Reversible inhibition of specific transcription factor-DNA interactions using CRISPR

Ali Shariati[1], Antonia Dominguez[2], Marius Wernig[3], Lei S. Qi[2,4,5] & Jan M. Skotheim[1,*]

[1] Department of Biology, Stanford University, Stanford, CA, 94305, USA

[2] Department of Bioengineering, Stanford University, Stanford, CA 94305, USA

[3] Department of Pathology, Stem Cell Institute, Stanford, CA 94305, USA

[4] Department of Chemical and Systems Biology, Stanford Univesity, Stanford, CA 94305, USA

[5] Stanford ChEM-H, Stanford University, Stanford, CA 94305, USA

*Address correspondence to skotheim@stanford.edu

Abstract

The control of gene expression by transcription factor binding sites frequently determines phenotype. However, it has been difficult to assay the function of single transcription factor binding sites within larger transcription networks. Here, we developed such a method by using deactivated Cas9 to disrupt binding to specific sites on the genome. Since CRISPR guide RNAs are longer than transcription factor binding sites, flanking sequence can be used to target specific sites. Targeting deactivated Cas9 to a specific Oct4 binding site in the *Nanog* promoter blocked Oct4 binding, reduced *Nanog* expression, and slowed division. Multiple guide RNAs allows simultaneous inhibition of multiple binding sites and conditionally-destabilized dCas9 allows rapid reversibility. The method is a novel high-throughput approach to systematically interrogate cis-regulatory function within complex regulatory networks.

Main text:

Binding of a transcription factors (TFs) to target regulatory sequences control when and where target genes are expressed. Recent technological advances have extensively mapped TF binding sites across the genome. However, these methods provide correlative information and the function of specific binding sites remains largely unknown. It is difficult to determine the function of specific binding sites by changing TF concentration. This is because TF concentration changes will not only affect the gene of interest, but also affect hundreds of additional genes regulated by the same TF that could also contribute to the phenotype of interest (Fig. 1A).

The difficulty of determining the function of specific regulatory sites on the genome may be alleviated using CRISPR-Cas9, which can be easily programmed to target specific genomic sequences. A number of CRISPR-Cas9-based approaches have been used to reveal the role of non-coding regulatory DNA[1]. Many approaches rely on recruitment of the catalytically deactivated Cas9 (dCas9) fused to chromatin modifying enzymes that control local gene expression[2–6]. Chromatin modification approaches are suitable for studying the effects of chromatin modifiers on gene expression and downstream effects of specific gene expression. However, they do not provide insight into the function of native transcription factor binding sites. The function of specific binding sites can be determined by introducing indel mutations using active Cas9[7–13]. However, Cas9-induced mutations are random, irreversible, and lack temporal control so that lethal mutations cannot be studied.

To better determine the function of specific transcription factor binding sites, we targeted dCas9 to sterically hinder TF binding (Fig. 1A). This works because CRISPR guide RNAs are longer than most transcription factor binding sites so that flanking sequence can be used to target specific sites over the other TF binding sites across the genome (Fig. 1B). Our approach is related to the CRISPR interference approaches targeting dCas9 to the transcription start site and gene body that have been developed for bacterial cells but not yet in mammalian cells. These methods reduce gene expression, but do not give insight into the function of specific transcription factor binding sites[4,14]. Additionally, in mammalian systems, dCas9 requires additional transcriptional repressors, such as KRAB, to efficiently reduce gene expression[15].

As a case study, we decided to target the Oct4 binding site 137 to 151 bp upstream of the *Nanog* transcription start site. We chose this site because of its putative role in the transcriptional positive feedback loop through which the pluripotency transcription factors, Oct4, Nanog and Sox2, directly promote each others transcription to maintain pluripotency in embryonic stem cells (ESCs) (Fig. 1C). This positive feedback loop may work through the direct binding of Oct4 and Nanog on each others promoter[16,17]. This hypothesis is supported by reporter assays using plasmids containing 406bp of the *Nanog* promoter with and without the Oct4 binding site[18], and by the fact that Oct4 binds to this site in ESCs (Supplementary Fig. 1A)[16,19]. However, standard knockdown or depletion experiments do not directly test the positive feedback hypothesis because these three TFs occupy about 40000 sites in the mouse genome forming a highly complex transcription network[20]. Thus, that Oct4 depletion results in decreased

Nanog expression could be an indirect consequence of the down-regulation of another Oct4 target gene.

To test the hypothesis that Oct4 directly regulates *Nanog* expression, we designed single guide RNAs (sgRNAs) that target dCas9 to the Oct4 binding site (Fig. 1D-E). We used lentiviral infection to express these sgRNAs in mouse ESCs expressing a Tet-inducible dCas9-mCherry fusion protein (Fig. 1F) and in which one allele of *Nanog* was tagged with Venus at the endogenous locus(*Nanog-Venus/dCas9-mCherry mESC*). We designed two sgRNAs, Oct4-Site Nanog-1 and Oct4-Site Nanog-2, that bind to the 14 bp Oct4 binding site and an additional 8 or 10 bp flanking sequence respectively. We note that the flanking sequence confers specificity to the Oct4 binding site in the *Nanog* promoter over other Oct4 binding sites in the genome (Fig. 1B and D). We also included 4 other control sgRNAs (Control1-4) that do not overlap with the targeted Oct4 binding site (Fig. 1E). Expression of the sgRNAs targeting the Oct4 site decreased Nanog-Venus protein and mRNA, reduced Oct4 binding and increased dCas9 binding to the targeted site (Fig. 1G-K). In contrast, expressing control sgRNAs, Control1-4, targeting dCas9 either upstream or downstream of the target site did not have this effect (Fig. 1G-K; and Supplementary Fig. 1B). To compare the effect of blocking Oct4 access to its binding site vs. deletion of the site, we used active Cas9 to delete the Oct4 binding site upstream of the *Nanog-Venus* allele. Deletion of this site reduced Nanog-Venus expression ~2-fold, which was comparable to inhibiting the site using dCas9 (Fig. 1L; Supplementary Fig. 1C).

To test the functional outcome of interfering with Oct4 binding to the *Nanog* promoter, we used time-lapse imaging to measure cell cycle progression in individual mESCs. Disrupting the Oct4 binding to the *Nanog* promoter elongated the cell cycle (median increased from 12h-13h to 16-17h, Fig. 1M, Supplementary Videos 1 and 2). This is consistent with the previously reported slow growth phenotype of Nanog deficient ESCs[21,22]. Thus, our experiment strongly supports the hypothesis that Oct4 directly promotes *Nanog* expression in mESCs to accelerate cell proliferation. Importantly, this example demonstrates that dCas9 can sterically hinder TF binding on specific sites on the genome in mammalian cells, which can be used to determine the contribution of individual TF binding sites to specific cellular functions.

