A Serial shRNA Screen for Roadblocks to Reprogramming Identifies the Protein Modifier SUMO2

Book Chapter

Author(s):
Borkent, Marti; Bennett, Brian D.; Lackford, Brad; Bar-Nur, Ori; Brumbaugh, Justin; Wang, Li; Du, Ying; Fargo, David C.; Apostolou, Effie; Cheloufi, Sihem; Maherali, Nimet; Elledge, Stephen J.; Hu, Guang; Hochedlinger, Konrad

Publication date:
2016-05-16

Permanent link:
https://doi.org/10.3929/ethz-b-000314651

Originally published in:
Stem Cell Reports 6(5), https://doi.org/10.1016/j.stemcr.2016.02.004

Stem Cell Reports
Article

OPEN ACCESS

A Serial shRNA Screen for Roadblocks to Reprogramming Identifies the Protein Modifier SUMO2

Marti Borkent,[1,2,3,4,8] Brian D. Bennett,[5,8] Brad Lackford,[6] Ori Bar-Nur,[1,2,3,4] Justin Brumbaugh,[1,2,3,4] Li Wang,[6] Ying Du,[5] David C. Fargo,[5] Effie Apostolou,[1,2,3,4] Sihem Cheloufi,[1,2,3,4] Nimet Maherali,[1,2,3,4] Stephen J. Elledge,[4,7] Guang Hu,[6,*] and Konrad Hochedlinger[1,2,3,4,*]

[1]Department of Molecular Biology, Center for Regenerative Medicine and Cancer Center, Massachusetts General Hospital, Boston, MA 02114, USA
[2]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA
[3]Harvard Stem Cell Institute, Cambridge, MA 02138, USA
[4]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA
[5]Integrative Bioinformatics, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA
[6]Epigenetics and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA
[7]Department of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA
[8]Co-first author
*Correspondence: hug4@niehs.nih.gov (G.H.), khochedlinger@mgh.harvard.edu (K.H.)
http://dx.doi.org/10.1016/j.stemcr.2016.02.004

SUMMARY

The generation of induced pluripotent stem cells (iPSCs) from differentiated cells following forced expression of OCT4, KLF4, SOX2, and C-MYC (OKSM) is slow and inefficient, suggesting that transcription factors have to overcome somatic barriers that resist cell fate change. Here, we performed an unbiased serial shRNA enrichment screen to identify potent repressors of somatic cell reprogramming into iPSCs. This effort uncovered the protein modifier SUMO2 as one of the strongest roadblocks to iPSC formation. Depletion of SUMO2 both enhances and accelerates reprogramming, yielding transgene-independent, chimera-competent iPSCs after as little as 38 hr of OKSM expression. We further show that the SUMO2 pathway acts independently of exogenous C-MYC expression and in parallel with small-molecule enhancers of reprogramming. Importantly, suppression of SUMO2 also promotes the generation of human iPSCs. Together, our results reveal sumoylation as a crucial post-transcriptional mechanism that resists the acquisition of pluripotency from fibroblasts using defined factors.

INTRODUCTION

The reprogramming of somatic cells into pluripotent cells using the classical set of transcription factors, OCT4, KLF4, SOX2, and C-MYC (OKSM) and conventional culture conditions (leukemia inhibitory factor, serum) usually takes several weeks and yields induced pluripotent stem cells (iPSCs) at extremely low frequencies (0.1%–3%) (Takahashi and Yamanaka, 2006). This observation suggests that reprogramming factors need to overcome undefined barriers that have been established by somatic cells to preserve cell identity and resist cell fate change. Identifying roadblocks to iPSC generation thus provides a valuable platform to dissect general principles of cell identity and cell fate change (Apostolou and Hochedlinger, 2013).

Previously identified barriers to reprogramming include regulators of cell cycle progression and senescence (e.g., P53, INK4A/ARF) (Krizhanovsky and Lowe, 2009), histone and DNA modifications (e.g., DNMT1, KDM2B, MBD3) (Mikkelsen et al., 2008; Rais et al., 2013; Wang et al., 2011), as well as signaling pathways and epigenetic processes that can be targeted by small compounds (e.g., ascorbic acid, GSK3 inhibitor, DOT1L inhibitor) (Bar-Nur et al., 2014; Esteban et al., 2010; Onder et al., 2012; Silva et al., 2008). However, suppression of some of these barriers may enhance iPSC formation only under specific culture conditions (e.g., MBD3) (dos Santos et al., 2014; Rais et al., 2013), potentially limiting its usefulness in different cellular contexts. Moreover, manipulation of certain barriers causes permanent aberrations of the epigenome (e.g., DNMT1) (Jackson-Grusby et al., 2001), complicating its applications in a therapeutic setting.

More recently, unbiased small hairpin RNA (shRNA) screens have been performed during iPSC formation, leading to the identification of novel roadblocks to reprogramming (Qin et al., 2014; Samavarchi-Tehrani et al., 2010; Yang et al., 2014). Surprisingly, individual suppression of hits that emerged from these screens showed rather modest effects (2- to 4-fold enhancement) compared with the simultaneous suppression of multiple hits (5- to 10-fold enhancement). Furthermore, there was little overlap among independent screening efforts, suggesting that reprogramming may be restrained by additional, yet to be identified barriers. Indeed, our lab recently discovered the histone chaperone CAF-1 as a novel barrier to iPSC generation using a chromatin-focused shRNA screen (Cheloufi et al., 2015).

The goal of this study was to identify potent roadblocks to reprogramming by performing a serial genome-wide shRNA enrichment screen in combination with a well-defined transgenic reprogramming system. Our screening

