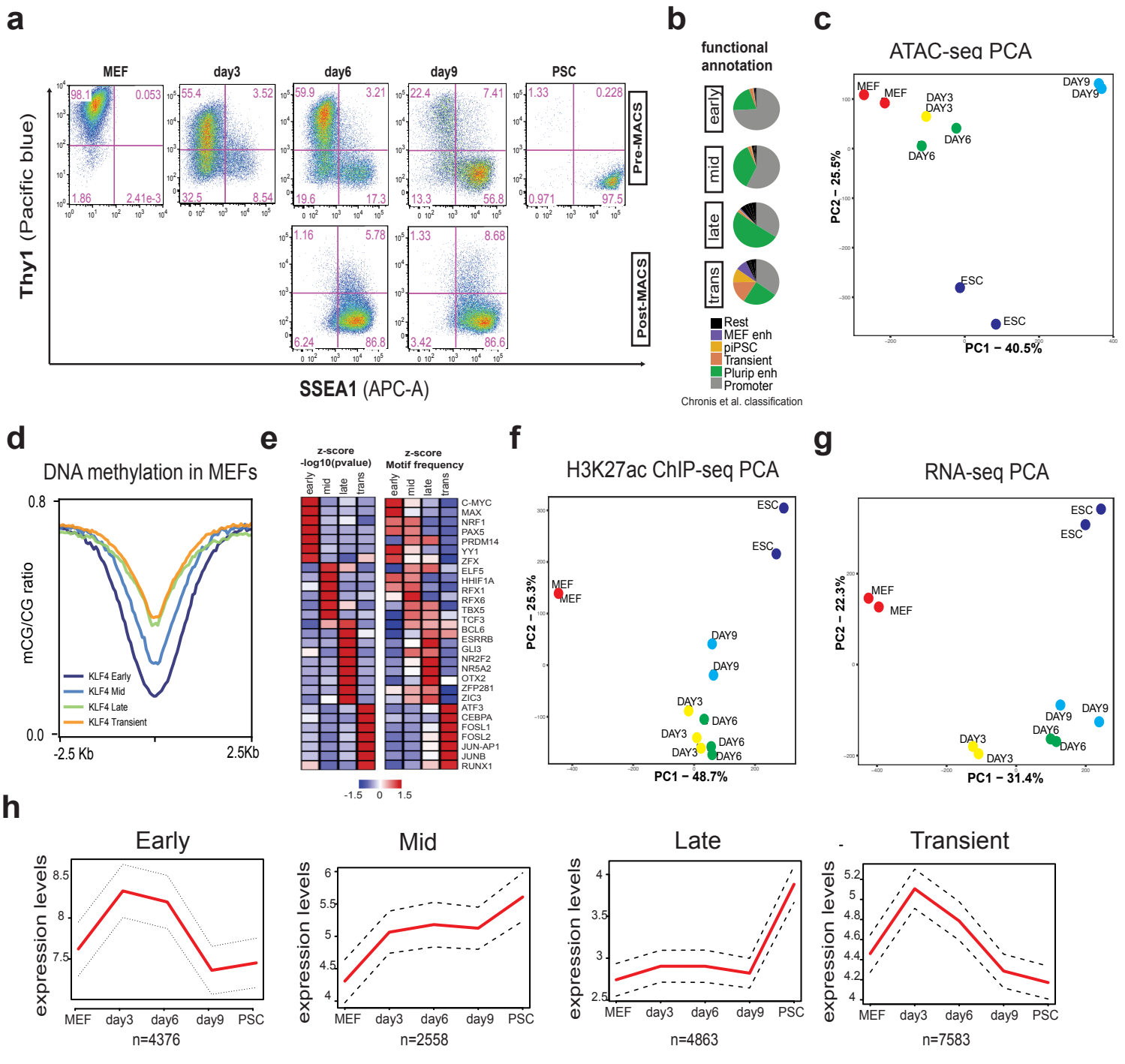
INVENTORY OF SUPPLEMENTAL INFORMATION

Supplementary Figures:

