Stem Cell Reports
Article

ISSCR

OPEN ACCESS

TINC— A Method to Dissect Regulatory Complexes at Single-Locus Resolution— Reveals an Extensive Protein Complex at the *Nanog* Promoter

Anja S. Knaupp,[1,2,3,10] Monika Mohenska,[1,2,3,10] Michael R. Larcombe,[1,2,3,10] Ethan Ford,[4,5] Sue Mei Lim,[1,2,3] Kayla Wong,[1,2,3] Joseph Chen,[1,2,3] Jaber Firas,[1,2,3] Cheng Huang,[6] Xiaodong Liu,[1,2,3] Trung Nguyen,[4,5] Yu B.Y. Sun,[1,2,3] Melissa L. Holmes,[1,2,3] Pratibha Tripathi,[1,2] Jahnvi Pflueger,[4,5] Fernando J. Rossello,[1,2,3] Jan Schröder,[1,2,3] Kathryn C. Davidson,[1,2,3] Christian M. Nefzger,[1,2,3] Partha P. Das,[1,2] Jody J. Haigh,[7,8,9] Ryan Lister,[4,5] Ralf B. Schittenhelm,[6,*] and Jose M. Polo[1,2,3,*]

[1]Department of Anatomy and Developmental Biology, Monash University, Clayton, VIC 3800, Australia
[2]Development and Stem Cells Program, Monash Biomedicine Discovery Institute, Clayton, VIC 3800, Australia
[3]Australian Regenerative Medicine Institute, Monash University, Clayton, VIC 3800, Australia
[4]Australian Research Council Centre of Excellence in Plant Energy Biology, The University of Western Australia, Perth, WA 6009, Australia
[5]Harry Perkins Institute of Medical Research, Nedlands, WA 6009, Australia
[6]Monash Proteomics and Metabolomics Facility, Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia
[7]Australian Centre for Blood Diseases, Monash University, Clayton, VIC 3004, Australia
[8]Department of Pharmacology and Therapeutics, University of Manitoba, Winnipeg, MB, Canada
[9]Research Institute in Oncology and Hematology, CancerCare Manitoba, Winnipeg, MB, Canada
[10]These authors contributed equally
*Correspondence: ralf.schittenhelm@monash.edu (R.B.S.), jose.polo@monash.edu (J.M.P.)
https://doi.org/10.1016/j.stemcr.2020.11.005

SUMMARY

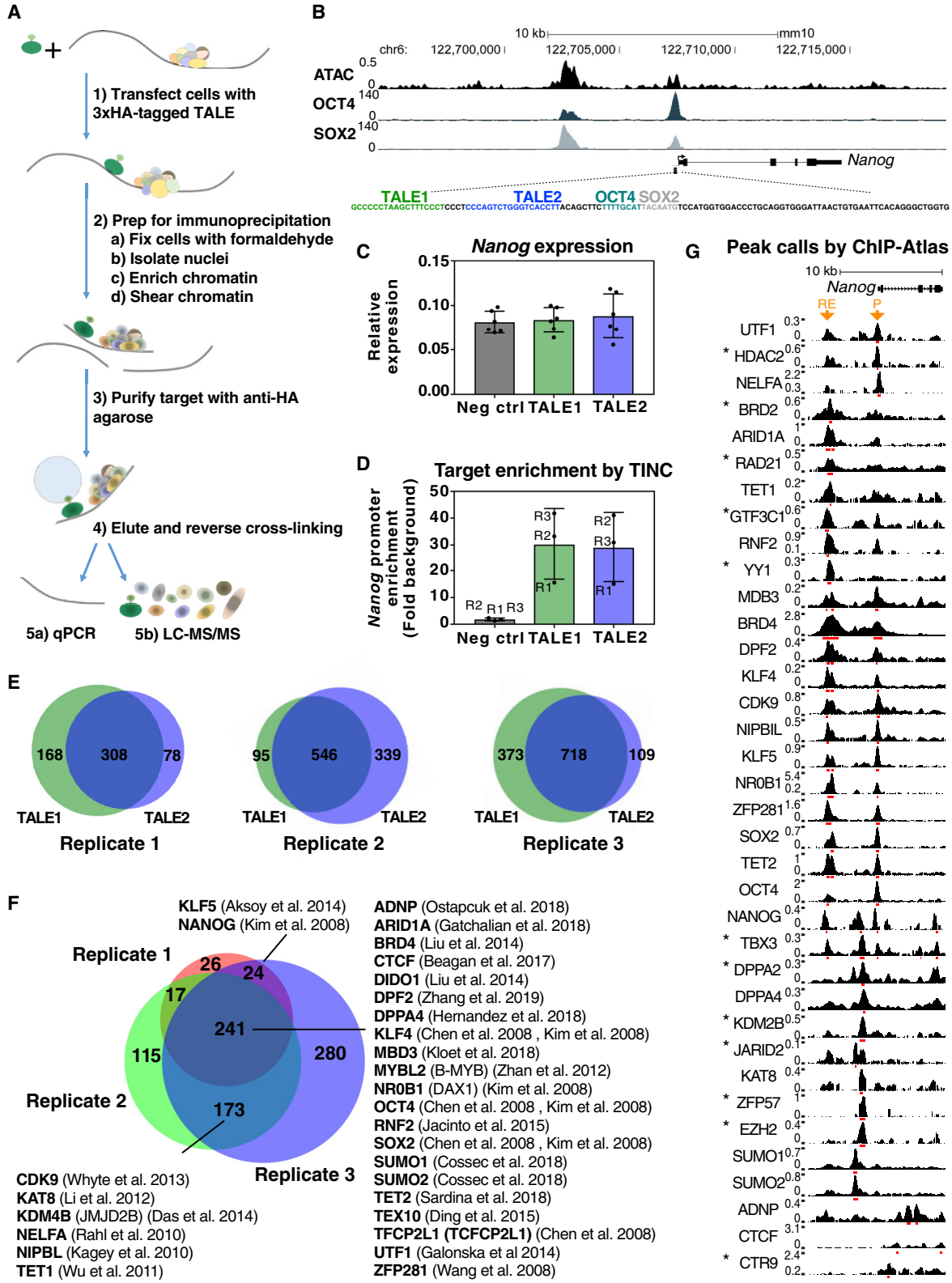
Cellular identity is ultimately dictated by the interaction of transcription factors with regulatory elements (REs) to control gene expression. Advances in epigenome profiling techniques have significantly increased our understanding of cell-specific utilization of REs. However, it remains difficult to dissect the majority of factors that interact with these REs due to the lack of appropriate techniques. Therefore, we developed TINC: TALE-mediated isolation of nuclear chromatin. Using this new method, we interrogated the protein complex formed at the *Nanog* promoter in embryonic stem cells (ESCs) and identified many known and previously unknown interactors, including RCOR2. Further interrogation of the role of RCOR2 in ESCs revealed its involvement in the repression of lineage genes and the fine-tuning of pluripotency genes. Consequently, using the *Nanog* promoter as a paradigm, we demonstrated the power of TINC to provide insight into the molecular makeup of specific transcriptional complexes at individual REs as well as into cellular identity control in general.

INTRODUCTION

Pluripotent stem cells (PSCs) carry immense therapeutic potential as they can produce any cell type of the body and can self-renew indefinitely. The two most common *in vitro* PSCs are blastocyst-derived embryonic stem cells (ESCs) (Kaufman et al., 1983; Martin 1981) and induced PSCs (iPSCs), which are obtained from somatic cells through expression of the transcription factors (TFs) OCT4, SOX2, KLF4, and C-MYC (Takahashi et al., 2007; Takahashi and Yamanaka 2006). Together with OCT4 and SOX2, NANOG forms the core transcriptional network in ESCs/iPSCs that mediates expression of self-renewal and pluripotency genes and repression of differentiation genes (Loh et al., 2006; Boyer et al., 2005; Chen et al., 2008; Marson et al., 2008; Kim et al., 2008). ESCs express *Nanog* heterogeneously and can be maintained upon *Nanog* deletion, although they are differentiation prone (Chambers et al., 2007). These NANOG fluctuations seem to play a role in lineage commitment with high levels impeding ESC differentiation (Abranches et al., 2014; Chambers et al., 2003; Kalmar et al., 2009).

Like any gene, *Nanog* expression is regulated by TFs that interact with regulatory elements (REs) and other factors. At least two REs control *Nanog* in ESCs: the promoter and the −5 kb enhancer (Apostolou et al., 2013; Kagey et al., 2010). NANOG binds to both of these REs (Chen et al., 2008; Kim et al., 2008) and mediates positive (Wu et al., 2006) and negative (Fidalgo et al., 2012; Navarro et al., 2012) feedback loops. Positive transcriptional regulation has been associated with binding of OCT4 and SOX2 to the *Nanog* promoter (Rodda et al., 2005). Conversely, *Nanog* auto-repression has been linked to the interaction of NANOG with ZFP281 and the NuRD repressor complex (Fidalgo et al., 2012). Nevertheless, while some of the proteins that occupy these REs are known, the full complex composition remains largely elusive.