The ease with which multiple sgRNAs can be expressed in cells suggests we can extend our approach to simultaneously target multiple TF binding sites. This multiplexing can be used to interrogate complex transcription networks in which phenotypes result from several TF target genes. To extend our approach, we chose to target another Oct4 binding site that was shown to regulate the *Utf1* gene in reporter assays[23]. *Utf1* itself is a transcription factor that regulates chromatin organization in mESCs and it shares many targets with Oct4 creating potential feedforward transcription motif[23]. Using ChIP-qPCR, we showed that Oct4 binds to its regulatory sequence downstream of the gene *Utf1*, which is consistent with the presence of an Oct4 ChIP-Seq binding peak (Fig. 2A, Supplementary Fig. 2A-B). Recruitment of dCas9 to this Oct4 site using three different sgRNAs (Oct4-Site Utf1-1-3) reduced *Utf1* mRNA expression, reduced Oct4 binding, and increased dCas9 binding to the targeted Oct4 site (Fig. 2B-E). Next, we infected cells with lentiviral particles to express two sgRNAs from a dual guide RNA construct. We showed that we can target dCas9 to the Oct4 sites in the regulatory sequence of both

Nanog and *Utf1* genes. As expected, *Nanog* mRNA and protein levels, and *Utf1* mRNA levels decreased (Fig. 2F-G). Importantly, these sgRNAs did not result in significant decrease in Oct4 binding to two other Oct4 binding sites with similar binding motifs, but different flanking sequence (Supplementary Fig. 1C-F). These results demonstrate that dCas9 can be used to simultaneously interfere with TF binding at multiple loci.

While targeting active Cas9 to TF binding sites can also be used to ascertain their function, such mutagenic approaches are not reversible. Reversibility is important because it allows the study of essential transcription factor binding sites, and can be used to generate controllable dynamics of inhibition of a specific site. For example, one could inhibit a site during a specific interval of development or cell cycle phase. Our approach can be modified to make it rapidly reversible. To do this, we employed a conditionally destabilizing domain (DD) that is rapidly degraded in the absence of the small molecule Shield1[24]. We generated mESCs expressing dCas9 fused to a DD domain and mCherry (ddCas9) (Fig. 2H). To measure the kinetics of ddCas9 degradation, we first grew ddCas9 expressing ESCs in the presence of Shield1. Next, we titrated out Shield1 by adding to the media excess amounts of the purified destabilized domain[25]. Most of the ddCas9-mCherry is degraded in less than one hour as measured by immunoblotting and live-cell imaging (Fig. 2I-J, Supplementary Video 3). To determine how rapidly transcription can be altered using our ddCas9 approach, we infected the *ddCas9-mCherry* ESC line with an sgRNA (Oct4-Site Nanog-1) that targets the Oct4 site in the *Nanog* promoter. After destabilizing ddCas9, Nanog mRNA increased more than 5-fold within one hour, which was closely followed by increases in Nanog protein (Fig. 2K-L). These results show that our ddCas9 can be used to rapidly and reversibly interfere with transcription factor binding.

Here, we showed that dCas9 can compete with endogenous transcription factors to disrupt their binding to specific target sites. This approach can be easily multiplexed to simultaneously target multiple TF sites and the fusion of a conditionally destabilized domain to dCas9 allows rapid and reversible exogenous control of TF binding to specific sites. We expect our approach to determine the function of specific TF binding sites within complex transcription networks via systematic perturbation using a library of sgRNAs.

Methods

Cloning:

sgRNAs were expressed using a lentiviral mouse U6 (mU6)-based expression vector that coexpressed Puro-T2A-BFP from an EF1α promoter (Fig. 1F). New sgRNA sequences were generated by PCR and introduced by InFusion cloning into the sgRNA expression vector digested with BstXI and XhoI. For multiplexing experiments, sgRNAs were expressed using a lentiviral dual sgRNA vector consisting of two sgRNA cassettes in tandem driven by the human U6 promoter and mouse U6 promoter, respectively, and a Puro-T2A-BFP cassettes (Fig. 2F). In the duel sgRNA vector, the mU6 vectors are cloned using InFusion to insert PCR products into a modified vector digested with BstXI and XhoI. The hU6 sgRNA vector was cloned by inserting PCR products with InFusion cloning into the parent vector digested with SpeI and XbaI. After sequence verification, the mU6 vector was digested with XbaI and XhoI and the mU6 sgRNA cassette was ligated into the hU6 vector digested with SpeI and SalI.

To assemble the doxycycline-inducible dCas9 construct (pSLQ1942), human codon-optimized *S. pyogenes* dCas9 was fused at the C-terminus with an HA tag and two SV40 nuclear localization signals (Fig. 1F). For visualization, mCherry was fused at the C-terminus following a P2A peptide. This cassette is driven by the TRE3G doxycycline-inducible promoter. Zeocin resistance and Tet-On 3G transactivator expression is driven by the Ef1α promoter. These cassettes were cloned into a Piggybac plasmid containing the 5' and 3' Piggybac homology arms. To assemble the ddCas9 plasmid (pSLQ2470), the destabilization domain that can be stabilized by Shield1 was amplified from pBMN FKBP(L106P)-YFP-HA and inserted into pSLQ1942. Then, the IRES HcRed-tandem was inserted using InFusion Cloning into KpnI digested pSLQ1942.

Cell culture:

All the ES lines were grown in ESGRO-2i medium (SF016-200, Millipore) supplemented with 100 units/ml streptomycin and 100 mg/ml penicillin on cell culture dishes coated with 0.1% gelatin (G1890, Sigma-Aldrich). The media was changed everyday and cells were passaged every 2 days using Accutase cell detachment solution (SCR005, Millipore). The HEK293T cells were grown in DMEM/F12 supplemented with 10% Fetal Bovine Serum (FBS), 100 units/ml streptomycin, and 100 mg/ml penicillin. All cell culture experiments were done at 37°C and 5% CO_2.