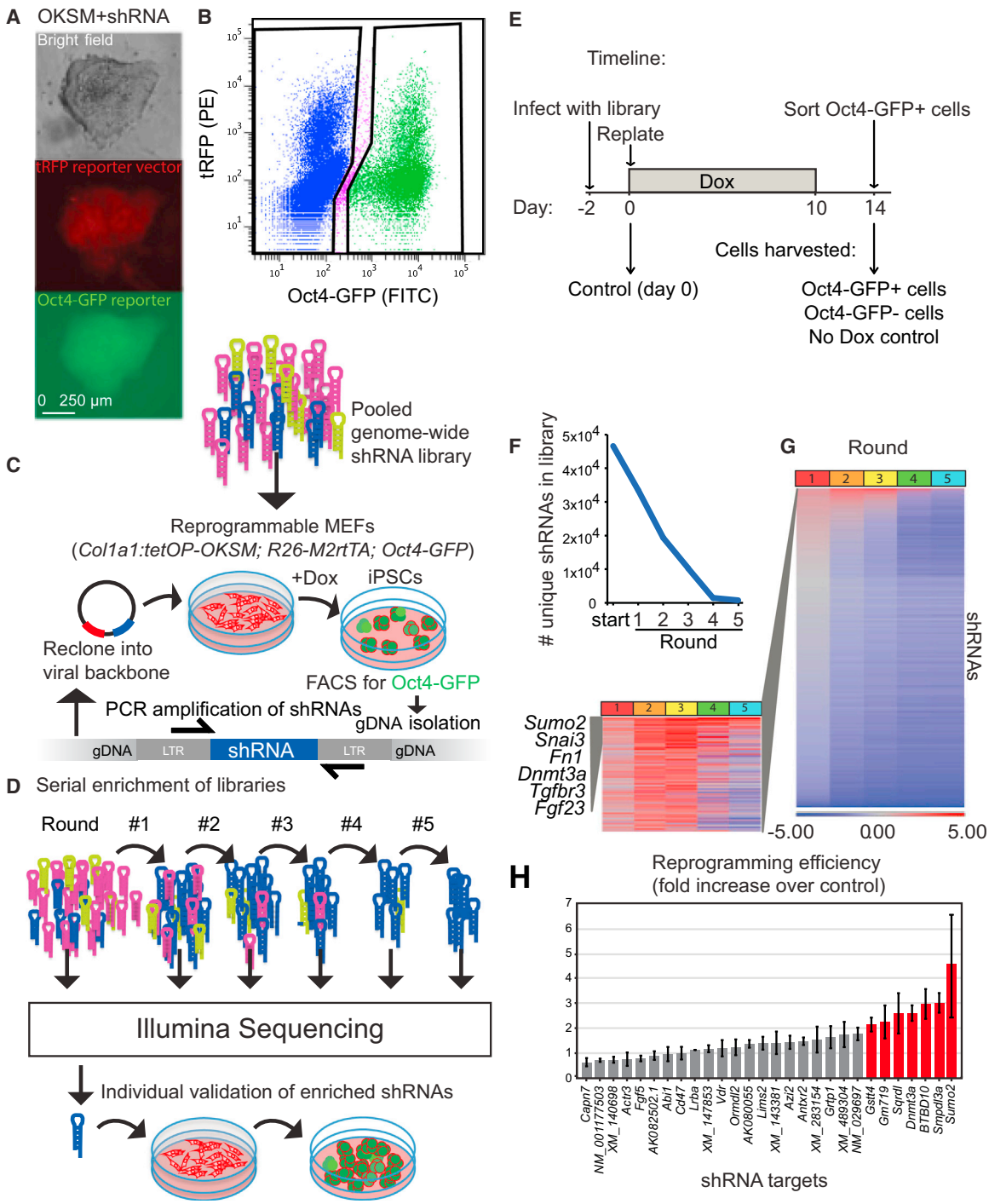

Figure 1. A Genome-Wide Serial shRNA Enrichment Screen during iPSC Generation

(A) Fluorescence microscopy image of a primary iPSC colony showing lentiviral tRFP (shRNA) expression and endogenous Oct4-GFP expression.

(B) Gating strategy to purify Oct4-GFP^{+} cells from lentivirally transduced cultures undergoing reprogramming.

(C) Schematic representation of one reprogramming/shRNA enrichment cycle.

(D) Overview of serial enrichment screen and validation experiments.

(E) Timeline of reprogramming experiments and strategy to collect control and experimental samples for subsequent analysis of shRNA library representation.

(F) Change in shRNA library complexity during enrichment screen, i.e., number of unique shRNAs at the start of rounds 1–5.

(legend continued on next page)

strategy uncovered SUMO2 as a top-scoring hit, thus implicating protein sumoylation as a mechanism that effectively resists transcription factor-induced pluripotency.

RESULTS

Serial shRNA Screen for Roadblocks to Reprogramming

To identify roadblocks to iPSC formation in an unbiased manner, we combined a well-defined transgenic reprogramming system with a genome-wide shRNA library targeting 18,464 genes with 60,642 hairpins. We utilized murine embryonic fibroblasts (MEFs) carrying a doxycycline (dox)-inducible polycistronic cassette encompassing the open reading frames for *Oct4*, *Klf4*, *Sox2*, and *c-Myc* (*OKSM*) in the *Col1a1* locus, the M2-rtTA transactivator in the *Rosa26* locus, and an EGFP reporter in the endogenous *Pou5f1* (*Oct4*) locus (Stadtfeld et al., 2010). We will refer to these transgenic MEFs as "reprogrammable cells" and the genotype as "*Col1a1-tetOP-OKSM*; *R26-M2rtTA*; *Oct4-GFP*." The shRNA library was generated by cloning shRNAs into the *pHAGE-Mir* lentiviral vector carrying a puromycin resistance gene and a turbo red fluorescent protein (tRFP) reporter (Meerbrey et al., 2011; Schlabach et al., 2008) (see Experimental Procedures for details). Transduction of reprogrammable MEFs with an identical empty vector gave rise to Oct4-GFP$^+$, tRFP$^+$ iPSC colonies upon exposure to dox, albeit at slightly lower frequencies than uninfected cells (Figures 1A and 1B; data not shown), demonstrating the feasibility of an shRNA screen using these cells and vector system.

To identify shRNAs that strongly enhance reprogramming with low background signal from passenger shRNAs, we devised a pooled screening strategy using serial enrichment of hairpin libraries. In brief, we infected reprogrammable MEFs with the pooled shRNA library 2 days before dox induction to ensure effective suppression of targets prior to initiation of reprogramming. After 10 days of OKSM expression, dox was withdrawn for 4 days to select for stably reprogrammed, transgene-independent colonies, followed by purification of emerging Oct4-GFP$^+$ cells by flow cytometry (Figure 1C). Enriched hairpins were amplified by PCR from genomic DNA and subsequently recloned into the original viral backbone before initiating another round of viral transduction and reprogramming. We performed five rounds of reprogramming and shRNA enrichment before focusing on candidates (Figures 1D and 1E). Parallel cultures of reprogrammable MEFs were exposed to dox alone or transduced with the viral library in the absence of dox before extracting genomic DNA (Figures 1E and S1); these samples served as controls for possible passenger hairpins that became passively enriched in expanding iPSC colonies or hairpins that merely affected the growth of uninduced reprogrammable MEFs. Library representation was then determined in all samples by deep (Illumina) sequencing of genomic DNA. Notably, we detected a gradual reduction of library complexity and a progressive accumulation of specific shRNA vectors during the five rounds of screening, suggesting enrichment of biologically meaningful hairpins (Figures 1F and 1G).

SUMO2 Emerges as a Top-Scoring Candidate Barrier to Reprogramming

We next applied stringent criteria to call hits based on the number of normalized reads and the cumulative fold change of shRNAs across all rounds of reprogramming relative to controls (Figure S1). Analysis of shRNAs that were consistently enriched revealed several candidate barriers to reprogramming including FGF5, SMPDL3A, and SUMO2 (Figures 1F–1H). Of note, some of the top-scoring shRNAs affected pathways that were previously shown to block reprogramming, including members of the FGF/FGFR and OLFR families (Dejosez et al., 2013; Qin et al., 2014). A complete list of candidates with associated cumulative enrichment scores is shown in Table S1. Of the 27 validated shRNAs, seven showed a more than 2-fold increase in iPSC formation (*Gstt4*, *Gm719*, *Sqrdl*, *Dnmt3a*, *BTBD10*, *Smpdl3a*, *Sumo2*) (Figure 1H), with *Sumo2* shRNA exhibiting the strongest phenotype.

Given the prominent effects on reprogramming of shRNAs targeting *Sumo2*, we will focus on this gene for the remainder of this article. SUMO2 (small ubiquitin-like modifier 2) is a ubiquitin-related protein that can be covalently attached to proteins as a lysine-linked monomer or polymer. This post-translational modification, lysine sumoylation, controls the stability, activity, and localization of hundreds of proteins and has been implicated in a number of biological processes including DNA replication and repair, heterochromatic gene silencing, and signal transduction (Flotho and Melchior, 2013).