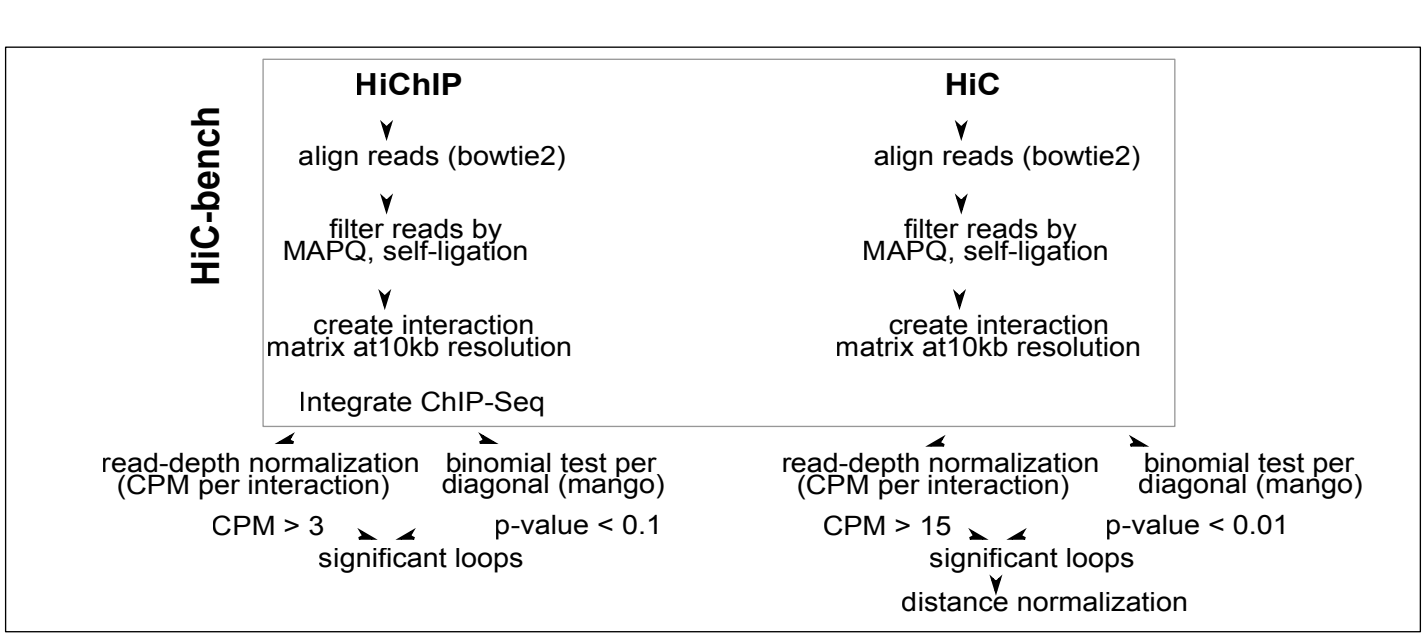
Supplementary Tables:

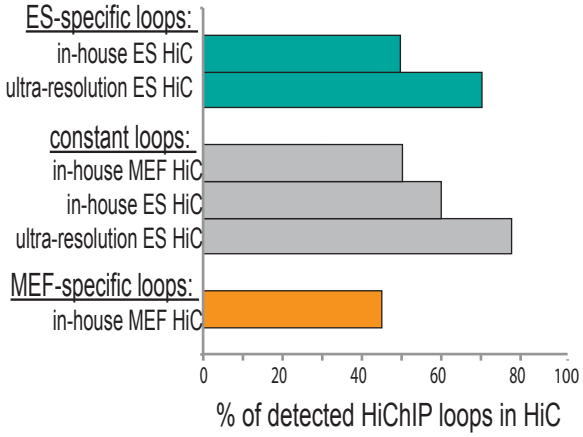
Figure S1

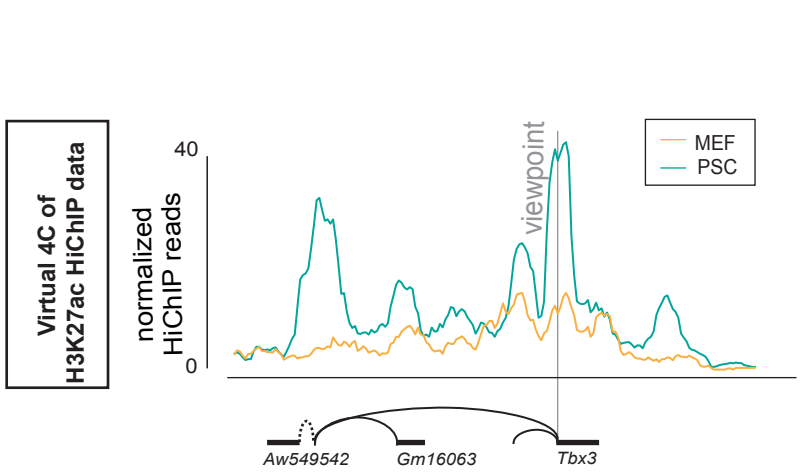

a
MEF
day3
day6
day9
PSC
Pre-MACS
Post-MACS
Thy1 (Pacific blue)
SSEA1 (APC-A)
b
functional annotation
early
mid
late
trans
Rest
MEF enh
piPSC
Transient
Plurip enh
Promoter
Chronis et al. classification
c
ATAC-seq PCA
PC2 – 25.5%
PC1 – 40.5%
d
DNA methylation in MEFs
mCG/CG ratio
KLF4 Early
KLF4 Mid
KLF4 Late
KLF4 Transient
-2.5 Kb
2.5Kb
e
z-score -log10(pvalue)
z-score Motif frequency
C-MYC
MAX
NRF1
PAX5
PRDM14
YY1
ZFX
ELF5
HHIF1A
RFX1
RFX6
TBX5
TCF3
BCL6
ESRRB
GLI3
NR2F2
NR5A2
OTX2
ZFP281
ZIC3
ATF3
CEBPA
FOSL1
FOSL2
JUN-AP1
JUNB
RUNX1
f
H3K27ac ChIP-seq PCA
PC2 – 25.3%
PC1 – 48.7%
g
RNA-seq PCA
PC2 – 22.3%
PC1 – 31.4%
h
Early
Mid
Late
Transient
expression levels
MEF day3 day6 day9 PSC
n=4376
n=2558
n=4863
n=7583

