Science Advances
AAAS

Supplementary Materials for

Endogenous retroviruses shape pluripotency specification in mouse embryos

Sergio de la Rosa *et al.*

Corresponding author: Nabil Djouder, ndjouder@cnio.es

Sci. Adv. **10**, eadk9394 (2024)
DOI: 10.1126/sciadv.adk9394

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

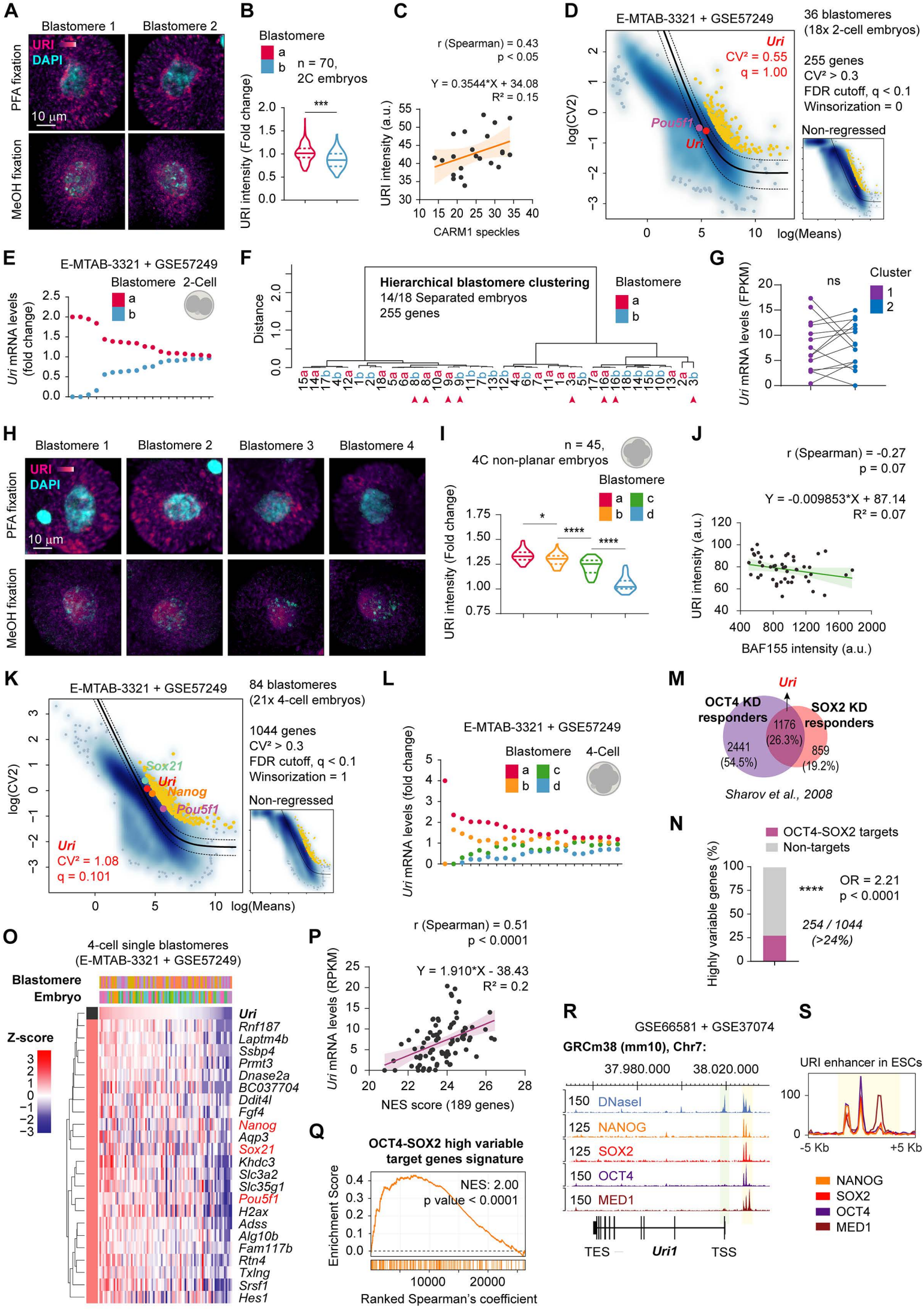

A
Blastomere 1
Blastomere 2
URI
DAPI
PFA fixation
MeOH fixation
10 µm
B
Blastomere
a
b
n = 70,
2C embryos
***
URI intensity (Fold change)
C
r (Spearman) = 0.43
p < 0.05
Y = 0.3544*X + 34.08
R² = 0.15
URI intensity (a.u.)
CARM1 speckles
D
E-MTAB-3321 + GSE57249
36 blastomeres
(18x 2-cell embryos)
Uri
CV² = 0.55
q = 1.00
255 genes
CV² > 0.3
FDR cutoff, q < 0.1
Winsorization = 0
Pou5f1
Uri
Non-regressed
log(CV2)
log(Means)
E
E-MTAB-3321 + GSE57249
Blastomere 2-Cell
Uri mRNA levels (fold change)
F
Hierarchical blastomere clustering
14/18 Separated embryos
255 genes
Blastomere
Distance
G
ns
Cluster
Uri mRNA levels (FPKM)
H
Blastomere 1
Blastomere 2
Blastomere 3
Blastomere 4
URI
DAPI
PFA fixation
MeOH fixation
10 µm
I
n = 45,
4C non-planar embryos
Blastomere
URI intensity (Fold change)
J
r (Spearman) = -0.27
p = 0.07
Y = -0.009853*X + 87.14
R² = 0.07
URI intensity (a.u.)
BAF155 intensity (a.u.)
K
E-MTAB-3321 + GSE57249
84 blastomeres
(21x 4-cell embryos)
1044 genes
CV² > 0.3
FDR cutoff, q < 0.1
Winsorization = 1
Non-regressed
Sox21
Uri
Nanog
Pou5f1
Uri
CV² = 1.08
q = 0.101
log(CV2)
log(Means)
L
E-MTAB-3321 + GSE57249
Blastomere 4-Cell
Uri mRNA levels (fold change)
M
Uri
OCT4 KD responders
SOX2 KD responders
1176 (26.3%)
2441 (54.5%)
859 (19.2%)
Sharov et al., 2008
N
OCT4-SOX2 targets
Non-targets
Highly variable genes (%)
****
OR = 2.21
p < 0.0001
254 / 1044 (>24%)
O
4-cell single blastomeres
(E-MTAB-3321 + GSE57249)
Blastomere
Embryo
Z-score
Uri
Rnf187
Laptm4b
Ssbp4
Prmt3
Dnase2a
BC037704
Ddit4l
Fgf4
Nanog
Aqp3
Sox21
Khdc3
Slc3a2
Slc35g1
Pou5f1
H2ax
Adss
Alg10b
Fam117b
Rtn4
Txlng
Srsf1
Hes1
P
r (Spearman) = 0.51
p < 0.0001
Y = 1.910*X - 38.43
R² = 0.2
Uri mRNA levels (RPKM)
NES score (189 genes)
Q
OCT4-SOX2 high variable target genes signature
NES: 2.00
p value < 0.0001
Enrichment Score
Ranked Spearman's coefficient
R
GSE66581 + GSE37074
GRCm38 (mm10), Chr7:
37.980.000
38.020.000
DNaseI
NANOG
SOX2
OCT4
MED1
TES
Uri1
TSS
S
URI enhancer in ESCs
-5 Kb
+5 Kb
NANOG
SOX2
OCT4
MED1

Fig. S1: URI is heterogeneously expressed and concurs with blastomere pluripotency bias in the early embryo. **A**, IF of URI in 2C embryos using paraformaldehyde (PFA) or methanol fixation (MeOH). Scale bar, 10 μm. **B**, URI intensity in grouped blastomeres from (A). Unpaired *t* test; ***$P < 0.001$. **C,** Linear regression and correlation analysis of URI intensity and CARM1 speckles in 2C embryos (See **Fig. 1D**, **E**). a.u. acronym referred arbitrary units. **D**, Highly variable gene analysis among single 2C embryo blastomeres. Single cell counts were regressed out for inter-embryo variability, non-regressed plot is also depicted. Mean (solid line) and 95% confidence intervals (dashed line) for the relationship between the square of the coefficient of variation (CV^2) and the average gene expression level (Means) are plotted. Yellow dots mark significant genes. Other color dots identify respective genes. **E**, Intra-embryo normalized *Uri* mRNA levels in 2C embryos from single blastomere obtained from indicated RNA-seq datasets. **F**, Hierarchical clustering analysis of single 2C embryo blastomeres using top highly variable gene candidates from (D). **G**, Paired normalized *Uri* mRNA levels in clustered blastomeres from (f). Paired *t* test; ns, non-significant. **H**, IF of URI in 4C embryos using PFA or MeOH fixation. Scale bar, 10 μm. Individual pictures for single blastomeres are shown. **I**, URI intensity for grouped blastomeres from (a). Unmatched one-way ANOVA analysis (Tukey post-hoc test); *$P < 0.05$, ****$P < 0.0001$. **J**, Linear regression and correlation analysis of URI and BAF155 intensity in non-planar shaped 4C embryos (See **Fig. 1, I and J**). a.u. acronym referred arbitrary units. **K**, Gene dispersion analysis among single 4C embryo blastomeres. Single cell counts were regressed out for inter-embryo variability, non-regression plot is also depicted. Mean (solid line) and 95% confidence intervals (dashed line) are plotted for the relationship between the square of the coefficient of variation and (CV^2) and the average gene expression level (Means). Yellow dots mark top 300 significant variable genes.

Other color dots identify respective genes. **L**, Intra-embryo normalized *Uri* mRNA levels in 4-cell embryos from single cell blastomere obtained from indicated RNA-seq datasets. **M**, Venn diagram depicting total and shared number of potential target genes for OCT4 and SOX2. **N**, Abundance of OCT4 and SOX2 target genes identified as highly variable genes across 4C embryo blastomeres analysis from (K). Fisher's test is used; ****$P < 0.0001$. **O**, Heat map showing *Uri*-ranked Z-scored mRNA expression of OCT4 and SOX2 highly variable potential target genes in 4C embryo single blastomeres. **P**, Linear regression and correlation analysis of normalized *Uri* mRNA levels and normalized enrichment score (NES) from Gene set enrichment analysis (GSEA) for OCT4 and SOX2 highly variable potential target gene signature. **Q**, GSEA of ranked gene correlation with *Uri*. **R**, Genomic read coverage of chromatin immunoprecipitation-sequencing analysis (ChIP-seq) in the *Uri* locus for different pluripotent core transcriptional factors. *Uri* locus is shown from the minus strand. Super-enhancer region (yellow band) is identified upstream of the transcription starting site (TSS, green band). **S**, Density plot from panel (R). Total number of embryos is referred to in each panel. Repository accession numbers for sequencing dataset analysis are indicated in respective panel and compiled in **table S1**.