Generation of Nanog-Venus/dCas9-mCherry cell line:

To generate an inducible dCas9 expressing ESC cell line, we inserted tetracycline inducible dCas9-mCherry into the genome of a Nanog-Venus-mESC line using the piggybac transposon system[26,27] (Fig. 1F). Cells were transfected with a tet-on dCas9 plasmid (pSLQ1942) and PiggyBac transposase using the Turbofect transfection reagent following the manufacturer's instructions (R0531, ThermoFisher Scientific). A clonal line was generated by manually selecting an mCherry positive colony and expanding the colony in 2i medium. Addition of 1μg/ml doxycyclin to the medium of the Nanog-Venus/dCas9-mCherry line resulted in a 65-fold increase in the dCas9-mCherry protein signal while the Nanog-Venus protein signal did not

change (Supplementary Fig. 3A-B). A similar strategy was used to infect Nanog-Venus line with destabilized dCas9 vector,pSLQ2470, to generate the Nanog-Venus/ddCas9-mCherry.

Deletion of Oct4 binding site:
We used Edit-R lentiviral inducible Cas9 (Dharmacon) to generate a Nanog-Venus-mESC line that expresses Cas9 by addition of doxycyclin. Next, we infected this line with lentiviral particles encoding an sgRNA targeting slightly downstream of the Oct4 binding site upstream of *Nanog*. The cells were grown in the presence of doxycyclin (1µg/ml) for three days. Individual colonies were grown in a 96 well plate and were genotyped using the following primers: Nanog-Genotype-F (5′-CTTCTTCCATTGCTTAGACGGC-3′), Nanog-Genotype-R (5′-GGCTCAAGGCGATAGATTTAAAGGGTAG-3′). We sequenced the PCR products of the genotyping reaction and a line with a ~230 bp deletion including the Oct4 binding site upstream of Nanog was used for analysis.

sgRNA lentiviral production:
Lentiviral particles containing sgRNA expression plasmids were generated by transfecting HEK293T cells with sgRNA plasmids, and with standard packaging constructs using the Turbofect transfection reagent as previously described[28] (R0531, ThermoFisher Scientific). One day after transfection, the HEK293T cell media was changed from DMEM/FBS to 2i (SF016-200, Millipore). The viral particles in the 2i media were collected after 48 hours, centrifuged, and filtered (0.45-µm syringe filter). The particles were then added to media of the Nanog-Venus/dCas9-mCherry ESC line. The sgRNA expressing cells were selected using puromycin. The expression of sgRNAs was also visually confirmed by microscopy of BFP expression. Supplementary Table 1 shows the sequences of all the sgRNAs used in this study.

Oct4 binding site identification:
To identify Oct4 binding sites, we used available ENCODE ChIP-seq data to find an Oct4 peak near the transcription start site (TSS) of *Nanog*. This broad Oct4 peak is between 500 and 33 bp upstream of the *Nanog* TSS (Fig. 1E). To find the exact location of the Oct4 binding site, we searched for transcription factor motifs using the JASPAR database[29]. This identified a consensus binding site between 137 and 151 bp upstream of the *Nanog* TSS. A similar strategy was used to identify the Oct4 binding site located 1825 bp downstream of the *Utf1* TSS (Fig. 2A). sgRNAs were designed based on available PAM sites (NGG) near the Oct4 binding sites. To obtain the vertebrate TF binding site length, we used the TFBSTools bioconductor package[30] to access the vertebrate TF binding site position frequency matrices of the JASPAR2018 library dataset[29,31]. The number of columns per matrix was used to obtain the distribution of the vertebrate TF binding site length.

Chromatin immunoprecipitation-qPCR (ChIP-qPCR):
ChIP-qPCR was performed using a SimpleChIP Enzymatic Chromatin IP kit following the manufacturer's protocol (#9002, Cell Signaling Technology). Briefly, up to 4 X 10^7 cells were fixed using 4% paraformaldehyde, and chromatin was prepared and fragmented using Micrococcal nuclease (Mnase). Duration and enzyme concentration was optimized to obtain chromatin fragments between 150 and 900 bp. Fragmented chromatin was incubated overnight

at 4°C with antibodies against HA or Oct4 to pull down dCas9-HA or Oct4 on ChIP grade agarose beads. Beads were washed several times, DNA-Protein cross-linking was reversed, and DNA was purified on a column. The purified DNA was used to quantify the binding of Oct4 or HA-dCas9 relative to input using quantitative PCR (qPCR). A list of primers used for ChIP-qPCR analysis is provided in Supplementary Table 2. qPCR was performed with two technical replicates and three or four biological replicates. All the reported enrichment values are normalized to the experiment done on a line expressing HA-dCas9, but no sgRNA. The enrichment was calculated by subtracting the Ct value from qPCR of the pull-down input from the Ct value from qPCR of the of chromatin input (ΔCt). The ΔCt for each sgRNA was subtracted from the ΔCt from the dCas9 only line (ΔΔCt) and relative enrichment was calculated as $2^{\Delta\Delta Ct}$. A goat polyclonal anti-Oct4 antibody (N19, sc-8628, Santa Cruz Biotechnology) was used to pull-down Oct4, and a rabbit polyclonal anti-HA tag (ab9110, Abcam) was used to pull down dCas9 tagged with HA and mCherry.

Quantitative RT-PCR:
Total RNA was harvested from cells using a PARIS RNA isolation kit 3 days after inducing the expression of dCas9 (AM 1921, ThermoFisher Scientific) using 1µg/ml of doxycycline. DNA contamination was removed by treating the isolated RNA with DNAase using a TurboDNA free kit (AM 1907, ThermoFisher Scientific). Quantitative RT-PCR was performed using an iTaq Universal SYBR green one-step kit and an iq-5 Bio-Rad instrument. The primer sequence is shown in Supplementary Table 2. qRT-PCR experiments were performed with two technical replicates and at least 3 biological replicates. mRNA fold change was calculated by subtracting the ΔCt (Ct of tested gene minus Ct for Actin) in samples from lines expressing dCas9 and different sgRNAs from the ΔCt of samples from the control line only expressing dCas9, but no sgRNA (ΔΔCt). The fold change was then calculated as $2^{-\Delta\Delta Ct}$.

Time-lapse microscopy:
For time lapse imaging, cells were plated on 35 cm glass bottom dishes (MatTek) coated with laminin (LN-521™ STEM CELL MATRIX). Imaging experiments were performed 3 days after the induction of dCas9 in a chamber at 37°C perfused with 5% CO_2 [28]. Images were taken every 30 minutes for cell cycle measurements and every 20 minutes for ddCas9 degradation measurements at up to 3 positions per dish for 3 dishes using a Zeiss AxioVert 200M microscope with an automated stage, and an EC Plan-Neofluar 5x/0.16NA Ph1 objective or an A-plan 10x/0.25NA Ph1 objective. Cell cycle duration was calculated by manually tracking cells from when they are born until when they complete mitosis. For ddCas9-mCherry degradation analysis, an excess amount of purified destabilized domain was added to the medium to titrate Shield1 and then the mCherry signal was measured. Cells were manually segmented to calculate the total amount of mCherry within each individual cell. Background signal from the area adjacent to the cell was measured and subtracted from the mCherry signal. The signal for each cell was then normalized to its value in the first frame of the movie.