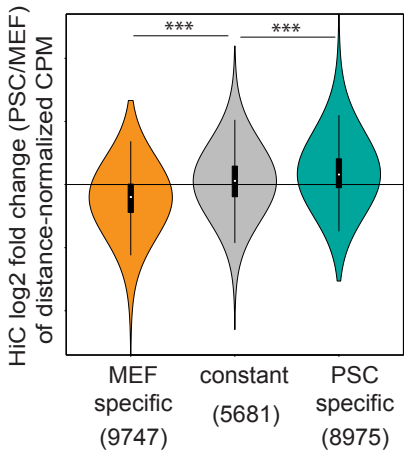
Supplementary Figure 1. a, FACS analysis plots showing expression of SSEA1 (early pluripotency marker) and Thy1 (somatic marker) at different stages of reprogramming, before and after SSEA1 enrichment by MACS isolation. **b**, Pie charts of functional classification of KLF4 Early, Mid, Late and Transient peaks (based on Chronis et al. 2017) (piPSC= partial iPSCs). **c**, PCA analysis of ATAC-seq peaks in MEF, PSC and different stages of reprogramming. **d**, Average line plot showing the methylated CG to non-methylated CG ratio from MEF data[12] centered (+/-2.5Kb) around different clusters of KLF4 binding sites (Early, Mid, Late or Transient KLF4 targets, Fig.2b)**. e**, Motif enrichment for Early, Mid, Late and Transient KLF4 binding sites. Selected factors are shown and their significance is expressed as Z-score of –log10(pvalue) (left) or z-score of motif frequency (right). **f**, PCA analysis of H3K27ac ChIP-seq peaks called in MEF, PSC and different stages of reprogramming **g**, PCA of RNA-seq in MEF, PSC and different stages of reprogramming. **h**, Line plots of the median expression (red line) of genes closest to Early, Mid, Late and Transient peaks, expressed as TPM (transcripts per million).

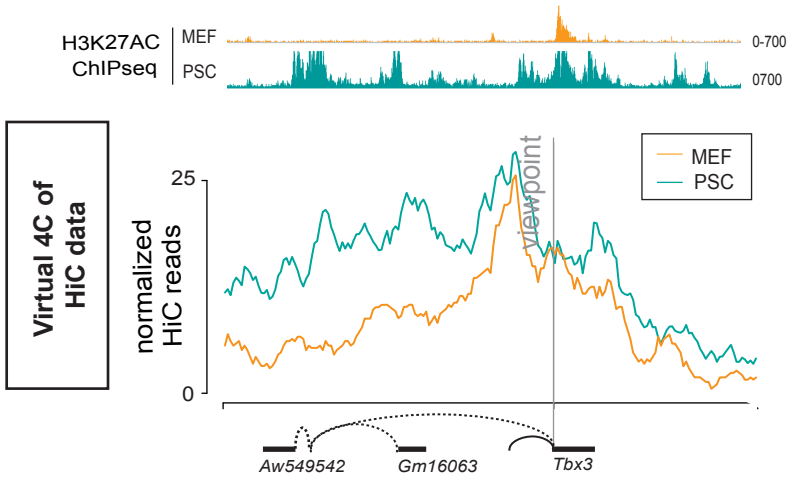
Figure S2


a
HiC-bench
HiChIP
align reads (bowtie2)
filter reads by MAPQ, self-ligation
create interaction matrix at10kb resolution
Integrate ChIP-Seq
HiC
align reads (bowtie2)
filter reads by MAPQ, self-ligation
create interaction matrix at10kb resolution
read-depth normalization (CPM per interaction)
binomial test per diagonal (mango)
CPM > 3
p-value < 0.1
significant loops
read-depth normalization (CPM per interaction)
binomial test per diagonal (mango)
CPM > 15
p-value < 0.01
significant loops
distance normalization


b
ES-specific loops:
in-house ES HiC
ultra-resolution ES HiC
constant loops:
in-house MEF HiC
in-house ES HiC
ultra-resolution ES HiC
MEF-specific loops:
in-house MEF HiC
0
20
40
60
80
100
% of detected HiChIP loops in HiC


c
Virtual 4C of H3K27ac HiChIP data
normalized HiChIP reads
40
0
viewpoint
MEF
PSC
Aw549542
Gm16063
Tbx3

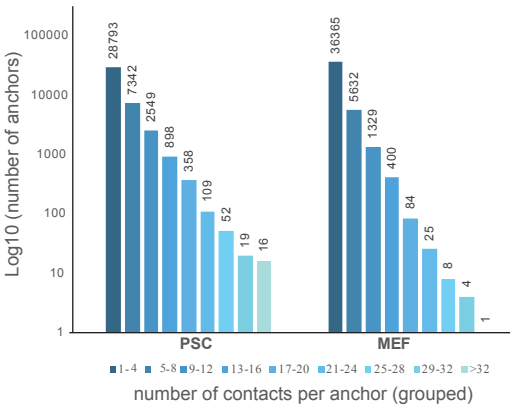

d
HiC log2 fold change (PSC/MEF) of distance-normalized CPM
***
***
MEF specific (9747)
constant (5681)
PSC specific (8975)


H3K27AC ChIPseq
MEF
PSC
0-700
0700
Virtual 4C of HiC data
normalized HiC reads
25
0
viewpoint
MEF
PSC
Aw549542
Gm16063
Tbx3