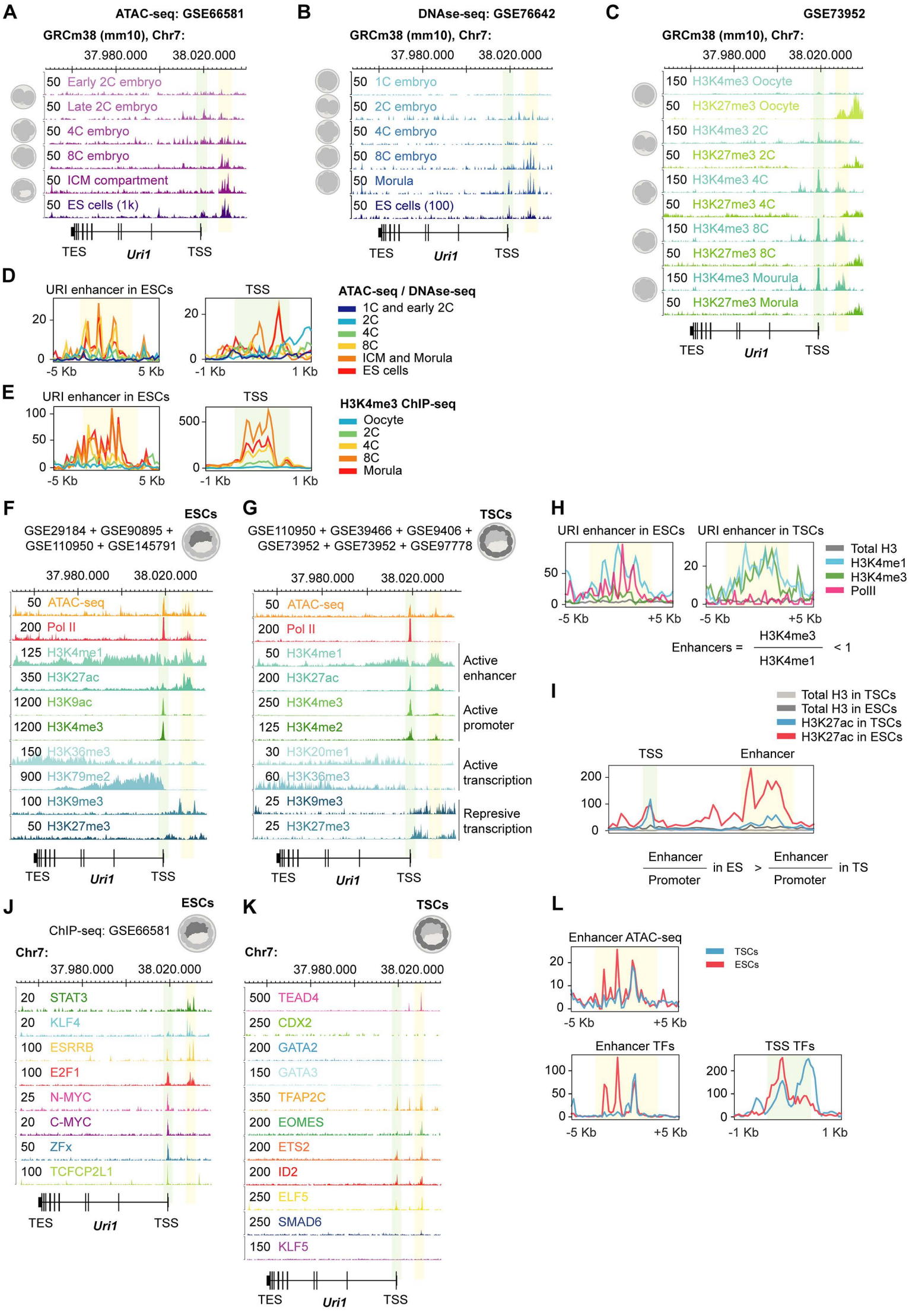

A
ATAC-seq: GSE66581
GRCm38 (mm10), Chr7:
37.980.000
38.020.000
50 Early 2C embryo
50 Late 2C embryo
50 4C embryo
50 8C embryo
50 ICM compartment
50 ES cells (1k)
TES
Uri1
TSS
B
DNAse-seq: GSE76642
GRCm38 (mm10), Chr7:
37.980.000
38.020.000
50 1C embryo
50 2C embryo
50 4C embryo
50 8C embryo
50 Morula
50 ES cells (100)
TES
Uri1
TSS
C
GSE73952
GRCm38 (mm10), Chr7:
37.980.000
38.020.000
150 H3K4me3 Oocyte
50 H3K27me3 Oocyte
150 H3K4me3 2C
50 H3K27me3 2C
150 H3K4me3 4C
50 H3K27me3 4C
150 H3K4me3 8C
50 H3K27me3 8C
150 H3K4me3 Mourula
50 H3K27me3 Morula
TES
Uri1
TSS
D
URI enhancer in ESCs
TSS
-5 Kb
5 Kb
-1 Kb
1 Kb
ATAC-seq / DNAse-seq
1C and early 2C
2C
4C
8C
ICM and Morula
ES cells
E
URI enhancer in ESCs
TSS
H3K4me3 ChIP-seq
Oocyte
2C
4C
8C
Morula
F
ESCs
GSE29184 + GSE90895 +
GSE110950 + GSE145791
37.980.000
38.020.000
50 ATAC-seq
200 Pol II
125 H3K4me1
350 H3K27ac
1200 H3K9ac
1200 H3K4me3
150 H3K36me3
900 H3K79me2
100 H3K9me3
50 H3K27me3
TES
Uri1
TSS
G
TSCs
GSE110950 + GSE39466 + GSE9406 +
GSE73952 + GSE73952 + GSE97778
37.980.000
38.020.000
50 ATAC-seq
200 Pol II
50 H3K4me1
200 H3K27ac
250 H3K4me3
125 H3K4me2
30 H3K20me1
60 H3K36me3
25 H3K9me3
25 H3K27me3
Active enhancer
Active promoter
Active transcription
Represive transcription
TES
Uri1
TSS
H
URI enhancer in ESCs
URI enhancer in TSCs
-5 Kb
+5 Kb
Total H3
H3K4me1
H3K4me3
PolII
Enhancers = H3K4me3 / H3K4me1 < 1
I
Total H3 in TSCs
Total H3 in ESCs
H3K27ac in TSCs
H3K27ac in ESCs
TSS
Enhancer
Enhancer / Promoter in ES > Enhancer / Promoter in TS
J
ESCs
ChIP-seq: GSE66581
Chr7:
37.980.000
38.020.000
20 STAT3
20 KLF4
100 ESRRB
100 E2F1
25 N-MYC
20 C-MYC
50 ZFx
100 TCFCP2L1
TES
Uri1
TSS
K
TSCs
Chr7:
37.980.000
38.020.000
500 TEAD4
250 CDX2
200 GATA2
150 GATA3
350 TFAP2C
200 EOMES
200 ETS2
200 ID2
250 ELF5
250 SMAD6
150 KLF5
TES
Uri1
TSS
L
Enhancer ATAC-seq
TSCs
ESCs
-5 Kb
+5 Kb
Enhancer TFs
TSS TFs
-1 Kb
1 Kb

Fig. S2: *Uri* super-enhancer region encompasses pluripotency onset and lineage segregation in mouse embryos. A, Genomic views of unique mapped reads for transposase-accessible chromatin-sequencing assay (ATAC-seq) across different pre-implanted mouse embryo stages. **B**, Unique read coverage from DNase I-hypersensitive site-sequencing (DNase-seq) along successive pre-implanted mouse embryo stages. **C**, View of mapped reads at *Uri* genomic region from epigenetic chromatin immunoprecipitation-sequencing analysis (ChIP-seq) markers in murine embryos. *Uri* locus is shown from the minus strand. Super-enhancer region (yellow band) is identified upstream of the transcription starting site (TSS, green band). **D**, Density plot from combine read coverages from (A, B). **E**, Density plot of H3K4me3 ChIP-seq from (C). **F** and **G**, Selected genomic views of chromatin immunoprecipitation-sequencing analysis (ChIP-seq) for epigenetic *Uri* locus in pluripotent mESCs (F) or trophoblast stem cells (TSC) (G). *Uri* locus is shown from the minus strand. Enhancer region (yellow band) is identified upstream of the transcription starting site (TSS, green band). **H**, Density plots for H3K27ac ChIP-seq from (F, G). Data showed higher enhancer/promoter ration in ESCs compared to TSCs. **I,** Density plots for indicated ChIP-seq in ESCs (left) or TSCs (right panel) from (F, G). Enhancers are defined to have high monomethylation at H3K4 (low H3K4me3/H3K4me1 ratio). **J** and **K**, View of unique mapped read coverages from essential transcriptional factor ChIP-seq datasets in ES (J) and TS (K) cells. **L**, Density plots for ATAC-seq and merged transcriptional factor ChIP-seq reads at promoter or enhancer locus in both murine TSCs and ESCs. Acronyms for 1C (1-cell embryo), 2C (2-cell embryo), 4C (4-cell embryo), 8C (8-cell embryo), ICM (Inner Cell Mass compartment) and ES (Embryonic Stem cell) are used. Repository accession numbers for sequencing dataset analysis are indicated in respective panel and compiled in **table S1**.

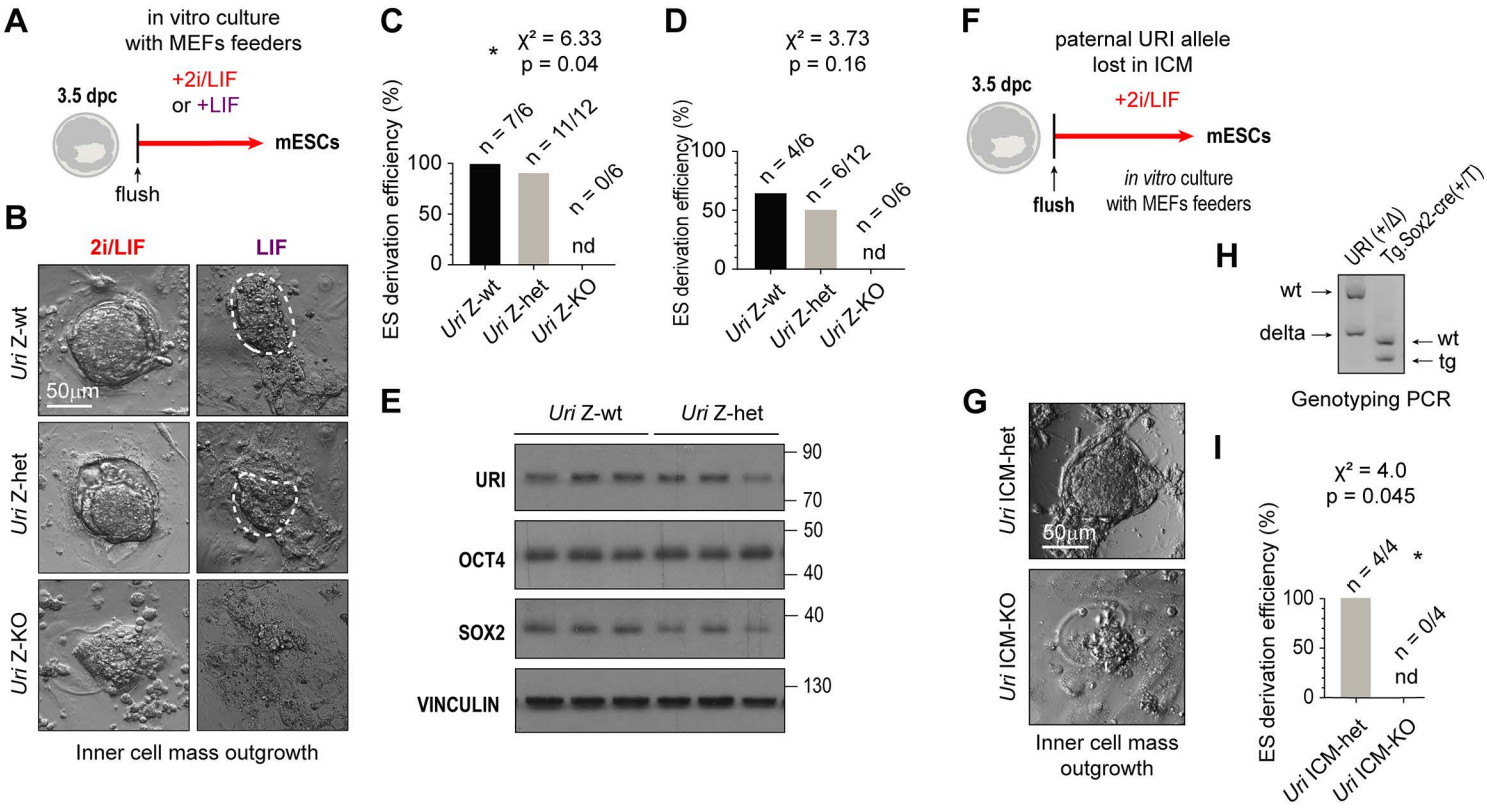

A
in vitro culture
with MEFs feeders
3.5 dpc
+2i/LIF
or +LIF
mESCs
flush
B
2i/LIF
LIF
Uri Z-wt
Uri Z-het
Uri Z-KO
50µm
Inner cell mass outgrowth
C
* χ² = 6.33
p = 0.04
ES derivation efficiency (%)
n = 7/6
n = 11/12
n = 0/6
nd
Uri Z-wt
Uri Z-het
Uri Z-KO
D
χ² = 3.73
p = 0.16
n = 4/6
n = 6/12
n = 0/6
nd
E
Uri Z-wt
Uri Z-het
URI
OCT4
SOX2
VINCULIN
90
70
50
40
40
130
F
paternal URI allele
lost in ICM
3.5 dpc
+2i/LIF
mESCs
flush
in vitro culture
with MEFs feeders
G
Uri ICM-het
Uri ICM-KO
50µm
Inner cell mass
outgrowth
H
URI (+/Δ)
Tg.Sox2-cre(+/T)
wt
delta
wt
tg
Genotyping PCR
I
χ² = 4.0
p = 0.045
ES derivation efficiency (%)
n = 4/4
*
n = 0/4
nd
Uri ICM-het
Uri ICM-KO

