cells were fixed (4% PFA) and permeabilized (0.1% Triton X-100), where necessary, at room temperature. After incubation with primary antibodies, cells were incubated (1h) with the appropriate secondary antibodies (Alexa Fluor 488 and/or 594 (1:200); Molecular Probes). Details and list of antibodies are in Table S4.

Chromatin Immunoprecipitation (ChIP)

Chromatin immunoprecipitation to detect H3K4me3, H3K27me3, Suz12 and Ezh2 enrichments was performed according to a previously validated protocol (Comes et al., 2013). Briefly, $1x10^6$ cells for NT and KD ESCs were fixed with formaldehyde at room temperature (RT, 10 min), followed by glycine (125 mM) to stop the crossing linking reaction (RT, 5 min). Nuclear extracts were sonicated using a Covaris S2 system sonicator according to manufacturer's instructions to achieve chromosome fragment lengths of 200-500 bp. After sonication, suitable amount of chromatin was incubated with the specific antibodies (Table S4). Immunoprecipitated complexes were recovered with protein A sepharose and samples were then washed with low and high salt buffers, reverse-crosslinked, and purified using the QIAquick PCR purification kit (Qiagen). Purified DNA was analysed by RT-qPCR using gene-specific primers (Table S3).

RNA Immunoprecipitation (RIP)

Native RNA immunoprecipitation experiments were performed using the Magna RIP™ RNA-Binding Protein Immunoprecipitation Kit (Millipore) according to the manufacturer's instructions. Briefly, $20x10^6$ NT ESCs were lysed to isolate nuclei in presence of protease and RNAse inhibitors, which were then treated with DNAseI. The complex magnetic beads-antibody was prepared and the immunoprecipitation followed for 4-16 hours or overnight. The antibodies anti-Suz12, anti-Ezh2 and anti-IgG used are reported in Table 4. RNAs from the immunoprecipitated and input fractions were purified, retrotranscribed using T-UCstem1-specific oligonucleotide and cDNAs were used for RT-qPCR.

Probes used for T-UCstem1 ChIRP

Biotin-labeled antisense T-UCstem1 DNA probes were designed using the suggested web tool (www.singlemoleculefish.com). We compared the probe

positions with the secondary structure prediction of the T-UCstem1. Only the probes located within a region with low probability to form stem were used [ODD probes: 5'-aggagtgtaggtagggattt -3', 5'- agctgggagagtgtgtgaaa -3'; EVEN probes: 5'-cctttccatggagaatctta -3', 5'- gcacttcaacacctttttcaa -3'].

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Microarray validation and bioinformatics analysis of uc.170+ locus (Related to Figure 1 and 2)

(A) RT-qPCR validation of T-UCEs and **(B)** microRNAs expression in ESCs and N/GCs. Relative RNA level was normalised to *U6* expression. Data are mean ± SEM (*n*=3 independent experiments); *p <0.01, **p <0.005, ***p <0.001. **(C)** Validation of *uc.170+* expression by semiquantitative RT-PCR (sqRT-PCR) performed on total RNA extracted from ESCs or N/GCs. –, RT minus control reactions. *Oct4 and bIII-tubulin* mRNAs were used as control markers of ESC neural differentiation. PCR amplifications were performed on biological triplicates, and the results of a representative experiment are shown. **(D)** *In silico* prediction alignment of *uc.170+* sequence base pairing with miR-9-5p/3p and relative ΔG, performed by *RNAHybrid* software. **(E)** Schematic representation of genomic location of uc.170 within Fam172a host-gene and its relative transcript. The TATA box, INR and TSS bioninformatically predicted are reported. In green are also reported the region targeted by the shRNAs used for the knockdown experiments: Sh1 targets a region within the uc.170, the Sh2 and Sh3 target regions at about 200bp and 1kb upstream the uc.170 respectively.

Figure S2. Expression analysis of T-UCstem1 (Related to Figure 2)

(A) Northern blotting analysis showing the expression of T-UCstem1 in Non Targeted (NT) and two independent T-UCstem1 KD ESC clones. Normalization was performed with U6. **(B)** T-UCstem1 sequence (1813bp): the red sequence was described by Bejerano et al. (Bejerano et al., 2004). **(C)** T-UCstem1 secondary structure prediction obtained by Minimun Energy Free (MEF) and Centroid algorithms. The binding sites for miR-9-3p/5p are indicated by black arrows.

