

Stem Cell Reports, Volume 18

Supplemental Information

Long noncoding RNA *Lnc530* localizes on R-loops and regulates R-loop formation and genomic stability in mouse embryonic stem cells

Daohua Gong, Lin Wang, Hu Zhou, Jing Gao, Weidao Zhang, and Ping Zheng

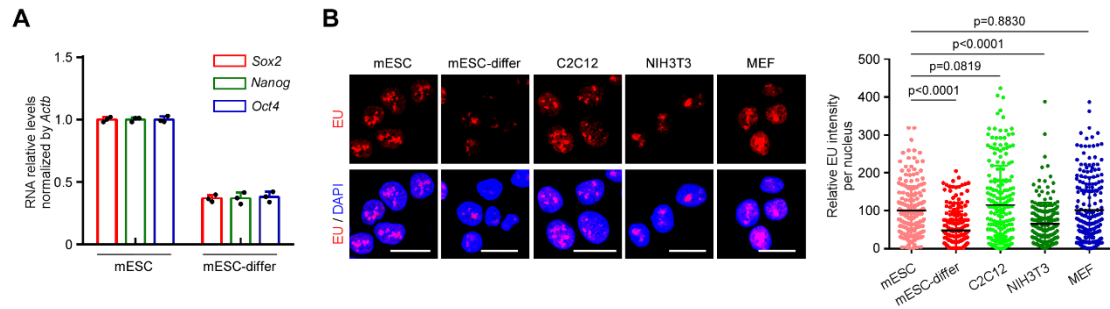


Figure S1. Evaluation of nascent RNA synthesis in mESCs and differentiated cells, related to Figure 1.

(A) Quantitative RT-PCR showed the relative mRNA expression of core pluripotency regulators *Sox2*, *Nanog* and *Oct4* in mESCs and mESC-differentiated (mESC-differ) cells. *Actb* was set as internal reference.

(B) Nascent RNA was monitored by EU (red). Immunostaining detected the RNA transcription in mESCs and differentiated cells. Left panel showed the representative immunostaining images. Right panel showed the quantification of EU intensity per nucleus. Scale bar, 20 μm . All experiments were independently repeated three times with similar results. Data were shown as mean \pm SD, two-tailed Student's *t*-test.

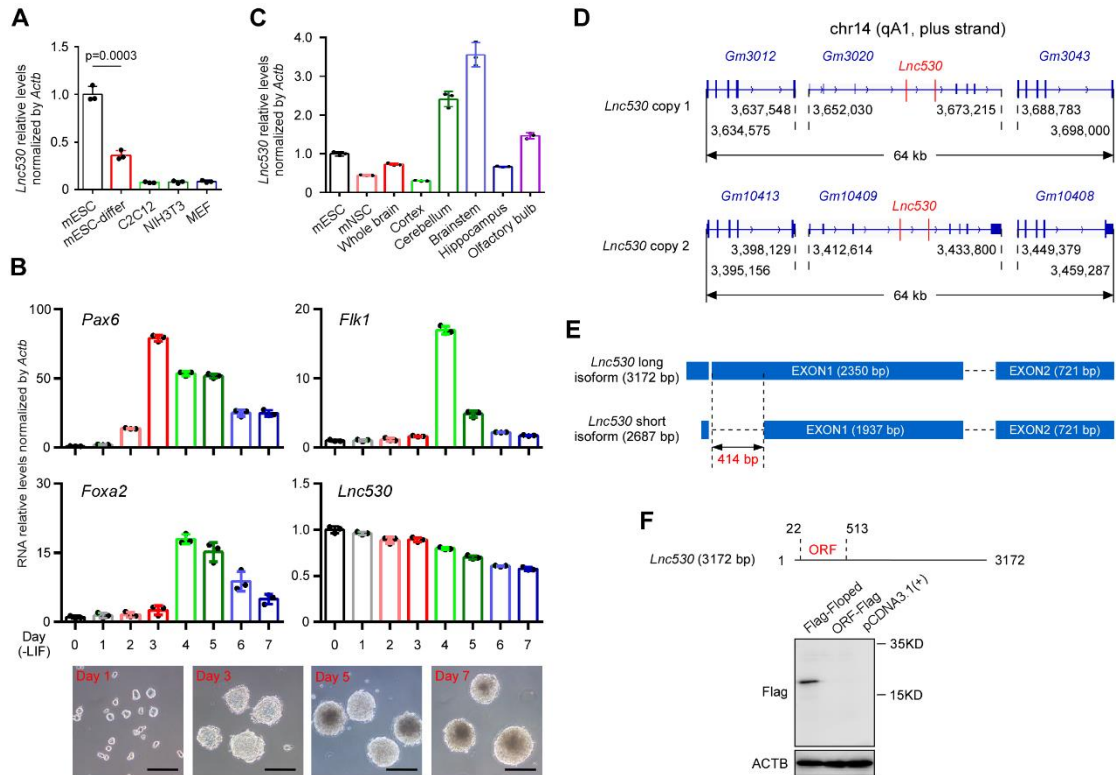


Figure S2. Characterization of *Lnc530* in mESCs.