Fig. S3: Zygotic URI-depleted embryos do not establish mESCs cultures. A, Pluripotent ESC derivation protocol from early pre-implanted blastocyst, placed *in vitro* in feeder gelatin-coated layer in the presence of leukemia inhibitory factor (LIF) and the addition of 2i (GSK3 and MEK inhibitors) from (See **Fig. 3, A and B**). **B**, Representative bright field images of ICM outgrowths 5 days after plating embryos *ex vivo* in the presence of 2i/LIF or LIF from (A). White dashed lines limit ICM outgrowth in LIF condition. Scale bars, 50 μm. **C**, Efficiency of mESCs lineage derivation from (B) under 2i/LIF condition. χ^2 test applied for the expected versus observed events; $*P < 0.05$. **D**, Efficiency of mESCs lineage derivation from (B) under LIF condition. χ^2 test applied for the expected versus observed events; ns, non-significant. **E**, Western blot analysis of established pluripotent cell lineages from (B). **F**, Pluripotent ESC derivation protocol of *Uri* ICM-KO embryos generated by crossing the *Sox2*-Cre mice with the *Uri* lox mouse model to generate URI heterozygous (ICM-het) and knockout (ICM-KO) embryos (See **Fig. 3, P and Q**). **G**, Representative bright field images of ICM outgrowths from embryos cultured *ex vivo* in the presence of 2i/LIF from (F). Scale bars, 50 μm. **H**, Representative DNA electrophoresis of PCR data for the genotyping of derived pluripotent cell lineage from (G). **I**, Efficiency of mESCs lineage derivation from (G). χ^2 test applied for the expected versus observed events. Total number of embryos is referred to in each panel.

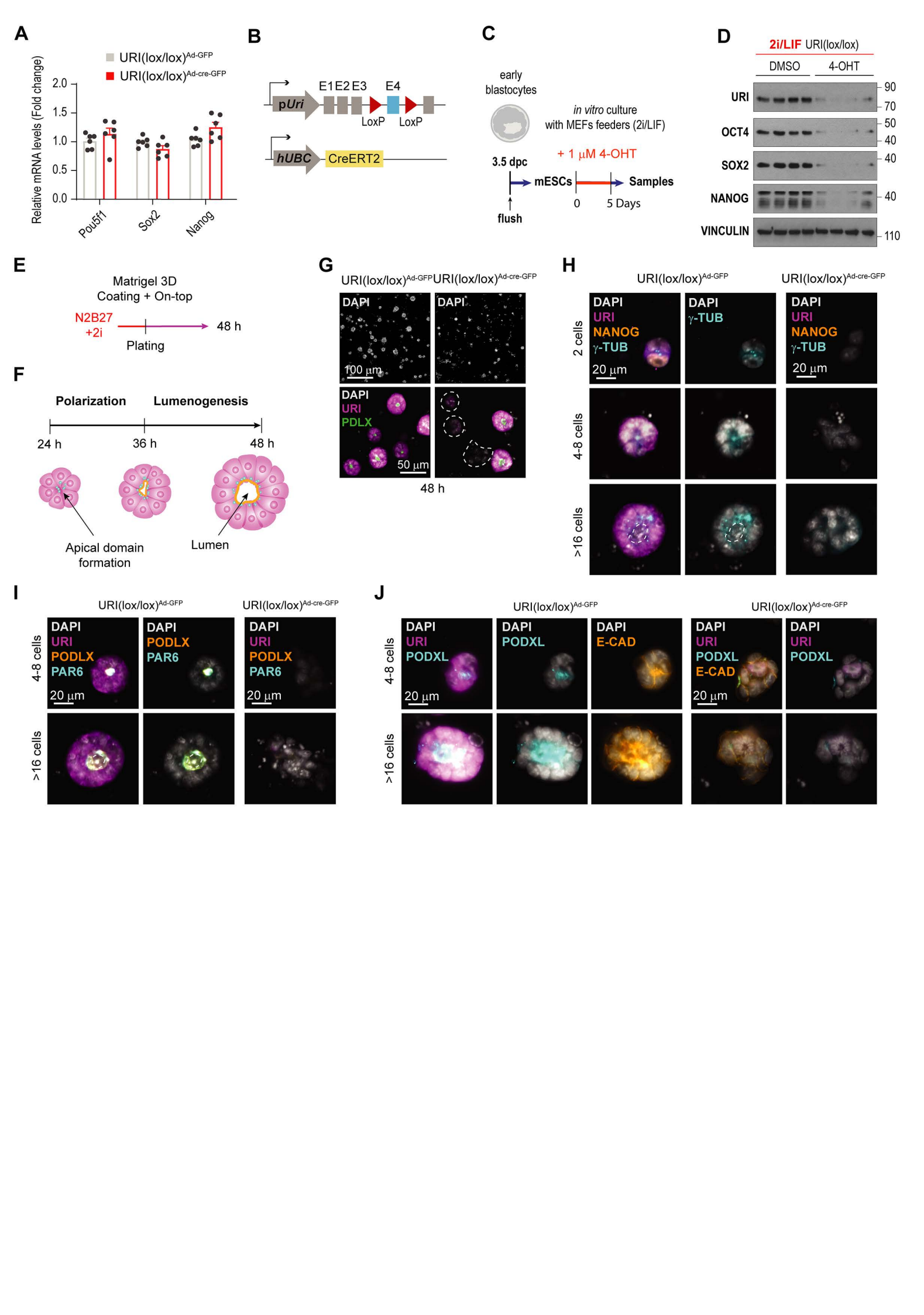

A
URI(lox/lox)Ad-GFP
URI(lox/lox)Ad-cre-GFP
Relative mRNA levels (Fold change)
2.0
1.5
1.0
0.5
0.0
Pou5f1
Sox2
Nanog
B
pUri
E1 E2 E3
E4
LoxP
LoxP
hUBC
CreERT2
C
early blastocytes
in vitro culture with MEFs feeders (2i/LIF)
+ 1 μM 4-OHT
3.5 dpc
mESCs
Samples
0
5 Days
flush
D
2i/LIF URI(lox/lox)
DMSO
4-OHT
URI
OCT4
SOX2
NANOG
VINCULIN
90
70
50
40
40
40
110
E
Matrigel 3D
Coating + On-top
N2B27
+2i
48 h
Plating
F
Polarization
Lumenogenesis
24 h
36 h
48 h
Apical domain formation
Lumen
G
URI(lox/lox)Ad-GFP
URI(lox/lox)Ad-cre-GFP
DAPI
100 μm
DAPI
URI
PDLX
50 μm
48 h
H
URI(lox/lox)Ad-GFP
URI(lox/lox)Ad-cre-GFP
2 cells
4-8 cells
>16 cells
DAPI
URI
NANOG
γ-TUB
20 μm
I
URI(lox/lox)Ad-GFP
URI(lox/lox)Ad-cre-GFP
4-8 cells
>16 cells
DAPI
URI
PODLX
PAR6
20 μm
J
URI(lox/lox)Ad-GFP
URI(lox/lox)Ad-cre-GFP
4-8 cells
>16 cells
DAPI
URI
PODXL
E-CAD
20 μm

Fig. S4: URI loss compromises the pluripotent potential of EPI cells. A, qRT-PCR of pluripotent core factor genes in mESCs treated with adenoviruses expressing either Cre recombinase combined with an enhanced green fluorescence protein EGFP (AdV-Cre-EGFP) or EGFP alone (AdV-EGFP) as control and cultured in presence of 2i/LIF. Data are represented as mean ± s.e.m. **B**, Scheme of *Uri* lox mice crossed with *hUBC*-CreERT2 mice. **C**, Murine ESCs derivation protocol from early pre-implanted blastocyst from (B), placed *in vitro* in feeder layer in the presence of 2i/LIF and treated with 4-hydroxytamoxifen (4-OHT) to deplete URI *in vitro*. **D**, Western blot analysis of 4-OHT-treated pluripotent mESCs from (C). **E**, Experimental scheme to generate rosette-like structure mimicking post-implanted epiblast (EPI) compartment. **F**, Scheme of the main morphological changes of the embryonic EPI compartment during implantation. **G**, IF of URI and PODLX in AdV-EGFP- or AdV-Cre-EGFP-infected rosette-like mESCs after 48 hours. Dashed white lines indicate URI negative colonies. Scale bar, 100 and 50 μm. **H**, Representative IF of URI, naïve pluripotent marker NANOG and the polarization marker ɣ-TUBULIN in different size rosette-like colonies. Scale bar, 20 μm. **I**, Representative IF of URI and lumenogenesis markers PODLX and PAR6 in different size rosette-like colonies. Scale bar, 20 μm. **J**, IF of URI, PODLX and the epithelial marker E-Cadherin in different size rosette-like colonies. Scale bar, 20 μm.