Suppression of SUMO2 Promotes Reprogramming without Compromising Proliferation or Pluripotency

Using more quantitative reprogramming assays, we found that suppression of SUMO2 increases the number of transgene-independent alkaline phosphatase-positive (AP$^+$)

(G) Heatmap depicting fold-change enrichment of shRNAs during five rounds of reprogramming. Blue bars represent lost shRNAs whereas red bars represent enriched shRNA relative to controls (see Experimental Procedures for details).
(H) Validation of candidates identified in serial shRNA screen (mean value from three biological replicates ± SD). Red bars depict shRNAs yielding a 2-fold or higher increase in reprogramming efficiency.

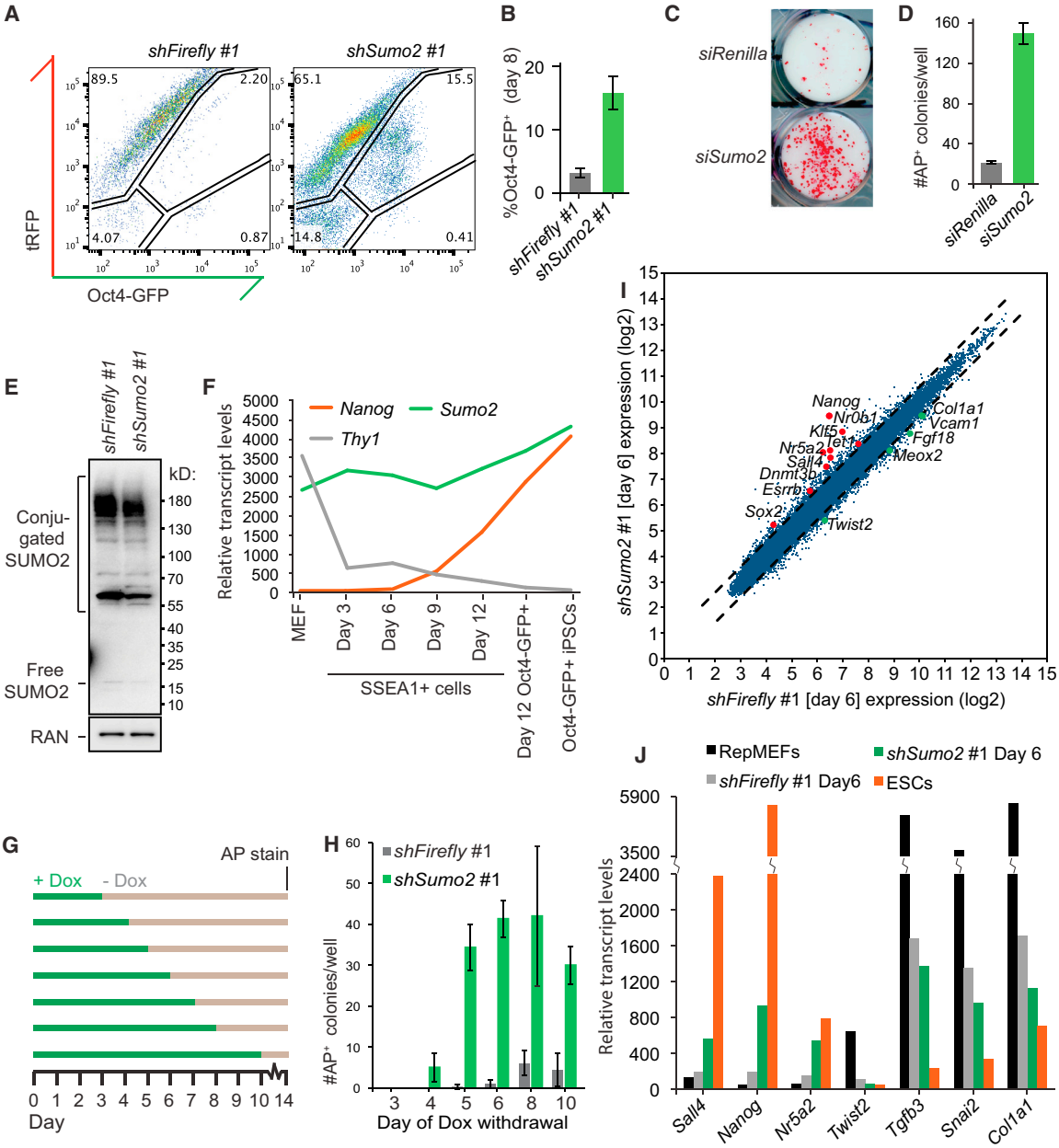

Figure 2. Suppression of SUMO2 Robustly Enhances and Accelerates Reprogramming

(A) Flow cytometric analysis of Oct4-GFP expression in reprogrammable MEFs after 8 days of OKSM expression in the presence of indicated shRNAs.

(B) Quantification of data shown in (A); shown is percentage of Oct4-GFP$^+$ cells per total number of cells using three biological replicates (mean ± SD).

(C) Alkaline phosphatase (AP) staining of iPSC colonies derived from reprogrammable MEFs transfected with indicated siRNAs.

(D) Quantification of data shown in (C); data represent mean from three biological replicates ± SD.

(E) Western blot analysis for SUMO2 expression in reprogrammable MEFs infected with indicated shRNA vectors and treated with doxycycline (dox) for 3 days.

(F) Expression dynamics of *Sumo2* mRNA in MEFs, iPSCs, and intermediate stages of reprogramming using a previously published expression time course (Polo et al., 2012). *Thy1*, fibroblast marker; *Nanog*, pluripotency marker.

(G) Scheme to determine minimal duration of OKSM expression (in days) required to achieve transgene-independent iPSC colonies.

(H) Data obtained from experiments depicted in (G) using indicated shRNAs and three biological replicates (mean ± SD).

(legend continued on next page)

iPSC-like colonies and the fraction of Oct4-GFP$^+$ cells by 7- to 8-fold (16% Oct4-GFP$^+$ cells with *Sumo2* shRNA; 2% with control shRNA at day 8) (Figures 2A–2D). We were able to recapitulate enhanced reprogramming with six independent shRNAs as well as small interfering RNA (siRNA) pools targeting *Sumo2*, documenting the consistency of the observed phenotype using either permanent or transient knockdown approaches (Figures 2C, 2D, and S2A). Expression of *Sumo2* shRNAs led to reduced *Sumo2* transcript levels as well as a decrease of free and conjugated SUMO2 protein levels, demonstrating specificity of the knockdown vector and diminution of global sumoylation levels (Figures 2E and S2B). Considering that sumoylated KLF4 reportedly inhibits cellular reprogramming (Tahmasebi et al., 2013), we next determined whether SUMO2 suppression leads to reduced KLF4 sumoylation. However, we failed to detect obvious differences in KLF4's sumoylation status in cells expressing OKSM and *Sumo2* shRNAs in a preliminary experiment, suggesting that the effect of SUMO2 on iPSC formation may involve other targets (Figure S2C). Moreover, forced expression of SUMO2 during iPSC generation (Figure S2D) was insufficient to block reprogramming (data not shown), indicating that downregulation of SUMO2 may not be required to attain pluripotency. Indeed, we found that endogenous *Sumo2* mRNA levels were comparable between MEFs and iPSCs and barely changed during the reprogramming process (Figure 2F). These results suggest that SUMO2 may play independent roles in somatic and pluripotent cell types. iPSCs generated with *Sumo2* shRNAs could be stably propagated over many passages and gave rise to well-differentiated teratomas, demonstrating that suppression of SUMO2 does not compromise the self-renewal or pluripotency of iPSCs (Figure S3A).