This knowledge gap is partly due to the difficulty in dissecting a specific regulatory complex. Chromatin immunoprecipitation (ChIP) has been invaluable to study *in vivo* DNA-protein interactions but only allows interrogation of one factor at a time, and requires *a priori* candidates and appropriate antibodies. Thus, being able to simultaneously analyze an entire protein complex is extremely advantageous but is challenging due to (1) a genomic region of low abundance (e.g., two copies per cell) has to be targeted specifically and (2) the interacting proteins have to be



(legend on next page)

enriched efficiently (e.g., mass spectrometry [MS] analyzes proteins without amplification).

Several groups have set out to overcome these limitations by developing locus-specific isolation or proximity labeling methods. These include methods based on nucleic acid hybridization (Antão et al., 2012; Déjardin and Kingston 2009; Ide and Dejardin 2015; Kennedy-Darling et al., 2014) or DNA-binding proteins, such as LexA (Byrum et al., 2012; Fujita and Fujii 2011), TetR (Pourfarzad et al., 2013), Cas9 (Waldrip et al., 2014; Fujita and Fujii 2013; Gao et al., 2018; Schmidtmann et al., 2016; Tsui et al., 2018; Qiu et al., 2019; Liu et al., 2017), or TALEs (Byrum et al. 2013; Fang et al., 2018; Fujita et al., 2013). Although nucleic acid hybridization-based approaches pioneered the field, they require intensive optimization (Antão et al., 2012; Déjardin and Kingston 2009; Ide and Dejardin 2015), which might be why they have not yet been widely adopted or translated to single-copy elements in mammalian cells. To interrogate such challenging regions, different groups have adapted the CRISPR-Cas9 technology (Fujita and Fujii 2013; Qiu et al., 2019; X. Liu et al., 2020, 2017). Unlike TetR and LexA, CRISPR-Cas9 does not require insertion of binding sites into the target sequence and can be customized relatively easily. However, there are increasing reports of high-frequency off-targeting events by CRISPR-Cas9 (Fu et al., 2013; Duan et al., 2014; Kuscu et al., 2014; Lin et al., 2014; X. Wang et al., 2015; Liu et al., 2018), which in turn potentially skew the results of such single-locus studies. Importantly, TALEs are associated with significantly lower off-targeting than CRISPR-Cas9 (X. Wang et al., 2015). Furthermore, evidence suggests that TALEs can be utilized to enrich a specific locus from the complex mammalian genome (Fang et al., 2018).

Taking advantage of these characteristics, we developed a TALE-based method that allows isolation of a specific genomic region from mammalian cells termed TINC (TALE-mediated isolation of nuclear chromatin). Using TINC, we interrogated the regulatory complex formed at the *Nanog* promoter in mouse ESCs.

RESULTS

Development of TINC to Identify the *Nanog* Regulatory Complex

To interrogate the protein complex formed at the *Nanog* promoter in ESCs we developed and applied TINC (Figure 1A). In brief, cells expressing a 3xHA-tagged TALE designed to bind to a region of interest are fixed and the nuclei and chromatin are isolated before sonication. The target region is then isolated using affinity purification of the TALE, and nucleic acids and proteins are further processed. To minimize selection of proteins enriched through TALE off-targeting, we performed TINC with two different TALEs targeting the *Nanog* promoter and considered only proteins enriched by both TALEs as genuine binders. To that end, four TALEs were designed (Figure S1A) and tested by ChIP-qPCR to identify the most efficient TALEs (Figure S1B). Consequently, TALE1 and TALE2 were selected (Figure 1B). Importantly, ESCs stably expressing either TALE showed unaltered *Nanog* expression (Figures 1C and S1C) and retained their pluripotency potential (Figure S1D).

Next, TINC was performed with both TALEs in three independent experiments. As negative control, empty vector-transfected ESCs were used. The DNA isolated in each TINC reaction was analyzed by qPCR, which indicated comparable *Nanog* promoter enrichment efficiencies by both TALEs (approximately 30-fold over background) (Figure 1D). To determine the genome-wide binding sites of TALE1 and TALE2, two replicates were sequenced. This revealed that TALE1 and TALE2 had 12 and 3 off-targets, respectively (Figures S1E and S1F; Table S1), which is considerably less than what had been shown for comparable CRISPR-Cas9 approaches (Liu et al., 2020, 2017). Importantly, this analysis also confirmed that the only

Figure 1. TINC Allows Isolation of the *Nanog* Regulatory Complex from ESCs

(A) Schematic of the TINC method.

(B) Two TALEs were designed to bind upstream of the OCT4-SOX2 motif in the *Nanog* promoter. ATAC-seq shows that the promoter and the −5 kb *Nanog* enhancer are located in open chromatin, and ChIP-seq indicates that both regions are targeted by OCT4 and SOX2.

(C) qRT-PCR of *Nanog* transcript levels (fold *Gapdh*) (mean ± SD; n = 6 independent experiments).

(D) TINC-qPCR confirmed strong enrichment of the *Nanog* promoter over the genomic background (*Sox2* regulatory region 2) by TALE1 and TALE2 (mean ± SD; n = 3 independent experiments, e.g., R1, R2, and R3).

(E) Overlap of proteins isolated by TALE1 and TALE2 upon subtraction of proteins enriched unspecifically from the empty vector controls (n = 3 independent experiments).

(F) Overlap of the intercept of proteins identified by TALE1 and TALE2 in three TINC runs shown in Figure 1E. The 455 proteins present in at least two of the three replicates (i.e., TINC proteins) contained many of the previously published *Nanog* promoter binders.

(G) ChIP-seq data processed by ChIP-Atlas (Oki et al., 2018) shown at the *Nanog* promoter. Proteins identified as *Nanog* promoter binders by ChIP-Atlas but not the original publication are marked with an asterisk. Below each BigWig track (black) are the respective peak calls by the ChIP-Atlas (red). The promoter (P) and the −5 kb enhancer (RE) of *Nanog* are indicated in orange.

See also Figure S1 and Tables S1 and S2.

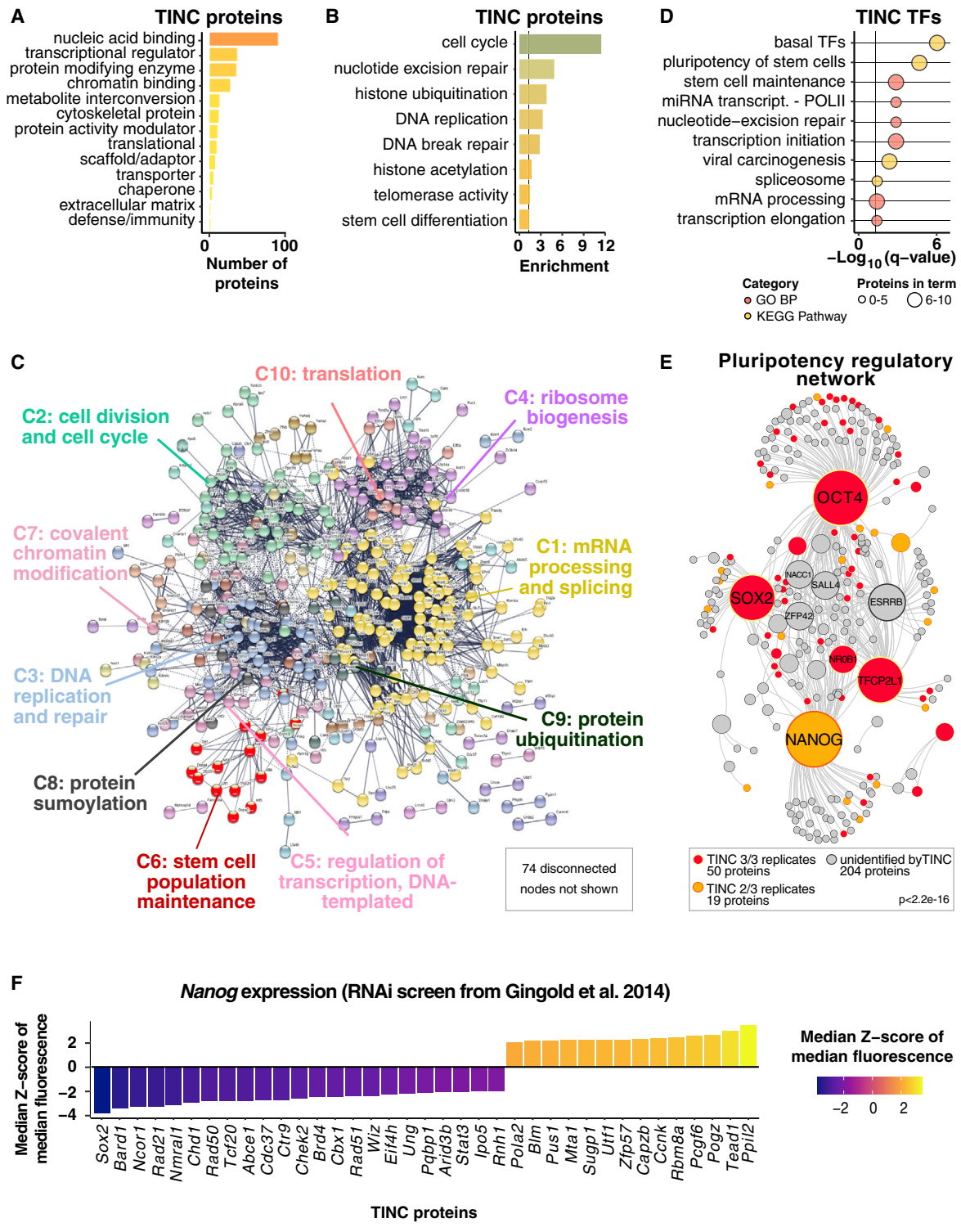


Figure 2. Characterization of the Proteins Identified at the *Nanog* Promoter in ESCs

(A) Protein class categories for all 455 TINC proteins. A total of 197 proteins were not assigned.