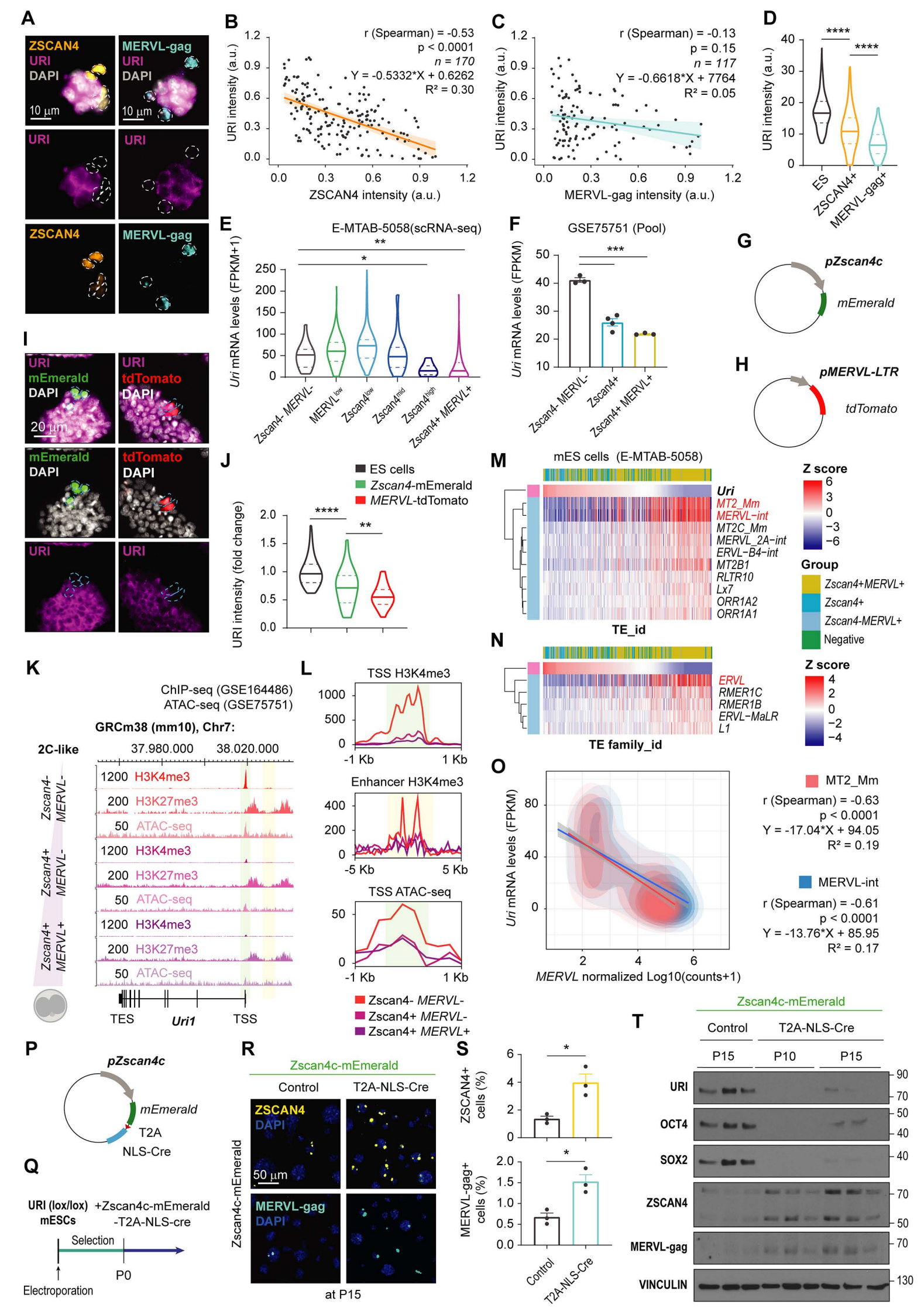

A
ZSCAN4
URI
DAPI
10 μm
MERVL-gag
URI
DAPI
10 μm
URI
URI
ZSCAN4
MERVL-gag
B
r (Spearman) = -0.53
p < 0.0001
n = 170
Y = -0.5332*X + 0.6262
R² = 0.30
URI intensity (a.u.)
ZSCAN4 intensity (a.u.)
C
r (Spearman) = -0.13
p = 0.15
n = 117
Y = -0.6618*X + 7764
R² = 0.05
URI intensity (a.u.)
MERVL-gag intensity (a.u.)
D
****
****
URI intensity (a.u.)
ES
ZSCAN4+
MERVL-gag+
E
E-MTAB-5058(scRNA-seq)
**
*
Uri mRNA levels (FPKM+1)
Zscan4- MERVL-
MERVL^low
Zscan4^low
Zscan4^mid
Zscan4^high
Zscan4+ MERVL+
F
GSE75751 (Pool)
***
Uri mRNA levels (FPKM)
Zscan4- MERVL-
Zscan4+
Zscan4+ MERVL+
G
pZscan4c
mEmerald
H
pMERVL-LTR
tdTomato
I
URI
mEmerald
DAPI
20 μm
URI
tdTomato
DAPI
mEmerald
DAPI
tdTomato
DAPI
URI
URI
J
ES cells
Zscan4-mEmerald
MERVL-tdTomato
****
**
URI intensity (fold change)
M
mES cells (E-MTAB-5058)
Uri
MT2_Mm
MERVL-int
MT2C_Mm
MERVL_2A-int
ERVL-B4-int
MT2B1
RLTR10
Lx7
ORR1A2
ORR1A1
TE_id
Z score
Group
Zscan4+MERVL+
Zscan4+
Zscan4-MERVL+
Negative
N
ERVL
RMER1C
RMER1B
ERVL-MaLR
L1
TE family_id
Z score
K
ChIP-seq (GSE164486)
ATAC-seq (GSE75751)
GRCm38 (mm10), Chr7:
2C-like
37.980.000
38.020.000
Zscan4- MERVL-
Zscan4+ MERVL-
Zscan4+ MERVL+
1200 H3K4me3
200 H3K27me3
50 ATAC-seq
TES
Uri1
TSS
L
TSS H3K4me3
Enhancer H3K4me3
TSS ATAC-seq
-1 Kb
1 Kb
-5 Kb
5 Kb
Zscan4- MERVL-
Zscan4+ MERVL-
Zscan4+ MERVL+
O
Uri mRNA levels (FPKM)
MERVL normalized Log10(counts+1)
MT2_Mm
r (Spearman) = -0.63
p < 0.0001
Y = -17.04*X + 94.05
R² = 0.19
MERVL-int
r (Spearman) = -0.61
p < 0.0001
Y = -13.76*X + 85.95
R² = 0.17
P
pZscan4c
mEmerald
T2A
NLS-Cre
Q
URI (lox/lox) mESCs
+Zscan4c-mEmerald
-T2A-NLS-cre
Selection
P0
Electroporation
R
Zscan4c-mEmerald
Control
T2A-NLS-Cre
ZSCAN4
DAPI
50 μm
MERVL-gag
DAPI
Zscan4c-mEmerald
at P15
S
ZSCAN4+ cells (%)
MERVL-gag+ cells (%)
Control
T2A-NLS-Cre
T
Zscan4c-mEmerald
Control
T2A-NLS-Cre
P15
P10
P15
URI
OCT4
SOX2
ZSCAN4
MERVL-gag
VINCULIN

Fig. S5. URI loss is a hallmark of the 2C-like cells. A, IF of wild-type mESCs for URI and the 2C-like markers ZSCAN4 or MERVL-gag. Dashed white outlines represent totipotent-like cells. Scale bar, 10 μm. **B** and **C**, Linear regression model and normalized correlation analysis of URI and ZSCAN4 (B) or MERVL-gag (C) intensity from (A). Results are plotted as arbitrary units (a.u.). **D**, URI levels of pluripotent mESCs and the two totipotent-like ZSCAN4+ or MERVL-gag+ cells from (B and C respectively). a.u. acronym referred arbitrary units. One-way ANOVA test (Tukey post-hoc test); ****$P < 0.0001$. **E**, Normalized *Uri* mRNA expression levels across different pluripotent or 2C-like cell populations from single cell RNA-seq datasets. *x* axis is sorted left-to-right from pluripotent population (*Zscan4-; MERVL-LTR-*) to totipotent-like cells (*Zscan4+; MERVL-LTR+*). Median and quartiles are shown. One-way ANOVA test (Tukey post-hoc test); *$P < 0.05$, ***$P < 0.001$. **F**, Normalized *Uri* mRNA expression across different pluripotent or 2C-like populations from a pooled RNA-seq dataset. One-way ANOVA test (Tukey correction); ***$P < 0.001$. **G**, Plasmid construct for labelling transient 2C-like cells with active transcription of the Zscan4c promoter. **H**, Plasmid construct for labelling transient 2C-like cells with active transcription of the MERVL-LTR promoter. **I**, IF of URI in 2C-like Zscan4-mEmerald or MERVL-LTR-tdTomato reporter mESCs. Scale bar, 20 μm. Median and quartiles are shown. **J**, URI intensity in wild-type ES, *Zscan4+* or *MERVL-LTR+* cells from (I). ANOVA with Tukey post hoc test; **$P < 0.01$, ****$P < 0.001$. **K**, Selected genomic views from chromatin-immunoprecipitation sequencing (ChIP-seq) of H3K4me3 and H3K27me3 and assay for transposase-accessible chromatin (ATAC-seq) in different pluripotent (Zscan4-; MERVL-LTR-) or totipotent-like (*Zscan4+; MERVL-LTR-* and *Zscan4+; MERVL-LTR+*) cell populations. *Uri* locus is shown from the minus strand. Super-enhancer region (yellow band) is identified upstream of the transcription starting site (TSS, green band). **L**, Density plots

for H3K4me3 ChIP-seq and ATAC-seq at the specified regions from (L). **M** and **N**, Heat map showing *Uri*-ranked Z-scored mRNA expression of transposable elements clustered per element (M) or family (N) repeats from single pluripotent and totipotent-like cell RNA-seq datasets. **O**, Linear regression model and correlation analysis of normalized URI and *MERVL-LTR* (MT2_Mm) (red) or *MERVL-int* (blue) counts from (M and N). **P**, Plasmid construct for depleting URI in 2C-like cells under the ZSCAN4c promoter. **Q**, Scheme of the timing for URI depletion in 2C-like cells by electroporating plasmid from (P). **R**, IF of mESCs for the totipotent-like marker ZSCAN4 or MERVL-gag in long term culture 2C-like dependent URI depleted cells after 15 passages (P15). Scale bar, 50 μm. **S**, Abundance of 2C-like cells from (R) after P15. Data is represented as mean ± s.e.m. *t* test analysis; $*P < 0.05$. **T**, Western blot analysis of post-electroporated mESCs at indicated number of passages as reported in panel (Q). Repository accession numbers for sequencing dataset analysis are indicated in respective panel and compiled in **table S1**.

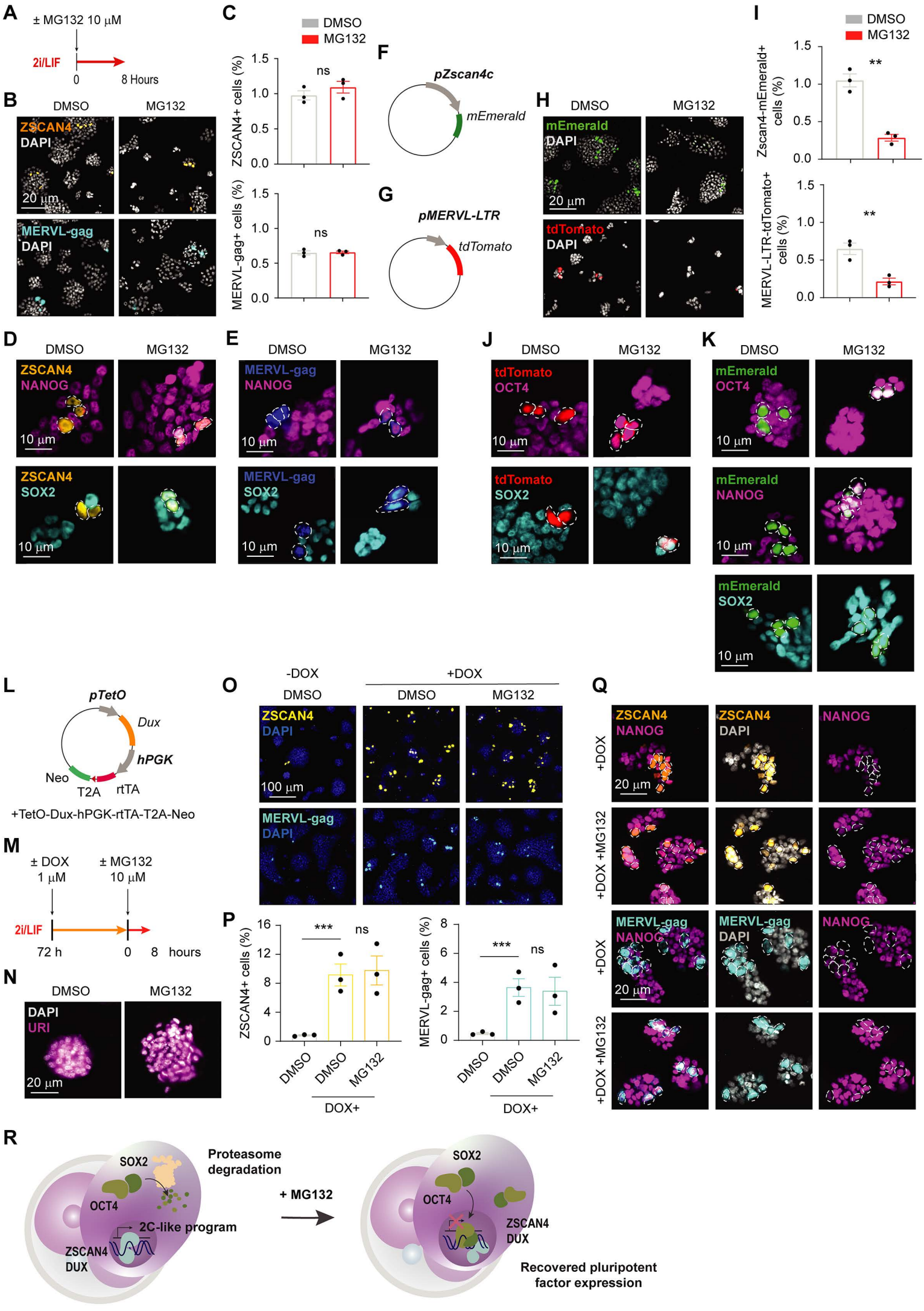

A
± MG132 10 μM
2i/LIF
0
8 Hours
B
DMSO
MG132
ZSCAN4
DAPI
20 μm
MERVL-gag
DAPI
C
DMSO
MG132
ns
ZSCAN4+ cells (%)
MERVL-gag+ cells (%)
F
pZscan4c
mEmerald
G
pMERVL-LTR
tdTomato
H
mEmerald
DAPI
tdTomato
I
Zscan4-mEmerald+ cells (%)
MERVL-LTR-tdTomato+ cells (%)
**
D
ZSCAN4
NANOG
SOX2
10 μm
E
MERVL-gag
J
tdTomato
OCT4
K
mEmerald
L
pTetO
Dux
hPGK
Neo
T2A
rtTA
+TetO-Dux-hPGK-rtTA-T2A-Neo
M
± DOX
1 μM
± MG132
10 μM
2i/LIF
72 h
0
8 hours
N
URI
O
-DOX
+DOX
100 μm
P
***
Q
+DOX
+DOX +MG132
R
SOX2
OCT4
2C-like program
ZSCAN4
DUX
Proteasome degradation
+ MG132
Recovered pluripotent factor expression

