UC Santa Cruz

UC Santa Cruz Previously Published Works

Title

Reversible Disruption of Specific Transcription Factor-DNA Interactions Using CRISPR/Cas9.

Permalink

https://escholarship.org/uc/item/089213z6

Journal

Molecular Cell, 74(3)

Authors

Shariati, S
Dominguez, Antonia
Xie, Shicong
et al.

Publication Date

2019-05-02

DOI

10.1016/j.molcel.2019.04.011

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Mol Cell. Author manuscript; available in PMC 2020 May 02.

Published in final edited form as:

Mol Cell. 2019 May 02; 74(3): 622–633.e4. doi:10.1016/j.molcel.2019.04.011.

Reversible Disruption of Specific Transcription Factor-DNA Interactions Using CRISPR/Cas9

S. Ali Shariati[1], **Antonia Dominguez**[2], **Shicong Xie**[1], **Marius Wernig**[3], **Lei S. Qi**[2,4,5], and **Jan M. Skotheim**[1,6,*]

[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 University, Stanford, CA 94305, USA

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

[6]Lead Contact

SUMMARY

The control of gene expression by transcription factor binding sites frequently determines phenotype. However, it is difficult to determine the function of single transcription factor binding sites within larger transcription networks. Here, we use deactivated Cas9 (dCas9) to disrupt binding to specific sites, a method we term CRISPRd. Since CRISPR guide RNAs are longer than transcription factor binding sites, flanking sequence can be used to target specific sites. Targeting dCas9 to an Oct4 site in the *Nanog* promoter displaced Oct4 from this site, reduced *Nanog* expression, and slowed division. In contrast, disrupting the Oct4 binding site adjacent to *Pax6* upregulated *Pax6* transcription and disrupting Nanog binding its own promoter upregulated its transcription. Thus, we can easily distinguish between activating and repressing binding sites and examine autoregulation. Finally, multiple guide RNA expression allows simultaneous inhibition of multiple binding sites, and conditionally destabilized dCas9 allows rapid reversibility.

In Brief

Interactions between transcription factors and their binding sites control gene transcription. Despite progress in mapping the binding sites of transcription factors across the genome, the function of the most binding sites remains largely unknown. We present a new method, termed CRISPRd, for the rapid functional analysis of specific binding sites.

*Correspondence: skotheim@stanford.edu.

AUTHOR CONTRIBUTIONS

All authors participated in experimental design. S.A.S. performed and analyzed ChIP-qPCR, qRT-PCR, microscopy, immunoblot, and flow cytometry experiments. S.X. and S.A.S. performed the bioinformatic analysis. M.W. helped set up the mESC culture. L.S.Q. participated in design and analysis of CRISPRd experiments. A.D. designed and cloned all the sgRNA, dCas9, and conditionally stable dCas9 vectors. J.M.S. and SA.S. wrote the paper.

DECLARATION OF INTERESTS

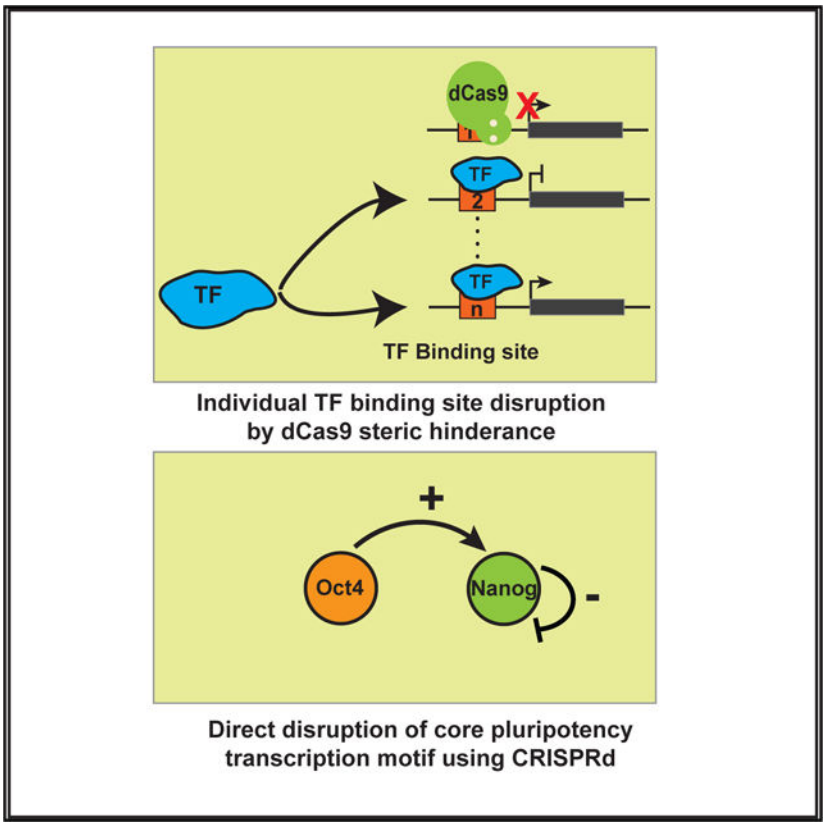
The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.molcel.2019.04.011.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Graphical abstract



INTRODUCTION

Binding of transcription factors (TFs) to specific regulatory sequences controls when and where target genes are expressed. While recent technological advances have extensively mapped TF binding sites across the genome, this provides only correlative information, and the function of specific binding sites remains largely unknown. The function of specific binding sites is difficult to determine by changing TF concentration because TF concentration changes will not only affect the gene of interest, but also hundreds of additional genes regulated by the same TF that could also affect phenotype (Figure 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 (Montalbano et al., 2017). Most commonly, CRISPR-Cas9 is used to target specific binding sites by introducing indel mutations. For example, a high-resolution tiling approach was used to systematically introduce indel mutations and identify functional elements across the enhancer region of BCL11A (Canver et al., 2015). Another deletion-based approach, termed CREST-seq, uses paired single guide RNAs (sgRNAs) to delete specific ~2 kb regions (Diao et al., 2017). Overlapping ~2 kb regions are then targeted to identify functional regulatory elements at higher resolution. However, Cas9-induced mutations are random and irreversible and lack temporal control so that lethal mutations

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

cannot be studied (Canver et al., 2015; Diao et al., 2017; Gasperini et al., 2017; Rajagopal et al., 2016; Sanjana et al., 2016).

One possibility to alleviate the drawbacks of using catalytically active Cas9 to target specific TF-DNA binding is to use a catalytically inactive Cas9 (dCas9). dCas9 has previously been used to downregulate transcription, but not to interrogate the function of a specific TF binding site. For example, in bacteria a method known as CRISPRi is based on recruiting dCas9 to interfere with transcription machinery directly (Bikard et al., 2013; Qi et al., 2013). In mammalian cells, a variation of this approach is based on targeting dCas9 fused with chromatin-modifying en-zymesthat inhibit gene expression, such as the histone deacety-lase complex KRAB (Dominguez et al., 2016; Klann et al., 2017; Korkmaz et al., 2016; Thakore et al., 2015; Xie et al., 2017). While such chromatin modification approaches are suitable for inhibiting gene expression, they cannot be used to determine the function of specific TF-DNA binding sites (Figure 1B) (Groner et al., 2010; Thakore et al., 2015). More specifically, if dCas9-KRAB is targeted to successfully inhibit expression of a gene, this does not give insight into how any TF-DNA binding sites in the gene promoter regulate expression (Table 1) (Dominguez et al., 2016; Qi et al., 2013). Similarly, CRISPRa approaches to activate expression of specific genes by using dCas9 to recruit transcriptional activators does not determine the function of specific TF-DNA binding sites (Konermann et al., 2015; Tanenbaum etal., 2014).

DESIGN

To better determine the function of specific TF binding sites, we developed a technique we named CRISPRd. This method disrupts specific TF-DNA interactions by using dCas9 to sterically hinder TF binding (Figure 1A). In contrastto CRISPRi approaches that only inhibit target gene expression without determining the function of specific TF-DNA interactions, interrupting a TF-DNA interaction using CRISPRd determines whether that interaction is activating, is repressing, or has no function (Figures 1B and 1C; Table 1). CRISPRd works because CRISPR guide RNAs are longer than most TF binding sites so that flanking sequence can be used to disrupt specific sites in the genome without affecting other sites (Figure 1D). We also hypothesized that dCas9 can outcompete TFs for binding to a specific site because the dCas9 dissociation rate is significantly slower (hours) (Jones et al., 2017) than the dissociation rate of typical TFs (seconds) (Chen et al., 2014). To test this, we designed sgRNAs that target dCas9 to specific TF-DNA binding sites (Figure 1E). We then used lentiviral infection to express these sgRNAs in mouse embryonic stem cells (ESCs) expressing a Dox-inducible dCas9-mCherry fusion protein (Figure 1E).

RESULTS AND DISCUSSION

CRISPRd Specifically Disrupts an Oct4 Binding Site Upstream of Nanog

As a first case study for our technology, 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 core pluripotency TFs, Oct4, Nanog, and Sox2, directly promote each other's transcription to maintain pluripotency in ESCs (Figure 2A). This positive feedback loop may work through

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the direct binding of Oct4 and Nanog to each other's promoter (Loh et al., 2006; Young, 2011). This hypothesis is supported by reporter assays using plasmids containing 406 bp of the *Nanog* promoter with and without the Oct4 binding site (Rodda et al., 2005) and by the fact that Oct4 binds to this site in ESCs (Figure S1A) (Kim et al., 2008; Loh et al., 2006). However, standard knock-down or depletion experiments do not directly test the positive feedback hypothesis because these three TFs occupy about 40,000 sites in the mouse genome, forming a highly complex transcription network (Sorrells and Johnson, 2015). Thus, that Oct4 depletion results in decreased *Nanog* expression could be an indirect consequence of the downregulation of another Oct4 target gene.