(B) Clustered GO biological processes and KEGG pathways enriched in the TINC protein dataset. Enrichment is the $-\log_{10}$(geometric mean) of p-values of clustered terms.

(legend continued on next page)

site co-bound by both TALEs is the *Nanog* promoter (Figure S1E). Hence, the protein content of each sample was analyzed by MS. Upon subtraction of proteins also detected in the negative controls, we obtained several hundred proteins from each sample and an overlap of approximately 60% between proteins enriched by TALE1 and TALE2 (Figure 1E). Further comparative analyses of the proteins revealed that 241 were present in all three and an additional 214 in two of the three replicates (Figure 1F; Table S1). Importantly, these 455 proteins, hereafter referred to as TINC proteins, contained 30 published binders (Figure 1F; Tables S1 and S2). In addition, ChIP-Atlas, which processes published ChIP sequencing (ChIP-seq) data via its own pipeline and can hence identify binding events not detected in the original publication (Oki et al., 2018), confirmed *Nanog* promoter targeting by an additional 12 TINC proteins (Figure 1G; Tables S1 and S2). Notably, TINC did not simply result in the detection of highly abundant chromatin-associated proteins, but also in the specific enrichment of lowly abundant candidates, which further supports the validity and sensitivity of TINC (Figure S1G). Altogether, this demonstrates that TINC enables interrogation of specific regulatory complexes and has the power to uncover hundreds of interacting proteins.

Analysis of the *Nanog* Regulatory Complex Reveals Proteins with Different Functions

Enrichment analyses of the TINC proteins showed overrepresentation of nucleic acid and chromatin binders, transcriptional regulators, and protein-modifying enzymes with roles in cell cycle, DNA replication and repair, histone modification, and stem cell regulation (Figures 2A and 2B). In accordance, protein network analysis identified protein clusters related to many of these processes (Figure 2C; Table S3). This shows that, while the cell cultures were not synchronized and therefore proteins associated with various stages of the cell cycle were detected, TINC identified a large number of proteins related to transcriptional regulation with a role in stem cell identity as expected from *Nanog*.

The TINC proteins included 56 TFs and 101 epigenetic modifiers (Figure S2A; Table S1). While the epigenetic modifiers were mainly associated with histone modifications (Figure S2B), the TFs showed enrichment for various processes related to transcriptional regulation and maintenance of pluripotency (Figure 2D). Intersection of the TINC proteins with the previously defined pluripotency regulatory network (Nefzger et al., 2017; Xu et al. 2013, 2014) revealed that 69 key network components form part of the *Nanog* regulatory complex (Figures 2E and S2C). We also identified ESRRB, SALL4, and NACC1 in our TALE pull-downs; however, these key regulators were also present in the negative controls (albeit with a lower number of peptide spectrum matches) and were therefore excluded from further analyses.

To determine the transcriptional output of the TINC proteins, we utilized a published RNAi screen conducted with a *Nanog*-GFP reporter ESC line (Gingold et al., 2014). Interestingly, while we identified 381 of the authors' small interfering RNA targets at the *Nanog* promoter, knockdown (KD) of only 23 resulted in a significant decrease in *Nanog* expression, indicative of a role of these proteins in maintenance of pluripotency (Figures 2F, S2D, and S2E). Conversely, KD of 14 TINC proteins resulted in a significant increase in *Nanog* expression, suggesting that these proteins act as repressors of *Nanog* and might play a role in silencing of this gene during differentiation (Figures 2F, S2D, and S2E). Together, these results show that various transcriptional regulators, including activators and repressors, form part of the *Nanog* regulatory complex in ESCs.

Different Requirements for *Nanog* Regulatory Complex Members in Pluripotency Acquisition and Maintenance

Cellular reprogramming is driven by reconfiguration of the epigenome, which leads to silencing of the somatic network and activation of the pluripotency program (Chronis et al., 2017; Knaupp et al., 2017; Maherali et al., 2007; Polo et al., 2012). To determine which TINC proteins are specific to pluripotency and hence upregulated during reprogramming of mouse embryonic fibroblasts (MEFs) into iPSCs, we performed differential gene expression (DGE) analysis. DGE analysis revealed that 61% of the TINC proteins are expressed significantly higher in iPSCs (versus 8% in MEFs) and that many of the known *Nanog* binders might not be pluripotency specific (Figure 3A).

Next, we selected proteins with a large change (e.g., RCOR2 and DNMT3L), an intermediate change (e.g.,

(C) Protein-protein interaction network analysis and MCL clustering of the TINC proteins. GO biological processes and KEGG pathways enriched in the top ten clusters are displayed.
(D) Enriched processes and pathways for TFs identified by TINC. Significance cutoff q = 0.05 (Benjamini-Hochberg correction) is shown (x axis).
(E) Overlap of the TINC proteins with the pluripotency regulatory network (Fisher's exact test, $<2.2 \times 10^{-16}$). Proteins identified in all three TINC replicates are depicted in red and proteins identified in two out of the three TINC replicates are shown in orange. The size of the node (protein) indicates the number of proteins interacting with that specific protein.
(F) KD of 37 TINC proteins resulted in a significant change in *Nanog* expression in a genome-wide RNAi screen (Gingold et al., 2014).
See also Figure S2 and Table S3.