Immunoblot:
Cells were lysed using an RIPA buffer supplemented with protease and phosphatase inhibitors. Proteins were separated on a 8% SDS-PAGE gel and were transferred to a Nitrocelluluse

membrane using an iBlot (IB21001, ThermoFisher Scientific). Membranes were incubated overnight at 4°C using the following primary antibodies: polyclonal rabbit anti-Nanog antibody (A300-397A, Bethyl Laboratories Inc.), polyclonal goat anti-Oct3/4 antibody (N19, sc-8628, Santa Cruz Biotechnology), and a rabbit polyclonal anti-HA (ab9110, abcam) to detect dCas9, Mouse Alpha-Tubulin(Sigma, T9026), Monoclonal Mouse Gapdh (MA5-15738, Pierce). The primary antibodies were detected using fluorescently labeled secondary antibodies (LI-COR) and were visualized using Licor Odyssey CLx.

Flow cytometry:
Cells were grown in the presence of doxycyclin (1µg/ml) for 3 days before exposure to an Accutase cell detachment solution. Cells were pelleted at 1000 rpm. The pellet was then resuspended in PBS and filtered using a 40µm Cell Strainer (Cornings, #352340). The flow cytometry measurements using a 488 nm Blue laser to detect Nanog-Venus were performed using a FACScan analyzer or a FACS ARIA at the Stanford University FACS facility. The flow cytometry measurements were repeated two times and for each experiment the Nanog-Venus amount of different sgRNAs was normalized to the median amount of dCas9 only expressing cells. For dual guide experiments cells were treated with doxycyclin (1µg/ml) for 5 days before flow cytometry analysis and the Nanog-Venus amounts were not normalized.

Acknowledgements:
The authors would like to thank members of the Skotheim, Qi and Wernig labs for providing feedback and reagents, and to thank Ben Topacio, Devon Brown and Dr. Evgeny Zatulovskiy for their help with flow cytometry analysis and cell culture. Flow cytometry analysis for this project was done on instruments in the Stanford Shared FACS Facility. We also thank Dr. Airlia Thompson for providing the purified DD protein, and Dr. Tom Wandless for providing the pBMN FKBP(L106P)-YFP-HA IRES HcRed-tandem plasmid. This work was supported by the NIGMS/NIH through an NRSA Award F32GM123576 (AS), ALS Association Milton Safenowitz Fellowship, Burroughs Wellcome Fund Postdoctoral Enrichment Program (AAD) and R01 GM092925 (JMS), and by Stanford University through a Bio-X Seed Grant (AS, AD, JMS, MW, LSQ).

Author Contribution:
All authors participated in experimental design. AS performed and analyzed ChIP-qPCR, qRT-PCR, microscopy, Immunoblot and flow cytometry experiments. MW helped set up the mESCs culture. AAD designed and cloned all the sgRNA, dCas9 and conditionally stable dCas9 vectors. JMS and AS wrote the paper.