To test the hypothesis that Oct4 directly regulates *Nanog* expression, we designed single sgRNAs that target dCas9 to the Oct4 binding site 137 to 151 bp upstream of the *Nanog* transcription start site (Figures 2B and 2C). We used lentiviral infection to express these sgRNAs in mouse ESCs expressing a Dox-inducible dCas9-mCherry fusion protein and in which one allele of *Nanog* was tagged with Venus at the endogenous locus (*Nanog-Venus/dCas9-mCherry*). 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. We also included four other control sgRNAs (Control1–4) that do not overlap with the targeted Oct4 binding site (Figure 2C).

Expression of the sgRNAs targeting this Oct4 site decreased the Nanog-Venus protein level as measured by flow cytometry (Figure 2D; Figure S1B). Similarly, the mRNA level of Nanog decreased by more than 50% when dCas9 was targeted to the Oct4 binding site (Figure 2E; Figure S1C). In contrast, expressing control sgRNAs, Control1–4, targeting dCas9 either upstream or downstream of the target site did not have this effect (Figures 2C–2E). Both wild-type Nanog and Nanog-Venus proteins show reduced expression, which implies that dCas9 can displace Oct4 from its binding site on both Nanog alleles (Figure 2F). To directly test whether dCas9 recruitment sterically hinders the binding of Oct4, we used chromatin immunoprecipitation (ChIP) to measure the binding of Oct4 and dCas9 in cell lines expressing targeting or control sgRNAs. Expression of targeting sgRNAs significantly reduced the binding of Oct4 to its target site while increasing the binding of dCas9 to this same site. This was not the case for control sgRNAs (Figures 2G and 2H; Figure S1D).

We next sought to compare the effect of using CRISPRd to block Oct4 access to its binding site in the *Nanog* promoter with the effect of deleting this site using active Cas9. Deletion of the Oct4 binding site upstream of the *Nanog-Venus* allele reduced expression ~2-fold, which was comparable to the effect of inhibiting Oct4 binding to the site with CRISPRd (Figure 2I; Figure S1E). These findings suggest that this specific Oct4 binding site is responsible for about 50% of *Nanog* mRNA transcription and that CRISPRd can efficiently counteract the effect of this site on *Nanog* expression by blocking Oct4 binding.

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

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Oct4 binding to the *Nanog* promoter elongated the cell cycle (median cell-cycle duration increased from 12–13 h to 16–17 h; Figure 2J; Videos S1 and S2). This is consistent with the previously reported slow growth phenotype of Nanog-deficient ESCs (Chambers et al., 2003; Mitsui et al., 2003). Thus, our experiment strongly supports the hypothesis that Oct4 directly promotes *Nanog* expression to accelerate cell proliferation in mESCs. Importantly, this example demonstrates that CRISPRd can be used to steri-cally 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.

Reversible Disruption of TF Binding Using Destabilized dCas9

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 TF binding sites and can be used to generate controllable dynamics of inhibition of a specific site at a specific time. For example, one could inhibit a site during a specific interval of development or cell-cycle phase.

Our approach can be engineered 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 Shieldl (Banaszynski et al., 2006). We generated mESCs expressing dCas9 fused to a DD and mCherry (ddCas9) (Figure 3A). To measure the kinetics of ddCas9 degradation, we first grew ddCas9-expressing ESCs in the presence of Shieldl. Next, we titrated out Shieldl by adding excess amounts of the purified destabilized domain to the media (Miyazaki et al., 2015). Most of the ddCas9-mCherry is degraded in less than 1 h as measured by immunoblotting and live-cell imaging (Figures 3B and 3C; Video S3). To determine how rapidly transcription can be altered using our ddCas9 approach, we infected the *ddCas9-mCherry* ESC line with a virus containing an sgRNA (Oct4-Site Nanog-1) that targets the Oct4 site in the *Nanog* promoter. After destabilizing ddCas9, Nanog mRNA increased about 5-fold within 1 h, which was closely followed by an increase in Nanog protein (Figures 3D and 3E). These results show that ddCas9 can be used to rapidly and reversibly interfere with TF binding.

CRISPRd Allows Simultaneous Targeting of Multiple TF Binding Sites

Next, we aimed to extend our approach to simultaneously target multiple TF binding sites. Such 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 (Kooistra et al., 2010). *Utf1* itself is a TF that regulates chromatin organization in mESCs and shares many targets with Oct4 (Kooistra et al., 2010). Oct4 binds to its regulatory sequence downstream of the gene *Utf1,* which is consistent with the presence of an Oct4 ChlP-seq binding peak (Figure 4A; Figures S2A and S2B). 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 (Figures 4B–4D; Figure S2C). Next, we transfected cells to express two sgRNAs from a dual guide RNA construct (Figure 4E). Targeting dCas9 to the Oct4 sites in the regulatory sequence of both *Nanog* and *Utf1* genes reduced *Nanog* mRNA and protein levels and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

reduced *Utf1* mRNA levels (Figures 4F and 4G). Importantly, these sgRNAs did not result in a significant decrease in Oct4 binding to two other Oct4 binding sites with similar binding motifs but with different flanking sequences (Figures S2D–S2G). These results demonstrate that dCas9 can be used to simultaneously disrupt TF binding at multiple loci.

CRISPRd Can Distinguish between Activating and Repressing Binding Sites

CRISPRd can be used to determine whether a specific TF-DNA binding interaction promotes or inhibits transcription. In contrast, CRISPRi solely aims to inhibit transcription by modifying local chromatin. Here, we examine a case in which an Oct4-DNA interaction was likely to inhibit transcription. In addition to binding to the promoters of genes involved in self-renewal, Oct4 binds to the promoters of genes regulating lineage specification, including the neuronal lineage (Loh et al., 2006). Interestingly, Oct4 also associates with transcriptional repressors in pluripotent cells (Liang et al., 2008), which suggests that Oct4 can both activate self-renewal genes and repress neuronal target genes to suppress neuronal fate in ESCs (Thomson et al., 2011). If this is the case, blocking an Oct4 repressive binding site near its neuronal targets should result in increased expression of the target gene. *Pax6* is a pan-neuronal TF that is not transcribed in mESCs but whose promoter is bound by Oct4 (Figure 5A) (Kim et al., 2008). Disruption of this binding site with CRISPRd resulted in a 5-± 1-fold increase in *Pax6* mRNA and an increase in Pax6 protein (Figures 5B and 5C). These findings show how CRISPRd can determine the function of a TF-DNA binding site that inhibits transcription.

Identification of an Autoinhibitory Element in the *Nanog* Promoter Using CRISPRd

To test whether CRISPRd can be broadly used to inhibit TF-DNA interactions, we sought to test it on an additional TF. We chose to target the binding of Nanog to its own promoter. While Nanog is considered to function as a negative regulator of its own transcription based on reporter assays (Fidalgo et al., 2012; Navarro et al., 2012), the mechanism responsible for this negative feedback regulation was not known. That a Nanog ChIP-seq peak exists 4.9 kb upstream of the transcription start site of the *Nanog* gene suggested that this direct interaction was the origin of the negative feedback loop (Figure 6A). We targeted this Nanog binding site using three independent sgRNAs. Two of these sgRNAs resulted in significant recruitment of dCas9 to this site and reduction of Nanog binding to its own promoter (Figures 6B and 6C). This, in turn, led to the increased expression of Nanog mRNA and protein (Figures 6D–6F). These results strongly suggest that Nanog binding to its own promoter forms a negative feedback loop (Figure 6G) and show that CRISPRd is applicable to different TFs.

Pluripotency Factor Binding Sites Can Be Disrupted Specifically across Genome

Using CRSIPRd to disrupt specific TF binding sites requires variable flanking sequences to design specific sgRNAs. In addition, dCas9 recruitment requires a protospacer-adjacent motif (PAM) sequence near the targeted binding site (Dominguez et al., 2016). To explore these requirements, we first performed an analysis of the 30 bp flanking sequences of Oct4:Sox2 binding sites genome-wide. This showed that the entire flanking region was random, which is ideal for designing specific sgRNAs (Figure 7A). In addition,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

approximately 98% of the pluripotency factor binding sites have a PAM available within 30 bp of the binding site (Figure 7B).

Next, we sought to address the possibility of off-target activity, in which dCas9 targeted for one specific TF binding site would inhibit another. The specificity of dCas9 binding to a TF binding site likely depends on the uniqueness of the targeting sgRNA. To assess off-target activity of sgRNAs targeting Oct4 binding sites, we calculated the minimum number of base pair changes required to match a pair of sgRNAs targeting two different Oct4 binding sites. First, we identified more than 35,173 possible sgRNAs that target 6,942 Oct4 binding sites genome-wide. For each site, we chose the sgRNAs that was the most different from the other sgRNAs. Then, we calculated the number of base pair changes needed to match this sgRNA to the closest other sgRNA for another Oct4 peak (Figures 7C and 7D). We find that more than 90% of the Oct4-DNA binding sites have a mismatch number that is greater than 4. This mismatch number of 4 is sufficient to yield specificity as indicated by ChIP analysis of potential off-target sites with mismatch numbers of 4 or 6 base pairs, which showed that binding of Oct4 and dCas9 to these off-target sites is comparable to the no sgRNA control (Figures 7E–7G). Similar to Oct4 sites, more than 95% of Nanog binding sites in the genome can also be targeted with an sgRNA with minimum mismatch number of 4 (Figure S3). While our off-target analysis is focused only on ChIP-seq peaks nearTF binding sites of interest, other studies have focused on off-target activity of CRISPR/Cas9 genome-wide. We refer readers to the following papers for a more extensive discussion (Boyle et al., 2017; Gilbert et al., 2014; Hilton et al., 2015; O'Geen et al., 2015; Wu etal., 2014).