(A) Quantitative RT-PCR showed that *Lnc530* was expressed in mESCs compared to several differentiated cell types under normal culture condition. *Actb* was set as internal reference.

(B) Quantitative RT-PCR showed the relative expression levels of four genes (*Pax6*, *Flk1*, *Foxa2* and *Lnc530*) in embryoid bodies (EBs) differentiated from mESCs for different time points. Lower panel showed the representative images of EBs. *Pax6* is an ectodermal marker, *Flk1* is a mesodermal marker, *Foxa2* is an endodermal marker. *Actb* was set as internal reference. Scale bar, 200 μ m.

(C) Quantitative RT-PCR showed the relative expression levels of *Lnc530* in different encephalic regions of C57BL/6 mice. *Actb* was set as internal reference.

(D) Based on mouse reference genome assembly version mm10, *Lnc530* gene has two genomic copies and is located on plus strand of chromosome 14.

(E) Two isoforms of *Lnc530*.

(F) The potential coding region of *Lnc530* did not encode protein or peptide. The Flag-tagged Floped protein was used as a positive control and pcDNA3.1 (+) was a negative control.

All experiments were independently repeated three times with similar results. Data were shown as mean \pm SD, two-tailed Student's *t*-test.

re-expression (DDX5-rescue). *Actb* was set as internal reference.

(H) Immunoblotting confirmed the KD and re-expression of TDP-43 (TDP-43-rescue) in mESCs.

All experiments were independently repeated three times with similar results. The relative protein levels in (E, F, H) were normalized by GAPDH or H3. Data were shown as mean \pm SD, two-tailed Student's *t*-test.

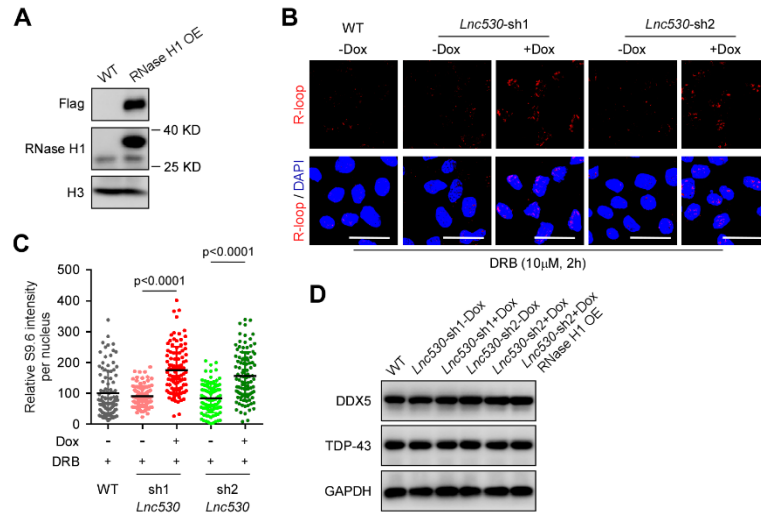


Figure S4. Loss of *Lnc530* perturbs R-loop homeostasis, related to Figure 5.

(A) Immunoblotting showed the over-expression (OE) of RNase H1 in wild-type mESCs.

(B and C) Immunostaining (B) and its quantification (C) revealed that *Lnc530* KD mESCs contained higher level of R-loops after DRB treatment. Scale bars, 30 μ m. DNA was counterstained with DAPI. At least 20 visual fields containing 500 cells were analyzed in three independent experiments. Scale bars, 30 μ m.

(D) Immunoblotting showed that the protein expressions of TDP-43 and DDX5 were not altered by *Lnc530* KD or/and RNase H1 over-expression (OE) in mESCs.

All experiments were independently repeated three times with similar results. Data were shown as mean \pm SD, two-tailed Student's *t*-test.