To complement the aforementioned marker-based assays of reprogramming with a functional readout, we determined whether suppression of SUMO2 could promote the formation of transgene-independent iPSC colonies after reduced pulses of OKSM expression (Figure 2G). Indeed, we found that knockdown of SUMO2 yielded transgene-independent iPSC colonies after only 4–5 days of OKSM expression, whereas stable iPSC colonies only emerged between days 6 and 8 in controls, consistent with previous findings (Stadtfeld et al., 2008) (Figure 2H). In agreement with this observation, we detected a significant upregulation of key embryonic stem cell (ESC)-associated transcripts (e.g., *Nanog*, *Nr5a2*, *Sall4*) and epigenetic regulators (e.g., *Dnmt3b* and *Tet1*) as well as a downregulation of MEF-associated transcripts (e.g., *Twist2*, *Fgf18*, *Meox2*) in cells expressing OKSM and *Sumo2* shRNAs at day 6 of reprogramming relative to a non-targeting shRNA control (Figures 2I and 2J). Critically, knockdown of SUMO2 had no discernible effect on cell proliferation or apoptosis of bulk cultures, thus excluding the possibility that the observed phenotype is due to accelerated growth or reduced cell death (Figures S3B and S3C). Together, these results demonstrate that both transient and constitutive suppression of SUMO2 markedly enhances and accelerates the formation of iPSCs from somatic cells.

SUMO2 Suppression Acts during Early-to-Mid Stages of Reprogramming

To understand how SUMO2 suppression influences the dynamics of iPSC formation, we utilized surface markers and a reporter allele to distinguish between early, mid, and late stages of reprogramming. We previously showed that cells undergoing successful reprogramming initially upregulate SSEA1 (early stage), followed by sequential activation of EPCAM (mid stage) and Oct4-GFP (late stage) (Polo et al., 2012; Stadtfeld et al., 2008) (Figure 3A). SUMO2 depletion had no pronounced effect on the earliest intermediates of reprogramming, as shown by comparable fractions of SSEA1$^+$ and EPCAM$^+$ cells at day 3 relative to controls (Figures 3B, 3C, and S4). However, we noticed a 3-fold increase in the fraction of SSEA1$^+$ cells and a 5-fold increase in the fraction of EPCAM$^+$ cells by day 5 as well as an 8-fold increase in the fraction of Oct4-GFP$^+$ cells by day 8 of reprogramming. These data indicate that SUMO2 suppression facilitates early-to-mid stages of iPSC formation based on surface marker expression in bulk cultures.

To corroborate the notion that SUMO2 suppression promotes early-to-mid stages of reprogramming, we determined iPSC colony formation efficiencies after transfecting reprogrammable MEFs with siRNAs against *Sumo2* either once (on day 0) or twice (on days 0 and 3) (Figures 3D and 3E). iPSC colony formation was essentially the same when SUMO2 was suppressed initially or continuously during a 6-day reprogramming period, suggesting that early SUMO2 suppression is sufficient to elicit enhanced reprogramming. Consistent with the acceleration of reprogramming upon suppression of SUMO2, we find that MEFs expressing OKSM and *Sumo2* shRNAs for only 6 days are molecularly most similar to advanced stages of reprogramming (day 9 and day 12 intermediates) using a previous expression time course (Polo et al., 2012), whereas MEFs expressing OKSM and a control shRNA are most similar to day-6 intermediates, as expected (Figure 3F).

(I) Scatterplot comparing microarray data (log2 values) of day-6 reprogramming intermediates expressing indicated shRNAs. Representative upregulated pluripotency genes are shown in red whereas downregulated somatic genes are shown in green.
(J) Normalized expression levels of representative pluripotency-associated and MEF-associated genes in indicated samples at day 6 of OKSM expression. Data obtained from one experiment. RepMEFs, reprogrammable MEFs; ESCs, embryonic stem cells.

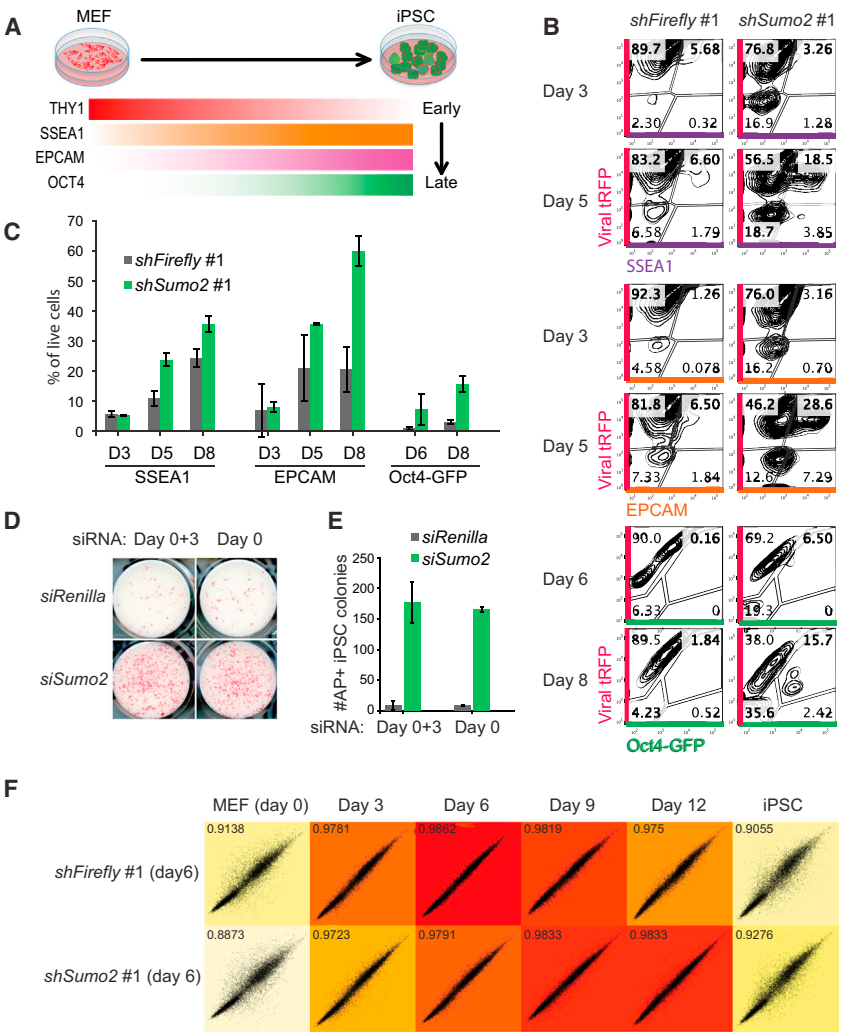

Figure 3. Effect of SUMO2 Suppression on Defined Reprogramming Intermediates

(A) Overview of surface markers and reporter alleles to distinguish between early, mid, and late stages of reprogramming.

(B) Flow cytometry analysis of these markers at intermediate stages of reprogramming in the presence of indicated shRNAs (tRFP$^+$ cells).

(C) Quantification of data shown in (B) and Figure S4 using three biological replicates ± SD; SSEA1 *shFirefly* sample contains only two biological replicates; see also Figure S4.

(D) AP$^+$ transgene-independent iPSC colonies obtained after transfection of reprogrammable MEFs with indicated siRNAs either once (day 0) or twice (day 0 and day 3) in the presence of dox for 6 days; iPSC colonies were scored after 4 days of dox withdrawal to capture stable iPSCs.

(E) Quantification of data shown in (D) using three biological replicates (mean ± SD).