Limitations

CRISPRd can be limited by the presence of PAM sequences near the TF binding site of interest. This limitation could be alleviated by using Cas9 or tholog with different PAM requirements (Dominguez et al., 2016; Shalem et al., 2015). We also note that our current methods do not allow us to use ddCas9 to reversibly control individual sites when targeting multiple sites using multiple guide RNAs. This limitation may be alleviated by using ddCas9 together with another RNA-guided nuclease, such as Cas12a, that has a different PAM requirement (Gao et al., 2017). This would allow the two nucleases to bind distinct guides so that their concentrations could be independently controlled.

It should also be noted that because of the slow search kinetics of dCas9 (Jones et al., 2017), disrupting a TF-DNA interaction is not immediate and can take hours. Protein engineering techniques to generate dCas9 variants with faster search kinetics could help achieve faster disruption of TF binding sites using CRISPRd. Finally, we note that CRISPRd should not be used if a TF binding site is too close to a site responsible for recruiting general transcription machinery, such as the transcription start site. In this case, the observed effect may be due to the disruption of the general transcription machinery rather than through the specific TF binding site.

Finally, we note that a closed chromatin state can impede binding of Cas9 to the genome (Daer et al., 2017; Horlbeck et al., 2016; Kuscu et al., 2014; Singh et al., 2015; Wu et al., 2014). The binding sites that were disrupted in this study are all located in open chromatin region as identified by ENOCDE Dnase-seq data (Experiment number: ENCSR000CMW),

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

which we expect to be typical because more than 95% of TF ChIP-seq peaks identified by ENCODE also fall within open chromatin (ENCODE Project Consortium, 2012; Thurman et al., 2012).

Conclusions

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 tested CRISPRd on Oct4 and Nanog, which are two key TFs that are expressed at different levels in single mESCs (Kolodziejczyk et al., 2015). While Nanog expression is among highly expressed TFs, Oct4 expression is close to the average expression of TFs expressed by mESCs (Figure S4), suggesting that CRISPRd works for both highly and normally expressed TFs.

The clear next step is to expand CRISPRd to allow screening binding sites genome-wide for cell biological phenotypes. We expect that our approach will be used to determine the function of specific TF binding sites within complex transcription networks via systematic perturbation using sgRNA libraries.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Goat polyclonal anti-Oct4	Santa Cruz Biotechnology	sc-8628
Rabbit polyclonal anti-HA tag	Abcam	Ab9110
Rabbit polyclonal Anti-Nanog	Bethyl	A300–397a
Mouse Anti-A-Tubulin	Sigma	T9026
Chicken Anti-Pax6	Developmental Studies Hybridoma Bank	AB-528427
Mouse Anti-Gapdh	Pierce	MA5–15738
Bacterial and Virus Strains		
E.Coli DH5Alpha	Agilent	200231
Chemicals, Peptides, and Recombinant Proteins		
Shield 1	Clontech	632189
Purified destabilized domain	Kopito Lab, Stanford University	N/A
ESGRO-2i media	Millipore	SF016–200
Gelatin	Sigma-Aldrich	G1890
TurboFect	Thermo-Fisher	R0532
Doxycycline, Hyclate, CAS 24390–14-5,	Calbiochem	324385–1GM
Critical Commercial Assays		

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SimpleChIP Enzymatic Chromatin IP Kit	Cell Signaling	9003,9002
One Step kit qRT-PCR	BioRad	1725150
In-Fusion HD Cloning Plus	Takara	638920
Experimental Models: Cell Lines		
Mouse: Nanog-Venus embryonic stem Cell, Male	Professor Nusse lab, Stanford	Generated by Filipczyk et al., 2013 from an R1 male murine ESC
Mouse: V6.5 embryonic stem cells, Male	Professor Wernig Lab, Stanford, Novus Biological	NBP1–41162, male murine ESCs
Human: HEK293T Cells, Female	ATCC	CRL-3216
Oligonucleotides		
Oligonucleotides are listed in Tables S1 and S2	This paper	N/A
Recombinant DNA		
See Method Details of the paper for constructs	This paper	N/A
Tet-On 3G Inducible Expression System-pCMV-Tet3G Regulator Plasmid & pTRE3G Response Plasmid	Clontech Laboratories	631168
pHR-mU6-sgRNA/EF1a-Puro-T2A-BFP	This Study	N/A
pHR-hU6-sgNanog / mU6-sgUtf1/ Ef1a-Puro-T2A-BFP	This Study	NA
Other		
ESGRO-2i media	Milipore	SF016–200
DMEM with 4.5g/L Glucose, Sodium Pyruvate; Without L-Glutamine, Phenol Red	Corning	17205CV
Penicillin Streptomycin 100X Solution	HyClone	SV30010
L-Glutamine, 200mM Solution	HyClone	SH30034.02

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Jan Skotheim (skotheim@stanford.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture—Mouse V6.5 ESC (male) and mouse Nanog-Venus ESC (male) 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 every day and cells were passaged every 2 days using Accutase cell detachment solution (SCR005, Millipore). The human HEK293T (female) 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. Cells were checked for mycoplasma contamination every three months and were never tested positive.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

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 (Ding et al., 2005; Filipczyk et al., 2013) (Figure 1E). Nanog-Venus-mESC line is originally created by Filipczyk et al. from an R1 mouse ESC background which are male cells (R1/E (ATCC SCRC-1036). 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 ESGRO-2i medium (SF016–200, Millipore). Addition of 1 mg/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 (Figures S5A and S5B). A similar strategy was used to infect Nanog-Venus line with destabilized dCas9 vector, pSLQ2470, to generate the Nanog-Venus/ddCas9-mCherry. ddCas9-mCherry construct has a Destabilizing Domain at the N-terminal followed by dCas9 fused to mCherry through a P2A peptide.

METHOD DETAILS

Cloning—sgRNAs were expressed using a lentiviral mouse U6 (mU6)-based expression vector that coexpressed Puro-T2A-BFP from an EF1α promoter (Figure 1E). 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 (Figure 4E). 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 (Figure 1E). 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 TetOn 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.

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

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

downstream of the Oct4 binding site upstream of *Nanog.* The cells were grown in the presence of doxycyclin (1 mg/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'-GGCTCAAGGCGATAGATTTAAAGGGT AG-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 (R0531, ThermoFisher Scientific) as previously described (Schwarz et al., 2018). 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 h, centrifuged, and filtered (0.45-mm 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. Table S1 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 (Figure 2C). To find the exact location of the Oct4 binding site, we searched for transcription factor motifs using the JASPAR database (Sandelin et al., 2004). 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 (Figure 4A). 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 to access the vertebrate TF binding site position frequency matrices of the JASPAR2018 library dataset (Khan et al., 2018; Tan and Lenhard, 2016). The number of columns per matrix was used to obtain the distribution of the vertebrate TF binding site length.

Chromatin immunoprecipitation-qPCR—ChIP-qPCR was performed using SimpleChIP Enzymatic Chromatin IP kit following the manufacturer's protocol (#9002, Cell Signaling Technology). Briefly, up to 4×10^7 cells were fixed using 4% paraformaldehyde, and chromatin was prepared and fragmented using Micrococcal nuclease (Mnase). Duration and enzyme concentration were 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 Table S2. 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

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

line expressing HA-dCas9, but no sgRNA. The enrichment was calculated by subtracting the Ct value from qPCR of the pull-down chromatin 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, a rabbit polyclonal anti-HA tag (ab9110, Abcam) was used to pull down dCas9 tagged with HA and mCherry, and a purified polyclonal rabbit anti-Nanog (Nanog Antibody, A300–397A, Bethyl) was used to pull down Nanog. We used Student's t test to measure the statistical significance of relative differences in Oct4, Nanog and dCas9 binding with p value < 0.05 considered significant.

Quantitative RT-PCR—Total RNA was harvested from cells using PARIS RNA isolation kit 3 days after inducing the expression of dCas9 (AM 1921, ThermoFisher Scientific) using 1 μg/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 Table S2. 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 Ctfor 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 $2^{-\Delta\Delta Ct}$. We used Student's t test to measure the statistical significance of RNA fold changes with p value < 0.05 considered significant.

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 [25]. Images were taken every 30 min for cell cycle measurements and every 20 min 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. 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.

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-Oct¾ antibody (N19, sc-8628, Santa Cruz Biotechnology), rabbit polyclonal anti-HA (ab9110, abcam) to detect dCas9, Mouse Alpha-Tubulin (Sigma, T9026) and Monoclonal Mouse Gapdh (MA5–15738, Pierce). The primary antibodies were detected using fluorescently labeled secondary antibodies (LI-COR) and were visualized using Li-Cor Odyssey CLx. For quantification of the results the integrated intensity of the band was quantified using ImageJ software and was normalized to the loading control.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Immunofluorescence staining—Cells were plated on 35 cm glass bottom dishes (MatTek) coated with laminin. Cells were grown in 2i media supplemented with doxycyclin (1 μg/mL) to induce dCas9 expression. The cells were washed with PBS two times and then fixed using 4% paraformaldehyde for 10 min. Cells were washed in PBS and were incubated in 1.5% Bovine Serum Albumin (BSA) solution for 1h. Cells were stained with a chicken Pax6 antibody (Developmental Studies Hybridoma Bank) for 2 h at room temperature. Alexa 633 secondary antibody (Invitrogen) was used to visualize the Pax6 signal and DAPI was used to visualize the nuclei. The intensity of Pax6 signal for each cell was measured by segmenting the nuclei and subtracting the background from an area adjacent to the cell.

Flow cytometry—Cells were grown in the presence of doxycyclin (1 μg/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 to three 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 (1 μg/mL) for 5 days before flow cytometry analysis and the Nanog-Venus amounts were not normalized. For statistical analysis of flow cytometry data the median signal of Nanog-Venus from three measurements were used to perform a t test with p value less than 0.05 considered significant.