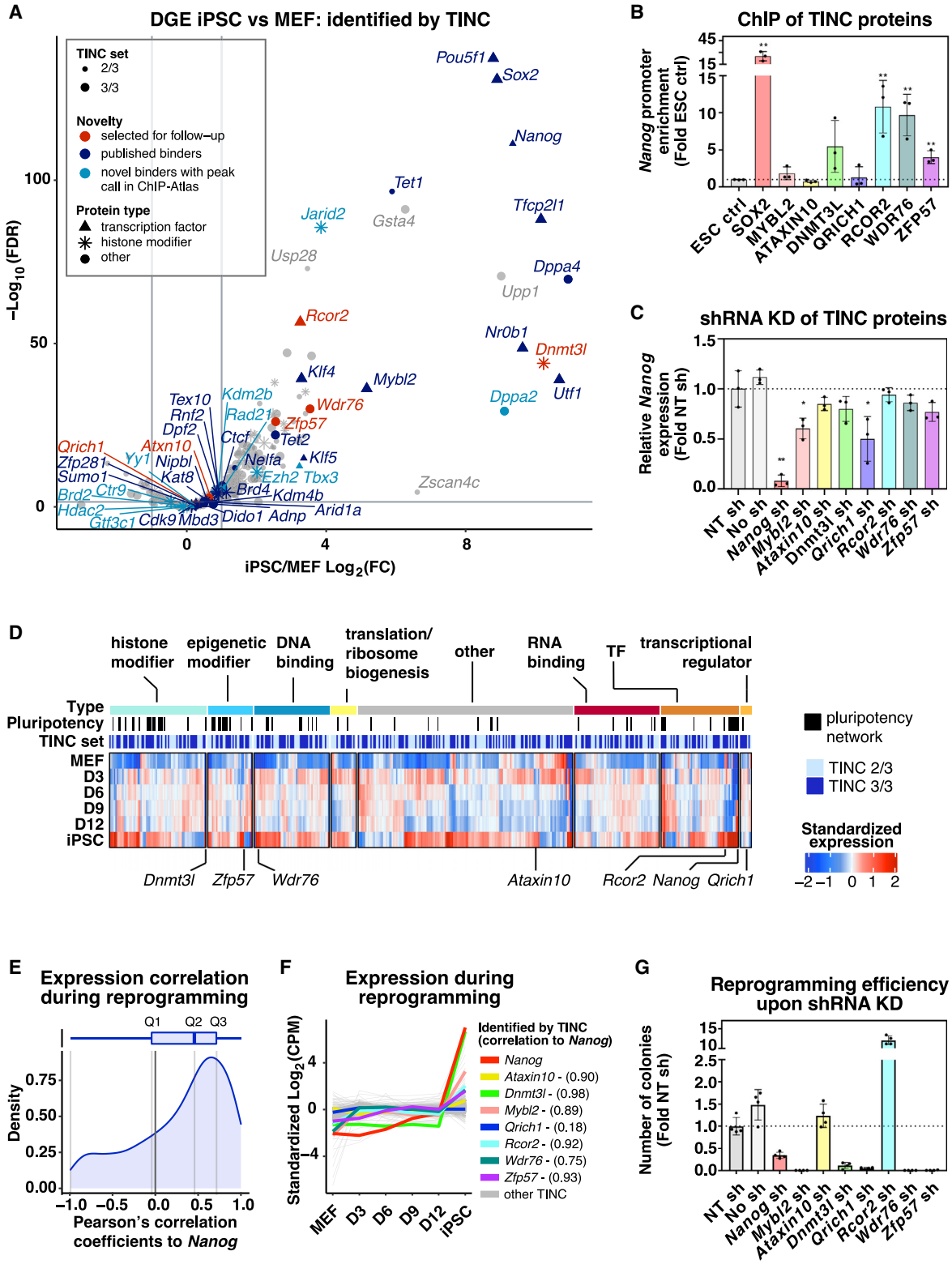


Figure 3. Interrogation of Novel *Nanog* Promoter Interactors Identified by TINC

(A) DGE analysis of TINC protein expression in iPSCs versus MEFs. Proteins selected for follow-up experiments are shown in red. Significance cutoff false discovery rate (FDR) of 0.05 is shown (y axis).

(legend continued on next page)

WDR76 and ZFP57), and no change (e.g., ATAXIN10 and QRICH1) in DGE for further investigation. To overcome the lack of ChIP antibodies, these six proteins as well as the positive controls MYBL2 (Zhan et al., 2012) and SOX2 were 3xHA tagged. This allowed us to confirm *Nanog* promoter binding by the majority of these candidates (Figure 3B), further validating TINC as well as endorsing ChIP-Atlas, which had identified ZFP57 as a *Nanog* binder (Figure 1G) while the original publication did not (Strogantsev et al., 2015). To determine whether these proteins directly control *Nanog* expression, we used small hairpin RNAs (shRNAs) for which we obtained approximately 75% target KD (Figures S3A and S3B). Depletion of several of these factors resulted in a change in ESC morphology, which was most pronounced for *Nanog* and *Rcor2* (Figure S3C). As shown previously, KD of *Nanog* led to a considerable amount of cell death (Chen et al.,2012), while loss of *Rcor2* triggered the cells to grow in a monolayer instead of dome-shaped colonies (Figure S3C). Despite the perceived morphological changes, only KD of *Mybl2* and *Qrich1* resulted in a significant change in *Nanog* expression (Figure 3C).

Analysis of TINC protein expression during MEF reprogramming revealed that many are transiently or permanently upregulated and that these proteins are mainly associated with chromatin organization and transcriptional regulation (Figures 3D, S3D, and S3E). Furthermore, 73% of the TINC proteins, including the majority of our candidates, are upregulated at the end of reprogramming and are strongly correlated with *Nanog* expression (Figures 3E, 3F, and S3D). Interestingly, while KD of *Mybl2*, *Dnmt3l*, *Qrich1*, *Wdr76*, and *Zfp57* resulted in a significant decrease in reprogramming efficiency, KD of *Rcor2* led to a significant increase (Figure 3G). Overall, these results suggest that many of the proteins identified at the *Nanog* promoter in the pluripotent state also play a role in establishing this state during reprogramming.

Since it has previously been proposed that *Rcor2* positively regulates MEF reprogramming (Yang et al., 2011), we set out to further explore its role during this conversion. Using a different reprogramming system (OKMS instead of OKSM) and two alternative *Rcor2* shRNAs individually and combined, we observed the same trend (Figures S3F and S3G). Flow cytometric analysis during reprogramming revealed that the total number of cells en route to becoming iPSCs (SSEA1+) approximately doubled due to an overall increase in cell numbers rather than an increase in the percentage of SSEA1+ cells (Figures S3H–S3J). Conversely, overexpression of *Rcor2* led to a decrease in iPSC formation, although non-significant, which indicates that RCOR2 may impair pluripotency acquisition (Figure S3K).

RCOR2 Is a Component of the Pluripotency Regulatory Network

To gain further insight into the targets and partners of RCOR2, we first performed ChIP-seq, which revealed extensive binding to intergenic and intronic regions in ESCs (Figure S4A; Table S4). Approximately 13% of the binding sites of RCOR2 are located in promoter regions, including the promoters of *Nanog*, *Oct4*, and *Sox2* (Figures 4A and S4A). Indeed, the majority of RCOR2 target genes are expressed (e.g., 8,514) and associated with various cell division- and pluripotency-related functions, while low or unexpressed targets (e.g., 3,030) are mainly associated with neuronal processes (Figure S4B). Intersection of the genome-wide binding sites and target genes of RCOR2 revealed extensive co-occupancy of RCOR2 with NANOG, OCT4, or SOX2 targets (Figures 4B and S4C; Table S4). A hypergeometric test validated that indeed RCOR2 co-occupancy with OCT4 ($p = 0$), SOX2 ($p = 1.01664 \times 10^{-310}$), or NANOG ($p = 3.96871 \times 10^{-246}$) targets is highly significant. Notably, genes occupied by all four factors have the highest median expression, include many key pluripotency regulators, and are enriched for various developmental processes (Figures 4B, 4C, S4D). In agreement, motif enrichment analysis showed several pluripotency TF motifs, including OCT4, SOX2, and NANOG in ESC enhancers (Figure 4D). In line with the observed binding of

(B) ChIP-qPCR analysis for *Nanog* promoter enrichment by 3xHA-tagged factors stably expressed in ESCs (mean ± SD; n = 3 independent experiments). $^{**}p \leq 0.01$ versus ESC control.

(C) qRT-PCR analysis of *Nanog* expression upon shRNA-mediated KD of selected factors and normalized to the levels of housekeeping gene *Ywhaz* (mean ± SD; n = 3 independent experiments). $^{*}p \leq 0.05$ and $^{**}p \leq 0.01$ versus non-targeting (NT) shRNA control.

(D) Heatmap showing standardized TINC protein expression during MEF reprogramming (grouped by protein categories). Black bars indicate that the proteins are a part of the pluripotency network from Figure 2E and blue bars whether they were identified in all three or in two out of the three TINC replicates.

(E) Distribution of Pearson's correlation coefficients of TINC protein expression in relation to the expression of *Nanog* during MEF reprogramming.

(F) Standardized gene expression dynamics of the TINC proteins during MEF reprogramming. Targets selected for further investigation are highlighted and their reprogramming correlation coefficients to *Nanog* are shown in brackets.

(G) MEF reprogramming efficiency upon shRNA-mediated KD of selected targets (mean ± SD; n = 4 technical replicates, representative of three independent experiments).

See also Figure S3.