(F) Correlation analysis between microarray data obtained in this study (day-6 reprogramming cultures expressing either *Firefly* or *Sumo2* shRNAs, single data points) and a previously reported reprogramming time course (Polo et al., 2012).

Collectively these phenotypic, functional, and molecular data suggest that SUMO2 suppression accelerates early-to-mid stages of reprogramming, ultimately leading to a strong increase in the number of Oct4$^+$ transgene-independent iPSCs.

SUMO2 Suppression Acts Independently of C-MYC and in Parallel with Small Molecules

We next investigated whether exogenous C-MYC expression is required for the enhancement of iPSC formation by *Sumo2* shRNAs. To this end, we derived reprogrammable MEFs from mice carrying the *Col1a1-tetOP-OKS-mCherry* allele (lacking the *c-Myc* transgene) in combination with the *R26-M2rtTA* allele (Figure 4A). Exposure of these MEFs to dox alone gave rise to extremely few, if any, AP$^+$ colonies after 9–21 days of OKS expression, and no Oct4-GFP$^+$ cells could be detected by day 9 of reprogramming (Figures 4B–4D). By contrast, depletion of SUMO2 in these cells using transient transfection of siRNA pools yielded iPSC colonies and stable Oct4-GFP$^+$ cells by flow cytometry as early as 9 days after OKS induction. We conclude that suppression of SUMO2 enhances reprogramming independently of exogenous C-MYC expression, thus enabling iPSC formation in the absence of this potent oncogene.

To determine whether SUMO2 suppression acts in parallel with small molecules that were previously shown to enhance reprogramming, we treated reprogrammable cells harboring shRNAs against *Sumo2* or *Firefly* luciferase with doxycycline in the presence or absence of ascorbic acid (AA) (Esteban et al., 2010), a DOT1L inhibitor (DOT1Li) (Onder et al., 2012), and a GSK3 inhibitor (GSK3i) (Silva et al., 2008). Consistent with previous reports, we found that exposure of reprogrammable cells to each of these compounds significantly enhanced the generation of AP$^+$ iPSC colonies, with combined AA/GSK3i treatment (AGi) exhibiting the strongest effect (Bar-Nur et al., 2014). Strikingly, SUMO2 depletion alone even surpassed the effect of AGi exposure based on AP$^+$ colony

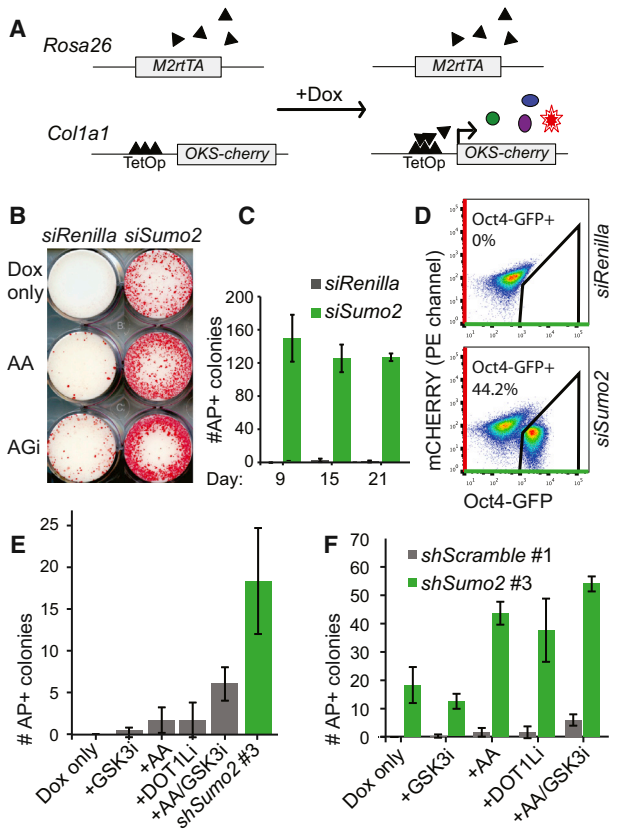

Figure 4. SUMO2 Suppression Acts Independently of C-MYC and in Parallel with Small Molecules

(A) Scheme depicting reprogrammable MEFs carrying *Col1a1-tetOP-OKS-mCherry* and *Rosa26-M2rtTA* alleles. Black triangles depict M2rtTA protein; green, purple, and blue ovals show OCT4, SOX2, and KLF4 proteins; red star symbolizes mCHERRY protein.

(B) Generation of AP$^+$ transgene-independent iPSC colonies obtained from these reprogrammable MEFs transfected with indicated shRNAs after exposure to dox and small molecules for 9 days.

(C) Quantification of data shown in (B), "dox only" samples and additional time points; data show mean from three biological replicates ± SD.

(D) Oct4-GFP expression of OKS reprogrammable MEFs treated with indicated siRNAs and dox for 9 days, followed by 5 days of dox-independent growth.

(E) Comparison of iPSC formation efficiencies from OKSM reprogrammable MEFs in the presence of either small molecules or siRNAs targeting *Sumo2*. Values show mean from three biological replicates ± SD.

(F) Combination treatment of reprogrammable MEFs with siRNA targeting *Sumo2* and indicated small molecule. Values show mean from three biological replicates ± SD.

numbers (Figure 4E). Moreover, suppression of SUMO2 further enhanced iPSC formation in the presence of either ascorbate, GSK3i, or DOT1Li (Figure 4F). These results underscore the strong effects of SUMO2 suppression alone on the reprogramming process and suggest that the sumoylation pathway may act in parallel to previously described facilitators of iPSC formation, including AA and inhibitors of H3K79 methylation and GSK3/WNT signaling.

Generation of iPSCs after as Little As 38 Hours of OKSM Expression

Given the additive effect between SUMO2 suppression and small molecule enhancers of reprogramming, we asked whether this combination treatment would allow us to further reduce the minimal time period of OKSM expression required to produce stable transgene-independent iPSCs. We used early passage reprogrammable MEFs (passage 2) carrying two copies of each of the *Col1a1-tetOP-OKSM* and *R26-M2rtTA* alleles to achieve optimal reprogramming efficiencies (Stadtfeld et al., 2010). MEFs exposed to DOT1Li, GSK3i, and AA required only 72 hr of OKSM expression to produce the first dox-independent AP$^+$ iPSCs, which is faster than any previously reported protocol (Figures 5A and 5B). Remarkably, suppression of SUMO2 further reduced this time window to 38 hrs. Emerging iPSC colonies activated the endogenous Oct4-GFP reporter, expressed endogenous NANOG, OCT4, and SOX2, gave rise to well-differentiated teratomas, and supported the formation of coat-color chimeras, indicating acquisition of an authentic iPSC state (Figures 5C–5F). Analysis of global transcriptional patterns of these iPSCs further revealed a remarkable similarity with either iPSCs generated after 10 days of OKSM expression or with an established ESC line, suggesting that abbreviated OKSM expression in the presence of chemicals and *Sumo2* siRNAs does not compromise the pluripotency program (Figures 5G–5I and S5). These results show that 1–2 days of OKSM expression are sufficient to produce stable, pluripotent iPSCs when SUMO2 expression is transiently suppressed under optimal culture conditions.