Computational analysis of flanking sequences—The ChIP-seq peaks for Oct4 binding were obtained from the GSM307137 dataset (Marson et al., 2008). The reads were scanned for the presence of an Oct4:Sox2 motif using theTFBSTools package in R (Tan and Lenhard, 2016). Next, 30 bp flanking sequences for each binding site were obtained using the R package IRange (Lawrence et al., 2013). The nucleotide distribution of the binding sites and their flanking sequences were calculated and visualized as a sequence logo using ggseqlogo R package (Wagih, 2017). PAM availability was calculated using the MatchPattern of Biostrings package to search for GG sequences on both strands of DNA (Pagès et al., 2019). All possible sgRNA sequences were obtained from the 20 bp sequences upstream of PAM sites within 30bp of all Oct4 biding sites across the genome. In total, 35,173 guides were obtained that flank Oct4 binding sites. For Nanog, the ± 30bp flanking the central 10bp of Nanog ChIP-seq were obtained from GSE11724 (Marson et al., 2008). We did not filter by position weight matrix as Nanog's matrix varies across databases and studies. 85,191 sgRNAs were obtained that flank Nanog binding sites. Mismatch number between pairs of sgRNAs were calculated using custom scripts and BioPython (Cock et al., 2009).

QUANTIFICATION AND STATISTICAL ANALYSIS

Cell cycle duration was calculated by time-lapse microscopy and manually tracking individual cells from when they are born until they complete mitosis. We used Student's t test to measure the statistical significance between the cell cycle speed of different sgRNAs groups.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

For ddCass9-mCherry degradation, cells were manually segmented to calculate the total intensity 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 time-point.

For Quantitative RT-PCR, we used Student's t test to measure the statistical significance of RNAfold changes with p value < 0.05 considered significant using ΔΔCT method. For ChIP-qPCR, we used Student's t test to measure the statistical significance of enrichment of pull-down for Oct4, Nanog and dCas9 over their respective input chromatin preparations with p value < 0.05 considered significant.

For quantification of the western blot results the integrated intensity of the band was quantified using ImageJ software and was normalized to the loading control. we used Student's t test to measure the statistical significance between protein concentration.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors would like tothank members ofthe Skotheim, Qi, and Wemig labs for providing feedback and reagents and Ben Topacio, Devon Brown, and Dr. Evgeny Zatulosvskiy 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. Wealsothank Dr. AirliaThompson for providingthe 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 and K99GM126027 (S.A.S.), ALS Association Milton Safenowitz Fellowship, Burroughs Welcome Fund Postdoctoral Enrichment Program (A.D.) and R01 GM092925 (J.M.S.), and by Stanford University through a Bio-X Seed Grant (S.A.S., A.D., J.M.S., M.W., L.S.Q.).

REFERENCES

Banaszynski LA, Chen LC, Maynard-Smith LA, Ooi AG, and Wandless TJ (2006).Arapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. Cell 126, 995–1004. [PubMed: 16959577]

Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, and Marraffini LA (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res. 41, 7429–7437. [PubMed: 23761437]

Boyle EA, Andreasson JOL, Chircus LM, Sternberg SH, Wu MJ, Guegler CK, Doudna JA, and Greenleaf WJ (2017). High-throughput biochemical profiling revealssequence determinantsofdCas9 off-target binding and unbinding. Proc. Natl. Acad. Sci. USA 114, 5461–5466. [PubMed: 28495970]

Canver MC, Smith EC, Sher F, Pinello L, Sanjana NE, Shalem O, Chen DD, Schupp PG, Vinjamur DS, Garcia SP, et al. (2015). BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature 527, 192–197. [PubMed: 26375006]

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, and Smith A (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell 113, 643–655. [PubMed: 12787505]

Chen J, Zhang Z, Li L, Chen B-C, Revyakin A, Hajj B, Legant W, Dahan M, Lionnet T, Betzig E, et al. (2014). Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. Cell 156, 1274–1285. [PubMed: 24630727]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, and de Hoon MJL (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25, 1422–1423. [PubMed: 19304878]

Daer RM, Cutts JP, Brafman DA, and Haynes KA (2017). The impact of chromatin dynamics on Cas9-mediated genome editing in human cells. ACS Synth. Biol. 6, 428–438. [PubMed: 27783893]

Diao Y, Fang R, Li B, Meng Z, Yu J, Qiu Y, Lin KC, Huang H, Liu T, Marina RJ, et al. (2017).A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. Nat. Methods 14, 629–635. [PubMed: 28417999]

Ding S, Wu X, Li G, Han M, Zhuang Y, and Xu T (2005). Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. Cell 122, 473–483. [PubMed: 16096065]

Dominguez AA, Lim WA, and Qi LS (2016). Beyond editing: repurposing CRlSPR-Cas9 for precision genome regulation and interrogation. Nat. Rev. Mol. Cell Biol. 17, 5–15. [PubMed: 26670017]

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74. [PubMed: 22955616]

Fidalgo M, Faiola F, Pereira C-F, Ding J, Saunders A, Gingold J, Schaniel C, Lemischka IR, Silva JCR, and Wang J (2012). Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming. Proc. Natl. Acad. Sci. USA 109, 16202–16207. [PubMed: 22988117]

Filipczyk A, Gkatzis K, Fu J, Hoppe PS, Lickert H, Anastassiadis K, and Schroeder T (2013). Biallelicexpression ofnanog protein in mouseembryonic stem cells. Cell Stem Cell 13, 12–13. [PubMed: 23827706]

Gao L, Cox DBT, Yan WX, Manteiga JC, Schneider MW, Yamano T, Nishimasu H, Nureki O, Crosetto N, and Zhang F (2017). Engineered Cpf1 variants with altered PAM specificities. Nat. Biotechnol. 35, 789–792. [PubMed: 28581492]

Gasperini M, Findlay GM, McKenna A, Milbank JH, Lee C, Zhang MD, Cusanovich DA, and Shendure J (2017). 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. [PubMed: 28712454]

Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, et al. (2014). Genome-scale CRlSPR-mediated control of gene repression and activation. Cell 159, 647–661. [PubMed: 25307932]

Groner AC, Meylan S, Ciuffi A, Zangger N, Ambrosini G, Denervaud N, Bucher P, and Trono D (2010). KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. PLoS Genet. 6, e1000869.

Hilton IB, D'lppolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, and Gersbach CA (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol. 33, 510–517. [PubMed: 25849900]

Horlbeck MA, Witkowsky LB, Guglielmi B, Replogle JM, Gilbert LA, Villalta JE, Torigoe SE, Tjian R, and Weissman JS (2016). Nucleosomes impede Cas9 access to DNA in vivo and in vitro. eLife 5, 2767.

Jones DL, Leroy P, Unoson C, Fange D, Curic V, Lawson MJ, and Elf J (2017). Kinetics of dCas9 target search in Escherichia coli. Science 357, 1420–1424. [PubMed: 28963258]

Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, Bessy A, Cheneby J, Kulkarni SR, Tan G, et al. (2018). JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Res. 46 (D1), D260–D266. [PubMed: 29140473]

Kim J, Chu J, Shen X, Wang J, and Orkin SH (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell 132, 1049–1061. [PubMed: 18358816]

Klann TS, Black JB, Chellappan M, Safi A, Song L, Hilton IB, Crawford GE, Reddy TE, and Gersbach CA (2017). CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. Nat. Biotechnol. 35, 561–568. [PubMed: 28369033]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Kolodziejczyk AA, Kim JK, Tsang JCH, Ilicic T, Henriksson J, Natarajan KN, Tuck AC, Gao X, Buhler M, Liu P, et al. (2015). Single cell RNA-sequencing of pluripotent states unlocks modular transcriptional variation. Cell Stem Cell 17, 471–485. [PubMed: 26431182]

Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H, et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature 517, 583–588. [PubMed: 25494202]

Kooistra SM, van den Boom V, Thummer RP, Johannes F, Wardenaar R, Tesson BM, Veenhoff LM, Fusetti F, O'Neill LP, Turner BM, et al. (2010). UTF1 regulates ES cell chromatin organization and gene expression. Stem Cells. 10.1002/stem.497.

Korkmaz G, Lopes R, Ugalde AP, Nevedomskaya E, Han R, Myacheva K, Zwart W, Elkon R, and Agami R (2016). Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. Nat. Biotechnol. 34, 192–198. [PubMed: 26751173]

Kuscu C, Arslan S, Singh R, Thorpe J, and Adli M (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nat. Biotechnol. 32, 677–683. [PubMed: 24837660]

Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, and Carey VJ (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol. 9, e1003118.

Liang J, Wan M, Zhang Y, Gu P, Xin H, Jung SY, Qin J, Wong J, Cooney AJ, Liu D, and Songyang Z (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat. Cell Biol. 10, 731–739. [PubMed: 18454139]

Loh Y-H, Wu Q, Chew J-L, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J, et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat. Genet. 38, 431–440. [PubMed: 16518401]

Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, Johnstone S, Guenther MG, Johnston WK, Wernig M, Newman J, et al. (2008). Connecting microRNAgenestothecoretranscriptional regulatory circuitry of embryonic stem cells. Cell 134, 521–533. [PubMed: 18692474]

Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, and Yamanaka S (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 113, 631–642. [PubMed: 12787504]

Miyazaki Y, Chen L-C, Chu BW, Swigut T, and Wandless TJ (2015). Distinct transcriptional responses elicited by unfolded nuclear or cytoplasmic protein in mammalian cells. eLife 4, 4.