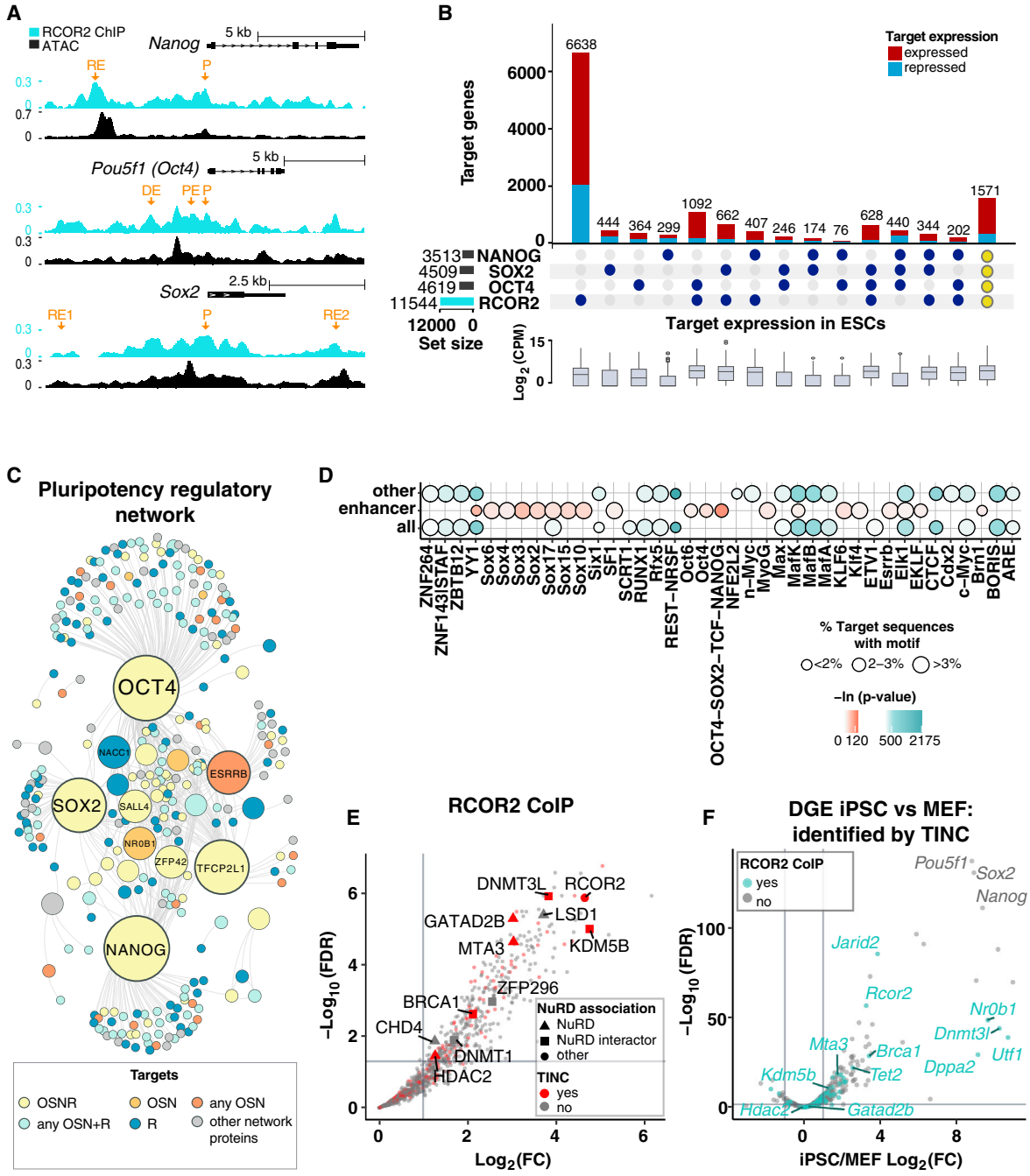


Figure 4. RCOR2 Is Part of the Pluripotency Regulatory Network in ESCs

(A) Tracks depicting RCOR2 ChIP-seq at *Nanog*, *Oct4*, and *Sox2*. Promoters (P) and known regulatory elements (REs), including the *Oct4* distal (DE) and proximal (PE) enhancers are indicated in orange. ATAC-seq shows regions of open chromatin.

(B) Intersection of NANOG, SOX2, OCT4, and RCOR2 target genes in ESCs. Expressed targets (red), repressed targets (blue), and the expression distributions are shown.

(C) Proteins of the pluripotency regulatory network are colored indicative of whether they are targets of NANOG (N), SOX2 (S), OCT4 (O), and/or RCOR2 (R). The size of the node (protein) indicates the number of proteins interacting with that specific protein.

(D) Motif enrichment analysis of RCOR2 ChIP-seq peaks. Where a point is present, a significant enrichment for the motif (x axis) was found at all sites, pluripotency enhancers, or at other sites (non-pluripotency enhancers) occupied by RCOR2 (y axis). Point size represents the proportion of sequences featuring the motif and color gradient the enrichment significance.

(legend continued on next page)

RCOR2 to neuronal genes (Figure S4B), we obtained motif enrichment for REST, a repressor associated with silencing of neuronal genes in non-neuronal cells (Chong et al., 1995; Schoenherr and Anderson 1995).

Next, to determine interactors of RCOR2, we conducted co-immunoprecipitations (CoIPs), which identified 368 proteins (Figures 4E and 4F; Table S4). Among those were 79 transcriptional regulators predominantly associated with histone modifications, including LSD1, which had previously been found to interact with RCOR2 in pluripotent and neural cells (Yang et al., 2011; Wang et al., 2016) (Figure S4E; Table S4). Upon close examination of the TINC results, we noticed that LSD1 had also been excluded due to traces in the negative controls. LSD1 forms part of various complexes, including CoREST/REST and NuRD (Andrés et al., 1999; Foster et al., 2010; Mosammaparast and Shi 2010). Notably, while RCOR2 CoIP resulted in enrichment of various NuRD complex components, CoREST and REST were not detected (Figures 4E; Table S4). Furthermore, integration of published ChIP-seq data confirmed extensive co-binding of RCOR2 and LSD1 within ESCs (Figures S4F and S4G; Table S4). Interestingly, while some of these sites are also bound by REST and CoREST or, to a larger extent by NuRD complex components, many sites are exclusively targeted by LSD1 and RCOR2 (Figures S4F and S4G). Notably, ESC enhancers did not show major occupancy by CoREST or REST but extensive targeting by LSD1, RCOR2, and NuRD members (Figure S4H). Inhibition of LSD1 has previously been shown to promote iPSC formation (Sun et al., 2016). Importantly, concomitant inhibition of LSD1 alleviated the inhibitory effect of RCOR2 overexpression, suggesting that RCOR2 exerts its function at least partly via LSD1 during reprogramming (Figure S3K). Together, our results revealed that RCOR2 binds extensively in ESCs and forms part of various regulatory complexes, including, LSD1, NuRD, and CoREST/REST complexes.

RCOR2 Is Required for Efficient ESC Differentiation

To further investigate the role of RCOR2, we created a CRISPR-Cas9 knockout (KO) ESC line (Figure 5A). *Rcor2* depletion was confirmed at the DNA, RNA, and protein levels (Figures S5A–S5C). Notably, *Rcor2* KO resulted in similar morphological changes as observed upon shRNA-mediated KD (Figures S3C and 5A). Subsequent RNA sequencing of the *Rcor2* KO line revealed transcriptional deregulation of 1,977 genes, approximately 74% of which are RCOR2 targets (Figure 5B; Table S4). Interestingly, many RCOR2 targets, which showed transcriptional deregulation upon *Rcor2* KO were related to cell cycle and cell division (Figure S5D). In agreement, growth rate analyses revealed a significant decrease in ESC and MEF doubling time upon *Rcor2* depletion, suggesting that RCOR2 is a negative regulator of cell-cycle progression (Figure S5E). Furthermore, this analysis revealed that RCOR2 is involved in transcriptional regulation of various pluripotency network components, including repression of *Oct4* and *Nanog* (Figures 5C and S5F).

Next, to assess the possible role of RCOR2 during differentiation, we subjected the KO and wild-type (WT) ESC lines to an embryoid body (EB) formation assay. *Rcor2* KO EBs showed a considerable amount of cell death and were significantly smaller than WT EBs (Figures 5D and 5E). Furthermore, differentiation into cardiac cells revealed that WT EBs generated beating colonies, while *Rcor2* KO EBs did not (Figures 5F; Video S1). Importantly, qRT-PCR revealed persistent expression of pluripotency genes in the *Rcor2* KO EBs (Figure 5G). In agreement with the EB assay, teratoma assays showed limited growth for *Rcor2* KO ESCs *in vivo*. Although *Rcor2* KO ESCs were able to eventually form teratomas with cell types from all three germ layers, it took an additional 5 days to reach a palpable size (Figures S5G–S5H). Together, these data suggest that RCOR2 plays a role in downregulating pluripotency genes during differentiation.

DISCUSSION

We developed TINC, which allows isolation of a specific genomic locus and applied it to interrogate the protein complex formed at the *Nanog* promoter in ESCs. To reduce the number of false positives, two TALEs were utilized and only proteins enriched by both were considered as true binders. Although our experimental setup allowed us to confirm the reproducibility of TINC, it prevented the usage of label-free quantitative software packages (such as MaxQuant [Cox and Mann 2008]). In future, such quantitative approaches could be used to decrease the amount of false negatives (e.g., despite higher peptide spectrum matches in the TALE samples LSD1, ESRRB, SALL4, and NACC1 had been excluded due to trace levels in the negative controls).

(E) Volcano plot showing the enrichment of proteins copurified by RCOR2 in comparison with the negative control (n = 3 technical replicates). Lines indicate the threshold above which proteins are significantly enriched (FDR < 0.05 and $\log_2$FC > 1). TINC proteins are shown in red and NuRD complex components and interactors are indicated.

(F) DGE analysis of TINC protein expression in iPSCs versus MEFs. Proteins also detected in RCOR2 CoIPs are colored in turquoise. Significance cutoff FDR = 0.05 is shown (y axis).

See also Figure S4 and Table S4.