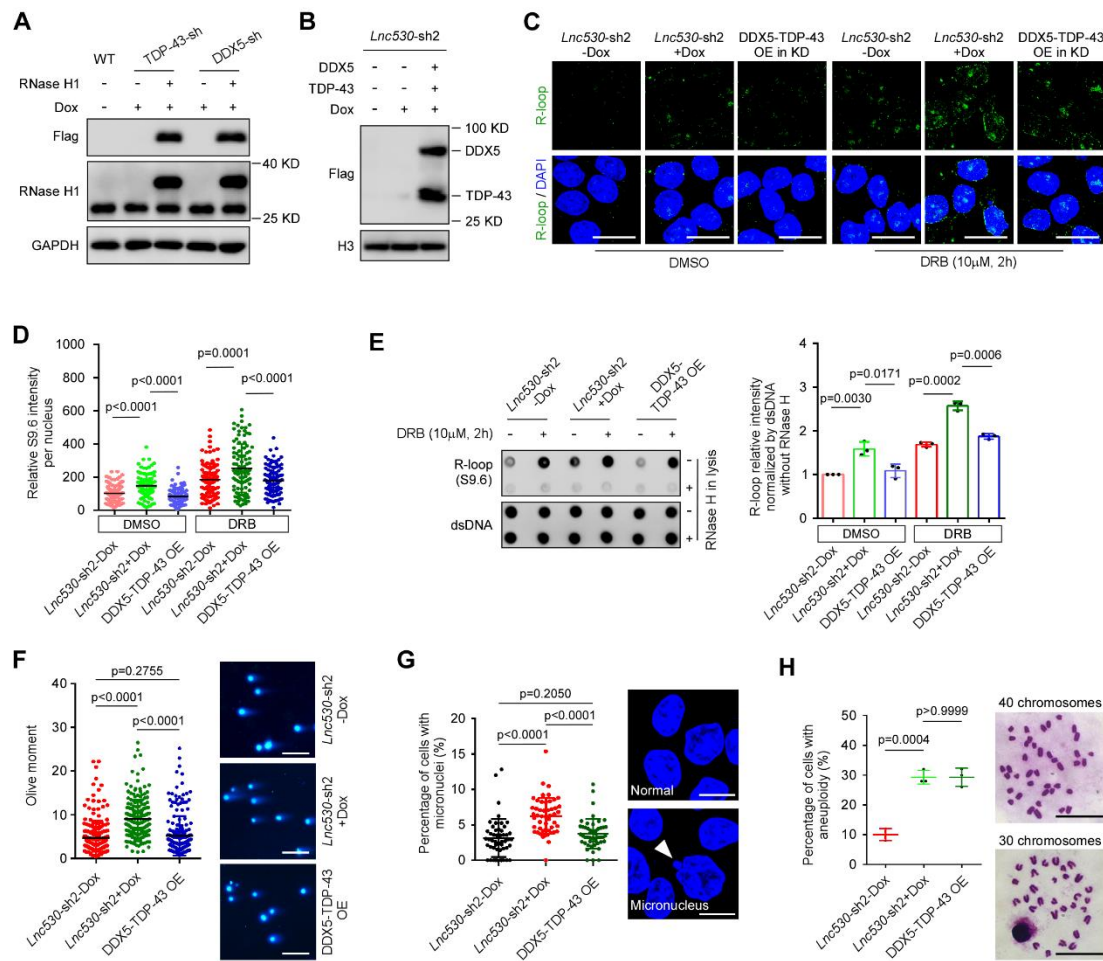


Figure S5. Over-expression of TDP-43 and DDX5 in *Lnc530* KD mESCs reduces the aberrant R-loop formation and DNA DSBs.

(A) Immunoblotting confirmed the forced expression of RNase H1 in TDP-43 KD or DDX5 KD mESCs.

(B) Immunoblotting verified the over-expression of DDX5 and TDP-43 proteins in *Lnc530* KD mESCs.

(C and D) Immunostaining showed that TDP-43 and DDX5 over-expression (OE) in *Lnc530* KD mESCs reduced R-loop formation (C). The quantification of R-loops per nucleus was shown in (D). At least 20 visual fields containing 500 cells were analyzed in three independent replications. Scale bar, 20 μ m.

(E) Dot blot analysis of R-loops in *Lnc530*-KD mESCs with TDP-43 and DDX5 over-expression (OE). Left panel showed the dot blot images. Right panel was the quantification of dot blot intensity. dsDNA was set as internal reference.

(F) Neutral comet assay showed that TDP-43 and DDX5 over-expression (OE) in *Lnc530* KD mESCs reduced the generation of DNA DSBs. Left panel was the quantification of comet tail length. Right panel showed the representative images. At least 200 tails were analyzed in each group. Scale bars, 400 μ m.

(G) Immunostaining showed that TDP-43 and DDX5 over-expression (OE) in *Lnc530* KD mESCs reduced the micronuclei formation. Left panel was the quantification. Right panel showed the representative images of micronucleus (arrow). At least 50 visual fields containing

1000 cells were analyzed in each group. Scale bars, 10 μm .