Montalbano A, Canver MC, and Sanjana NE (2017). High-throughput approaches to pinpoint function within the noncoding genome. Mol. Cell 68, 44–59. [PubMed: 28985510]

Navarro P, Festuccia N, Colby D, Gagliardi A, Mullin NP, Zhang W, Karwacki-Neisius V, Osorno R, Kelly D, Robertson M, and Chambers I (2012). OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. EMBO J. 31, 4547–4562. [PubMed: 23178592]

O'Geen H, Henry IM, Bhakta MS, Meckler JF, and Segal DJ (2015). A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. Nucleic Acids Res. 43, 3389–3404. [PubMed: 25712100]

Pagès H,Aboyoun P, Gentleman R, and DebRoy S (2019). Biostrings: efficient manipulation of biological strings. R package version https://rdrr.io/bioc/Biostrings/.

Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, and Lim WA (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173–1183. [PubMed: 23452860]

Rajagopal N, Srinivasan S, Kooshesh K, Guo Y, Edwards MD, Banerjee B, Syed T, Emons BJM, Gifford DK, and Sherwood RI (2016). High-throughput mapping of regulatory DNA. Nat. Biotechnol. 34, 167–174. [PubMed: 26807528]

Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng H-H, and Robson P (2005). Transcriptional regulation of nanog by OCT4 and SOX2. J. Biol. Chem. 280, 24731–24737. [PubMed: 15860457]

Sandelin A, Alkema W, Engström P, Wasserman WW, and Lenhard B (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res. 32, D91–D94. [PubMed: 14681366]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Sanjana NE, Wright J, Zheng K, Shalem O, Fontanillas P, Joung J, Cheng C, Regev A, and Zhang F (2016). High-resolution interrogation of functional elements in the noncoding genome. Science 353, 1545–1549. [PubMed: 27708104]

Schwarz C, Johnson A, Koivomagi M, Zatulovskiy E, Kravitz CJ, Doncic A, and Skotheim JM (2018). A precise Cdk activity threshold determines passage through the restriction point. Mol. Cell 69, 253–264.e5. [PubMed: 29351845]

Shalem O, Sanjana NE, and Zhang F (2015). High-throughput functional genomics using CRISPR-Cas9. Nat. Rev. Genet. 16, 299–311. [PubMed: 25854182]

Singh R, Kuscu C, Quinlan A, Qi Y, and Adli M (2015). Cas9-chromatin binding information enables more accurate CRISPR off-target prediction. Nucleic Acids Res. 43, e118–e118. [PubMed: 26032770]

Sorrells TR, and Johnson AD (2015). Making sense of transcription net-works. Cell 161, 714–723. [PubMed: 25957680]

Tan G, and Lenhard B (2016). TFBSTools: an R/bioconductor package for transcription factor binding site analysis. Bioinformatics 32, 1555–1556. [PubMed: 26794315]

Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, and Vale RD (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. Cell 159, 635–646. [PubMed: 25307933]

Thakore PI, D'Ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, Reddy TE, Crawford GE, and Gersbach CA (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat. Methods 12, 1143–1149. [PubMed: 26501517]

Thomson M, Liu SJ, Zou L-N, Smith Z, Meissner A, and Ramanathan S (2011). Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. Cell 145, 875–889. [PubMed: 21663792]

Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, et al. (2012). The accessible chromatin landscape of the human genome. Nature 489, 75–82. [PubMed: 22955617]

Wagih O (2017). ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics 33, 3645–3647. [PubMed: 29036507]

Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, Cheng AW, Trevino AE, Konermann S, Chen S, et al. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat. Biotechnol. 32, 670–676. [PubMed: 24752079]

Xie S, Duan J, Li B, Zhou P, and Hon GC (2017). Multiplexed engineering and analysis of combinatorial enhancer activity in single cells. Mol. Cell 66, 285–299.e5. [PubMed: 28416141]

Young RA (2011). Control of the embryonic stem cell state. Cell 144, 940–954. [PubMed: 21414485]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights

- CRISPRd disrupts transcription factor DNA interactions using dCas9
- Single TF binding sites can be targeted because sgRNAs are longer than binding sites
- CRISPRd distinguishes between inhibitory and activating TF binding sites
- CRISPRd uncovers regulatory principles of the pluripotency transcription network

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

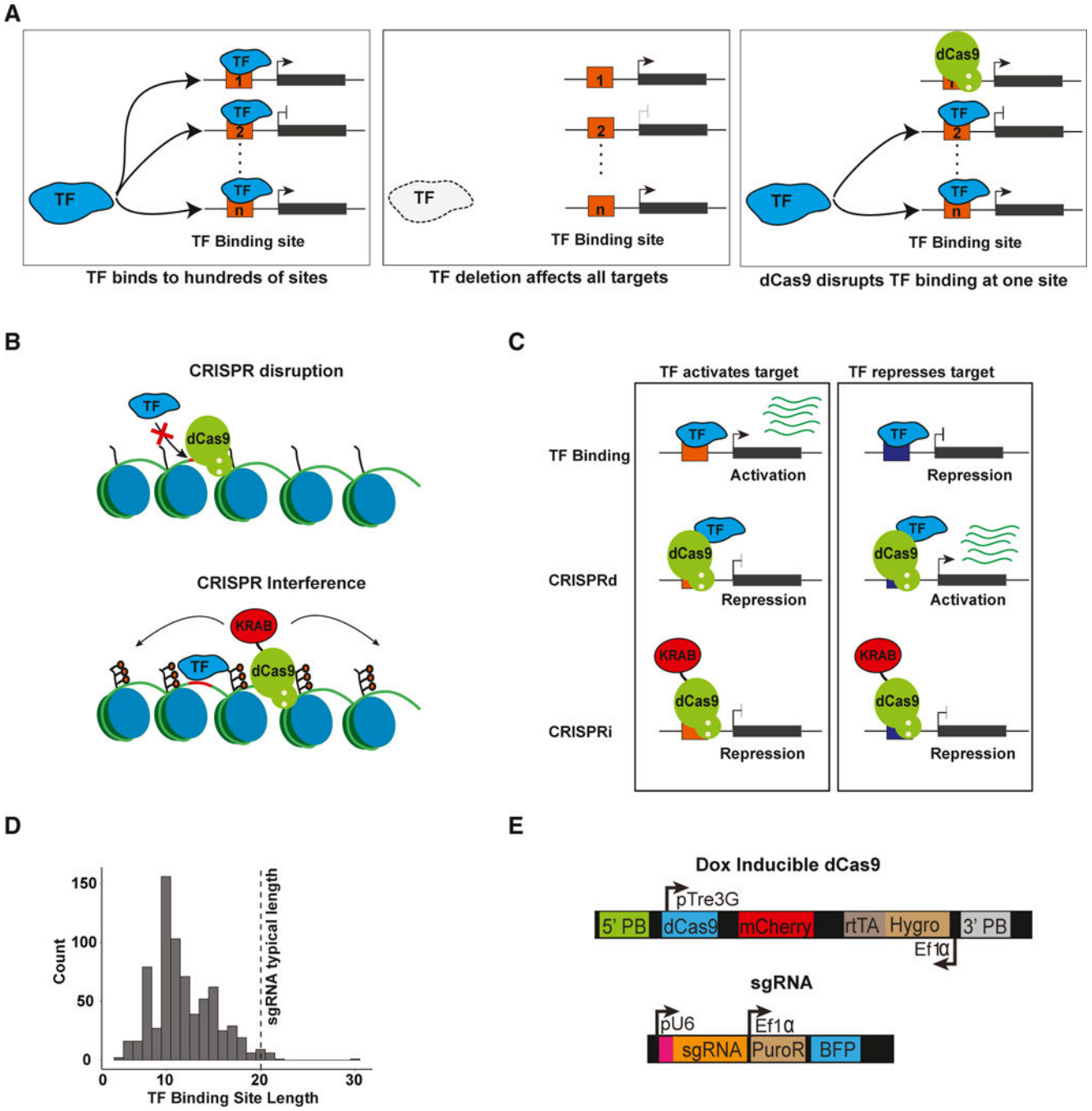


Figure 1. Experimental Design and Comparison with Existing Methods

(A) Schematic showing a TF binding to hundreds of sites across the genome (left). Function of individual binding sites cannot be determined by deleting the TF as this will not only affect the binding site of interest but also hundreds of other targets (middle). CRISPRd targets dCas9 to sterically inhibit transcription factorbinding at a specific site (right).

(B) Schematic of CRISPRd and CRISPRi approaches. CRISPRi aims to downregulate gene expression by inducing chromatin modifications, while CRISPRd aims to disrupt a specific TF-DNA interaction.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(C) Schematic showing the expected effects on gene expression using CRISPRd and CRISPRi. A TF binding site interaction can activate or repress target genes (top). CRISPRd can distinguish activating from repressing functions (middle). CRISPRi represses the expression of the targeted locus without distinguishing between activating or repressing TF-DNA binding sites (bottom).

(D) Distribution of TF consensus binding site lengths in vertebrate genomes. The vertical line shows the typical size of an sgRNA, which is longer than most TF binding sites so that the flanking sequence can be used to target individual binding sites.