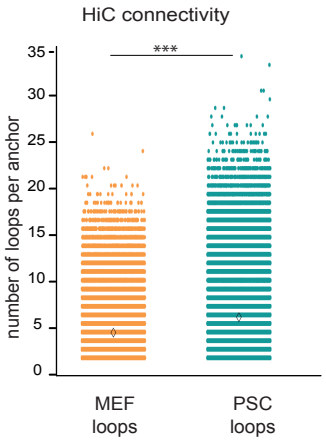
Supplementary Figure 2. a, Schematic work-flow for HiChIP and HiC analysis. **b**, Percentages of PSC-specific, constant or MEF-specific H3K27ac HiChIP loops that were detected in HiC experiments (either generated in-house or published ultra-resolution HiC in PSC[38]). **c**, Normalized HiChiP (top) and HiC (bottom) signals in MEF and PSC are illustrated in a virtual 4C format around the indicated viewpoint (*Tbx3* promoter). H3K27ac ChIP-seq tracks are shown in MEF and PSC. **d**, Violin plot representing log2 fold change of distance-normalized HiC signal in PSCs versus MEFs of MEF-specific, constant and PSC-specific loops as called by H3K27ac HiChIP. Only contacts that were detected as significant in HiC data are considered. Numbers of considered loops per category are shown in parenthesis. Unpaired two-sided t-test was used to determine the p-value.