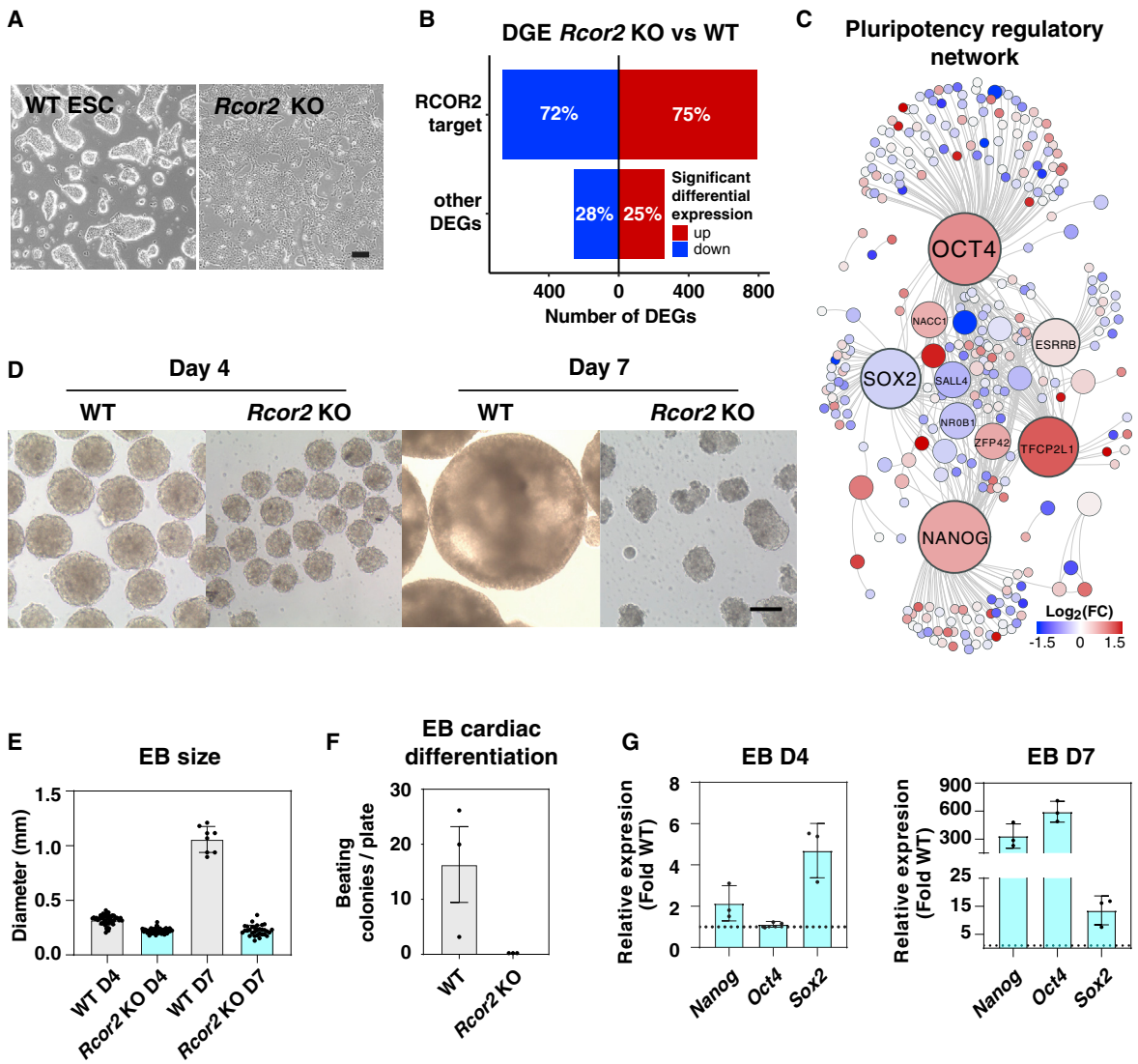


Figure 5. *Rcor2* KO ESCs Have a Differentiation Impairment

(A) Representative bright-field images of WT and *Rcor2* KO ESCs. Scale bar, 250 μm.

(B) DGE analysis of *Rcor2* KO versus WT ESCs separated into genes occupied by RCOR2 and genes that are not RCOR2 targets (y axis). Red indicates genes that show an increase and blue indicates genes that show a decrease in expression upon *Rcor2* KO (n = 2 independent experiments).

(C) Proteins of the pluripotency regulatory network colored according to their change in expression upon *Rcor2* KO.

(D) Representative bright-field images of WT and *Rcor2* KO EBs on days 4 and 7 of culture. Scale bar, 250 μm.

(E) EB sizes on day 4 (D4) and day 7 (D7) of culture as measured by diameter (mean ± SD; n ≥ 8 EBs, representative of two independent experiments).

(F) Frequency of contractile cardiac colonies obtained in an EB cardiac differentiation assay (mean ± SD; n = 3 technical replicates).

(G) qRT-PCR analysis to examine pluripotency marker expression in *Rcor2* KO EBs. Transcript levels were normalized to the levels of the housekeeping gene *Ywhaz* and then to corresponding WT EBs (mean ± SD; n = 3 technical replicates, representative of 2 independent experiments).

See also Figure S5 and Table S4 and Video S1.

TINC allowed us to identify 455 proteins at the *Nanog* promoter in ESCs (i.e., TINC proteins). While only a fraction of these proteins were known binders, many had previously been associated with a role in pluripotency. For example, TINC revealed direct *Nanog* promoter targeting by JMJD1A, RCOR2, and ZFP57, all of which had been

linked to a change in *Nanog* expression upon KD (Loh et al., 2007; Riso et al., 2016; Yang et al., 2011). Many of the TINC proteins are upregulated during reprogramming and KD of several of them impaired iPSC formation, suggesting that they play a role in establishing the pluripotent state. Conversely, KD of *Rcor2* resulted in a significant increase in colony numbers. Further investigation revealed that *Rcor2* depletion leads to increased *Nanog* levels, a decrease in MEF/ESC doubling time, and increased cell numbers; all processes that facilitate reprogramming. Furthermore, our data suggest that RCOR2 mediates its function at least partially through an interaction with LSD1, inhibition of which also promotes cellular reprogramming (Sun et al., 2016). We validated our result in different reprogramming systems; however, since previous work observed RCOR2 as a positive regulator of cellular reprogramming (Yang et al., 2011), its mechanisms might be expression level and/or reprogramming system dependent. Further studies may need to be performed to address these differences in observations.

Interestingly, RCOR2 seems to be non-essential for ESC maintenance despite occupying many expressed genes. Indeed, approximately 70% of its targets are transcriptionally active and are associated with maintenance of pluripotency, while inactive RCOR2 targets are enriched for various neuron-related functions. Similarly, REST has been implicated in the repression of neuronal genes in non-neuronal cells (Chong et al., 1995; Schoenherr and Anderson 1995). Importantly, our data revealed that RCOR2 forms part of the REST complex in ESCs and occupies approximately 40% of REST binding sites. While many of these regions are also targeted by LSD1, they account for less than 5% of RCOR2 or LSD1 binding sites. Notably, RCOR2 and LSD1 have a similar number of targets (approximately 12,000 genes), over 70% of which are shared. This suggests that, while in ESCs RCOR2 forms part of the REST complex, it also interacts with alternative LSD1 complexes, including NuRD. Previous work has shown that LSD1 is significantly more likely to occupy pluripotency enhancers with NuRD complex components than the CoREST/REST complex (Whyte et al., 2018) and our data revealed that RCOR2 does so too. This suggests that RCOR2 may be involved in the fine-tuning of pluripotency genes and the repression of lineage-specific genes in ESCs as part of various LSD1 complexes.

Similar to RCOR2, LSD1 does not play a major role in ESC maintenance; however, LSD1 inhibition leads to incomplete repression of many ESC genes during differentiation (Whyte et al., 2018). Similarly, *Rcor2* KO ESCs were characterized by an inefficiency to form EBs, which in turn showed prolonged expression of pluripotency genes, as well as poor differentiation potential *in vitro* and *in vivo*. This further supports a functional interaction of LSD1 and RCOR2 during pluripotency exit.

In summary, TINC allowed us to interrogate the regulatory complex formed at the *Nanog* promoter in an unbiased manner and revealed that transcriptional regulation of this TF occurs at many different levels. Furthermore, our data suggest that many factors that aid in downregulating *Nanog* during differentiation, such as RCOR2, already reside at its promoter in the pluripotent state. Together, this implies a highly complex and coordinated interplay of multiple factors that ensures the correct NANOG levels so that ESCs can self-renew and exert their full differentiation potential, rapidly and on-cue.

EXPERIMENTAL PROCEDURES

Generation of TALE-Expressing ESC Lines

The *Nanog* targeting TALEs were created as described (Briggs et al., 2012) and stable ESC lines were generated as described in the Supplemental Information.

TINC

Approximately 1×10^9 cells were fixed with formaldehyde, and nuclei and chromatin were isolated as described previously (Knaupp et al., 2017; Kustatscher et al., 2014). The chromatin was then sonicated using a Bioruptor NextGen device (Diagenode) and the TALEs immunoprecipitated using Anti-HA Agarose Resin (Pierce 26182) as described in the Supplemental Information. *Nanog* promoter enrichment was confirmed by qPCR before liquid chromatography-tandem MS (LC-MS/MS) analysis.