(H) TDP-43 and DDX5 over-expression (OE) in *Lnc530* KD mESCs failed to reduce the aneuploidy rate. Left panel was the quantification. Right panel showed the representative images. At least 150 metaphase spreads were examined in three replications in each group. Scale bars, 20 μm .

All experiments were repeated three times with similar results. Data were shown as mean \pm SD, two-tailed Student's *t*-test.

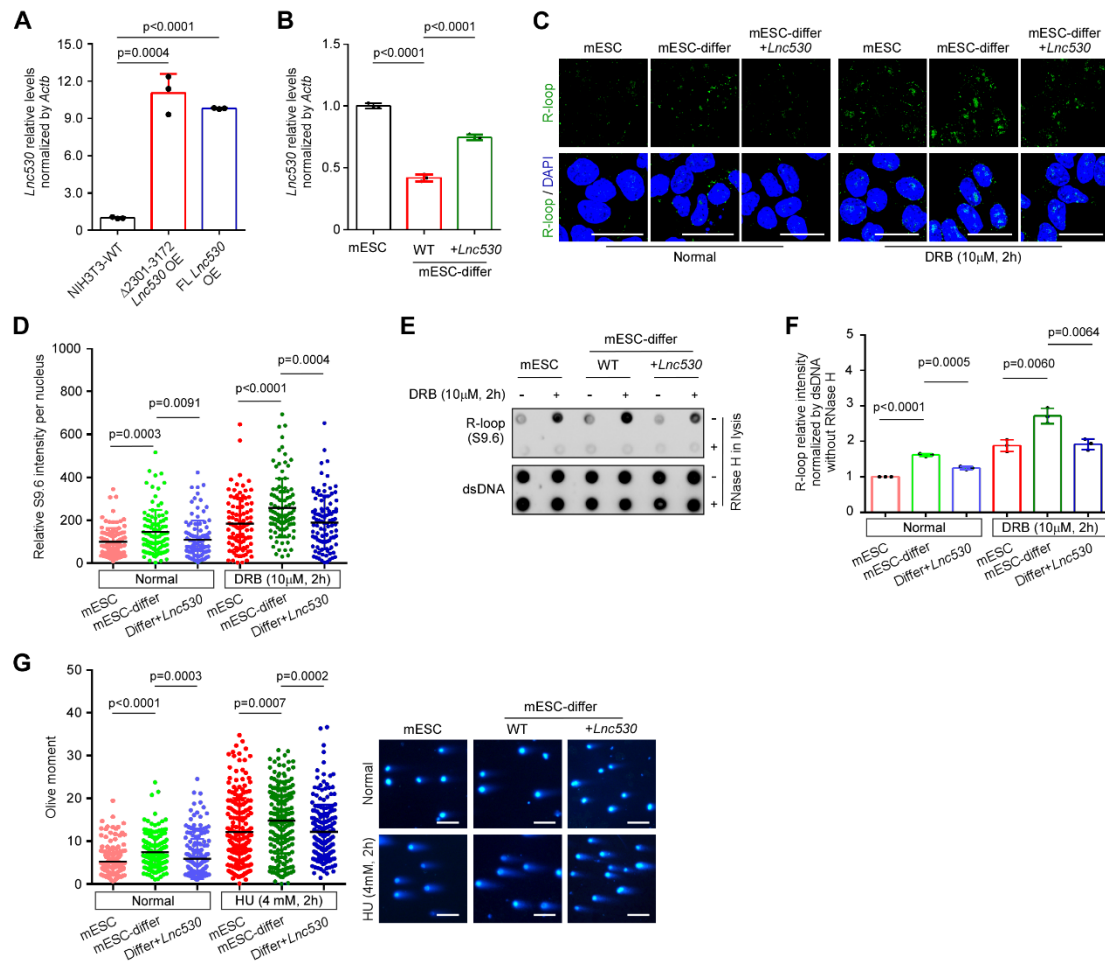


Figure S6. Over-expression of *Lnc530* in mESC-differentiated cells reduces R-loop formation and DNA damage.

(A) Quantitative RT-PCR confirmed the over-expression (OE) of full length (FL) *Lnc530* and $\Delta 2301-3172$ *Lnc530* in NIH3T3 cells. *Actb* was set as internal reference.

(B) The relative expression levels of *Lnc530* in mESCs, isogenic differentiated cells, and isogenic differentiated cells with *Lnc530* over-expression. *Actb* was set as internal reference.

(C and D) Immunostaining showed that over-expression of *Lnc530* in mESC-differentiated cells (mESC-differ) decreased the R-loop formation (C). The quantification of R-loops was shown in (D). At least 20 visual fields containing 500 cells were analyzed in three independent replications in each group. Scale bar, 20 μm .