Figure S3

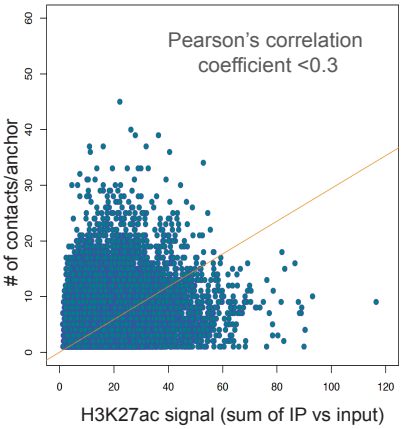
a



b

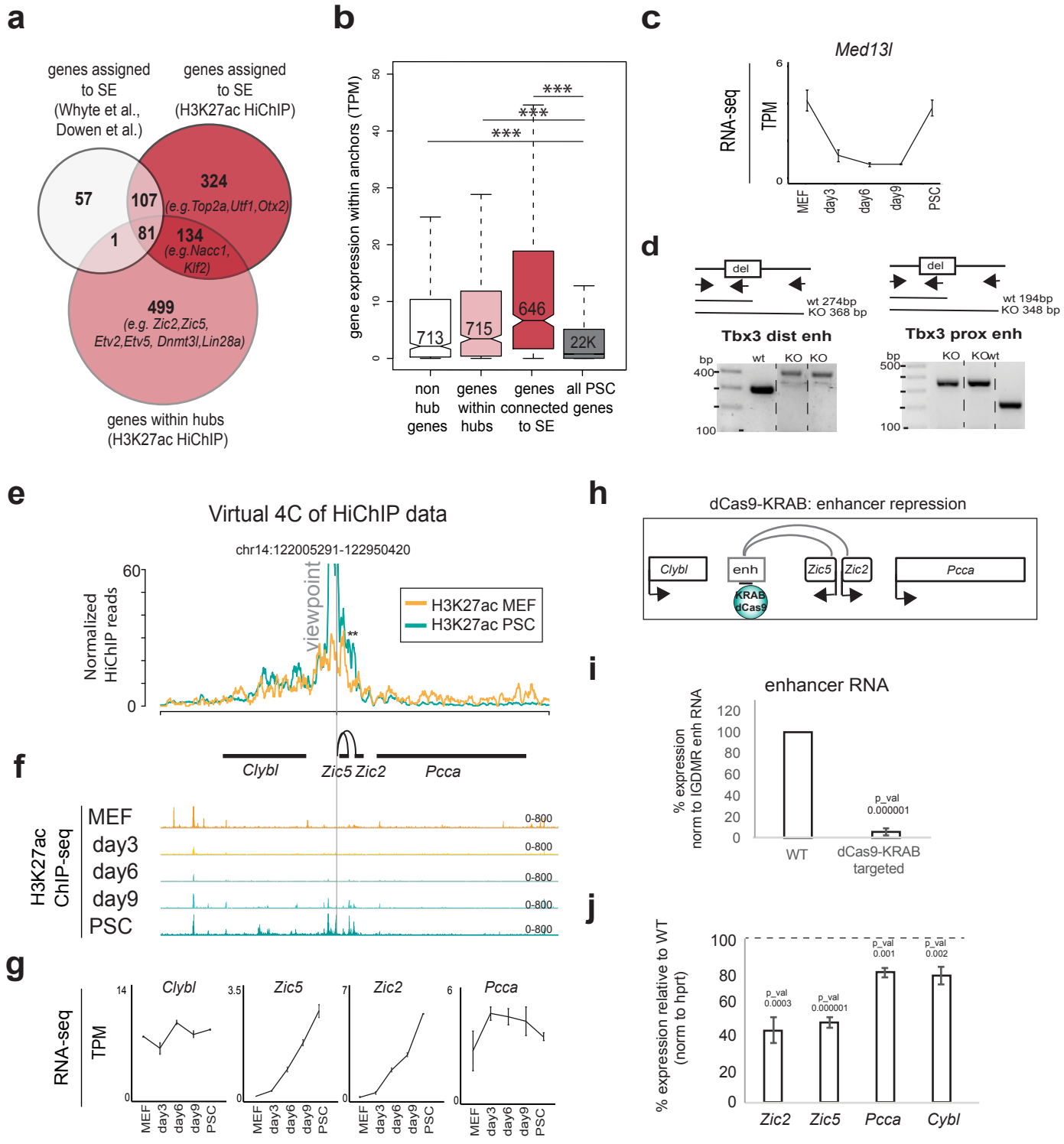


c



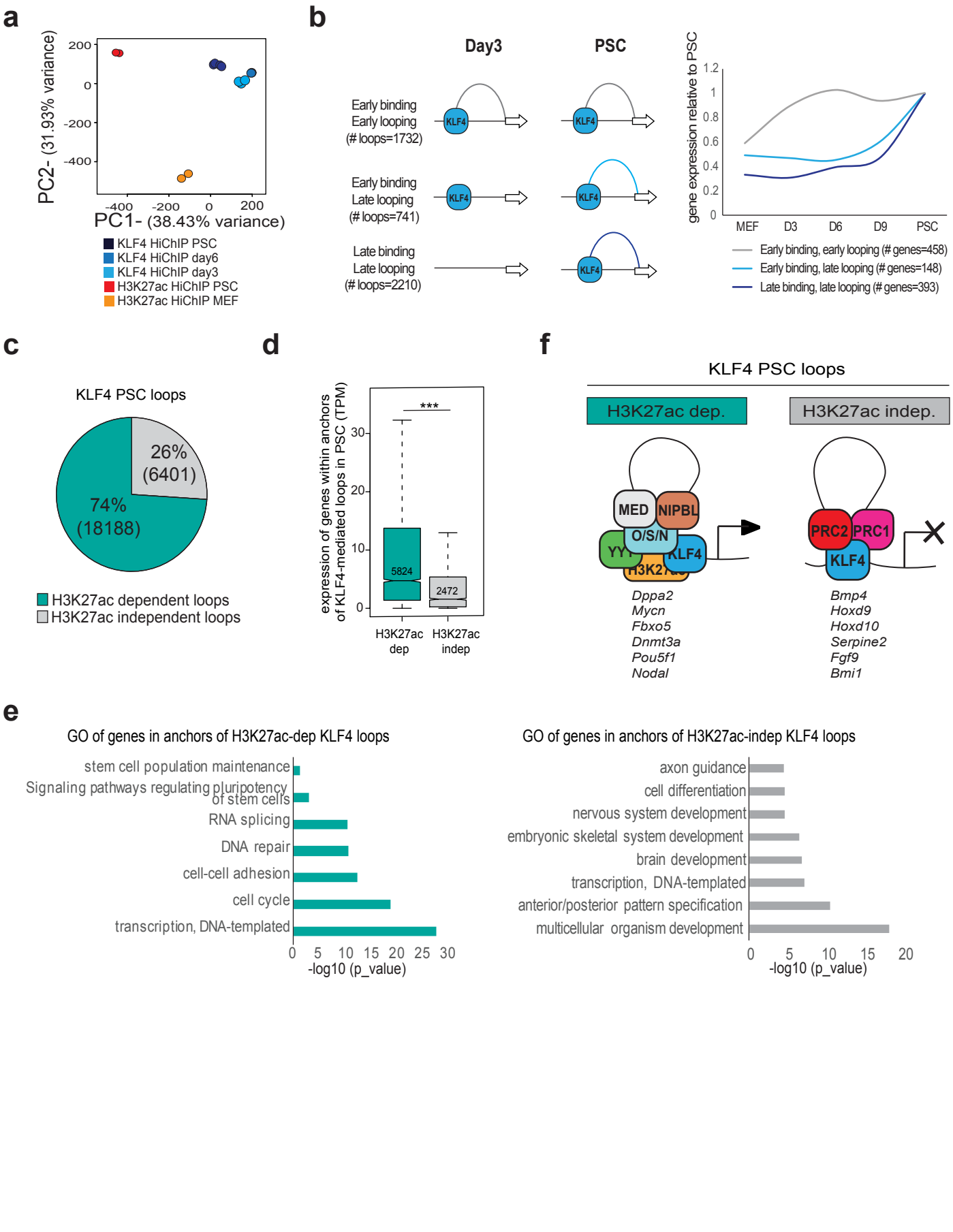
Supplementary Figure 3. a, Histogram of anchor connectivity based on H3K27ac MEF and PSC HiChIP called loops. The numbers of contacts per anchor are grouped as shown in the bottom and the actual number of anchors is depicted on top of each bar. **b,** Connectivity of MEF or PSC anchors based on HiC-called loops represented as number of high-confidence contacts around each 10kb anchor. Wilcoxon rank sum test was used to compare connectivity and asterisks indicate significant difference with $p<0.001$. **c**, Scatter plot showing the correlation of H3K27ac ChIP-seq strength (sum of H3K27ac ChIP/input of all peaks within the anchor) with the number of H3K27ac HiChIP contacts per anchor in PSCs.