Figure S3. Functional characterization of T-UCstem1 KD ESCs (Related to Figure 3)

(A) Cell viability of Control (NT) and two independent T-UCstem1 KD ESC clones measured by the CyQuant® assay and expressed as relative fluorescence units (RFU). Data are mean ± SEM (*n*=3 independent experiments); **p <0.005. **(B)** Time-course analysis of automated cell counting of FBS/LIF/Feeders Control (NT) and T-UCstem1 KD ESCs. Data are mean ± SEM (*n*=3 independent experiments); ***p <0.001. **(C-D)** FACS-based analysis of cell division quantification **(C)** by dye dilution (CFSE) in Control (NT) and T-UCstem1 KD ESCs at different time points and **(D)** in two independent T-UCstem1 KD clones at 72hrs**.** Data are mean ± SEM (*n*=3 independent experiments); *p <0.01, **p <0.005. **(E)** Schematic representation of the experimental procedure. **(F)** RT-qPCR analysis of uc170+ in hESCs transfected with siRNA uc.170+ or scr (100 nmol) for 48h. **(G)** Representative photomicrographs of scrambled and siRNA uc.170+ colonies. Scale bar, 200 μm. **(H)** Automated cell counting of hESCs transfected with siRNA uc.170+ or scr (100 nmol) for 48h. **(I)** RT-qPCR analysis of T-UCstem1 level in KD ESCs transfected with uc170+ cDNA expression vector. NT and KD ESCs were used as positive and negative control, respectively. *U6* was used as a loading control. Data are mean ± SEM (*n*=3 independent experiments); **p <0.005. **(J)** RT-qPCR of miR9-5p and its target genes in KD ESCs transfected with uc170+ cDNA expression vector. NT and KD ESCs were used as positive and negative control, respectively. Data are mean ± SEM (*n*=3 independent experiments); **p <0.005, *p<0.01. **(K)** Representative photomicrographs of FBS/LIF KD colonies and KD colonies transfected with uc170+ cDNA expression vector. Scale bar, 100 μm. **(L)** FACS-based analysis of cell proliferation quantification by EdU incorporation in NT, KD and KD ESCs upon uc170+ cDNA expression vector or empty vector transfection (48hrs). Representative FACS plots of biological triplicates are shown.

Figure S4. Functional characterization of T-UCstem1 KD ESCs (Related to Figure 4)

(A-B) RT-qPCR analysis of **(A)** miR-9 and **(B)** its target genes (*Lin28b* and *Tlx1*) in NT and T-UCstem1 KD ESCs transfected with antagomiR-9 5p/3p or scr (100 nmol) at 48hrs after transfection. Data are mean ± SEM (*n*=3 independent experiments); **p <0.005, *p<0.01. **(C)** Automated cell counting of NT and KD ESCs transfected with antagomiR-9 5p/3p or scr (100 nmol) at 48hrs after transfection. Data are mean ± SEM (*n*=3 independent experiments); **p <0.005, ***p <0.001. **(D)** Western blot analysis of Lin28B-Flag in NT, KD and in KD Lin28B overexpressing cells. GAPDH was used as a loading control. **(E)** RT-qPCR analysis of Lin28b in NT, in KD and in KD ESCs upon Lin28B-Flag/empty vector transfection. **(F-G)** Representative pictures **(F)** of NT and KD ESCs during clonogenic assay **(G)** with relative colony number of NT and KD colonies stained with crystal violet. Scale bar, 200 μm. Data are mean ± SEM (*n*=3 independent experiments); **p <0.005. **(H)** RT-qPCR analysis of pluripotency-associated genes (*Nanog*, *Sox2* and *Oct4*) in Control (NT) and T-UCstem1 KD in High density/Feeders culture conditions and plated at low density. Data are mean ± SEM (*n*=3 independent experiments); *p <0.01, **p <0.005. **(I)** Representative pictures of NESTIN, BRA and SOX17 (scale bars, 75 μm) immunostaining in NT ESC colonies. Nuclei were stained with DAPI. **(J)** Representative immunofluorescence of SSEA1 in NT and KD ESC colonies at day 6 after plating in clonogenic assay. Nuclei were stained with DAPI. Scale bars, 75 μm. **(K)** RT-qPCR analysis of pluripotency-associated genes (*Nanog* and *Sox2*) in Control (NT), KD and KD + *2i* (CHIR99021+ PD0325901) ESCs, at day 6 after plating. Data are mean ± SEM (*n*=3 independent experiments); **p <0.005. **(L)** Automated cell counting of NT and KD ESCs and KD + *2i* (CHIR99021+ PD0325901) ESCs, at day 6 after plating. Data are mean ± SEM (*n*=3 independent experiments); ***p <0.001.

Figure S5. Analysis of T-UCstem1 KD ESCs pluripotency *in vitro* and *in vivo* (Related to Figure 4 and 5)

(**A**) Representative photomicrographs by Discovery.V12 Zeiss microscopy of chimeric embryos from EGFP-labelled KD and WT ESCs injected into morula and dissected at E9.5. **(B)** Time-course expression profiles of mesoderm (*Bra*) and cardiac (*Nkx2.5* and α*MHC*) markers in Control (NT) and two independent T-UCstem1 KD clones. Relative RNA level was normalised to *Gapdh* expression. Data are mean ± SEM (*n*=3 independent experiments); *p <0.01, **p <0.005, ***p <0.001. **(C)** FACS-based quantification of BRACHIURY (day 8) and MF20 (day 10) positive cells in Control (NT) and KD ESC cardiac differentiation. Data are mean ± SEM (n=3 independent experiments).