(E and F) Dot blot analysis of R-loops in mESCs, mESC-differentiated cells (mESC-differ), and differentiated cells over-expressing *Lnc530* (E). The quantification of dot blot intensity was shown in (F). dsDNA was used as internal reference.

(G) Neutral comet assay revealed that over-expression of *Lnc530* in mESC-differentiated cells reduced the level of DNA DSBs. Left panel was the quantification of comet tail length. Right panel showed the representative images. At least 200 tails were analyzed in each group. Scale bars, 400 μm .

All experiments were repeated three times with similar results. Data were shown as mean \pm SD, two-tailed Student's *t*-test.

Table S1. Protein candidates interacting with *Lnc530*.

Protein	Unique peptide counts					
	Negative control		Normal culture		4mM HU treatment	
	replicate 1	replicate 2	replicate 1	replicate 2	replicate 1	replicate 2
DDX5	0	0	0	1	2	3
TDP-43	0	0	0	0	1	2
Rplp0	0	0	0	0	2	1
Rpl3	0	0	0	0	2	2
Rpl4	0	0	0	0	5	5
Rpl7a	0	0	0	0	1	3
Rpl9	0	0	0	0	1	1
Rpl18	0	0	0	0	2	1
Rpl19	0	0	0	0	2	2
Rpl26	0	0	0	0	1	2
Rpl28	0	0	0	0	1	1
Rpl29	0	0	0	0	1	1
Hist1h1d	0	0	0	0	1	1
Rps8	0	0	0	0	1	3
Rps4x	0	0	0	0	2	2
GAPDH	0	0	0	0	1	1
Jup	0	0	0	0	3	1
Ctnnb1	0	0	0	0	1	1

Table S2. Primers for RT-qPCR.

Gene	RT-PCR primers	
<i>Lnc530</i>	Forward	5'-TCAGATAGCTCAGGTCGTCAGT-3'
	Reverse	5'-AAGGATTGTTGTAGAATCATGTCTG-3'
<i>Actb</i>	Forward	5'-TTGCTGATCCACATCTGCTGGAAGG-3'
	Reverse	5'-GTGTGACGTTGACATCCGTAAAGAC-3'
<i>Sox2</i>	Forward	5'-GACCGGCGGCAACCAGAAG-3'
	Reverse	5'-GCGCTTGCTGATCTCCGAGT-3'
<i>Nanog</i>	Forward	5'-AAAGGATGAAGTGCAAGCGGTG-3'
	Reverse	5'-TTCCAGATGCGTTCACCAGATA-3'
<i>Oct4</i>	Forward	5'-AGGAGCTAGAACAGTTTGCCAA-3'
	Reverse	5'-TTCTCATTGTTGTCGGCTTCCT-3'
<i>Pax6</i>	Forward	5'-TACCAGTGTCTACCAGCCAAT-3'
	Reverse	5'-TGCACGAGTATGAGGAGGTCT-3'
<i>Flk1</i>	Forward	5'-TTTGGCAAATACAACCCTTCAGA-3'
	Reverse	5'-GCAGAAGATACTGTCACCACC-3'
<i>Foxa2</i>	Forward	5'-CCCTACGCCAACATGAACTCG-3'
	Reverse	5'-GTTCTGCCGGTAGAAAGGGA-3'

Table S3. Primers for PCR cloning.