Figure S4


a
genes assigned to SE (Whyte et al., Dowen et al.)
genes assigned to SE (H3K27ac HiChIP)
57
107
324
(e.g.Top2a,Utf1,Otx2)
1
81
134
(e.g.Nacc1, Klf2)
499
(e.g. Zic2,Zic5, Etv2,Etv5, Dnmt3l,Lin28a)
genes within hubs (H3K27ac HiChIP)
b
gene expression within anchors (TPM)
***
***
***
713
715
646
22K
non hub genes
genes within hubs
genes connected to SE
all PSC genes
c
Med13l
RNA-seq
TPM
MEF
day3
day6
day9
PSC
d
del
wt 274bp
KO 368 bp
del
wt 194bp
KO 348 bp
Tbx3 dist enh
Tbx3 prox enh
bp
wt
KO
KO
400
100
500
100
e
Virtual 4C of HiChIP data
chr14:122005291-122950420
Normalized HiChIP reads
viewpoint
H3K27ac MEF
H3K27ac PSC
Clybl
Zic5 Zic2
Pcca
f
H3K27ac ChIP-seq
MEF
day3
day6
day9
PSC
0-800
g
RNA-seq
TPM
Clybl
Zic5
Zic2
Pcca
h
dCas9-KRAB: enhancer repression
Clybl
enh
Zic5
Zic2
Pcca
i
enhancer RNA
% expression norm to IGDMR enh RNA
p_val 0.000001
WT
dCas9-KRAB targeted
j
% expression relative to WT (norm to hprt)
p_val 0.0003
p_val 0.000001
p_val 0.001
p_val 0.002
Zic2
Zic5
Pcca
Cybl

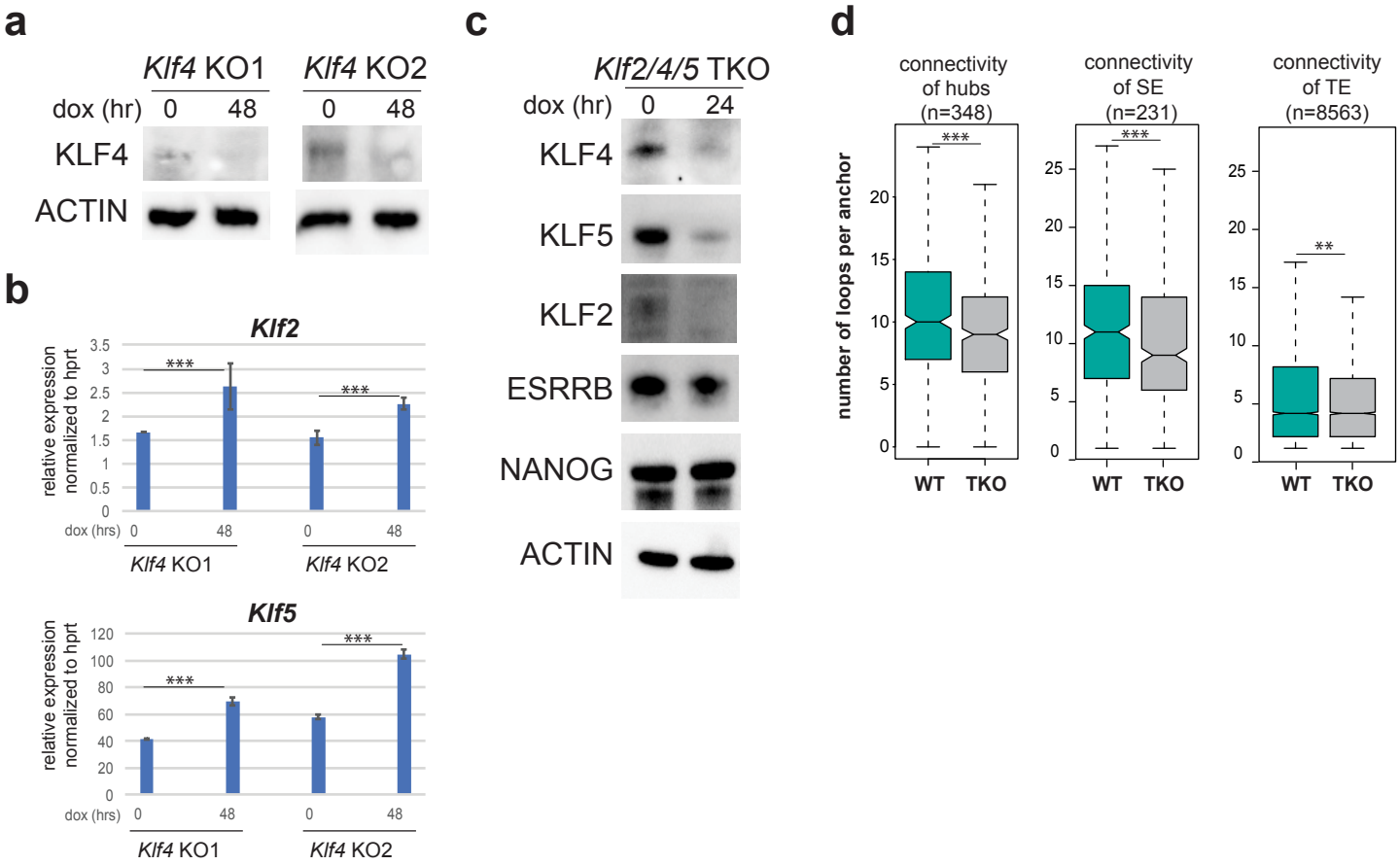
Supplementary Figure 4. a, Venn diagram showing overlap between previously assigned target genes for super-enhancers (SE), newly identified SE target genes based on H3K27ac HiChIP contacts in PSCs, and genes connected to PSC-specific enhancer hubs, which represent enhancers contacting more than one gene according H3K27ac HiChIP (see also Fig.4a). **b**, RNA levels of hub genes, non-hub genes or genes connected to SE in PSC samples as measured by RNA-seq and expressed as transcripts per million (TPM). All genes that are not connected to enhancer hubs, but are still detected within PSC-specific HiChIP loops were considered. Expression of all genes expressed in PSC (>1TPM) is shown as reference. **c**, RNA-seq signal (TPM) of *Med13l* -which is not part of the *Tbx3* enhancer hub (see Fig.4b)- during reprogramming **d**, Genotyping strategy and results confirming the homozygous deletion of the distal (left) or the proximal (right) *Tbx3* enhancers. **e**, Example of a newly identified enhancer hub in PSCs. Normalized HiChIP signal around the viewpoint is illustrated as a virtual 4C plot. **f**, H3K27ac ChIP-seq IGV tracks during reprogramming. **g**, RNA-seq signal of genes within the hub (*Zic2* and *Zic5*), or nearby genes (*Clybl* and *Pcca*), are shown for each reprogramming stage to highlight concordance with H3K27ac HiChIP data and coordinated upregulation of genes within the hub. **h**, Schematic illustration of the CRSIPRi (dCas9-KRAB) targeting strategy for inactivation of the *Zic2/Zic5* enhancer hub. **i**, RT-qPCR showing relative levels of the enhancer RNA (normalized to an unaffected enhancer RNA (IGDMR)) in wild type (WT) or dCas9-KRAB-targeted ESCs. P-values were calculated using unpaired one-tailed t-test. Error bars indicate standard deviation from n=2 biological replicates. **j**, RT-qPCR showing expression changes of genes within the hub (*Zic2* and *Zic5*) and nearby genes (*Clybl* and *Pcca),* calculated as percentage relative to WT after normalization to *Hprt* expression**.** P-values were calculated using unpaired one-tailed t-test. Error bars indicate standard deviation from n=2 biological replicates.