Fig. S6. Proteasome inhibition reinstates the expression of the pluripotent core factors in the totipotent-like cells. A, Experimental scheme of mESCs treated with 10 μM of MG132 for 8 hours. **B**, IF of ZSCAN4 or MERVL-gag in MG132-treated mESCs. Red arrowheads mark positive 2C-like cells. Scale bar, 20 μm. **C**, Abundance of 2C-like cells in MG132-treated ZSCAN4 (up) or MERVL-gag (bottom) positive cells from (B). Data is represented as mean ± s.e.m. *t* test was applied; ns, non-significant. **D** and **E**, IF of pluripotent NANOG and SOX2 in ZSCAN4 (D) or MERVL-gag (E) positive cells before and after MG132 treatment. Dashed white outlines show 2C-like cells. Scale bar, 10 μm. **F** and **G**, Plasmid constructs and experimental scheme of mESCs stably transfected with *Zscan4*-mEmerald (F) or *MERVL-LTR*-tdTomato (G) constructs and treated with the proteasome inhibition MG132. **H**, Fluorescence of endogenous signal for the respective reporter in mESCs treated as described in (A). Scale bar, 20 μm. **I**, Abundance of 2C-like cells in MG132-treated *Zscan4*-mEmerald (top) or *MERVL-LTR*-tdTomato (bottom) models from (H). Data is represented as mean ± s.e.m. *t* test was applied; $^{**}P < 0.01$. **J** and **K**, IF of pluripotent OCT4, SOX2 and NANOG in MG132-treated 2C-like MERVL-LTR-tdTomato (J) or Zscan4c-mEmerald (K) reporter cell lines. **L**, Plasmid construct for inducible DUX expression to induce 2C-like state after electroporation in mESCs. **M**, Experimental scheme for mESCs cultured in 2i/LIF and treated first with doxycycline (DOX) to induce expression of DUX, and then with MG132 to inhibit proteasome. **N**, IF of URI in MG132-treated mESCs. Scale bar, 20 μm. **O**, IF of ZSCAN4 and MERVL-gag in mESCs from (M). Scale bar, 100 μm. **P**, Abundance of ZSCAN4 (left) and MERVL (right) positive cells from (O). **Q**, IF of pluripotent NANOG in DUX-induced 2C-like cells treated with MG132 from (M). Scale bar, 20 μm. **R**, Scheme depicting the pluripotent core factors reinstation in totipotent-like cells after proteasome inhibition. Total number of embryos is referred in each panel.

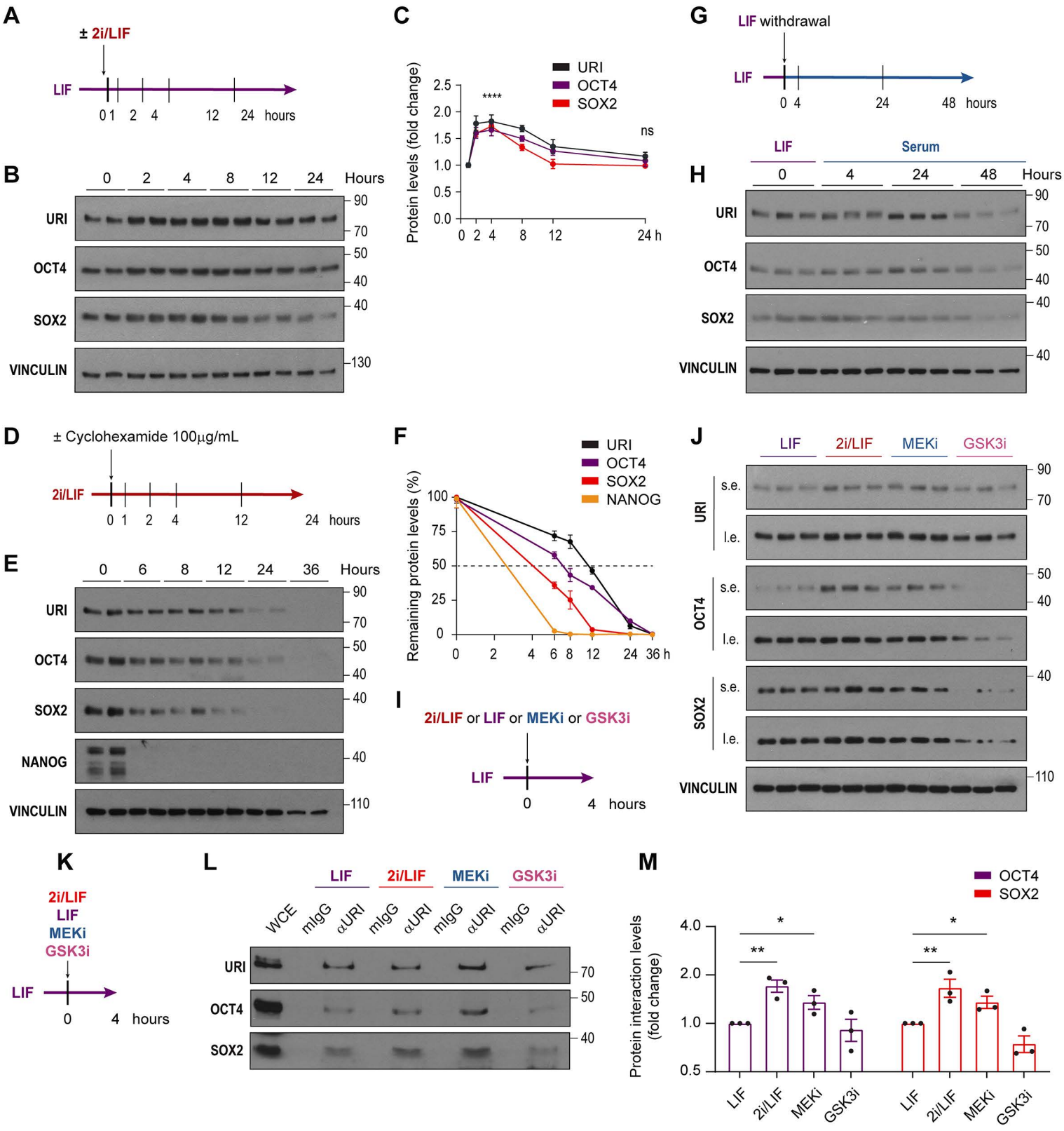

A
± 2i/LIF
LIF
0 1 2 4 12 24 hours
B
0 2 4 8 12 24 Hours
URI
OCT4
SOX2
VINCULIN
90
70
50
40
40
130
C
URI
OCT4
SOX2
Protein levels (fold change)
****
ns
0 2 4 8 12 24 h
D
± Cyclohexamide 100µg/mL
2i/LIF
0 1 2 4 12 24 hours
E
0 6 8 12 24 36 Hours
URI
OCT4
SOX2
NANOG
VINCULIN
90
70
50
40
40
40
110
F
URI
OCT4
SOX2
NANOG
Remaining protein levels (%)
100
75
50
25
0
0 2 4 6 8 12 24 36 h
G
LIF withdrawal
LIF
0 4 24 48 hours
H
LIF
Serum
0 4 24 48 Hours
URI
OCT4
SOX2
VINCULIN
90
70
50
40
40
40
I
2i/LIF or LIF or MEKi or GSK3i
LIF
0 4 hours
J
LIF
2i/LIF
MEKi
GSK3i
URI
OCT4
SOX2
s.e.
l.e.
VINCULIN
90
70
50
40
40
110
K
2i/LIF
LIF
MEKi
GSK3i
LIF
0 4 hours
L
LIF
2i/LIF
MEKi
GSK3i
WCE
mIgG
αURI
URI
OCT4
SOX2
70
50
40
M
OCT4
SOX2
Protein interaction levels (fold change)
4.0
2.0
1.0
0.5
*
**
LIF
2i/LIF
MEKi
GSK3i

Fig. S7. URI interacts with OCT4 and SOX2 under naïve pluripotency conditions. **A**, Experimental scheme of induction of pluripotent naïve state by addition of 2i/LIF overtime in mESCs cultured *in vitro*. **B**, Western blot (WB) analysis of naïve pluripotency induced following 2i/LIF treatment from (A). **C**, Protein levels of URI, OCT4 and SOX2 normalized to starting timepoint from (B). Data is represented as mean ± s.e.m. one-way ANOVA (Tukey post-hoc test); ****$P < 0.0001$; ns, non-significant. **D**, Experimental scheme of protein turnover analysis by treating mESC with cyclohexamide overtime. **E**, WB analysis from experiment described in (D). **F**, Protein levels of URI, OCT4 and SOX2 normalized to starting timepoint from (E). **G**, Experimental scheme of LIF withdrawal over time in mESCs cultured *in vitro* to induce differentiation or pluripotency shutdown. **H**, WB analysis from the experiment described in (G). **I**, Experimental scheme of treatment with either 2i/LIF, LIF alone, MEKi/LIF or GSK3i/LIF in mESCs during 4 hours. **J**, WB analysis from experiment described in (I). **K**, Experimental scheme of mESCs culture in presence of either 2i, 2i/LIF, MEK inhibitor (MEKi) or GSK3 inhibitor (GSK3i) during 4 hours. **L**, Co-immunoprecipitation of URI in mESCs treated as described in (K). **M**, Interaction levels of URI with OCT4 or SOX2 in different medium conditions normalized to pulldown abundancy. One way ANOVA test (Tukey post-hoc test); *$P < 0.05$, **$P < 0.01$; ns, non-significant.