Gene	RACE and ORF or CDS primers	
<i>Lnc530</i> RACE	GSP1	5'-AAACCGAATCCAAGAACACATC-3'
	GSP2	5'-GGTCGTCAGTCCTATAACCAGGT-3'
<i>Lnc530</i> -full length (pCDH-CMV)	Forward	5'-TGACCTCCATAGAAGATTCTAGATAGGGCGATTGGG CCCTCTAGAT-3'
	Reverse	5'-AGATCCTTCGCGGCCGCGGATCCCACTAAGGGTGG CTTGTGTGT-3'
<i>Lnc530</i> - Δ 2301-3172 (pCDH-CMV)	Forward	5'-TGACCTCCATAGAAGATTCTAGATAGGGCGATTGGG CCCTCTAGAT-3'
	Reverse	5'-AGATCCTTCGCGGCCGCGGATCCGGGAATGAATGA CTTCTACTTG-3'
<i>Lnc530</i> -ORF	Forward	5'-GAATTCATGCATGCTCGAGCGGCCGCCA-3'
	Reverse	5'-GAATTCCTACTTATCGTCGCATCCTTGTAATCCCTAT TCAACATTGTACTTGAAGT-3'
Δ 2301-3172 (pEasy-Blunt)	Forward	5'-TAGGGCGATTGGGCCCTCTAGATGCA-3'
	Reverse	5'-GGGAATGAATGACTTCTACTTGGCT-3'
Δ 1-2300 (pEasy-Blunt)	Forward	5'-AGCACCTCAGCACCTTTCTCCAGGA-3'
	Reverse	5'-CACTAAGGGTGGCTTGTGTGTTTGC-3'
Δ 2601-3172 (pEasy-Blunt)	Forward	5'-TAGGGCGATTGGGCCCTCTAGATGCA-3'
	Reverse	5'-GGTCTCACGTAATCAACTCATCATGA-3'
Δ 2301-2600 (pEasy-Blunt)	Forward-1	5'-TAGGGCGATTGGGCCCTCTAGATGCA-3'
	Reverse-1	5'-CATACAGACATTAGTTAACTTGTATGGGAATGAATGA CTTCTACTTGGCT-3'
	Forward-2	5'-AGCCAAGTAGAAGTCATTCATTCCCATACAAGTTAAC TAATGTCTGTATG-3'
	Reverse-2	5'-CACTAAGGGTGGCTTGTGTGTTTGC-3'
<i>Tdp-43</i> - Δ RRM1	Forward-1	5'-TCTAGAGCCACCATGGACTACAAGGACGACGATGAC AAGATGTCTGAATATATTCGG-3
	Reverse-1	5'-TGGGCTTTGCTTAGAGTTGGGAAGGAGGTCAGATG TTTTCTGGACT-3
	Forward-2	5'-AGTCCAGAAAACATCTGACCTCCTTCCCAACTCTAA GCAAAGCCCA-3
	Reverse-2	5'-CGCGGATCCCTACATTCCCAGCCAGAAG-3
<i>Ddx5</i> -CDS	Forward	5'-CTAGCTAGCTCGAGTTATTCTAGTGACCGAGA-3'
	Reverse	5'-CGCGGATCCTTATTGAGAATACCCTGTTGGCAT-3'
<i>Rnase H1</i> -CDS	Forward	5'-CTAGCTAGCTTCTATGCGGTGAGGAGAGGCCG-3'
	Reverse	5'-CGCGGATCCTCAGTCCTCAGACTGCTTCGCT-3'

Table S4. Short hairpin RNA sequences used for gene knockdown.

Gene	shRNA mir
<i>Lnc530-shRNA1</i>	5'-TGCTGTTGACAGTGAGCGCGGTGAGAATTTGAAGAAAGAGTA GTGAAGCCACAGATGTACTCTTTCTTCAAATTCTCACCATGCCTA CTGCCTCGGA -3'
<i>Lnc530-shRNA2</i>	5'-TGCTGTTGACAGTGAGCGCCTTGTCTTTCTGATTCTCAAGTAG TGAAGCCACAGATGTACTTGAGAATCAGAAAGACAAGATGCCTAC TGCCTCGGA-3'
<i>Tdp-43-shRNA</i>	5'-TGCTGTTGACAGTGAGCGCTCGCTGTGTCTCATTATCTAGTAG TGAAGCCACAGATGTACTAGATAATGAGACACAGCGATTGCCTAC TGCCTCGGA-3'
<i>Ddx5-shRNA</i>	5'-TGCTGTTGACAGTGAGCGACCTGTGGTTGTCTTAACTAATTAG TGAAGCCACAGATGTAATTAGTTAAGACAACCACAGGCTGCCTAC TGCCTCGGA-3'

Table S5. Antibody and drug information.