Finally, we tested whether suppression of SUMO2 also enhances human reprogramming. To this end, we infected human dermal fibroblasts with vectors expressing *OKSM* and either *SUMO2* or *FIREFLY* (FF) shRNAs and measured the formation of AP$^+$ colonies at day 21 (Figures 5J–5L). Consistent with our observations in the mouse system, we find that suppression of SUMO2 increases the formation of human iPSC-like colonies by 4- to 6-fold. We also noticed that human iPSC-like colonies expressing *SUMO2* shRNAs formed earlier than controls, suggesting acceleration of reprogramming (data not shown). Thus, sumoylation is a conserved reprogramming barrier across murine and human somatic cells.

DISCUSSION

Here, we identified SUMO2 as a potent roadblock to iPSC generation by combining a well-defined transgenic

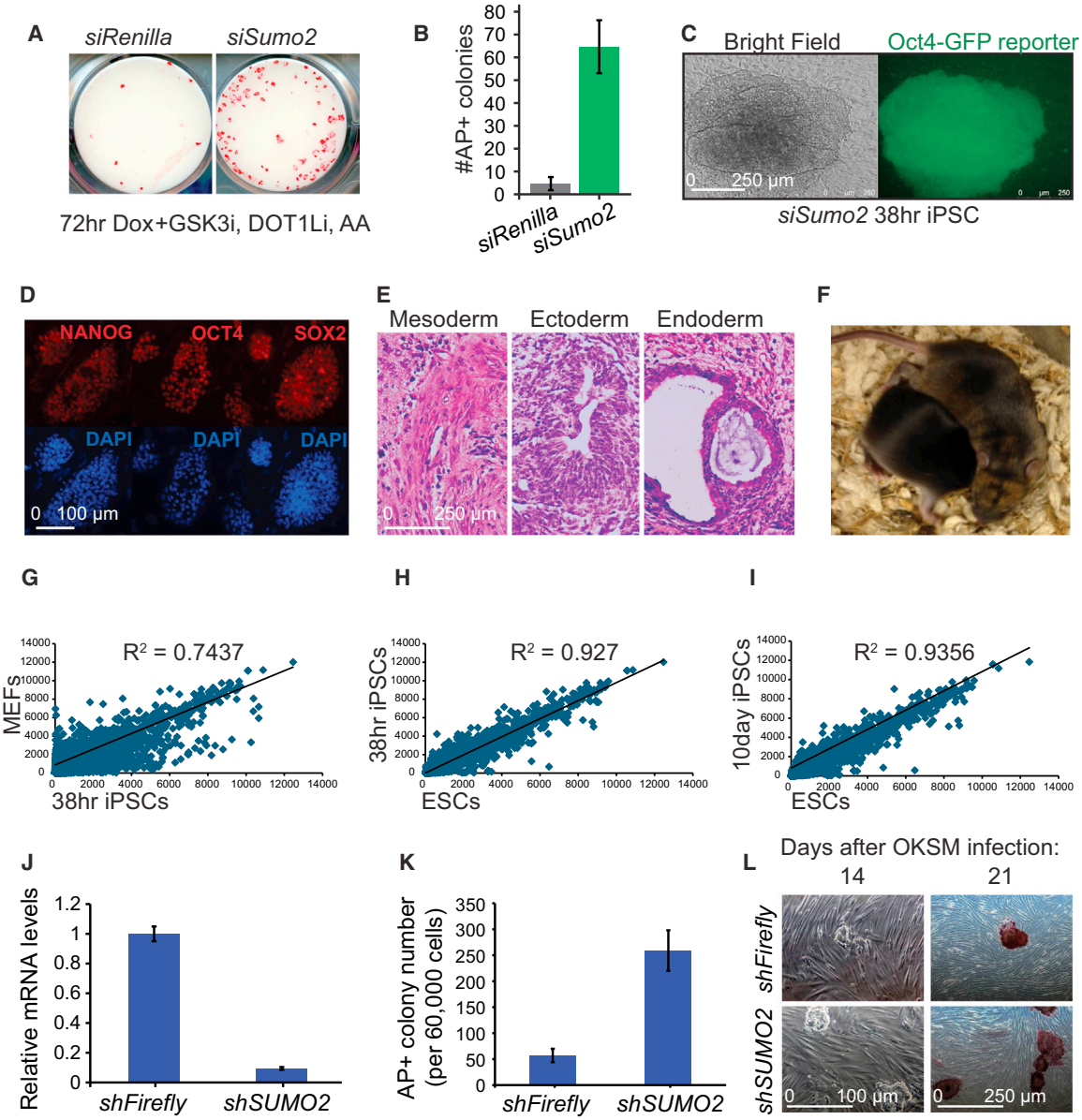

Figure 5. Generation of iPSCs after as Little as 38 Hours of OKSM Expression and Human Reprogramming Experiments

(A) Treatment with ascorbic acid (AA), DOT1L inhibitor (DOT1Li), and GSK3 inhibitor (GSK3i) facilitates the recovery of transgene-independent AP$^+$ iPSC colonies from control and *Sumo2* siRNA transfected MEFs after 72 hr of OKSM expression.

(B) Quantification of data shown in (A) using three biological replicates (mean ± SD).

(C) Suppression of SUMO2 enables generation of Oct4-GFP$^+$ transgene-independent iPSCs after 38 hr of OKSM expression in the presence of the indicated small molecules.

(D) Representative immunofluorescence images showing expression of endogenous OCT4, NANOG, and SOX2 in iPSCs generated after 38 hr of OKSM expression and Sumo2 siRNA transfection.

(E) iPSCs shown in (C) are pluripotent as determined by their potential to differentiate into the three germ layers in teratomas.

(F) iPSCs produced after 38 hr of OKSM expression give rise to coat-color chimeras.

(G–I) Correlation analysis of microarray data (single data points) between MEFs and iPSCs derived after 38 hr of OKSM expression and *Sumo2* siRNA transfection (38hr iPSCs) (G), 38hr iPSCs and ESCs (H), and iPSCs derived after 10 days of OKSM and *Sumo2* shRNA expression (10 day iPSCs) and ESCs (I).

(J) Suppression of *SUMO2* mRNA levels in foreskin fibroblasts using *SUMO2* or *FIREFLY* shRNA vectors as determined by qPCR. RNA from three independent experiments was pooled for qPCR analysis; values represent means from three technical replicates ± SD.

(legend continued on next page)

reprogramming system with a genome-wide shRNA screening approach. In contrast to previous shRNA or siRNA screens conducted during iPSC formation, we employed a serial shRNA enrichment strategy, which may reduce the number of false-positive hits and allow for selection of shRNAs with stronger phenotypes. Indeed, suppression of a top candidate, SUMO2, markedly enhanced and accelerated iPSC formation compared with individual hits that emerged from previous large-scale screens or candidates that were selected based on gene-expression differences between somatic and pluripotent cells. In agreement with the notion that the expression of barrier genes does not necessarily have to be different between MEFs and iPSCs, we found that *Sumo2* mRNA levels do not dramatically change during reprogramming. To our knowledge, iPSC formation after 38 hr of OKSM expression represents the shortest time period that has been reported to obtain stably reprogrammed cells from fibroblasts. In addition to SUMO2, our screen uncovered a number of other candidate barriers to iPSC formation, which provide a useful resource for future mechanistic studies of the reprogramming process.