Figure S6. Epigenetic analysis of T-UCstem1 KD ESCs (Related to Figure 6)

(A) RT-qPCR of selected developmental genes (*Nestin*, *Gata6* and *Foxa2*) in Control (NT) and two independent T-UCstem1 KD ESC clones. Relative RNA level was normalised to *Gapdh.* Data are mean ± SEM (*n*=3 independent experiments); ***p <0.001. **(B)** RT-qPCR analysis of pluripotency and differentiation-associated genes in hESCs transfected with siRNA uc.170+ or scr (100 nmol) for 48h Control (NT). Relative RNA level was normalised to *Gapdh* expression. Data are mean ± SEM (*n*=3 independent experiments); *p <0.01, **p <0.005, ***p <0.001. **(C)** H3K4me3/H3K27me3 occupancy at bivalent-associated promoters (Nestin, Gata6 and Foxa2) analyzed by Chip-qPCR in Control (NT) and KD ESC clones. Data are mean ± SEM (*n*=3 independent experiments); **p <0.01, ***p <0.005. **(D)** Native RNA immunoprecipitation (n-RIP) of *T-UCstem1* in ESCs, using antibodies against SUZ12, EZH2 or IgG as control. Data are mean ± SEM (n=3 independent experiments); ** p<0,01. **(E)** ChIP-qPCR of SUZ12 and EZH2 binding at Nr2f1 promoter in KD and Control (NT) cells. Myo6 promoter has been reported as control. Data are mean ± SEM (n=3 independent experiments); **p <0.01. **(F)** Western blot analysis of cell

SUZ12 and EZH2 in Control (NT) and two independent T-UCstem1 KD ESC clones.

GAPDH was used as a loading control.

SUPPLEMENTAL REFERENCES

Bedzhov, I., Leung, C.Y., Bialecka, M., and Zernicka-Goetz, M. (2014). In vitro culture of mouse blastocysts beyond the implantation stages. Nature protocols *9*, 2732-2739.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., and Haussler, D. (2004). Ultraconserved elements in the human genome. Science *304*, 1321-1325.

Cabianca, D.S., Casa, V., Bodega, B., Xynos, A., Ginelli, E., Tanaka, Y., and Gabellini, D. (2012). A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. Cell *149*, 819-831.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230-1234.

Comes, S., Gagliardi, M., Laprano, N., Fico, A., Cimmino, A., Palamidessi, A., De Cesare, D., De Falco, S., Angelini, C., Scita, G.*, et al.* (2013). L-Proline induces a mesenchymal-like invasive program in embryonic stem cells by remodeling H3K9 and H3K36 methylation. Stem Cell Reports *1*, 307-321.

D'Aniello, C., Fico, A., Casalino, L., Guardiola, O., Di Napoli, G., Cermola, F., De Cesare, D., Tate, R., Cobellis, G., Patriarca, E.J.*, et al.* (2015). A novel autoregulatory loop between the Gcn2-Atf4 pathway and (L)-Proline [corrected] metabolism controls stem cell identity. Cell Death Differ *22*, 1094-1105.

D'Aniello, C., Fiorenzano, A., Iaconis, S., Liguori, G.L., Andolfi, G., Cobellis, G., Fico, A., and Minchiotti, G. (2013). The G-protein-coupled receptor APJ is expressed in the second heart field and regulates Cerberus-Baf60c axis in embryonic stem cell cardiomyogenesis. Cardiovasc Res *100*, 95-104.

Fico, A., Manganelli, G., Simeone, M., Guido, S., Minchiotti, G., and Filosa, S. (2008). High-throughput screening-compatible single-step protocol to differentiate embryonic stem cells in neurons. Stem Cells Dev *17*, 573-584.

Gruber, A.R., Lorenz, R., Bernhart, S.H., Neubock, R., and Hofacker, I.L. (2008). The Vienna RNA websuite. Nucleic acids research *36*, W70-74.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development *136*, 1063-1069.

Kirkeby, A., Nolbrant, S., Tiklova, K., Heuer, A., Kee, N., Cardoso, T., Ottosson, D.R., Lelos, M.J., Rifes, P., Dunnett, S.B.*, et al.* (2017). Predictive Markers Guide Differentiation to Improve Graft Outcome in Clinical Translation of hESC-Based Therapy for Parkinson's Disease. Cell stem cell *20*, 135-148.

Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic acids research *42*, D68-73.

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

Nolbrant, S., Heuer, A., Parmar, M., and Kirkeby, A. (2017). Generation of high-purity human ventral midbrain dopaminergic progenitors for in vitro maturation and intracerebral transplantation. Nature protocols *12*, 1962-1979.

Sloutskin, A., Danino, Y.M., Orenstein, Y., Zehavi, Y., Doniger, T., Shamir, R., and Juven-Gershon, T. (2015). ElemeNT: a computational tool for detecting core promoter elements. Transcription *6*, 41-50.