(E) 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.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

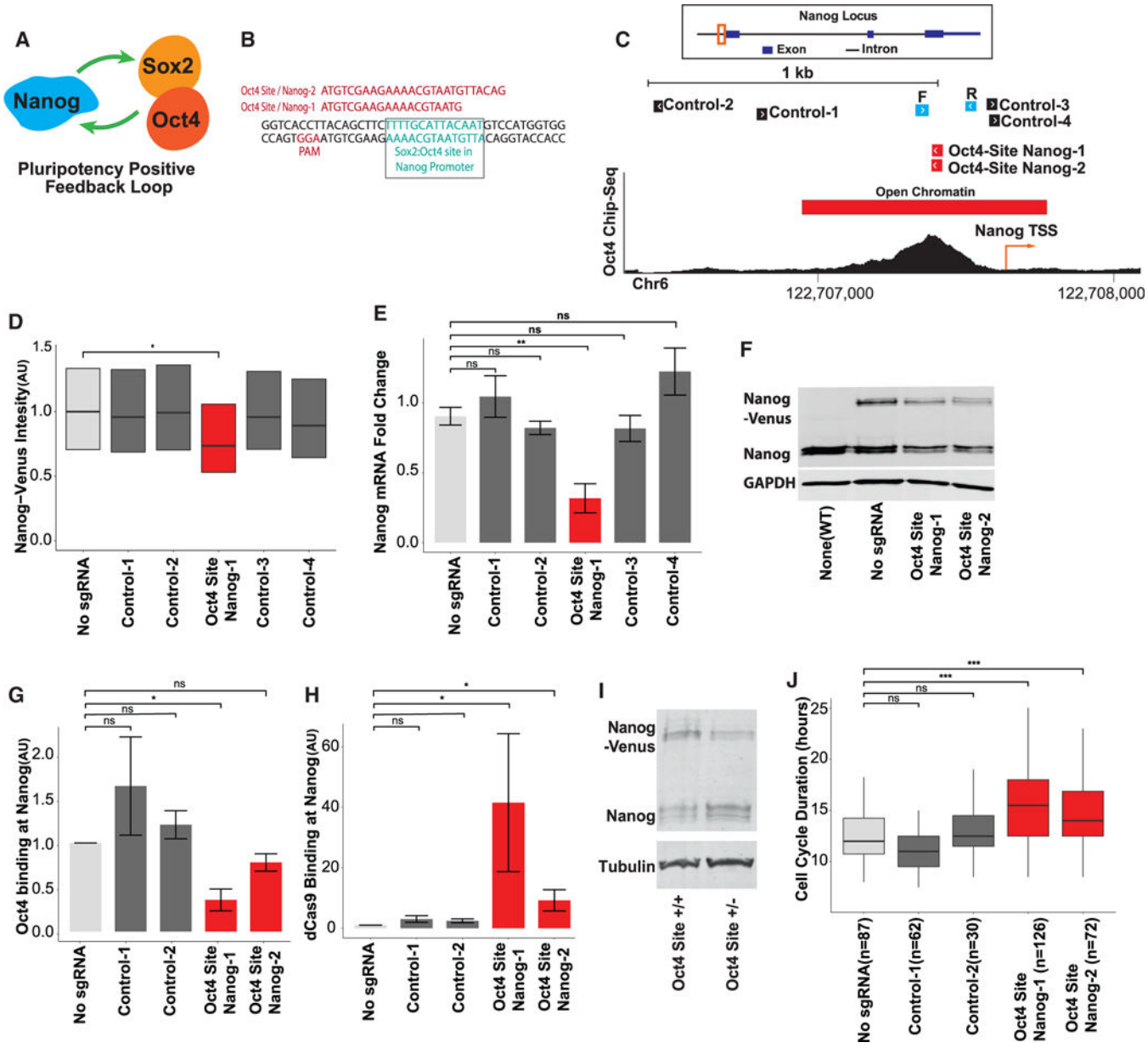


Figure 2. CRISPRd Disrupts Oct4 Binding in the *Nanog* Promoter, Decreases *Nanog* Expression, and Elongates the Cell Cycle

(A) Schematic of the transcriptional positive feedback loop that maintains pluripotency in mouse embryonic stem cells.

(B) The Oct4 binding motif upstream of *Nanog* is shown in green, and the targeting sgRNAs and PAM sequences are shown in red.

(C) Top:adiagram of the Nanog locus with the genomic region shown below indicated by the orange box. Bottom: ChlP-seq dataforOct4 show the chromosomal position of an Oct4 peak upstream of *Nanog.* Also shown are the positions oft argeting sgRNAs (red), Control sgRNAs (black), and qPCR primers (blue). The red box denotes open chromatin identified by ENCODE DNase-Seq (Experiment number: ENCSR000CMW)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(D) Flow cytometry of Nanog-Venus shows that targeting an upstream Oct4 binding site reduces Nanog protein. Control sgRNAs targeting dCas9 upstream or downstream of the target site has no effect.
(E) qRT-PCR measurement of *Nanog* mRNA shows a decrease when the targeting sgRNA is expressed, but not when the control sgRNAs are expressed.
(F) lmmunoblot of Nanog shows that both *Nanog* alleles decrease their expression.
(G) Oct4 ChlP-qPCR shows that Oct4 binding upstream of Nanog is reduced by the indicated targeting sgRNA.
(H) dCas9 ChlP-qPCR measurements for cells expressing dCas9 and the indicated sgRNA.
(I)Immunoblot of Nanog protein in wild-type (+/+) and heterozygous (+/−) Oct4 binding site deletion mESCs.
(J) Distributions of cell-cycle durations for cells expressing dCas9 and the indicated sgRNA.
For all panels, the targeting sgRNAs are shown in red, the control sgRNAs are shown in dark gray, and the no sgRNA control is shown in light gray. Bottom and upper lines of boxplots show the first and third interquartile range and the middle line shows the median. Bar plots show mean and associated standard error.
ns, *, and *** denote $p > 0.05$, $p < 0.05$, and $p < 0.001$, respectively.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

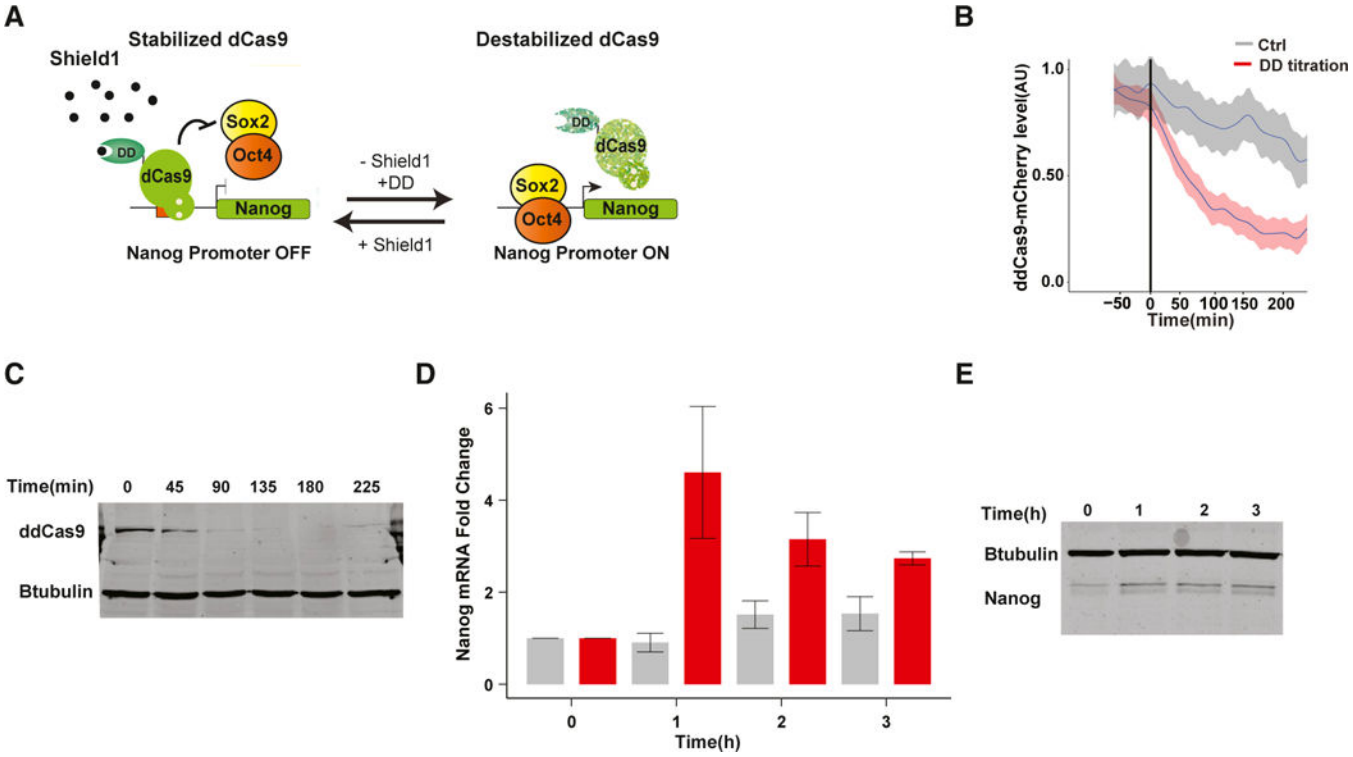


Figure 3. Reversible Disruption of TF-DNA Interactions with ddCas9

(A) Schematic of conditionally destabilized dCas9 (ddCas9) used for reversible inhibition.

(B and C) Time lapse microscopy (B) and immunoblot (C) of ddCas9 degradation after Shieldl titration by excess competitive inhibitor protein (DD).