Figure S5


a
PC2- (31.93% variance)
PC1- (38.43% variance)
KLF4 HiChIP PSC
KLF4 HiChIP day6
KLF4 HiChIP day3
H3K27ac HiChIP PSC
H3K27ac HiChIP MEF
b
Day3
PSC
Early binding
Early looping
(# loops=1732)
Early binding
Late looping
(# loops=741)
Late binding
Late looping
(# loops=2210)
KLF4
gene expression relative to PSC
MEF
D3
D6
D9
PSC
Early binding, early looping (# genes=458)
Early binding, late looping (# genes=148)
Late binding, late looping (# genes=393)
c
KLF4 PSC loops
26%
(6401)
74%
(18188)
H3K27ac dependent loops
H3K27ac independent loops
d
expression of genes within anchors
of KLF4-mediated loops in PSC (TPM)
***
5824
2472
H3K27ac
dep
H3K27ac
indep
f
KLF4 PSC loops
H3K27ac dep.
H3K27ac indep.
MED
NIPBL
O/S/N
YY1
KLF4
PRC2
PRC1
Dppa2
Mycn
Fbxo5
Dnmt3a
Pou5f1
Nodal
Bmp4
Hoxd9
Hoxd10
Serpine2
Fgf9
Bmi1
e
GO of genes in anchors of H3K27ac-dep KLF4 loops
stem cell population maintenance
Signaling pathways regulating pluripotency
of stem cells
RNA splicing
DNA repair
cell-cell adhesion
cell cycle
transcription, DNA-templated
-log10 (p_value)
GO of genes in anchors of H3K27ac-indep KLF4 loops
axon guidance
cell differentiation
nervous system development
embryonic skeletal system development
brain development
transcription, DNA-templated
anterior/posterior pattern specification
multicellular organism development
-log10 (p_value)

Supplementary Figure 5. a, PCA analysis of loops called as significant by H3K27ac and KLF4 HiChIP in different samples. **b**, Left: Chromatin loops that were detected by both KLF4 and H3k27ac HiChIP in PSCs were clustered based on the timing of KLF4 binding and looping during reprogramming. Right: Line plot showing expression changes of genes that belong to each of the indicated loop categories during reprogramming (median values are plotted relative to PSC). **c**, Pie chart showing the percentage of KLF4 PSC loops that were also detected by H3K27ac HiChIP in PSCs (H3K27ac-dependent) or not (H3K27ac-independent). **d**, Boxplot showing expression of genes within all anchors of KLF4-mediated loops that are either H3K27ac-dependent or independent. **e**, Gene ontology for genes within anchors of H3K27ac-dependent or -independent KLF4 loops. **f**, Proposed model for different categories of chromatin loops mediated by KLF4 and cofactors. Example genes are reported for each category.