Antibodies	Source	Dilution
DDX5 Rabbit monoclonal antibody	Abcam; Cat# ab126730	1:1000 (IF) 1:1000 (IB) 1:50 (IP)
TDP-43 Rabbit monoclonal antibody	Abcam; Cat# ab190963	1:200 (IF) 1:1000 (IB) 1:50 (IP)
S9.6 Mouse monoclonal antibody	Kerafast; Cat# ENH001	1:100 (IF)
S9.6 Mouse monoclonal antibody	ATCC; Cat# HB-8730	1:50 (IP) 1:500 (Dot blot)
dsDNA Mouse monoclonal antibody	Santa Crus Biotechnology; Cat# sc-58749	1:500 (Dot blot)
BrdU Rat monoclonal antibody, also used for immunostaining of CldU	Abcam; Cat# ab6326	1:1000 (IF)
γH2AX Rabbit polyclonal antibody	Cell Signaling Technology; Cat# 9718	1:1000 (IB) 1:1000 (IF)
RNase H1 Mouse monoclonal antibody	Santa Crus Biotechnology; Cat# sc-376326	1:500 (IB)
GAPDH Mouse monoclonal antibody	Sangon Biotech; Cat# D190090	1:1000 (IB)
Flag Rabbit polyclonal antibody	Sigma; Cat# F7425	1:1000 (IB)
SOX2 Rabbit polyclonal antibody	Merck Millipore; Cat# AB5603	1:500 (IB)
NANOG Rabbit polyclonal antibody	Abcam; Cat# ab70482	1:1000 (IB)
OCT4 Rabbit polyclonal antibody	Abcam; Cat# ab19857	1:1000 (IB)
Fibrillarin Rabbit monoclonal antibody	Cell Signaling Technology; Cat# 2639	1:1000 (IF)
Histone H3 Mouse monoclonal antibody	TransGen Biotech; Cat# HL102	1:1000 (IB)
ACTB Mouse monoclonal antibody	Abclonal; Cat# AC004	1:1000 (IB)
Goat anti-Rat IgG secondary antibody, Alexa Fluor Cy3	Thermo Fisher Scientific; Cat# A-10522	1:500 (IF)
Goat anti-Mouse IgG (H+L) secondary antibody, Alexa Fluor 488	Thermo Fisher Scientific; Cat# A-11029	1:500 (IF)
Donkey anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluor 555	Thermo Fisher Scientific; Cat# A-31570	1:500 (IF)
Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific; Cat# A-11034	1:500 (IF)
Donkey anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor 555	Thermo Fisher Scientific; Cat# A-31572	1:500 (IF)
Goat anti-Mouse IgG (H+L)	Thermo Fisher Scientific;	1:5000 (IB)

secondary antibody, HRP	Cat# 31430	
Goat anti-Rabbit IgG (H+L) secondary antibody, HRP	Thermo Fisher Scientific; Cat# 31460	1:5000 (IB)
5-Chloro-2'-deoxyuridine (CldU)	Sigma-Aldrich; Cat# C6891	20 μ M
Hydroxylurea (HU)	Sigma-Aldrich; Cat# H8627	4 mM
5,6-Dichlorobenzimidazole 1- β -D- Ribofuranoside	Tokyo Chemical Industry; Cat# ; D4292	10 μ M

Supplemental experimental procedures

Mouse ESC differentiation

Mouse embryoid bodies (EBs) formation and mESCs differentiation were performed as previously described (Fu et al., 2022). Briefly, mESCs were cultured in mESC medium without LIF in an ultra-low attachment dish for EBs formation. EBs were collected at the indicated time points. For regular differentiation, mESCs were cultured in mESCs medium without LIF for 14 days in culture dish pre-coated with gelatin.

Rapid amplification of cDNA ends (RACE)

The RACE protocol is based on published Smart-seq2 protocol (Picelli et al., 2014). The oligo-dT primer, template-switching oligos (TSO) and ISPCR primers were synthesized as described in the Smart-seq2 method. 5' end and 3' end of *Lnc530* were amplified using ISPCR primer and gene specific primer 1 (GSP1) or gene specific primer 2 (GSP2), respectively. The products were cloned into pEasy-Blunt plasmid (Transgene, CB101-01). The primers were listed in [Table S3](#).

Neutral comet assay

The neutral comet assay was performed as previously described (Zhao et al., 2015). Briefly, cells were dissociated and mixed with low-melting agarose, and spread onto the slides. The cells on slides were lysed in lysis buffer [2.5 M NaCl, 100 mM EDTA, 10 mM Tris (pH = 8.3), 1% N-lauroylsarcosine, and 1% Triton X-100] for 1 h. Electrophoresis was performed for 30 min in buffer [300 mM sodium acetate, 100 mM Tris (pH = 8.3)]. The slides were washed and fixed with 100% ethanol, followed by staining with DAPI. The comet tails were analyzed by Komet 7 comet assay software (Andor Technology).

Karyotyping

The cultured cells were treated with 120 ng/mL colcemid (Gibco, 15212-012) for 2 h, digested and re-suspended with hypotonic buffer (75 mM KCl) for 15 min at 37 °C. The hypotonic cells were fixed with methanol/glacial acetic acid (3:1), dropped onto glass slides, and air-dried. Samples were stained with 3% Giemsa buffer (Gibco, 10092013) at 37 °C for 30 min. Images were captured using a Leica TCS SP5 confocal microscope system (Leica Microsystems).

Cell proliferation assay

The cell proliferation assay was performed with Click-iT™ EdU Kit (Thermo, C10337). Briefly, cells were labeled with 20 μM EdU (5-ethynyl-2'-deoxyuridine) for 2 h, digested and suspended in 4% PFA for 30 min, followed by treatment with 0.2% Triton X-100 for 20 min. Cells were incubated with reaction buffer prepared according to the manufacturer's protocol and sorted by flow cytometry.