While SUMO2 has not previously been recognized as a roadblock to reprogramming, a recent report suggested that knockdown of the upstream SUMO-conjugating enzyme UBC9 blocks iPSC formation (Tahmasebi et al., 2014). Several reasons may account for the apparent discrepancy between our studies. For example, UBC9 knockdown is expected to inhibit SUMO1, SUMO2, and SUMO3, which may be toxic to cells, whereas SUMO2 depletion may be tolerated by cells owing to compensation by SUMO1 and/or SUMO3. In support of this idea, *Ubc9* deletion in vivo causes a more severe phenotype compared with individual deletion of *Sumo1*, *Sumo2*, or *Sumo3* (Wang et al., 2014). Furthermore, knockdown of UBC9 in ESCs reportedly triggers differentiation (Tahmasebi et al., 2014). Considering that the same constitutive knockdown vectors were used to assess UBC9's role in ESC self-renewal and reprogramming, it is possible that a potential increase in the number of iPSCs was overlooked due to their immediate loss through differentiation. Indeed, a recent shRNA screen conducted during iPSC formation with retroviral vectors, which become silenced in iPSCs, identified UBC9 as a potent reprogramming barrier (Cheloufi et al., 2015).

Mechanistically, SUMO2 depletion may enhance iPSC formation by derepressing epigenetically silenced pluripotency loci (Poleshko et al., 2014; Yang et al., 2015) as well as histone and protein biosynthesis genes important for cellular growth and proliferation (Neyret-Kahn et al., 2013). Consistently, we find that key pluripotency genes and epigenetic regulators are expressed more robustly in SUMO2-depleted reprogramming intermediates compared with controls. In addition, SUMO2 suppression may contribute to iPSC formation by directly modulating certain pluripotency factors during iPSC generation. In agreement with this view, sumoylation of SOX2 and KLF4 reportedly impairs transcriptional activity and compromises pluripotency (Tahmasebi et al., 2013; Wu et al., 2012). Moreover, overexpression of sumoylation-deficient variants of these transcription factors slightly enhances reprogramming into iPSCs, although the reported effects were subtle compared with the phenotype reported here. Considering these observations and our preliminary finding that KLF4 is not differentially sumoylated upon suppression of SUMO2, we surmise that SUMO2 acts at multiple levels to resist the acquisition of pluripotency rather than to control the activity or stability of a single protein target. It should be informative to identify relevant SUMO2 targets in MEFs and iPSCs using recently developed proteomics and chromatin immunoprecipitation sequencing (Becker et al., 2013; Neyret-Kahn et al., 2013) approaches.

Our findings may have practical implications for basic science and cell therapy. The ease with which SUMO2 can be inhibited using transient siRNA delivery should facilitate the mechanistic dissection of the reprogramming process in more homogeneous cell cultures. The observation that SUMO2 depletion increases human iPSC generation, cooperates with small-molecule enhancers of reprogramming, and negates the need for exogenous C-MYC expression, may further facilitate the efficient and safe generation of patient-specific iPSCs from rare donor cells.

EXPERIMENTAL PROCEDURES

Tissue Culture and Virus Production

Reprogrammable mouse embryonic fibroblasts (RepMEFs) were derived from embryonic day 13.5–15.5 embryos carrying either the *Col1a1-tetOP-OKSM* or *Col1a1-tetOP-OKS-mCherry* alleles in combination with the *Rosa26-M2rtTA* alleles (Bar-Nur et al., 2014; Stadtfeld et al., 2010). Reprogramming was initiated by adding 20 ng/ml doxycycline (dox) (Sigma, catalog #D9891-25G) to RepMEFs and, where indicated, 25 μg/ml L-AA (Sigma, catalog #A4544-25G), 3 μM GSK3 inhibitor (CHIR99021, Stemgent, catalog #04-0004), ALK5 inhibitor (Calbiochem, catalog #616452),

(K) Reprogramming efficiency of cells characterized in (J) as measured by the number of AP^+ colonies per 60,000 plated fibroblasts. Values represent means from three biological replicates ± SD.

(L) Representative images of bright-field and AP^+ colonies obtained with *SUMO2* and *FIREFLY* shRNA vectors after 14 and 21 days of OKSM expression.

1 μM MEK inhibitor (PD0325901, Stemgent, catalog #04–0006), or DOT1L inhibitor (generous gift from Dr. Peter Brown). Viral transductions, using pHage or pGIPZ vectors in combination with packaging plasmids psPax2 and pDM2.G, were performed by spin infection for 30 min at 2,150 rpm at room temperature. The following shRNA seed sequences were used for mouse experiments: pHage shRNA (shSumo2 #1), ATAAGAGCTGAATGAGCATGCC; pGIPZ (shSumo2 #2–7; Dharmacon); #2 (V2LMM_2701), TTCTGGAG TAAAGTAGCAG; #3 (V2LMM_5114), TAAGAGCTGAATGAG CATG; #4 (V3LMM_496391), TAGTAGACACCTCCAGTCT; #5 (V3LMM_496392), AAACTGCACCACAGAACCA; #6 (V3LMM_ 496393), TGTTCTCAGTCTTGACTCC; #7 (V3LMM_496396), AATCTTAAACTGCACCACA. Transfections of RepMEFs with 1.5 μl siRNAs (esiRNA technology, Sigma; Sumo2: EMU095391, Renilla: EHURLUC) were performed in 12-well plates using Lipofectamine 2000 and Opti-MEM (Life Technologies) according to the manufacturer's instructions. For human reprogramming experiments, BJ fibroblasts were first infected with control or SUMO2 pGIPZ shRNA (V3LHS_388696 and V3LHS412780; GE Dharmacon) viruses. Two days later, cells were infected with the human pHAGE-STEMCCA virus (Sommer et al., 2012) to initiate reprogramming. Cells were replated in fibroblast medium (DMEM and 10% FBS) and cultured for 4 days. Medium was then switched to a 1:1 mixture of fibroblast medium and human ESC medium (E8; STEMCELL Technologies) for another 2 days, and finally to 100% human ESC medium. Cells were stained for alkaline phosphatase activity (Stemgent), and AP^+ iPSCs were counted ~3 weeks after STEMCCA infection.

Flow Cytometry

For isolation of Oct4-GFP^+ iPSCs, $SSEA1^+$ cells were first enriched by MACS sorting using SSEA1 antibody-coated magnetic beads (Miltenyi). Oct4-GFP^+ cells were then purified within the enriched $SSEA1^+$ fraction by FACS. Gates were set based on uninfected MEFs, virally transduced MEFs ($tRFP^+$), and Oct4-GFP^+ iPSCs. Intermediates of reprogramming were analyzed by flow cytometry using the following antibodies: THY1-Viogreen (Becton Dickinson, catalog #561616) or THY1-Pacific Blue (eBioscience, catalog #48-0902-82), SSEA1-APC (Biolegend, catalog #125608) or SSEA1-PE-Cy7 (Miltenyi Biotec, catalog #130-100-426), and EPCAM-PE-Cy7 (eBioscience, catalog #25-5791-80) (1:200 for 30 min at 4°C). BD's Annexin V kit was used to measure apoptotic cells. All cytometry data were analyzed and plotted using FlowJo software.

Quantification of Reprogramming Efficiencies

For macroscopic detection of iPSC colonies, AP staining was carried out according to manufacturer's instructions using the Vector Labs AP staining kit (catalog #5100). AP staining was always performed 2–4 days after dox withdrawal to eliminate partially reprogrammed colonies and to score for transgene-independent iPSCs. Colonies were counted manually or by custom-made Nikon software (CL-Quant).

Teratoma and Chimera Formation

For teratoma generation, iPSC lines (passage 6 or higher) were harvested and resuspended in 600 μl medium per confluent six wells. Mice were anesthetized with Avertin and injected with 150 μl cell suspension subcutaneously. Tumors were harvested 3–4 weeks after injection and analyzed histologically. For chimera production, iPSC lines were injected as a single-cell suspension into day-3.5 blastocysts isolated from intercrosses of C57Bl/6×BDF1 mice. Blastocysts were transferred into pseudopregnant Swiss Webster recipient animals.