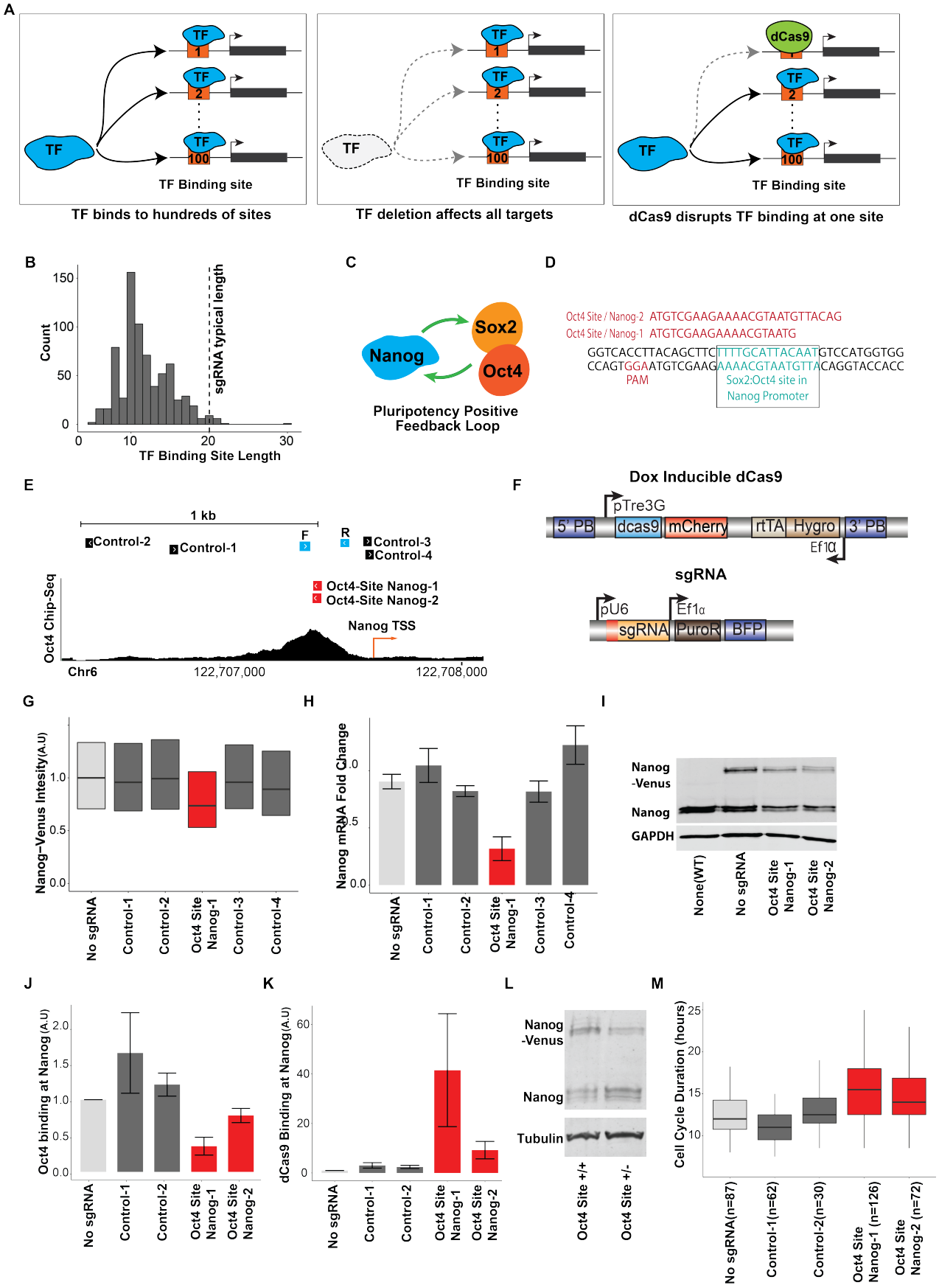

A
TF
TF 1
TF 2
TF 100
TF Binding site
TF binds to hundreds of sites
TF deletion affects all targets
dCas9
dCas9 disrupts TF binding at one site
B
Count
sgRNA typical length
TF Binding Site Length
C
Nanog
Sox2
Oct4
Pluripotency Positive Feedback Loop
D
Oct4 Site / Nanog-2 ATGTCGAAGAAAACGTAATGTTACAG
Oct4 Site / Nanog-1 ATGTCGAAGAAAACGTAATG
GGTCACCTTACAGCTTCTTTTGCATTACAATGTCCATGGTGG
CCAGTGGAATGTCGAAGAAAACGTAATGTTACAGGTACCACC
PAM
Sox2:Oct4 site in Nanog Promoter
E
1 kb
Control-2
Control-1
F
R
Control-3
Control-4
Oct4 Chip-Seq
Oct4-Site Nanog-1
Oct4-Site Nanog-2
Nanog TSS
Chr6
122,707,000
122,708,000
F
Dox Inducible dCas9
pTre3G
5' PB
dcas9
mCherry
rtTA
Hygro
3' PB
Ef1α
sgRNA
pU6
Ef1α
sgRNA
PuroR
BFP
G
Nanog-Venus Intesity(A.U)
No sgRNA
Control-1
Control-2
Oct4 Site Nanog-1
Control-3
Control-4
H
Nanog mRNA Fold Change
No sgRNA
Control-1
Control-2
Oct4 Site Nanog-1
Control-3
Control-4
I
Nanog -Venus
Nanog
GAPDH
None(WT)
No sgRNA
Oct4 Site Nanog-1
Oct4 Site Nanog-2
J
Oct4 binding at Nanog (A.U)
No sgRNA
Control-1
Control-2
Oct4 Site Nanog-1
Oct4 Site Nanog-2
K
dCas9 Binding at Nanog (A.U)
No sgRNA
Control-1
Control-2
Oct4 Site Nanog-1
Oct4 Site Nanog-2
L
Nanog -Venus
Nanog
Tubulin
Oct4 Site +/+
Oct4 Site +/-
M
Cell Cycle Duration (hours)
No sgRNA(n=87)
Control-1(n=62)
Control-2(n=30)
Oct4 Site Nanog-1 (n=126)
Oct4 Site Nanog-2 (n=72)

Figure 1: A. Schematic showing dCas9 can be targeted to sterically inhibit transcription factor binding at a specific site (RHS).
B. Distribution of transcription factor binding site length in vertebrate genomes.
C. Schematic of the transcriptional positive feedback loop that maintains pluripotency in mouse embryonic stem cells.
D. Oct4 binding motif upstream of *Nanog* and targeting sgRNAs.
E. Oct4 ChIP-seq data, sgRNA, and qPCR primer sequences near the *Nanog* TSS.
F. The doxycyclin inducible vector contains dCas9 under the control of a *TRE3G* promoter and another cassette with an *EF1α* promoter driving hygromycin resistance and an rtTA transactivator. The sgRNA vector contains an sgRNA cassette with customizable guide sequence expressed from the U6 promoter and an expression cassette containing an *EF1α* promoter driving expression of a puromycin-resistance gene and BFP.
G. Flow cytometry of Nanog-Venus, H, qRT-PCR of *Nanog* mRNA, I, immunoblot , J, Oct4 ChIP-qPCR, and K, dCas9 ChIP-qPCR measurements for cells expressing dCas9 and the indicated sgRNA.
L. Immunoblot of Nanog protein in wild type (+/+) and heterozygous (+/-) Oct4 binding site deletion.
M. Distributions of cell cycle durations for cells expressing dCas9 and the indicated sgRNA. Targeting sgRNAs are shown in red, control sgRNAs in dark gray, and the no sgRNA control in light gray. Bottom and upper lines of box plots show the first and third interquartile range and the middle line shows the median. Bar plots show mean and associated standard error.