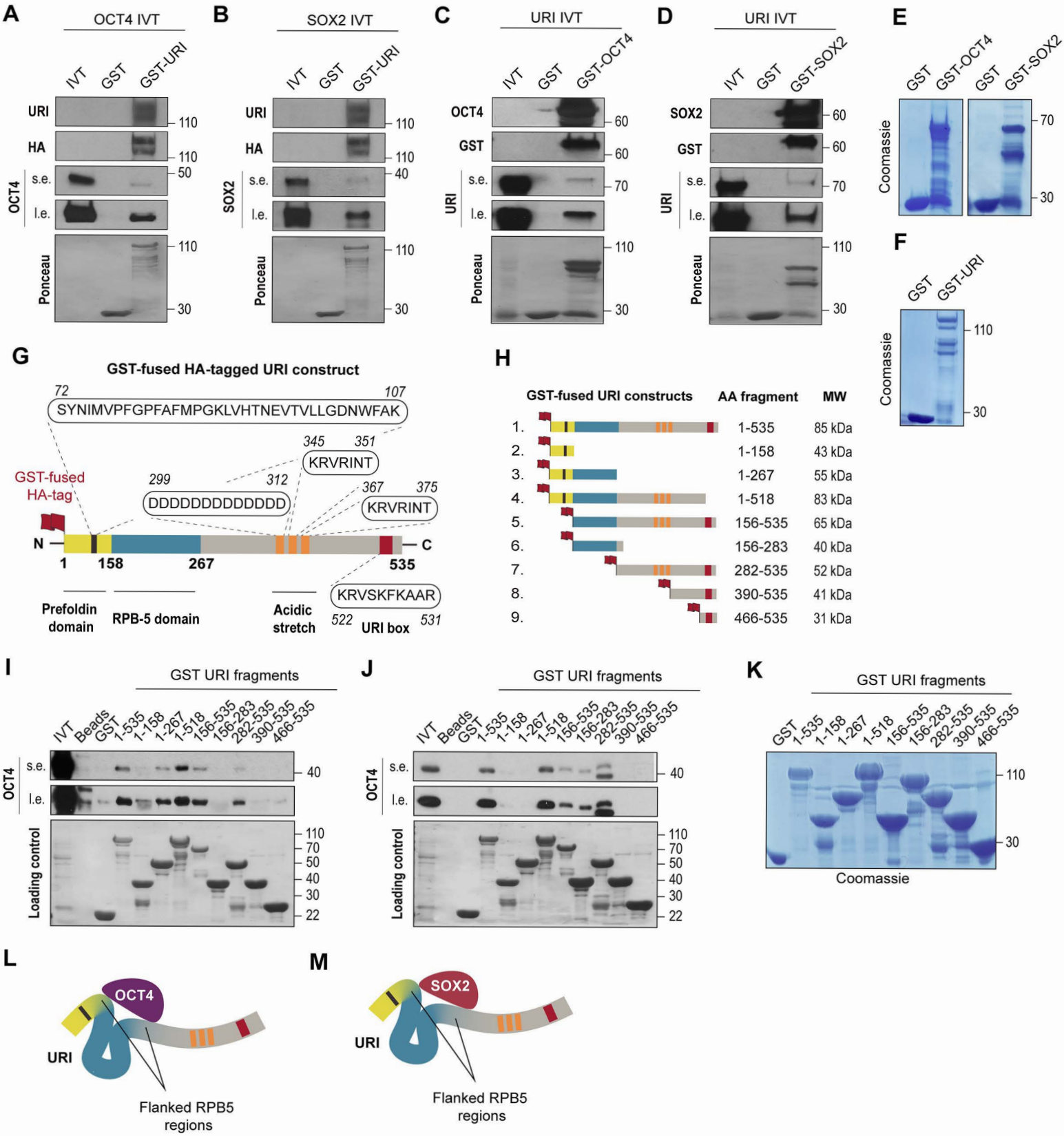
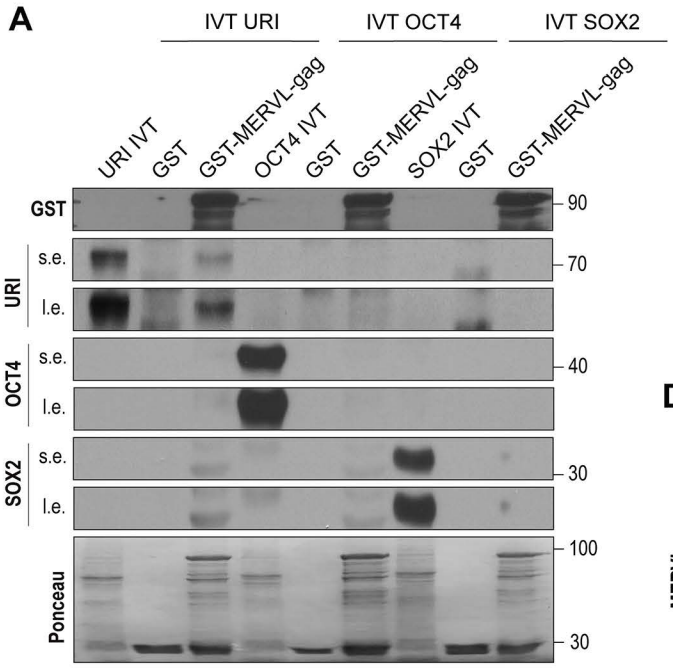
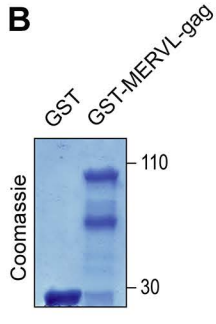
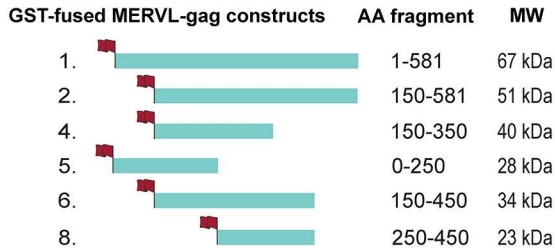
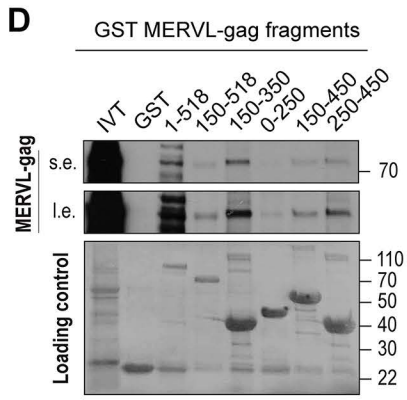

A
OCT4 IVT
IVT
GST
GST-URI
URI
HA
OCT4
s.e.
l.e.
Ponceau
110
110
50
110
30
B
SOX2 IVT
IVT
GST
GST-URI
URI
HA
SOX2
s.e.
l.e.
Ponceau
110
110
40
110
30
C
URI IVT
IVT
GST
GST-OCT4
OCT4
GST
URI
s.e.
l.e.
Ponceau
60
60
70
110
30
D
URI IVT
IVT
GST
GST-SOX2
SOX2
GST
URI
s.e.
l.e.
Ponceau
60
60
70
110
30
E
GST
GST-OCT4
GST
GST-SOX2
Coomassie
70
30
F
GST
GST-URI
Coomassie
110
30
G
GST-fused HA-tagged URI construct
72
107
SYNIMVPFGPFAFMPGKLVHTNEVTVLLGDNWFAK
345
351
KRVRINT
GST-fused
HA-tag
299
312
DDDDDDDDDDDDD
367
375
KRVRINT
N
C
1
158
267
535
Prefoldin
domain
RPB-5 domain
Acidic
stretch
522
URI box
531
KRVSKFKAAR
H
GST-fused URI constructs
AA fragment
MW
1.
1-535
85 kDa
2.
1-158
43 kDa
3.
1-267
55 kDa
4.
1-518
83 kDa
5.
156-535
65 kDa
6.
156-283
40 kDa
7.
282-535
52 kDa
8.
390-535
41 kDa
9.
466-535
31 kDa
I
GST URI fragments
IVT
Beads
GST
1-535
1-158
1-267
1-518
156-535
156-283
282-535
390-535
466-535
OCT4
s.e.
l.e.
Loading control
40
110
70
50
40
30
22
J
GST URI fragments
IVT
Beads
GST
1-535
1-158
1-267
1-518
156-535
156-283
282-535
390-535
466-535
OCT4
s.e.
l.e.
Loading control
40
110
70
50
40
30
22
K
GST URI fragments
GST
1-535
1-158
1-267
1-518
156-535
156-283
282-535
390-535
466-535
110
30
Coomassie
L
OCT4
URI
Flanked RPB5
regions
M
SOX2
URI
Flanked RPB5
regions

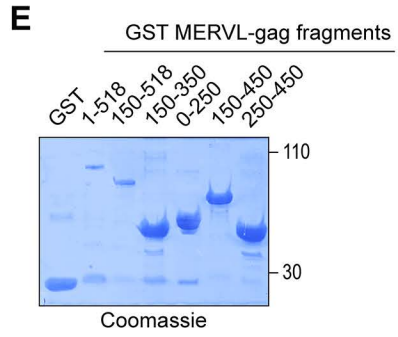
Fig. S8. URI interacts with OCT4 and SOX2 trough RPB5 flanking binding sites. A and **B**, Pulldown of *in vitro* translation (IVT) OCT4 (A) and SOX2 (B) protein using GST or GST-URI. s.e. and l.e. exposures are shown. **C** and **D**, Pulldown of IVT URI protein using GST, GST-OCT4 (C) or GST-SOX2 (D). s.e. and l.e. exposures are shown. **E**, Coomassie blue staining for purified OCT4 and SOX2 GST-fused proteins. **F**, Coomassie blue staining for purified URI GST-fused protein. **G**, Scheme of URI protein domains and amino acid sequences. **H**, GST-fused URI constructs strategy for pull down experiments. **I**, and **J**, Pulldown of OCT4 (I) or SOX2 (J) IVT using GST-fused URI fragments from (H). **K,** Coomassie blue staining for purified URI GST-fused fragments from (H). **L** and **M**, Scheme representing the binding regions of OCT4 (L) and SOX2 (M) on URI protein.











F

> MERVL-gag
> HERVL-gag

```
MNLLK....YWNWLVDPPALSTIETSPDSLPPGSSENFEDPWLKLYSELKEANDLDFLN.......ELGDSVHKAFYKMGKKSENDFTGWLLLVSVEKMMNERKELCDKIERLQTQVNDLK  110
MEFFRWFWGFWSWLLN...MIRAKNAKNSIS.NNMENTDSPWHDDFRELCKINAFDTPDSPLMRGKQFSDSIHNTFDHMWRTKEHNEAGWLLPSSLERVMKENNELRDSVSQLQKQILSLK  117

VAKCVLEENLLSCSNRAQVAENQTETLIVRLAELQRKFKSQPQSVSTVKVRALIGKEWDPTTWDGDVWEDHVEAENFESSDSQGFAPPEEVVPSAPP...LEIMPSPHPEINFAESDKPAM  227
SAMIALS..LISCRERAEIMEKQTQALMR.VADLQQQVHAQPHQVSTVKVRTLIGKEWDHATWNGDIWAS.....PSPVVATSPPLPMLPSAFPPLSEEINPALPEATVMASPEAVAR  226

TFTTDVSQGPP...IVSSRPVTRLKAKQAPRGEVESVVHEEIRYTTKELNEFANSFKQKPGEYVWEWILRVWDKGGRNIKLEQAEFIDMGPLSRDSRFNTEARIVKKGVKSLFEWLAEVF.  344
QDNVDSPQEPPPTPLFASRPITRLKSWQAPRGEVQSVTHEEVHYTQKELLEFSNLYKQKSGEQAWEWILRVWDNGGRNIELDQAEFIDLGPLSRDSAFNVAAQGVKKGSNSLFAWLAEIWE  346

IKRWPTGNDLEMPDIPWLSVDEGILRLREIAMLEWIYCVKHNCPQWEGPEDMPFTSSIRRKLVRGAPAHLKGFVLSLFLVPDLSIGDASAQLDELNSLGLVCFRGNKGQVAALNHRRQ..G  463
IKRWPTVSELEMPDLPWFNVEEGIQRLREIGMVEWISHFRPTHPSWEGPEDIPLTNALQNRFVR.APASLKSPVIALLCMSDLTVGTTVTQLQNLNTMGIICSRGGRGQVAALNHQRQIKG  465

DSSYYNGQRRQKNVYNNIPSNGQHRRGEIYNGMARLDLWYWLTNHGVSRNEIHRKPTAYLFDLYKQKNSQTNERKATLDCGKQPNERKATLDRGKQQSRPVNQFPDLRQFADPEPLE  581
GHSYCNGQQRQ........SSNQ.......NSLTHVELWHWLINHGVPRSEIDRKPTAFLLNLYKQKTSRSNGQKTNLNYKNR.............ESWPLNQFPDLSQFTDPEPLE  557
```

Non-conserved
Similar
Identical

47.47% Aa sequence identical in 1-518 fragment
72.73% Aa sequence identical in 250-350 fragment

Fig. S9. URI interacts with MERVL-gag through a highly conserved region to the human counterpart transposon, HERVL.

A, Pulldown of IVT URI, OCT4 and SOX2 proteins using GST or GST-MERVL-gag. s.e. and l.e. exposures are shown. **B**, Coomassie blue staining for purified MERVL-gag GST-fused protein. **C**, GST-fused MERVL-gag constructs strategy for pull down experiments. **D**, Pulldown of URI IVT using GST-fused MERVL-gag fragments from (C). **E,** Coomassie blue staining for purified MERVL-gag GST-fused fragments from (C). **F**, Sequence alignment between MERVL-gag and the reconstituted HERVL-gag protein.

Supplementary tables

Table S1. Accession numbers and references for NGS analysis.