LC-MS/MS Analysis of Proteins Enriched by TINC

To determine the proteins isolated by TINC, in-gel tryptic digests were analyzed using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) as described in the Supplemental Information.

All other experimental procedures are detailed in Supplemental Information.

Data and Code Availability

The accession numbers for the data reported in this paper are available in GEO: GSE160816 and ProteomeXchange: PXD022088 (Perez-Riverol et al., 2019).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.stemcr.2020.11.005.

AUTHOR CONTRIBUTIONS

J.M.P. conceptualized the study. A.S.K. and J.M.P. conceived and designed the experiments. R.B.S. and A.S.K. designed the proteomics approach. A.S.K. and M.R.L. performed the experiments with support from S.M.L., K.W., J.C., J.F., X.L., Y.B.Y.S., M.L.H., J.P., P.T., and K.C.D. M.M. performed the bioinformatic analyses

with assistance from F.J.R. and J.S. under supervision of J.M.P. TALEs were created by E.F. with support from T.N. and J.J.H. LC-MS/MS analyses were performed by R.B.S. with support from C.H. R.L., P.P.D., C.M.N., and K.C.D. helped with experimental design and manuscript editing. A.S.K., M.M., M.R.L., R.B.S., and J.M.P. wrote the manuscript and all authors approved of the final manuscript.

CONFLICTS OF INTERESTS

J.M.P. is a founder and member of the SAB of Mogrify; however, the work presented in this manuscript is not related to this company. All other authors have no conflict of interest or financial interest to declare.

ACKNOWLEDGMENTS

This work was supported by a National Health and Medical Research Council (NHMRC) grant (APP1069830) to R.L. and J.M.P. J.M.P. and R.L. were supported by the Australian Research Council (ARC) Stem Cells Australia Special Initiative, an NHMRC CDF (APP1036587 to J.M.P.), an ARC Future Fellowship (FT120100862 to R.L. and FT180100674 to J.M.P.), and a Silvia and Charles Viertel Senior Medical Research Fellowship. R.L. was also supported by the ARC Center of Excellence program in Plant Energy Biology (CE140100008), HHMI International Research Scholarship, and NHMRC Investigator grant GNT1178460. A.S.K. was supported by an NHMRC ECF (APP1092280). We thank Flowcore, the MHTP Medical Genomics Facility, Micromon, the Monash Bioinformatics and Histology Platforms. This study used BPA-enabled/NCRIS-enabled infrastructure located at the Monash Proteomics and Metabolomics Facility led by R.B.S. The Australian Regenerative Medicine Institute is supported by grants from the State Government of Victoria and the Australian Government. The PiggyBac vector and the OKMS MEFs were kindly provided by Andras Nagy.

Received: June 25, 2020
Revised: November 8, 2020
Accepted: November 9, 2020
Published: December 8, 2020

REFERENCES

Abranches, E., Guedes, A.M.V., Moravec, M., Maamar, H., Svoboda, P., Raj, A., and Henrique, D. (2014). Stochastic NANOG fluctuations allow mouse embryonic stem cells to explore pluripotency. Development *141*, 2770–2779.

Andrés, M.E., Burger, C., Peral-Rubio, M.J., Battaglioli, E., Anderson, M.E., Grimes, J., Dallman, J., Ballas, N., and Mandel, G. (1999). CoREST: a functional corepressor required for regulation of neural-specific gene expression. Proc. Natl. Acad. Sci. U S A *96*, 9873–9878.

Antão, J.M., Mason, J.M., Déjardin, J., and Kingston, R.E. (2012). Protein landscape at *Drosophila melanogaster* telomere-associated sequence repeats. Mol. Cell Biol. *32*, 2170–2182.

Apostolou, E., Ferrari, F., Walsh, R.M., Bar-Nur, O., Stadtfeld, M., Cheloufi, S., Stuart, H.T., Polo, J.M., Ohsumi, T.K., Borowsky, M.L., et al. (2013). Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. Cell Stem Cell *12*, 699–712.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Briggs, A.W., Rios, X., Raj, C., Yang, L., Zhang, F., Mali, P., George, M., and Church. (2012). Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res. *40*, e117.

Byrum, S.D., Raman, A., Taverna, S.D., and Tackett, A.J. (2012). ChAP-MS: a method for identification of proteins and histone posttranslational modifications at a single genomic locus. Cell Rep. https://doi.org/10.1016/j.celrep.2012.06.019.

Byrum, S.D., Taverna, S.D., and Tackett, A.J. (2013). Purification of a specific native genomic locus for proteomic analysis. Nucleic Acids Res. *41*, e195.

Chambers, I., Douglas, C., 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.

Chambers, I., Jose, S., Douglas, C., Nichols, J., Nijmeijer, B., Robertson, M., Jan, V., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230–1234.

Chen, T., Du, J., and Lu, G. (2012). Cell growth arrest and apoptosis induced by Oct4 or Nanog knockdown in mouse embryonic stem cells: a possible role of Trp53. Mol. Biol. Rep. *39*, 1855–1861.

Chen, X., Han, X., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106–1117.

Chong, J.A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J.J., Zheng, Y., Boutros, M.C., Altshuller, Y.M., Frohman, M.A., Kraner, S.D., and Mandel, G. (1995). REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. Cell https://doi.org/10.1016/0092-8674(95)90298-8.

Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Shan, S., Ernst, J., and Plath, K. (2017). Cooperative binding of transcription factors orchestrates reprogramming. Cell *168*, 442–459.e20.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. https://doi.org/10.1038/nbt.1511.

Déjardin, J., and Kingston, R.E. (2009). Purification of proteins associated with specific genomic loci. Cell https://doi.org/10.1016/j.cell.2008.11.045.

Duan, J., Lu, G., Xie, Z., Lou, M., Luo, J., Guo, L.i., and Zhang, Y. (2014). Genome-wide identification of CRISPR/Cas9 off-targets in human genome. Cell Res. *24*, 1009–1012.

Fang, F., Xia, N., Angulo, B., Carey, J., Cady, Z., Durruthy-Durruthy, J., Bennett, T., Sebastiano, V., Renee, A., and Pera, R. (2018). A distinct isoform of ZNF207 controls self-renewal and pluripotency of human embryonic stem cells. Nat. Commun. *9*, 4384.

Fidalgo, M., Faiola, F., Pereira, C., Ding, J., Saunders, A., Gingold, J., Schaniel, C., Lemischka, I.R., Silva, J.C.R., and Wang, J. (2012). Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming. Proc. Natl. Acad. Sci. U S A *109*, 16202–16207.

Foster, C.T., Dovey, O.M., Lezina, L., Luo, J.L., Gant, T.W., Barlev, N., Bradley, A., and Cowley, S.M. (2010). Lysine-specific demethylase 1 regulates the embryonic transcriptome and CoREST stability. Mol. Cell Biol. *30*, 4851–4863.

Fujita, T., Asano, Y., Ohtsuka, J., Takada, Y., Saito, K., Ohki, R., and Fujii, H. (2013). Identification of telomere-associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP). Sci. Rep. *3*, 3171.

Fujita, T., and Fujii, H. (2011). Direct identification of insulator components by insertional chromatin immunoprecipitation. PLoS One *6*, e26109.

Fujita, T., and Fujii, H. (2013). Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. Biochem. Biophys. Res. Commun. *439*, 132–136.

Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Keith Joung, J., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. *31*, 822–826.

Gao, X.D., Tu, L.C., Mir, A., Rodriguez, T., Ding, Y., Leszyk, J., Dekker, J., Shaffer, S.A., Zhu, L.J., Wolfe, S.A., and Sontheimer, E.J. (2018). "C-BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9-APEX2. Nat. Methods https://doi.org/10.1038/s41592-018-0006-2.

Gingold, J.A., Fidalgo, M., Guallar, D., Lau, Z., Sun, Z., Zhou, H., Faiola, F., Huang, X., Lee, D.F., Waghray, A., et al. (2014). A genome-wide RNAi screen identifies opposing functions of Snai1 and Snai2 on the Nanog dependency in reprogramming. Mol. Cell *56*, 140–152.

Ide, S., and Dejardin, J. (2015). End-targeting proteomics of isolated chromatin segments of a mammalian ribosomal RNA gene promoter. Nat. Commun. *6*, 6674.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., Nynke, L., van Berkum, Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430–435.

Kalmar, T., Lim, C., Hayward, P., Muñoz-Descalzo, S., Nichols, J., Garcia-Ojalvo, J., and Arias, A.M. (2009). Regulated fluctuations in Nanog expression mediate cell fate decisions in embryonic stem cells. PLoS Biol. *7*, e1000149.

Kaufman, M.H., Robertson, E.J., Handyside, A.H., and Evans, M.J. (1983). Establishment of pluripotential cell lines from haploid mouse embryos. J. Embryol. Exp. Morphol. *73*, 249–261.