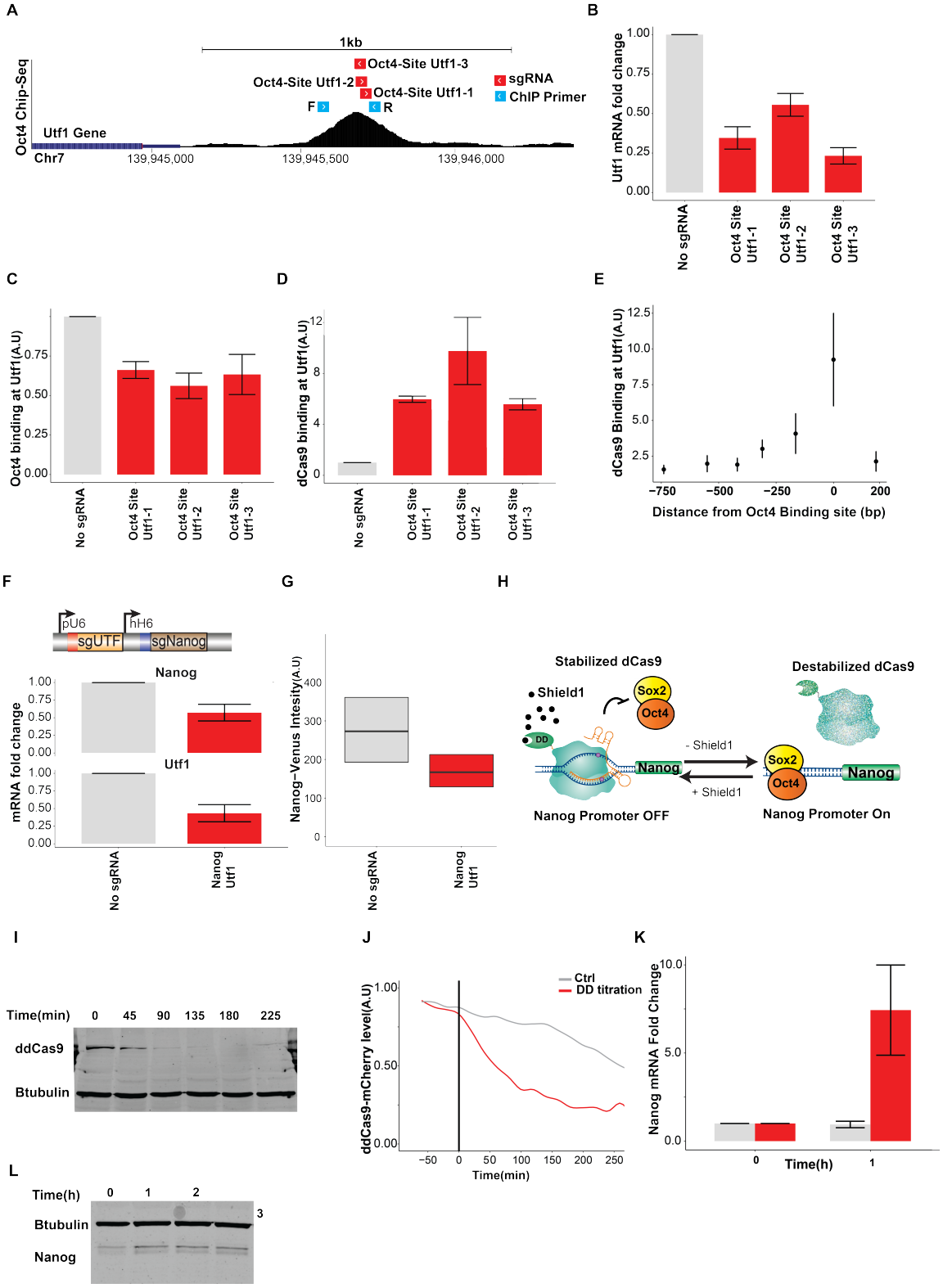
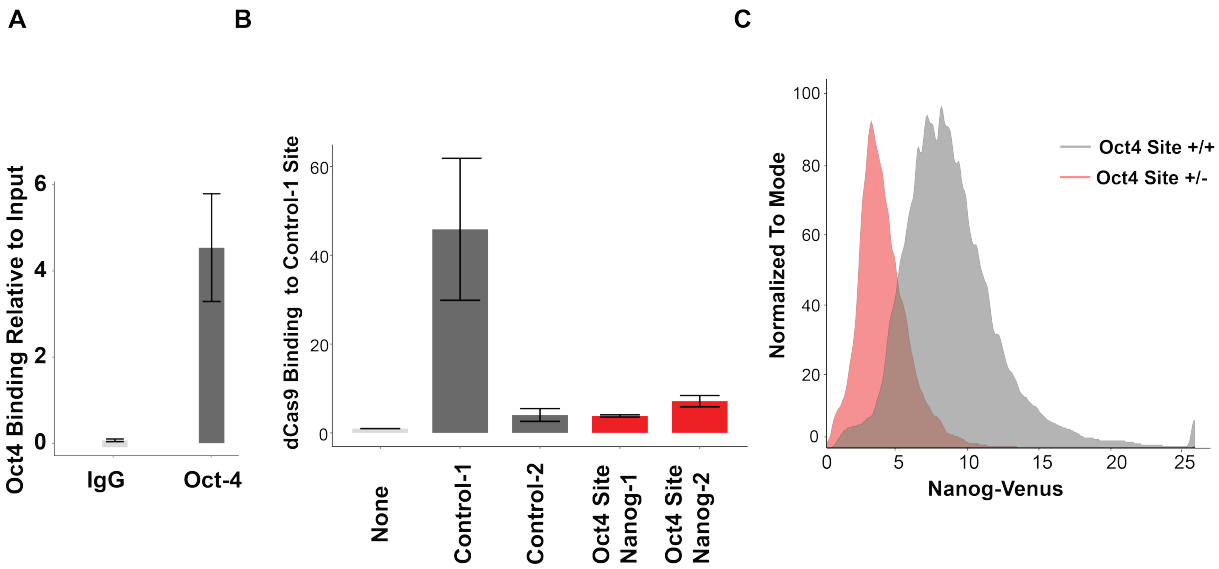
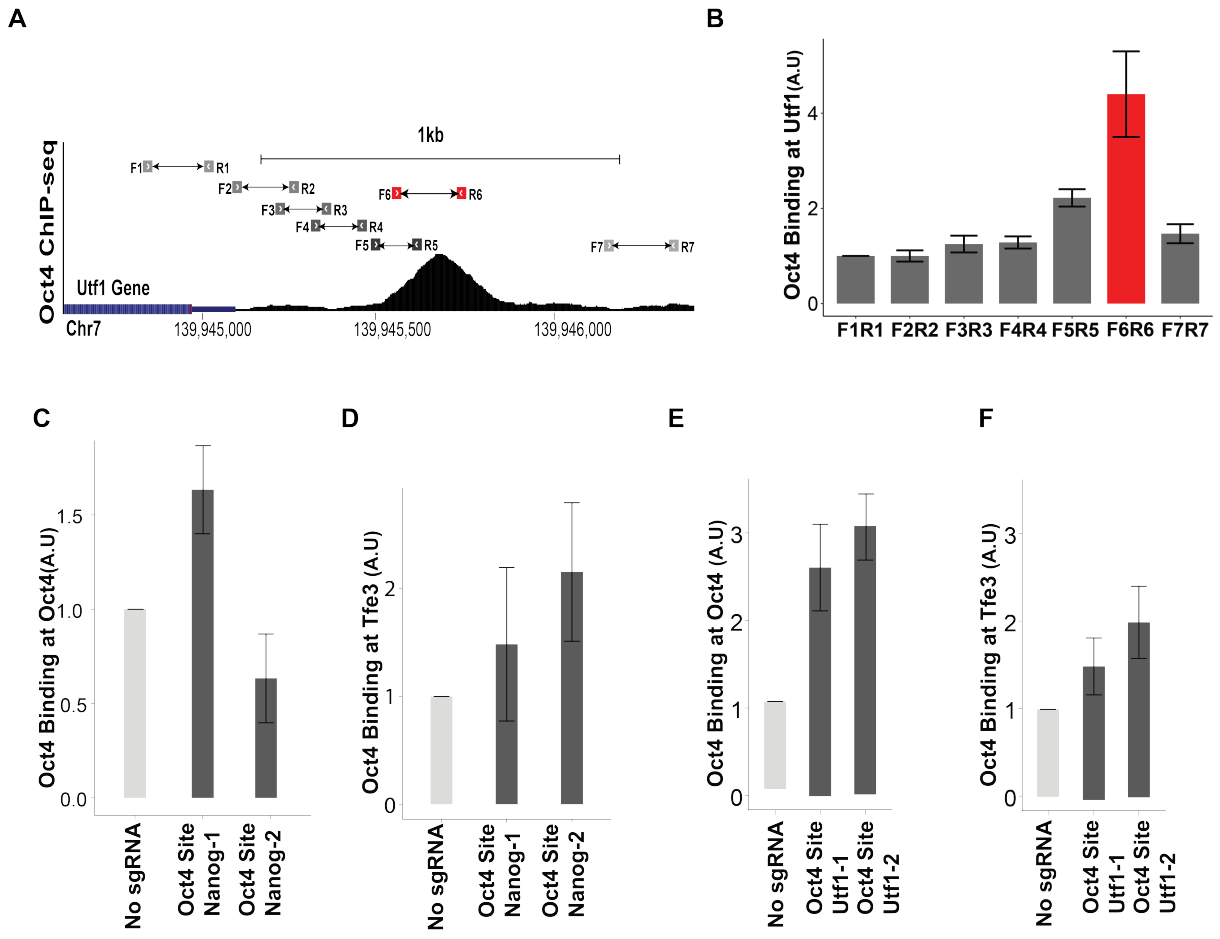