Figure	NGS	Dataset	Accession	PMID
Fig. 1C	RNA-seq	Pre-implanted mouse embryos (bulk)	GSE45719	24408435
fig, S, 1D to G	RNA-seq	Single cell dataset for 2-cell embryos	E-MTAB-3321	27015307
			GSE57249	25096407
fig. S, 1K to Q	RNA-seq	Single cell dataset for 4-cell embryos	E-MTAB-3321	27015307
			GSE57249	25096407
fig. S, 1R and S	ChIP-seq	Pluripotent core factors: NANOG, SOX2, OCT4 and MED1 in mESCs	GSE44286	23582322
	DNAse-seq	Chromatin accessibility in ES cells by DNAse I hypersensitivity	GSE37074	ENCODE
fig. S2A	ATAC-seq	Chromatin accessibility in embryos	GSE66581	27309802
fig. S2B	DNAse-seq	Chromatin accessibility in embryos by DNAse I hypersensitivity	GSE76642	27259149
fig. S2C	ChIP-seq	H3K4me3 and 27me3 in embryos	GSE73952	27626379
Fig. 2, A and 2, D to H	RNA-seq	Single cells datasets for morula and blastocysts	GSE45719	24408435
fig S2F	ChIP-seq	H3K9me3, K27ac, K27me3, K79me2, K36me3, K4me1, K4me3 in ES cells	GSE90895	28111071
	ATAC-seq	Chromatin accessibility in ES cells	GSE110950	31628347
	ChIP-seq	PolII in ES cells	GSE145791	32616013
	ChIP-seq	H3K9ac in ES cells	GSE29184	22763441
fig. S2G	ChIP-seq	H3K4me1, K27ac, K20me1 and K36me3 in TS cells	GSE39406	23178118

	ChIP-seq	H3K4me3, K4me2 and 27me3 in TS cells	GSE73952	27626379
	ChIP-seq	H3K9me3 in TS cells	GSE97778	29686265
	ATAC-seq	Chromatin accessibility in TS cells	GSE110950	31628347
	ChIP-seq	PolII in TS cells	GSE39406	23178118
fig. S2J	ChIP-seq	TFs in ES cells: STAT3, KLF4, ESRRB, E2F1, NMYC, CMYC, ZFx and TCFCP2L1	GSE66581	27309802
fig. S2K	ChIP-seq	CDX2 in TS cells	GSE42207	23396136
	ChIP-seq	GATA2 and GATA3 in TS cells	GSE92287	28232602
	ChIP-seq	KLF5 in TS cells	GSE109250	31777916
	ChIP-seq	ELF5, EOMES, ETS2, ID2, SMAD6, TFAP2C and TEAD4 in TS cells	GSE110950	31628347
Fig. 6, A to C, and fig. S5, E and M to O	RNA-seq	Single cell isolated by Zscan4 reporter	E-MTAB-5058	27681430
Fig. 6, D to F	RNA-seq	Reporter for Dppa2 (GFP)	GSE120950	30692203
		Reporter for Dppa4 (GFP)	GSE120950	30692203
		Knockdown of Ncl (shRNA)	GSE100939	29937225
		Reporter for Sp110 (GFP)	GSE120950	30692203
		Knockout of METTL3	GSE146467	33658714
		Knockdown of UBE2i (shRNA)	GSE70863	26365490
		Reporter for Usp3 (GFP)	GSE120950	30692203
		Knockdown of UBA2 (shRNA)	GSE70863	26365490
		Knockout of SMCHD1	GSE126467	33523915
		Treatment with 2-DG	GSE113671	31932739
		Knockdown of SUMO2 (shRNA)	GSE70863	26365490

Knockdown of TRIM28 (shRNA)	GSE70863	26365490
Reporter for Bahd1 (GFP)	GSE120950	30692203
Reporter for Hdac9 (GFP)	GSE120950	30692203
Knockdown of SAE1 (shRNA)	GSE70863	26365490
Knockout of SMG7 (sgRNA)	GSE133234	32523982
Knockdown of LINE1 (ASO)	GSE100939	9937225
Knockout of ZMYM2	GSE119819	32032525
Overexpression DPPA2	GSE127811	31226106
Overexpression DCAF11	GSE132746	33357405
Knockout of KAP1	GSE74278	27003935
Knockdown of UBC9 (shRNA)	GSE99009	30401455
Knockout of YTHDC1	GSE146467	33658714
Reporter for Eya1 (GFP)	GSE120950	30692203
Knockout of TRF2	GSE156534	33239785
Knockout of LSD1	GSE93952	33414108
Knockout of ZFP57	GSE123942	31399135
Mutant for Tip60/Ep400	GSE85505	28445719
Knockout (triple) of H1	GSE153620	34875212
Knockout of DNMT1 (sgRNA)	GSE121459	31209294
Knockout of SUV39h	GSE57092	24981170
Reporter for Trp63 (GFP)	GSE120950	30692203
Reporter for Irf1 (GFP)	GSE120950	30692203
Treatment with aphidicolin	PRJNA415135	32163370
Knockout of SETDB1	PRJNA544540	31914391
Knockdown of SENP6 (shRNA)	GSE70863	26365490
Overexpression of NEFA	GSE113671	31932739
Knockout of miR-34a	GSE69484	28082412

Knockdown of CAF1 (p60) (shRNA)	E-MTAB-2684	26237512
Reporter for Nefa (GFP)	GSE113671	31932739
Knockout of KAP1	GSE41903	23233547
Reporter for Gata3 (GFP)	GSE120950	30692203
Knockdown of RIF1a	GSE98255	29040764
Knockout of MYC (sgRNA)	GSE121459	31209294
Reporter for Zscan4 (Emerald)	GSE75751	27681430
Knockdown of CAF1 p150 (shRNA)	E-MTAB-2684	26237512
Reporter for Tox3 (GFP)	GSE120950	30692203
Knockdown of RYBP (PRC1)	PRJNA604675	32203418
Reporter for Zscan4 and MERVL-LTR	GSE75751	27681430
Overexpression of DUX	GSE85627	28459457
Knockdown of LSM4 (siRNA)	GSE168728	33991488
Reporter for Zscan4 (Emerald)	GSE85627	28459457
Knockdown of CAF1 p60 (shRNA) vs. GFP negative	E-MTAB-2684	26237512
Knockdown of ISY1 (siRNA)	GSE168728	33991488
Knockdown of CHAF1b p60 (shRNA)	GSE70863	26365490
Reporter for 2C-like cells	E-MTAB-2684	26237512
Reporter for 2C-like cells	GSE133234	32523982
Knockdown of EFTUD2 (siRNA)	GSE168728	33991488
Reporter for 2C-like cells	GSE121459	31209294
Reporter for Zscan4 and MERVL-LTR	GSE119819	32032525
Knockdown of SNRPB (siRNA)	GSE168728	33991488

		Knockdown of SNRPD2 (siRNA)	GSE168728	33991488
		Knockdown of CAF1 p150 (shRNA) vs. GFP negative	E-MTAB-2684	26237512
		Knockdown of CHAF1a p150 (shRNA)	GSE70863	26365490
fig. S5F	RNA-seq	Pooled 2C-like cells isolated by Zscan4 and MERVL-LTR reporters	GSE75751	27681430
fig. S5, K and L	ChIP-seq	Epigenetic markers in 2C-like cells: H3K4me3 and K27me3	GSE164486	33636112
	ATAC-seq	Chromatin accessibility in ES and 2C-like	GSE75751	27681430
Fig. 8, A and D	RNA-seq	Pooled microinjected MERVL-LTR (MT2_Mm) CRISPR KO embryos at 4C stages	GSE242123	37781606
Fig. 8, B, C, E and F	RNA-seq	Pooled microinjected MERVL KD embryos at 4C stages	GSE196520	36864102
Fig. 8, G and H	ATAC-seq	Pooled microinjected MERVL KD embryos from at 2C and 4C stages	GSE196520	36864102
	ChIP-seq peaks	Pluripotent core factors: SOX2, OCT4 in mESCs	GSE44286	23582322

Table S2. Gene signatures (excel file)

List of genes used for enrichment analysis (provided as both ENSEMBL id and gene name)

Table S3. Reagents used

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-CARM1 (3H2)	NovusBio	NBP2-37645
Anti-CDX2	BioGenex	MU392A-UC
Anti-cleaved Caspase-3 (CC3) (Asp175)	Cell Signaling	9661
Anti-E-CADHERIN (36)	BD Bio	610182
Anti-goat Alexa Fluor 555	Invitrogen	A21432
Anti-GFP	Genetex	GTX113617
Anti-H3 pan Ac	Abcam	ab47915
Anti-H3 Total	Abcam	ab1791
Anti-H3K4me3	Abcam	ab8580
Anti-H3K9me3	Abcam	ab8898
Anti-MERVL-gag	NovusBio	NBP2-66963
Anti-mouse Alexa Fluor 488	Invitrogen	A11001
Anti-mouse Alexa Fluor 555	Invitrogen	A21422
Anti-mouse Alexa Fluor 647	Invitrogen	A21235
Anti-NANOG (D2A3)	Cell Signaling	8822
Anti-NANOG (SER211)	In-home made (CNIO)	-
Anti-OCT4	Abcam	ab19857
Anti-PARD6B	SantaCruz	sc-166405
Anti-PODXL	R&DSystems	MAB1556
Anti-rabbit Alexa Fluor 488	Invitrogen	A11008
Anti-rabbit Alexa Fluor 555	Invitrogen	A21429
Anti-rabbit Alexa Fluor 647	Invitrogen	A21244
Anti-rat Alexa Fluor 555	Invitrogen	A21434
Anti-SMARCC1 (BAF155)	Sigma	HPA024352
Anti-SOX2	Abcam	ab97959
Anti-SOX2	EBio	14-9811-80
Anti-URI mAb or RbAb	(*26*)	-
Anti-VINCULIN (hVIN-1)	Sigma	V9131
Anti-ZSCAN4	Sigma	AB4340
Anti-γ-TUBULIN (GTU-88)	Sigma	T6557
Chemicals, Peptides, and Recombinant Proteins		
4-hydroxitamoxifen (4-OHT)	PeproTech	6833585

B27 supplement	Gibco	17504044
CHIR99021 (GSK3i)	Axon Medchem	1386
Doxycycline	Sigma	D9891
Gelatin	Sigma	G1890
Knock-out serum replacement (KSR)	Gibco	10828010
Matrigel	BD Bioscience	356230
MG132 Proteasome inhibitor	Sigma	M8699
Triptolide	Sigma	T3652
N2 supplement	Gibco	17502048
PD0325901 (MEKi)	Axon Medchem	1408
Reduced serum media Opti-MEM	ThermoFisher	51985034
hCG	Sigma	CG5
Hyaluronidase type IV-S	Sigma	H3884
KSOM media	Sigma	MR-106-D
M16 media	Sigma	M7292
M2 medium	Sigma	M7167
Paraffin mineral oil	Nidacon	NO-100
PMSG	ProspecBio	hor-272
DAPI	Sigma	D9542
Hoechst	Sigma	B2261
Mowiol 4-88	Sigma	81381
BL21 CodonPlus (DE3)-RIPL	Agilent	230265
Bradford	Biorad	5000006
Dynabeads slurry protein A	GE Healthcare	17-1279-01
Dynabeads slurry protein G	GE Healthcare	17-0618-01
Glutathione Sepharose beads	Cytiva	GE17-5132-02
TRIzol	Sigma	15596026
BL21 CodonPlus (DE3)-RIPL	Agilent	230265
Control antisense oligonucleotide (Antisense LNA GapmeR control)	Qiagen	339515 LG00000002-DDA
Uri antisense oligonucleotide (URI1_1 Antisense LNA GapmeR)	Qiagen	339511 LG00800668-DDA
Uri antisense oligonucleotide (URI1_9 Antisense LNA GapmeR)	Qiagen	339511 LG00800676-DDA
Human adenovirus type 5 carrying Cre and EGFP (AdV5-CMV-Cre-EGFP)	VectorBioLabs	1700