Kennedy-Darling, J., Guillen-Ahlers, H., Shortreed, M.R., Scalf, M., Frey, B.L., Kendziorski, C., Olivier, M., Gasch, A.P., and Smith, L.M. (2014). Discovery of chromatin-associated proteins via sequence-specific capture and mass spectrometric protein identification in *Saccharomyces cerevisiae*. J. Proteome Res. *13*, 3810–3825.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049–1061.

Knaupp, A.S., Buckberry, S., Pflueger, J., Lim, S.M., Ford, E., Larcombe, M.R., Rossello, F.J., de Mendoza, A., Alaei, S., Firas, J., et al. (2017). Transient and permanent reconfiguration of chromatin and transcription factor occupancy drive reprogramming. Cell Stem Cell *21*, 834–845.e6.

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.

Kustatscher, G., Wills, K.L.H., Furlan, C., and Rappsilber, J. (2014). Chromatin enrichment for proteomics. Nat. Protoc. *9*, 2090–2099.

Lin, Y., Cradick, T.J., Brown, M.T., Deshmukh, H., Ranjan, P., Sarode, N., Wile, B.M., Vertino, P.M., Stewart, F.J., and Bao, G. (2014). CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. Nucleic Acids Res. https://doi.org/10.1093/nar/gku402.

Liu, X., Chen, Y., Zhang, Y., Liu, Y., Liu, N., Botten, G.A., Cao, H., Orkin, S.H., Zhang, M.Q., and Xu, J. (2020). Multiplexed capture of spatial configuration and temporal dynamics of locus-specific 3D chromatin by biotinylated dCas9. Genome Biol. *21*, 59.

Liu, X.S., Wu, H., Krzisch, M., Wu, X., Graef, J., Muffat, J., Hnisz, D., Li, C.H., Yuan, B., Xu, C., et al. (2018). Rescue of fragile X syndrome neurons by DNA methylation editing of the FMR1 gene. Cell *172*, 979–992.e6.

Liu, X., Zhang, Y., Chen, Y., Li, M., Zhou, F., Li, K., Cao, H., Ni, M., Liu, Y., Gu, Z., et al. (2017). In situ capture of chromatin interactions by biotinylated dCas9. Cell *170*, 1028–1043.e19.

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., 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.

Loh, Y.H., Zhang, W., Chen, X., George, J., and Ng, H.H. (2007). Jmjd1a and Jmjd2c histone H3 lys 9 demethylases regulate self-renewal in embryonic stem cells. Genes Dev. *21*, 2545–2557.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., et al. (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell *1*, 55–70.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Tobias, B., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., et al. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521–533.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. U S A *78*, 7634–7638.

Mosammaparast, N., and Shi, Y. (2010). Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. Annu. Rev. Biochem. *79*, 155–179.

Navarro, P., Nicola, F., Douglas, C., Gagliardi, A., Mullin, N.P., Zhang, W., Karwacki-Neisius, V., 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.

Nefzger, C.M., Rossello, F.J., Chen, J., Liu, X., Knaupp, A.S., Firas, J., Paynter, J., Buckberry, S., Lim, S.M., Williams, B., et al. (2017). Cell type of origin dictates the route to pluripotency. Cell Rep. *21*, 2649–2660.

Oki, S., Ohta, T., Go, S., Hatanaka, H., Ogasawara, O., Okuda, Y., Kawaji, H., Nakaki, R., Sese, J., and Meno, C. (2018). ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. EMBO Rep. *19*. https://www.embopress.org/doi/abs/10.15252/embr.201846255.

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Suresh, H., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. *47*, D442–D450.

Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J., et al. (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. Cell *151*, 1617–1632.

Pourfarzad, F., Ali, A., de Boer, E., Ten Have, S., van Dijk, T.B., Kheradmandkia, S., Stadhouders, R., Thongjuea, S., Solar, E., Gillemans, N., et al. (2013). Locus-specific proteomics by TChP: targeted chromatin purification. Cell Rep. https://doi.org/10.1016/j.celrep.2013.07.004.

Qiu, W., Xu, Z., Zhang, M., Zhang, D., Fan, H., Li, T., Wang, Q., Liu, P., Zhu, Z., Du, D., et al. (2019). Determination of local chromatin interactions using a combined CRISPR and peroxidase APEX2 system. Nucleic Acids Res. *47*, e52.

Riso, V., Cammisa, M., Kukreja, H., Anvar, Z., Verde, G., Sparago, A., Acurzio, B., Lad, S., Lonardo, E., Sankar, A., et al. (2016). ZFP57 maintains the parent-of-origin-specific expression of the imprinted genes and differentially affects non-imprinted targets in mouse embryonic stem cells. Nucleic Acids Res. *44*, 8165–8178.

Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H., Wang, B., Ng, H.H., and Robson, P. (2005). Transcriptional regulation of Nanog by OCT4 and SOX2. J. Biol. Chem. *280*, 24731–24737.

Schmidtmann, E., Anton, T., Rombaut, P., Herzog, F., and Leonhardt, H. (2016). Determination of local chromatin composition by CasID. Nucleus *7*, 476–484.

Schoenherr, C.J., and Anderson, D.J. (1995). The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. Science *267*, 1360–1363.

Strogantsev, R., Krueger, F., Yamazawa, K., Shi, H., Gould, P., Goldman-Roberts, M., McEwen, K., Sun, B., Pedersen, R., Anne, C., and Ferguson-Smith. (2015). Allele-specific binding of ZFP57 in the epigenetic regulation of imprinted and non-imprinted monoallelic expression. Genome Biol. *16*, 112.

Sun, H., Liang, L., Yuan, L., Feng, C., Li, L., Zhang, Y., He, S., Pei, D., Guo, Y., and Zheng, H. (2016). Lysine-specific histone demethylase 1 inhibition promotes reprogramming by facilitating the expression of exogenous transcriptional factors and metabolic switch. Sci.Rep. *6*, 30903.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Tsui, C., Inouye, C., Levy, M., Lu, A., Florens, L., Washburn, M.P., and Tjian, R. (2018). dCas9-targeted locus-specific protein isolation method identifies histone gene regulators. Proc. Natl. Acad. Sci. U S A *115*, E2734–E2741.

Waldrip, Z.J., Byrum, S.D., Storey, A.J., Gao, J., Byrd, A.K., Mackintosh, S.G., Wahls, W.P., Taverna, S.D., Raney, K.D., and Tackett, A.J. (2014). A CRISPR-based approach for proteomic analysis of a single genomic locus. Epigenetics *9*, 1207–1211.

Wang, X., Wang, Y., Wu, X., Wang, J., Wang, Y., Qiu, Z., Chang, T., Huang, H., Lin, R.J., and Yee, J.K. (2015). Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. Nat. Biotechnol. *33*, 175–178.

Wang, Y., Wu, Q., Yang, P., Wang, C., Liu, J., Ding, W., Liu, W., Bai, Y., Yang, Y., Wang, H., et al. (2016). LSD1 Co-repressor Rcor2 orchestrates neurogenesis in the developing mouse brain. Nat. Commun. *7*, 10481.

Whyte, W.A., Bilodeau, S., Orlando, D.A., Hoke, H.A., Frampton, G.M., Foster, C.T., Cowley, S.M., and Young, R.A. (2018). Author correction: enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature *562*, E24.

Wu, Q., Chen, X., Zhang, J., Loh, Y.H., Low, T.Y., Zhang, W., Zhang, W., Sze, S.K., Lim, B., and Ng, H.H. (2006). Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. J. Biol. Chem. *281*, 24090–24094.

Xu, H., Ang, Y.S., Sevilla, A., Lemischka, I.R., and Ma'ayan, A. (2014). Construction and validation of a regulatory network for pluripotency and self-renewal of mouse embryonic stem cells. PLoS Comput. Biol. *10*, e1003777.

Xu, H., Baroukh, C., Dannenfelser, R., Chen, E.Y., Tan, C.M., Yan, K., Kim, Y.E., Lemischka, I.R., and Ma'ayan, A. (2013). ESCAPE: database for integrating high-content published data collected from human and mouse embryonic stem cells. Database (Oxford) *2013*, bat045.

Yang, P., Wang, Y., Chen, J., Hong, L., Lan, K., Zhang, Y., Chen, S., Zhu, B., and Gao, S. (2011). RCOR2 is a subunit of the LSD1 complex that regulates ESC property and substitutes for SOX2 in reprogramming somatic cells to pluripotency. Stem Cells *29*, 791–801.

Zhan, M., Riordon, D.R., Yan, B., Tarasova, Y.S., Bruweleit, S., Tarasov, K.V., Li, R.A., Wersto, R.P., and Boheler, K.R. (2012). The B-MYB transcriptional network guides cell cycle progression and fate decisions to sustain self-renewal and the identity of pluripotent stem cells. PLoS One *7*, e42350.