A
1kb
Oct4-Site Utf1-3
Oct4-Site Utf1-2
Oct4-Site Utf1-1
sgRNA
ChIP Primer
F
R
Oct4 Chip-Seq
Utf1 Gene
Chr7
139,945,000
139,945,500
139,946,000
B
Utf1 mRNA fold change
No sgRNA
Oct4 Site Utf1-1
Oct4 Site Utf1-2
Oct4 Site Utf1-3
C
Oct4 binding at Utf1(A.U)
D
dCas9 binding at Utf1(A.U)
E
dCas9 Binding at Utf1(A.U)
Distance from Oct4 Binding site (bp)
F
pU6
hH6
sgUTF
sgNanog
Nanog
Utf1
mRNA fold change
G
Nanog−Venus Intesity(A.U)
H
Stabilized dCas9
Destabilized dCas9
Shield1
Sox2
Oct4
DD
Nanog
- Shield1
+ Shield1
Nanog Promoter OFF
Nanog Promoter On
I
Time(min)
0
45
90
135
180
225
ddCas9
Btubulin
J
Ctrl
DD titration
ddCas9-mCherry level(A.U)
Time(min)
K
Nanog mRNA Fold Change
Time(h)
L
Time(h)
Btubulin
Nanog

Figure 2: A. Oct4 ChIP-seq data, sgRNA, and qPCR primer sequences downstream of the *Utf1* gene.
B. *Utf1* mRNA qRT-PCR measurement, C, Oct4 ChIP-qPCR, and, D, dCas9 ChIP-qPCR for cells expressing dCas9 and the indicated sgRNAs.
E. dCas9 ChIP-qPCR measurement using overlapping primers near the Oct4 binding site downstream of the *Utf1* gene.
F. Top: schematics of the construct used to express two sgRNAs. Bottom: *Nanog* and *Utf1* mRNA qRT-PCR measurement in cells expressing either the dCas9 alone or with the sgRNAs targeting the indicated Oct4 sites near both genes (Nanog Utf1).
G. Flow cytometry measurement of Nanog-Venus protein in control and Nanog Utf1 sgRNA expressing cells.
H. Schematic of conditionally-destabilized dCas9 (ddCas9) used for reversible inhibition.
I. Time-course Immunoblot, and, J, microscopy of ddCas9 degradation after Shield1 titration by excess competitive inhibitor protein (DD).
K. Nanog mRNA qRT-PCR, and, L, Immunoblot measurement of Nanog protein after ddCas9 degradation by Shield1 titration.



Supplementary Figure 1: A. Oct4 ChIP-qPCR using primers adjacent to the Oct4 binding site upstream of the *Nanog* gene following pull-down with an Oct4 antibody or IgG control antibody. B. dCas9 ChIP-qPCR using primers next to the binding site of the control-1 sgRNA (see Fig. 1E).

C. Flow cytometry measurement of Nanog-Venus in wild type (Oct4 site +/+) and heterozygous Oct4 site deletion (Oct4 site +/-) upstream of Nanog-Venus.

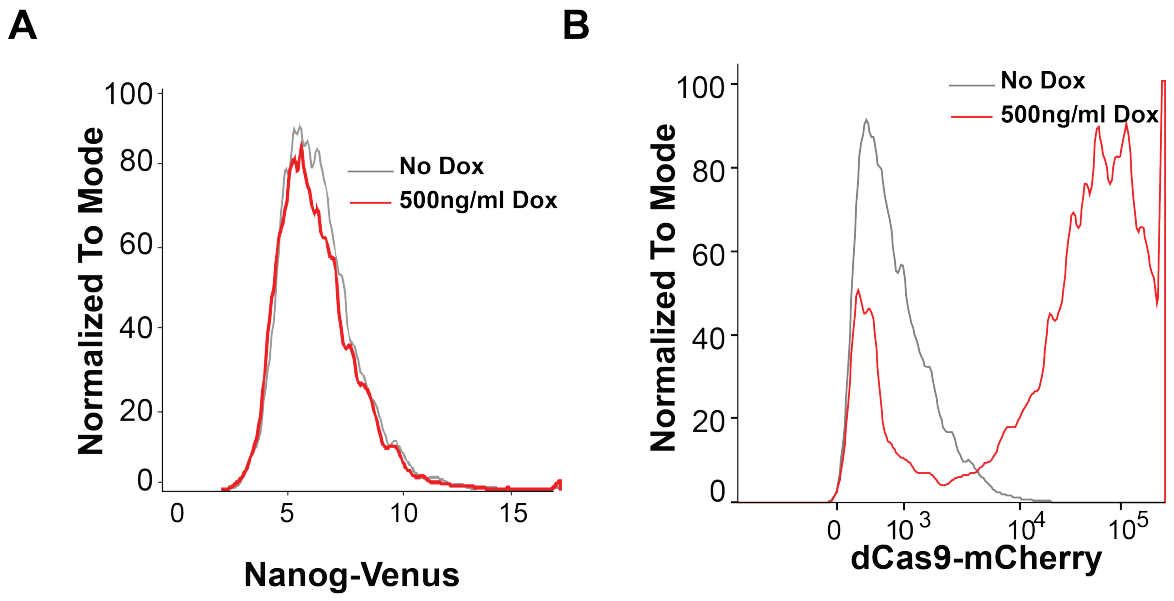


Supplementary Figure 2: A. Position of ChIP-qPCR primers used to identify the Oct4 binding site in the regulatory sequences downstream of the *Utf1* gene.