(D and E) Nanog mRNAqRT-PCR (D) and immunoblot measurement (E) ofNanog protein afterddCas9 degradation by Shieldl titration. $p<0.05$forcomparison of Nanog mRNAfold change in (D).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

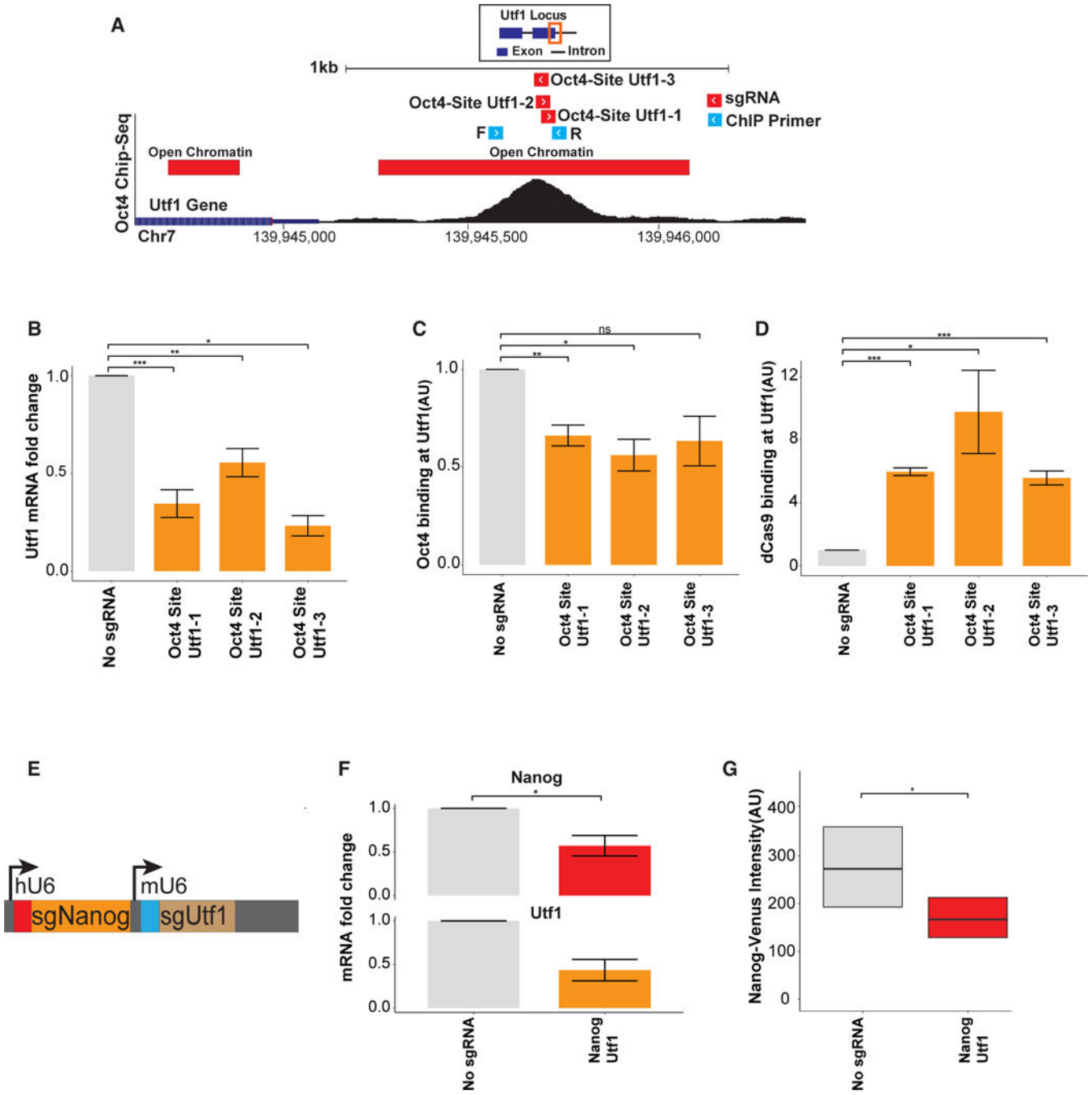


Figure4. Multiplexing CRISPRd

(A) Top: a diagram ofthe Utf1 locuswith the genomic region of interest shown below indicated by the orange box. Bottom: Oct4 ChlP-seq data near Utf1.Also shown are sgRNA (red) and qPCR primer sequences (blue). The red box denotes open chromatin identified by ENCODE DNase-seq (Experiment number: ENCSR000CMW).

(B-D) *Utf1* mRNA qRT-PCR measurement (B), Oct4 ChlP-qPCR (C), and dCas9 ChlP-qPCR (D) for cells expressing dCas9 and the indicated sgRNAs.

(E) Schematic of the construct used to express two sgRNAs.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(F) *Nanog* and *Utf1* mRNAqRT-PCR measurement in cells expressing dCas9 either alone or with the sgRNAs targeting the indicated Oct4 sites near *Nanog* and *Utf1* genes.
(G) Flow cytometry measurement of Nanog-Venus protein in control and Nanog Utf1 sgRNA-expressing cells.
Bottom and upper lines of boxplots show the first and third interquartile range and the middle line shows the median. Bar plots show mean and associated standard error. ns, *, and *** denote $p > 0.05$, $p < 0.05$, and $p < 0.001$, respectively.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

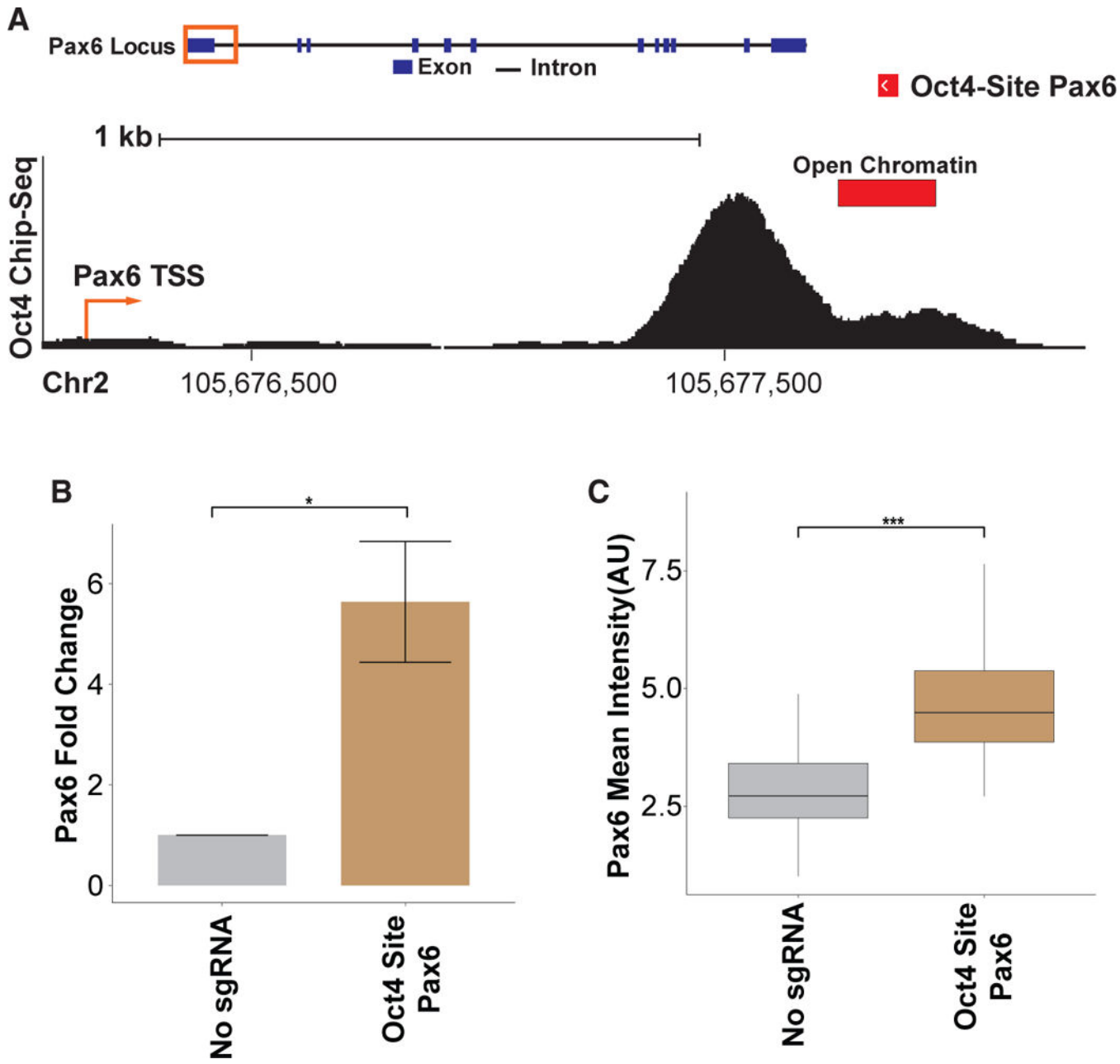


Figure 5. CRISPRd Disrupts Oct4 Inhibition of *Pax6* Expression

(A) Top: a diagram of the Pax6 locus with the genomic region shown below indicated by the orange box. Bottom: ChIP-seq data for Oct4 show the chromosomal position of an Oct4 peak in the regulatory region of the *Pax6* gene. Also shown is the position of the targeting sgRNA (red). The red box denotes open chromatin identified by ENCODE DNase-seq (Experiment number: ENCSR000CMW).

(B) qRT-PCR of *Pax6* mRNA in cells expressing the sgRNA targeting the nearby Oct4 site.