In vitro RNA pull down

RNA was *in vitro* transcribed with T7 RNA polymerase (Fermentas, EP0111), and labeled with biotin using Pierce™ RNA 3' End Desthiobiotinylation Kit (Thermo, 20163). *In vitro* RNA

pull down assay was performed as previously described (Wang et al., 2014). Briefly, labeled RNA in RNA structure buffer (10 mM Tris pH = 7.0, 0.1 M KCl, 10 mM MgCl₂) was heated to 95 °C for 2 min, then left it at room temperature to allow proper secondary structure formation. Samples were subjected to RNA pull down using Pierce™ Magnetic RNA-Protein Pull-Down Kit (Thermo, 20164). The proteins were eluted and detected by immunoblotting analysis.

Mass spectrometry

1) In-Gel digestion

Gel section was cut into approx. 1 mm³ pieces and destained using 30 mM K₃Fe(CN)₆:100 mM Na₂S₂O₃=1:1 solution. The reaction was stopped by removing supernatant, added 100 uL of 100 mM ammonium bicarbonate for 20 min, followed by dehydration in 100% acetonitrile for 5 min. Gel slice was then reduced by 20 mM DTT at 56 °C for 30 min and alkylated with 100 mM IAA at room temperature for 20 min. Gel slice was then incubated in a 50 mM ammonium bicarbonate solution containing 10 ng/mL trypsin (Promega Biotech Co., Madison, WI, USA.) overnight. Peptides were extracted with 0.1% TFA/80% acetonitrile, dried by vacuum centrifugation.

2) LC-MS/MS

The reverse phase high-performance liquid chromatography (RP-HPLC) separation was achieved on the Easy nano-LC system (Thermo Fisher Scientific) using a self-packed column (75 µm×150 mm; 3 µm ReproSil-Pur C18 beads, 120 Å, Dr.Maisch GmbH, Ammerbuch, Germany) at a flow rate of 300 nL/min. The mobile phase A of RP-HPLC was 0.1% formic acid in water, and B was 0.1% formic acid in acetonitrile. The peptides were eluted using a gradient (2–90% mobile phase B) over a 60 min period into a nano-ESI orbitrap Q Exactive mass spectrometer (Thermo Fisher Scientific). The mass spectrometer was operated in data-dependent mode with each full MS scan (m/z 300-1600) followed by MS/MS for the 10 most intense ions with the parameters: ≥ +2 precursor ion charge, 2 Da precursor ion isolation window and 27 normalized collision energy of HCD. Dynamic Exclusion™ was set for 45 s. The full mass and the subsequent MS/MS analyses were scanned in the Orbitrap analyzer with R = 70,000 and R= 17,500, respectively.

3) Data analysis

The MS data were analyzed using the software MaxQuant (<http://maxquant.org/>, version 1.6.5.0). Proteins were identified by searching MS and MS/MS data of peptides against a Mouse proteome database (downloaded from Uniprot). Trypsin/P was selected as the digestive enzyme with two potential missed cleavages. The search included variable modifications of methionine oxidation and N-terminal acetylation, and fixed modification of cysteine carbamidomethylation. Peptides of minimum 6 amino acids and maximum of two missed cleavages were allowed for the analysis. For peptide and protein identification, false discovery rate was set to 0.01.

Supplemental references

Fu, H., Wang, T., Kong, X., Yan, K., Yang, Y., Cao, J., Yuan, Y., Wang, N., Kee, K., Lu, Z.J., *et al.* (2022). A Nodal enhanced micropeptide NEMEP regulates glucose uptake during mesendoderm differentiation of embryonic stem cells. *Nat Commun* 13, 3984. 10.1038/s41467-022-31762-x

Picelli, S., Faridani, O.R., Bjorklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* 9, 171-181. 10.1038/nprot.2014.006

Wang, P., Xue, Y., Han, Y., Lin, L., Wu, C., Xu, S., Jiang, Z., Xu, J., Liu, Q., and Cao, X. (2014). The STAT3-binding long noncoding RNA Inc-DC controls human dendritic cell differentiation. *Science* 344, 310-313. 10.1126/science.1251456

Zhao, B., Zhang, W.D., Duan, Y.L., Lu, Y.Q., Cun, Y.X., Li, C.H., Guo, K., Nie, W.H., Li, L., Zhang, R., *et al.* (2015). Filia Is an ESC-Specific Regulator of DNA Damage Response and Safeguards Genomic Stability. *Cell stem cell* 16, 684-698. 10.1016/j.stem.2015.03.017