B. Oct4 ChIP-qPCR data using a set of overlapping primers to confirm the location of the Oct4 binding site downstream of the *Utf1* gene.

C. ChIP-qPCR analysis of Oct4 binding at an Oct4 site near the *Oct4* gene, or, D, an Oct4 site near the *Tfe3* gene in ESCs expressing the Oct4-Site Nanog-1 sgRNA or the Oct4-Site Nanog-2 sgRNA.

E. ChIP-qPCR analysis of Oct4 binding at an Oct4 site near the *Oct4* gene, or, D, an Oct4 site near the *Tfe3* gene in ESCs expressing the Oct4-Site Utf1-1 sgRNA or the Oct4-Site Utf1-2 sgRNA.



Supplementary Figure 3: A. Flow cytometry analysis of Nanog-Venus, and, B, dCas9-mCherry expression in the Nanog-Venus/dCas9-mCherry ESC line.

Supplementary Video 1, 2: Phase contrast time lapse microscopy of the dCas9/Nanog-Venus control cell line without (Video 1) and with expression of the Oct4-Site Nanog-1 sgRNA (Video 2).

Supplementary Video 3: Time lapse phase contrast and fluorescent microscopy of degradation of ddCas9-mCherry after Shield1 titration by excess DD protein as a competitive inhibitor (Video 3).

References:

1. Montalbano, A., Canver, M. C. & Sanjana, N. E. High-Throughput Approaches to Pinpoint Function within the Noncoding Genome. *Mol. Cell* **68,** 44–59 (2017).
2. Thakore, P. I., Black, J. B., Hilton, I. B. & Gersbach, C. A. Editing the epigenome: technologies for programmable transcription and epigenetic modulation. *Nat. Methods* **13,** 127–137 (2016).
3. Korkmaz, G. *et al.* Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat. Biotechnol.* **34,** 192–198 (2016).
4. Dominguez, A. A., Lim, W. A. & Qi, L. S. Beyond editing: Repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nature Reviews Molecular Cell Biology* **17,** 5–15 (2016).
5. Klann, T. S. *et al.* CRISPR–Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nat. Biotechnol.* **35,** 561–568 (2017).
6. Xie, S., Duan, J., Li, B., Zhou, P. & Hon, G. C. Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells. *Mol. Cell* **66,** 285–299.e5 (2017).
7. Canver, M. C. *et al.* BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* **527,** 192–197 (2015).
8. Canver, M. C. *et al.* Variant-aware saturating mutagenesis using multiple Cas9 nucleases identifies regulatory elements at trait-associated loci. *Nat. Genet.* **49,** 625–634 (2017).
9. Sanjana, N. E. *et al.* High-resolution interrogation of functional elements in the noncoding genome. *Science (80-.).* **353,** 1545–1549 (2016).
10. Diao, Y. *et al.* A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. *Nat. Methods* **14,** 629–635 (2017).
11. Diao, Y. *et al.* A new class of temporarily phenotypic enhancers identified by CRISPR/Cas9-mediated genetic screening. *Genome Res.* **26,** 397–405 (2016).
12. Gasperini, M. *et al.* CRISPR/Cas9-Mediated Scanning for Regulatory Elements Required for HPRT1 Expression via Thousands of Large, Programmed Genomic Deletions. *Am. J. Hum. Genet.* **101,** 192–205 (2017).
13. Rajagopal, N. *et al.* High-throughput mapping of regulatory DNA. *Nat. Biotechnol.* **34,** 167–174 (2016).
14. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152,** 1173–83 (2013).
15. Gilbert, L. a *et al.* CRISPR-Mediated Modular RNA-Guided Regualtion of Transcription in Eukaryotes. *Cell* **154,** 442–451 (2013).
16. Loh, Y.-H. *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38,** 431–440 (2006).
17. Young, R. A. Control of the embryonic stem cell state. *Cell* **144,** 940–954 (2011).
18. Rodda, D. J. *et al.* Transcriptional regulation of Nanog by OCT4 and SOX2. *J. Biol. Chem.* **280,** 24731–24737 (2005).
19. Kim, J., Chu, J., Shen, X., Wang, J. & Orkin, S. H. An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. *Cell* **132,** 1049–1061 (2008).

20. Sorrells, T. R. & Johnson, A. D. Making sense of transcription networks. *Cell* **161,** 714–723 (2015).

21. Mitsui, K. *et al.* The homeoprotein nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113,** 631–642 (2003).

22. Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113,** 643–655 (2003).

23. Kooistra, S. M. *et al.* UTF1 Regulates ES Cell Chromatin Organization and Gene Expression. *Stem Cells* (2010).

24. Banaszynski, L. A., Chen, L. chun, Maynard-Smith, L. A., Ooi, A. G. L. & Wandless, T. J. A Rapid, Reversible, and Tunable Method to Regulate Protein Function in Living Cells Using Synthetic Small Molecules. *Cell* **126,** 995–1004 (2006).

25. Miyazaki, Y., Chen, L., Chu, B. W., Swigut, T. & Wandless, T. J. Distinct transcriptional responses elicited by unfolded nuclear or cytoplasmic protein in mammalian cells. *Elife* **4,** (2015).

26. Filipczyk, A. *et al.* Biallelic expression of nanog protein in mouse embryonic stem cells. *Cell Stem Cell* **13,** 12–13 (2013).

27. Ding, S. *et al.* Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* **122,** 473–483 (2005).

28. Schwarz, C. *et al.* A Precise Cdk Activity Threshold Determines Passage through the Restriction Point. *Mol. Cell* **69,** 253–264.e5 (2018).

29. Sandelin, A. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* **32,** 91D–94 (2004).

30. Tan, G. & Lenhard, B. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. *Bioinformatics* **32,** 1555–6 (2016).

31. Khan, A. *et al.* JASPAR 2018: Update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* **46,** D260–D266 (2018).