(C) Immunofluorescence staining of Pax6 protein.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Bottom and upper lines of boxplots show the first and third interquartile range and the middle line shows the median. Bar plots show mean and associated standard error. * and *** denote $p < 0.05$ and $p < 0.001$, respectively.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

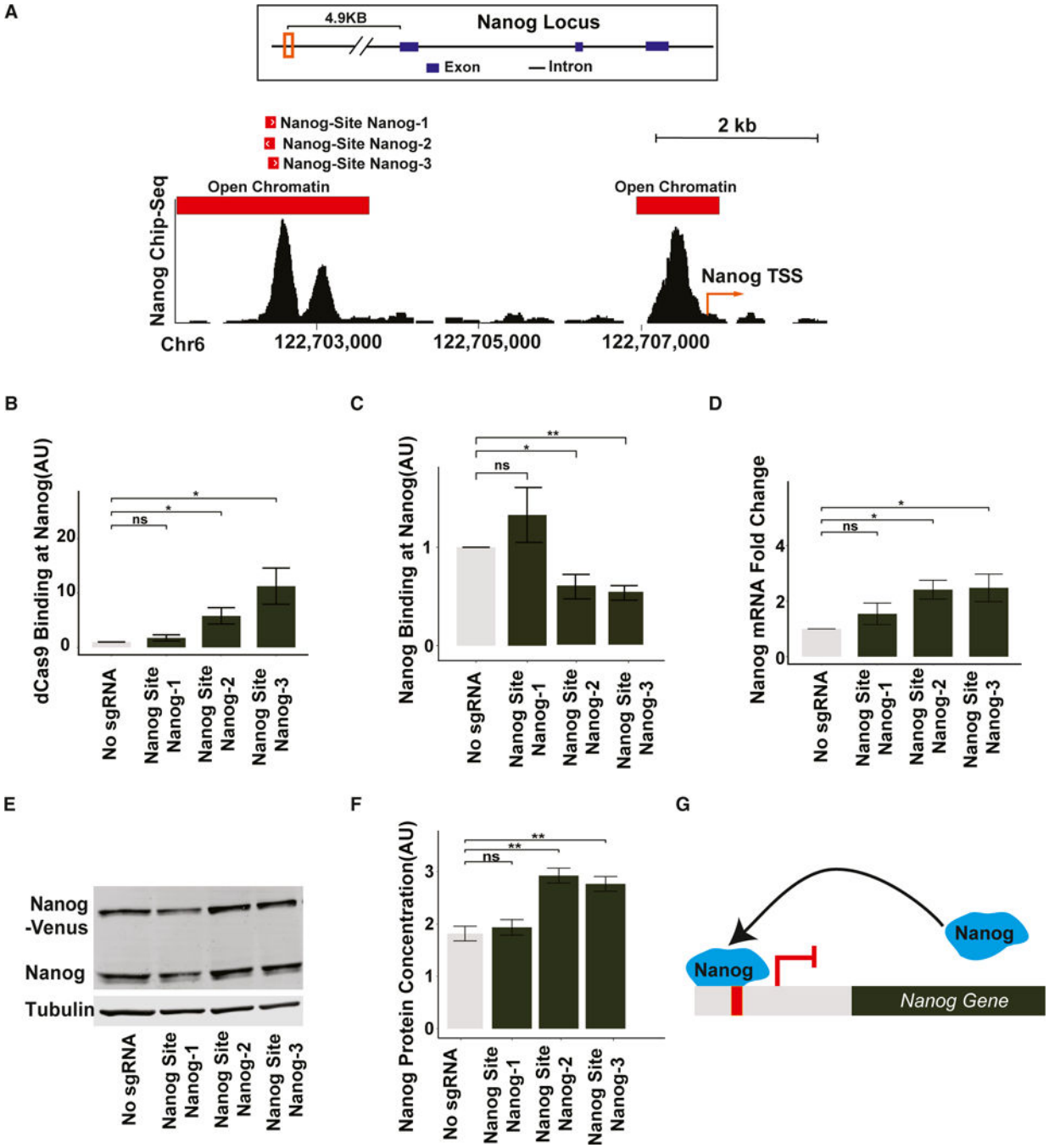


Figure 6. CRISPRd Disrupts Nanog DNA Binding and Identifies an Autoinhibitory Element in the Nanog Promoter

(A) Top: a diagram ofthe Nanog locus with the genomic region shown below indicated by the orange box. Bottom: ChIP-seq data for Nanog show a binding site located 4.9 kb upstream of its own transcription start site. Position of the targeting sgRNA shown in red. The red box denotes open chromatin identified by ENCODE DNase-seq (Experiment number: ENCSR000CMW).

(B) dCas9 ChIP-qPCR measurements for cells expressing dCas9, and the indicated sgRNA shows recruitment of dCas9 to the Nanog binding site.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(C) Nanog ChIP-qPCR shows that Nanog binding to its own promoter is decreased by 2 out of 3 targeting sgRNAs.
(D) qRT-PCR of *Nanog* in cells expressing the sgRNA targeting the Nanog binding site shows increased expression of Nanog mRNA.
(E)Immunoblot of Nanog in cells expressing the indicated sgRNA.
(F)Quantification of three independent Nanog immunoblots as in (E).
(G)Schematic representation of the negative feedback loop formed by Nanog binding to its own promoter.
Bar plots show mean and associated standard error. ns, *, and *** denote $p > 0.05$, $p < 0.05$, and $p < 0.001$, respectively.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

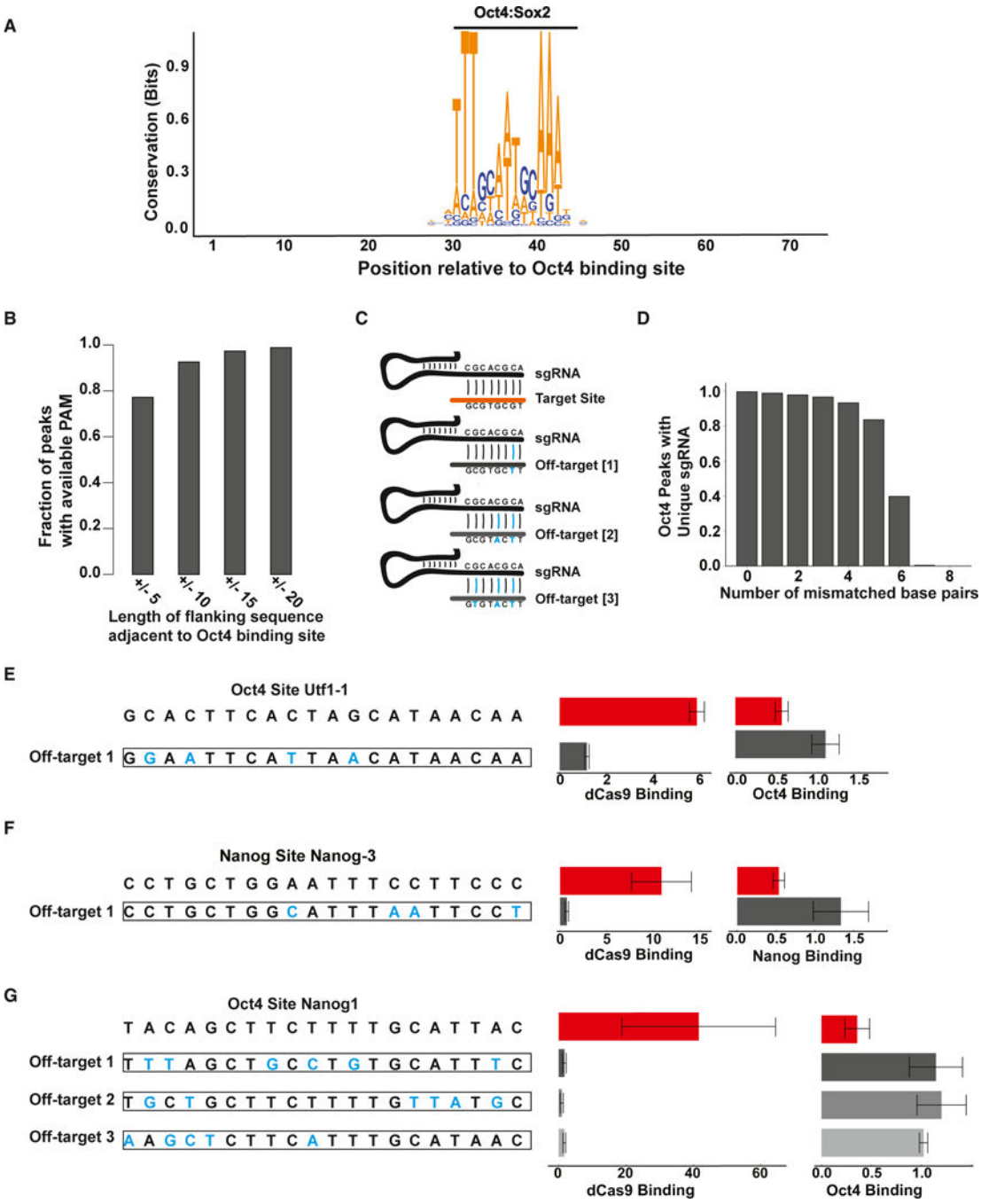


Figure 7. Computational and Experimental Analysis of CRISPRd Specificity

(A) The sequence logo of Oct4:Sox2 binding sites ± 30 bp flanking regions shows that the flanking regions are random. The height of base pairs signifies the degree of conservation for each position.

(B) Fraction of Oct4 binding sites that have a PAM within the indicated distance from the binding site.

(C) A schematic depiction of mismatches between pairs of DNA sequences and sgRNAs. The values inside the brackets show the mismatch number between sgRNA and target site.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(D) Fraction of Oct4 binding sites that can be targeted with at least with one sgRNA as a function of minimum mismatch number, i.e., the number of base pair changes required to match two sgRNAs. 35,173 sgRNAs for 6,942 Oct4 binding sites were analyzed.
(E) Oct4 and dCas9 ChIP-qPCR for an off-target Oct4 site with mismatch number of 4 normalized to dCas9-expressing cells.
(F) Nanog and dCas9 ChIP-qPCR for an off-target Nanog site with 4 mismatches normalized to dCas9-expressing cells.
(G) Oct4 and dCas9 ChIP-qPCR for three off-target Oct4 sites with 6 mismatches normalized to dCas9-expressing cells.
The on-site values for (E), (F), and (G) are replotted for comparison from previous figures for each indicated sgRNA.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1.

Comparison of CRISPRd with Previously Used CRISPR Technologies in Mammalian Cells

Method	Application	Mechanism	Device Example
CRISPRi	Repression of Transcription	Transcription block or chromatin modifications	dCas9-KRAB
CRISPRa	Activation of Transcription	Recruitment of transcription activators	dCas9-VPR
CRISPR/Cas9	Mutation in the regulatory sequences	Double strand break and repair	Cas9
CRISPRd	Disturbing TF-DNA interaction	Competition for specific TF target site and steric hindrance	dCas9