Human adenovirus type 5 enconding EGFP (AdV5-CMV-EGFP)	VectorBioLabs	1060
Plasmids		
pcDNA3.3-HA-OCT4	Addgene	26816
pcDNA3.3-HA-SOX2	Addgene	26817
pcDNA3.1-EF1α-mURI	This study	-
pcDNA3.1-EF1α-MERVL-gag	This study	-
pGEX4T-1-OCT4	Addgene	40633
pGEX4T-1-SOX2	This study	-
pGEX4T-1-MERVL-gag (1-581 Aa)	This study	-
pGEX4T-1-MERVL-gag (150-581 Aa)	This study	-
pGEX4T-1-MERVL-gag (150-350 Aa)	This study	-
pGEX4T-1-MERVL-gag (0-250 Aa)	This study	-
pGEX4T-1-MERVL-gag (150-450 Aa)	This study	-
pGEX4T-1-MERVL-gag (250-450 Aa)	This study	-
pGEX4T-1-URI (1-158 Aa)	(*30*)	-
pGEX4T-1-URI (1-267 Aa)	(*30*)	-
pGEX4T-1-URI (1-518 Aa)	(*30*)	-
pGEX4T-1-URI (1-535 Aa)	(*30*)	-
pGEX4T-1-URI (156-283 Aa)	(*30*)	-
pGEX4T-1-URI (156-535 Aa)	(*30*)	-
pGEX4T-1-URI (282-535 Aa)	(*30*)	-
pGEX4T-1-URI (390-535 Aa)	(*30*)	-
pGEX4T-1-URI (466-535 Aa)	(*30*)	-
pMERVL-LTR-tdTomato	Addgene	40281
pMERVL-LTR-tdTomato-T2A-Cre	This study	This study
pZscan4-mEmerald-T2A-Cre	This study	This study
pSBbi-GN	Addgene	60517
pZscan4-mEmerald	(*56*)	
tetO-FLAG-mDUX-hPGK-rtTA-T2A-Neo	Addgene	138320
tetO-mDUX-hPGK-Puro-T2A-rtTA	Addgene	99284
Critical Commercial Assays		
M-MLV Reverse Transcriptase	Thermofisher	28025013
GoTaq qPCR Master Mix	Promega	A6002
Mycoplasma PCR test kit	Biontex	6833585

Vectastain ABC HRP Kit (Peroxidase, Mouse IgG)	Vector Laboratories	PK-4002
Vectastain ABC HRP Kit (Peroxidase, Rabbit IgG)	Vector Laboratories	PK-4001
Liquid DAB + Substrate Chromogen System	Dako	K3468
Reticulocyte lysate system	Promega	L5010
GoTaq SYBR green master mix	Promega	A6002
Experimental Models: Cell Lines		
293A cell line	ThermoFisher	R70507
Mouse embryonic fibroblast (MEFs)	This study	-
Mouse embryonic stem cells (mESCs)	This study	-
Experimental Models: Organisms/Strains		
Mouse: C57BL/6	CNIO Animal Facility	N/A
Mouse: Tg. CAG-Cre	(*37*)	MGI: 3586452
Mouse: Sox2-Cre, Edil3Tg(Sox2-Cre)1Amc	(*38*)	MGI: 2656539
Mouse: UBQ-CreERT2, NdorTg(UBC-Cre/ERT2)1Ejb	(*50*)	MGI: 123200
Mouse: Zp3-Cre, Tg(Zp3-Cre)93Knw	(*40*)	MGI: 2176187
Mouse: Uri_lox	(*30*)	NR
Oligonucleotides		
Primers For Genotyping	This Paper, Table S4	N/A
Primers For qRT-PCR	This Paper, Table S4	N/A
Software and Algorithms		
ATACseqQC 1.26	PMID: 29490630	https://bioconductor.org/packages/release/bioc/html/ATACseqQC.html
Bowtie2 2.4.4	PMID: 19261174	https://github.com/BenLangmead/bowtie2
ClusterProfiler 4.4.2	PMID: 34557778	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
DeepTools 3.0.2	PMID: 27079975	https://deeptools.readthedocs.io/en/develop/

edgeR 4.0.2	PMID: 19910308	https://bioconductor.org/packages/release/bioc/html/edgeR
csaw 1.36	PMID: 26578583	https://bioconductor.org/packages/release/bioc/html/csaw
Cutadapt 3.4.1	PMID: 28715235	https://github.com/marcelm/cutadapt
DESeq2 1.32	PMID: 25516281	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
FastICA 1.2.3	PMID: 10946390	https://CRAN.R-project.org/package=fastICA
FastQC 0.11.8	Babraham bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
FIJI/ImageJ 1.53i	PMID: 22743772	https://imagej.net/software/fiji
Galaxy project	PMID: 35446428	https://galaxyproject.org/
Ggplot2 3.3.6	H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016	https://CRAN.R-project.org/package=ggplot2
GraphPad Prism 9.4.0	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
HISAT2 2.1	PMID: 25751142	http://daehwankimlab.github.io/hisat2/
HOMER 4.11	PMID: 20513432	http://homer.ucsd.edu/homer/index.html
Illustrator 26.3.1	Adobe	https://.adobe.com/products/illustrator.html
Imaris 8.4.1	Bitplane	http://www.bitplane.com/imaris/imaris
MACS2 2.1.1	PMID: 18798982	https://github.com/macs3-project/MACS

msa 1.34		https://bioconductor.org/packages/release/bioc/html/msa.html
PcaMethods 1.88	PMID: 17344241	https://www.bioconductor.org/packages/release/bioc/html/pcaMethods.html
Photoshop 23.4.0	Adobe	https://adobe.com/products/photoshop.html
Picard 2.18.2	Broad Institute	https://broadinstitute.github.io/picard/
Psych 2.2.5	Revelle, W.R. Procedures for Psychological, Psychometric, and Personality Research. *Northwestern University, Evanston, Illinois*, 2022	https://CRAN.R-project.org/package=psych
pyGenomeTracks 3.7	PMID: 32745185	https://github.com/deeptools/pyGenomeTracks
R 4.1.0	The R Foundation for Statistical Computing	https://www.r-project.org/
RStudio 1.4.1717	Integrated Development for R. RStudio	https://www.rstudio.com/
SAMtools 1.9	PMID: 19505943	https://github.com/samtools/samtools
Seurat v5	PMID: 37231261	https://cran.r-project.org/web/packages/Seurat/
Sra-tools 2.11	PMID: 21062823	https://github.com/ncbi/sra-tools
ssGSEA 2.0	PMID: 16199517	https://github.com/broadinstitute/ssGSEA2.0
Statmod 1.4.36	PMID: 21044043	https://CRAN.R-project.org/package=statmod
StringTie2 2.2.2	PMID: 31842956	https://ccb.jhu.edu/software/stringtie/

Sva 3.44	PMID: 16632515	https://bioconductor.org/packages/release/bioc/html/sva.html
Trim Galore 0.6.3	Babraham bioinformatics	https://github.com/FelixKrueger/TrimGalore
Trimmomatic 0.38	PMID: 24695404	http://usadellab.org/cms/?page=trimmomatic
Other		
Dialysis cassette (10000 MWCO)	ThermoFisher	66380
Glass bottom 35mm dish	Ibidi	81158
In vitro fertilization (IVF) multidishes	Sigma	Z688754
Digital sonifier	Branson	S450D
Electroporation Cuvettes (2mm gap)	NepaGene	EC-002S
Fluorescence confocal microscopy	Leica	TCS SP5 WLL
Fluorescence confocal microscopy	Leica	TCS SP8 MP
Fluorescence microscopy	Olympus	BX61
Inverted phase-contrast microscope	Leica	DMIRE2
Irradiator	Theratronics	Gammacell 1000
Manual pneumatic microinjector	Eppendorf	CellTram 4r Oil
Microinjector for embryos	Eppendorf	Femtojet 4i
Micromanipulator for embryos	Eppendorf	TransferMan 4r
NEPA21 electroporator	NepaGene	CU500
Ultracentrifuge (SW28 rotor)	Beckman	Avanti J-25
Low protein-biding tubes	ThermoFisher	90410

Table S4. Primer sequences for genotyping and qRT-PCR

Gene	Technique	Sequence (5'-3')
Cre(Cpxm1)-F	Genotyping	CCATCTGCCACCAGCCAG
Cre(Cpxm1)-R	Genotyping	TCGCCATCTTCCAGCAGG
Cre-F	Genotyping	ACTGGGATCTTCGAACTCTTTGGAC
Cre-R	Genotyping	GATGTTGGGGCACTGCTCATTCACC
GFP-F	Genotyping	TGACCCTGAAGTTCATCTGCA
GFP-R	Genotyping	TCACGAACTCCAGCAGGACCA
Rosa26-rtTA-F	Genotyping	AAAGTCGCTCTGAGTTGTTAT
Rosa26-rtTA-R1	Genotyping	GGAGCGGGAGAAATGGATATG
Rosa26-rtTA-R2	Genotyping	GCGAAGAGTTTGTCCTCAACC
Sox2-Cre-F1	Genotyping	CTTGTGTAGAGTGATGGCTTGA
Sox2-Cre-F2	Genotyping	TAGTGCCCCATTTTTGAAGG
Sox2-Cre-R	Genotyping	CCAGTGCAGTGAAGCAAATC
URI-delta-lox-F1	Genotyping	CGTGAAGAGAGGTGAAGAAC
URI-delta-lox-F2	Genotyping	CCCTCTTGCCTTCATGCC
URI-delta-lox-R	Genotyping	AAACACAAGTGTAAAATGTCCC
mActin-F	qRT-PCR	CACAGCTGAGAGGGAAATCG
mActin-R	qRT-PCR	AGTTTCATGGATGCCACAGG
mGapdh-F	qRT-PCR	CGTCCCGTAGACAAAATGGT
mGapdh-R	qRT-PCR	TCAATGAAGGGGTCGTTGAT
Gm4340-F	qRT-PCR	CGAGGCACTGGGTCTAAGAG
Gm4340-R	qRT-PCR	CCAATGAACAGGTCATGCTG
mDub1-F	qRT-PCR	GGAGACATGGTGGTTGCTCT
mDub1-R	qRT-PCR	CTCTCCCAACTCAGACTGTGC
mDux-F	qRT-PCR	ACTTCTAGCCCCAGCGACTC
mDux-R	qRT-PCR	CCATGCTGCCAGGATTTCTA

MERVL-LTR-F	qRT-PCR	CTTCCATTCACAGCTGCGACTG
MERVL-LTR-R	qRT-PCR	CTAGAACCACTCCTGGTACCAAC
MERVL-pol-F	qRT-PCR	CCCATCATGAGCTGGGTACT
MERVL-pol-R	qRT-PCR	CGTGCAGAGCCATCAGTAAA
mIAPEz-F	qRT-PCR	CAGACTGGGAGGAAGAAGCA
mIAPEz-R	qRT-PCR	ATTGTTCCCTCACTGGCAAA
mLINE1-F	qRT-PCR	TTTGGGACACAATGAAAGCA
mLINE1-R	qRT-PCR	CTGCCGTCTACTCCTCTTGG
MMERVK10C-F	qRT-PCR	CAAATAGCCCTACCATATGTCAG
MMERVK10C-R	qRT-PCR	GTATACTTTCTTCTTCAGGTCCAC
mNanog-F	qRT-PCR	AGGGTCTGCTACTGAGATGCTCTG
mNanog-R	qRT-PCR	CAACCACTGGTTTTTCTGCCACCG
mPou5f1-F	qRT-PCR	CTGTAGGGAGGGCTTCGGGCACTT
mPou5f1-R	qRT-PCR	CTGAGGGCCAGGCAGGAGCACGAG
mSox2- F	qRT-PCR	GGCAGCTACAGCATGATGCAGGAGC
mSox2- R	qRT-PCR	CTGGTCATGGAGTTGTACTGCAGG
mSp110-F	qRT-PCR	AAGGATCCAGGAACCCCTTA
mSp110-R	qRT-PCR	GCATAGGCGATGTTCACCTT
mTcstv1-F	qRT-PCR	TGAACCCTGATGCCTGCTAAGACT
mTcstv1-R	qRT-PCR	AGATGGCTGCAAAGACACAACTGC
mTcstv3-F	qRT-PCR	AGAAAGGGCTGGAACTTGTGACCT
mTcstv3-R	qRT-PCR	AAAGCTCTTTGAAGCCATGCCCAG
mZfp352-F	qRT-PCR	AAAGCCTTGATCCTCAGGTG
mZfp352-R	qRT-PCR	GCCGAAGAGTTTTTCTGAGG
mZscan4-F	qRT-PCR	GAGATTCATGGAGAGTCTGACTGATGAGTG
mZscan4-R	qRT-PCR	GCTGTTGTTTCAAAAGCTTGATGACTTC