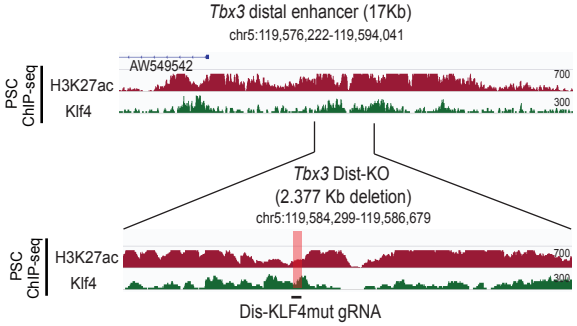
Figure S6


a
Klf4 KO1
Klf4 KO2
dox (hr) 0 48 0 48
KLF4
ACTIN
b
Klf2
relative expression normalized to hprt
***
***
dox (hrs) 0 48 0 48
Klf4 KO1
Klf4 KO2
Klf5
relative expression normalized to hprt
***
***
dox (hrs) 0 48 0 48
Klf4 KO1
Klf4 KO2
c
Klf2/4/5 TKO
dox (hr) 0 24
KLF4
KLF5
KLF2
ESRRB
NANOG
ACTIN
d
connectivity of hubs (n=348)
connectivity of SE (n=231)
connectivity of TE (n=8563)
number of loops per anchor
***
***
**
WT TKO
WT TKO
WT TKO

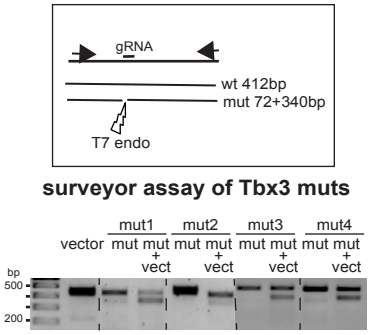
Supplementary Figure 6. a, Western blot analysis showing KLF4 protein levels before (0) and after (48hr) dox-induction in two ESC clones that harbor dox-inducible CRISPR-Cas9 and gRNAs that target the *Klf4* gene (KLF4 KO1 and KLF4 KO2). **b**, RT-qPCR showing elevated levels of *Klf2* and *Klf5* genes in dox-induced KLF4 KO ESCs. **c**, Western blot showing levels of indicated proteins in a clonal population of ESCs containing an inducible CRISPR-Cas9 construct and gRNAs that target the *Klf2*, *Klf5* and *Klf4* genes. Cells were either untreated (0, wild type or WT cells) or treated with dox for 24 hours (triple knock-out or TKO). **d**, Boxplot showing the connectivity of H3K27ac HiChIP anchors that contain hubs, supoerenhancers (SE) or typical enhancers (TE) in WT or TKO ESCs. Asterisks indicate significance as calculated using Wilcoxon rank sum test.

Figure S7

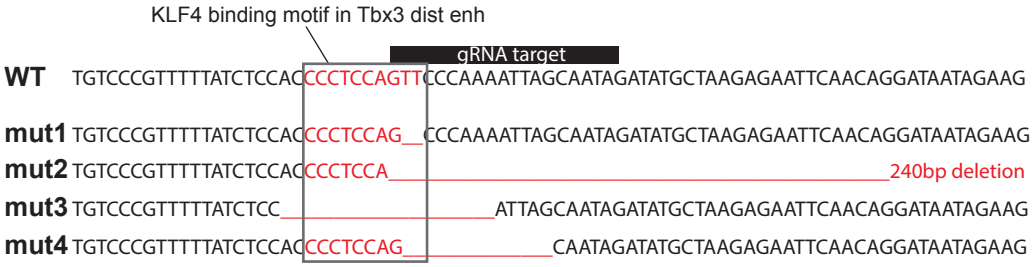
a



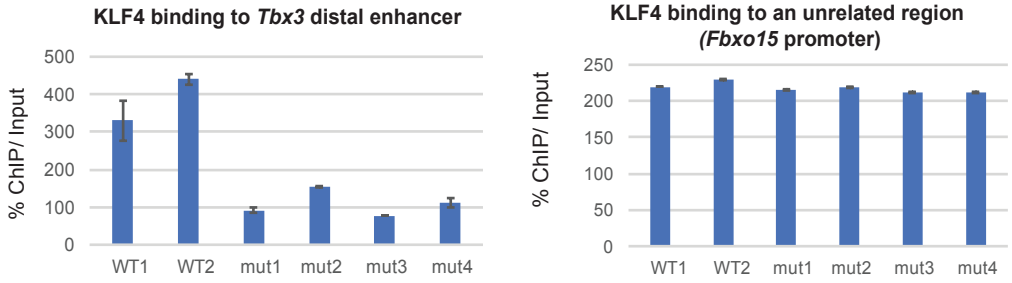
b



c



d



Supplementary Figure 7. a, IGV tracks of H3K27ac and KLF4 ChIP-seq in PSCs showing the whole *Tbx3* distal enhancer (top), the region that was deleted by CRISPR/Cas9 (Dist-KO, bottom, see Fig.4f) and the location of the gRNA used to mutate a specific KLF4 binding motif (Dis-KLF4mut gRNA). **b**, Genotyping strategy of the surveyor assay used to detect mutation/indel at the target KLF4 binding site within the distal *Tbx3* enhancer (Dis-KLF4mut*)*. The results for 4 homozygously mutated clones (mut1-4) are shown. **c**, Sequencing results of the four Mut clones compared to the wild type (WT). **d**, ChIP-qPCR showing the relative levels of KLF4 binding to *Tbx3* distal enhancer in two WT clones and four Mut clones (left panel). Values show percentage of ChIP signal over input. As control, binding of KLF4 to an unaffected region (*Fbxo15* promoter) was tested (right panel).