Immunofluorescence Analysis

iPSC lines (passage 6 or higher) were seeded in a 24-well plates at a low density. Once small colonies emerged, wells were washed with 1× PBS and fixed by a 5- to 10-min incubation in 10% formalin at room temperature. After washes in 1× PBS, cells were blocked in 1× PBS containing 2% BSA and 0.1% Triton X-100. Primary and secondary antibodies were diluted in blocking solution at a concentration of 1:200 and added for 1 hr at room temperature or overnight at 4°C. Primary antibodies were anti-NANOG (Abcam, catalog #AB80892), anti-SOX2 (Santa Cruz Biotechnology, catalog #Sc-17320), anti-OCT4 (Santa Cruz, catalog #Sc-8628); secondary antibodies were donkey anti-goat immunoglobulin G (IgG) or donkey anti-rabbit IgG Alexa Fluor 546-conjugated antibodies (Life Technologies). After two washes in 1× PBS, cells were immobilized on slides in mounting medium containing DAPI (Vectashield, Vector Labs) and analyzed.

RNA Expression Analysis

For global gene-expression analysis, total RNA was isolated from indicated samples using an RNeasy Mini Kit (Qiagen) and analyzed by Affymetrix microarray chips. The raw microarray expression signals in the CEL files were normalized using the Affymetrix Expression Console software and robust multiarray averaging normalization. For qPCR analysis, Brilliant III SYBR-Green-based master mix was used according to the manual (Agilent), following RNA isolation (RNeasy kit, Qiagen) and reverse transcription (Transcriptor First Strand cDNA Synthesis Kit, Roche) of Sumo2 or control knockdown samples 2 days after initiation of reprogramming. Samples were run in triplicate on the Lightcycler 480 (Roche). Primer sequences for qPCR were as follows: Sumo2 (forward), AAG GAAGGAGTCAAGACTGAGAA; Sumo2 (reverse), CGGAATCT GATCTGCCTCATTG; GAPDH (forward), AGGTCGGTGTGAACG GATTTG; GAPDH (reverse), TGTAGACCATGTAGTTGAGGTCA.

Scatterplot

Microarray gene-expression measurements for samples containing Sumo2 or Firefly shRNA for 6 days were obtained. For genes with more than one probe, the average of all probes was used. These expression measurements were plotted (on a log2 scale), with Firefly expression on the x axis, Sumo2 knockdown expression on the y axis, and each point representing a gene. Genes falling outside of the dashed lines have a fold change of greater than 1.5 between Sumo2 knockdown and Firefly. Some key genes were highlighted, with those being upregulated in Sumo2 knockdown colored red, and those downregulated colored green.

Correlation Matrix

Previously published microarray gene-expression measurements representing stages of reprogramming were obtained from GEO (GSE42379). The following samples were used: KH2-MEF_m2,

KH2-MEF_m3, Day3_SSEA1+ M2, Day3_SSEA1+ M3, Day6_SSEA1+ M2 and Day6_SSEA1+ M3, Day9_SSEA1+ M2, Day12_SSEA1+ M2, iPS_KH2-SC_MEF_1–5, iPS_KH2-SC_MEF_1–6, and iPS_KH2-SC_MEF_1–3 (GSM1038591, GSM1038592, GSM1038595, GSM1038598, GSM1038601, GSM1038604, GSM1038607, GSM1038611, GSM1038612, GSM1038613, GSM1038614), representing day 0 of reprogramming (MEF), day 3, day 6, day 9, day 12, and fully reprogrammed (iPS). For time points with multiple samples, the average expression measurement for each probe was used. For genes with more than one probe, the average of all probes was used.

The gene-expression measurements of the Sumo2 knockdown and Firefly samples were compared with those of the samples representing stages of reprogramming. Only the genes that were present in both datasets were considered. To remove the batch effect we used the "Remove Batch Effect" tool in the Partek Genomics Suite, which fits a linear model to the data which includes batch as a component, then subtracts that component from the data.

A matrix was constructed with pairwise comparisons between the Sumo2 knockdown and Firefly samples (as the rows in the matrix) and each of the samples representing stages of reprogramming (as the columns in the matrix). A scatterplot was generated and a Pearson's R correlation coefficient was calculated for each pairwise comparison. Scatterplots that are shaped closer to a straight line and have a higher correlation coefficient indicate that the global expression patterns of the two samples more closely match.

Western Blot Analysis

Protein lysates were run on 4%–20% Mini Protean TGX gels (Bio-Rad), blotted onto Immobilon-P membrane (EMD Millipore), and incubated with anti-SUMO2 antibody ab3742 (Abcam) and anti-RAN antibody 610341 (BD) for western blot analysis. We used anti-KLF4 antibody AF3158 (R&D Systems), anti-OCT4 antibody 11,263-1-AP (ProteinTech), and anti-SUMO2 antibody ab81371 (Abcam) for immunoprecipitation experiments.

Generation of pHAGE-Mir Vector

The pHAGE lentiviral backbone was released from pHAGE-EF1a-eGFP-W vector (kindly provided by Dr. Richard Mulligan's laboratory) using NotI + BamHI. tRFP was PCR amplified from pTurboRFP-N (Evrogen) and blunt-end cloned into the above pHAGE backbone to generate the intermediate pHAGE-EF1a-tRFP vector. The Mir30-shRNA cassette and PGK-Puro selection marker was digested from MSCV-PM using BgIII + ClaI, and blunt-end cloned into the BamHI digested pHAGE-EF1a-tRFP intermediate to generate the pHAGE-Mir vector.

shRNA Screen and Identification of Hits

RepMEFs were expanded until passage 4 in 4% oxygen, switched to atmospheric oxygen, and infected with the pooled shRNA library as described above. For each shRNA (621,000 shRNAs in total), 1–2 × 10^3 cells were infected to achieve good coverage. Infected cells were passaged onto gelatinized 15-cm cell culture dishes (Falcon) in reprogramming medium (ESC medium supplemented with AA and dox) for 10 days, and in dox/ascorbic acid-free ESC media for an additional 4 days. Cells were harvested, pooled, and purified with SSEA1-linked magnetic beads using an AutoMACS sorter (Miltenyi). SSEA1-enriched cells were then FACS sorted for endogenous Oct4-GFP expression.

Genomic DNA was extracted from collected Oct4-GFP$^+$ cells by lysing the cells in 10 mM Tris (pH 8.0), 10 mM EDTA, 10 mM NaCl, and 0.5% sarkosyl. Lysates were treated with 0.1 mg/ml RNase A at 37°C for 30 min and 0.5 mg/ml proteinase K at 55°C for 1–2 hr, then phenol-chloroform extracted, ethanol precipitated, and resuspended in 10 mM Tris-HCl (pH 8.0). For each sample, all the genomic DNA was used as template for shRNA PCR, usually in multiple PCR reactions. Each 50 μl PCR reaction contained 2.5 μg genomic DNA template, 200 μM dinucleotide triphosphates (dNTPs), 400 nM of each PCR primer (pHAGE-Mir-PCR, 5′-GCAAACTGGGGCACAGATGATGCGG; BC1R-L, 5′-CGCCTCCCCTACCCGGTAGA), 1× Q5 reaction buffer, 1× Q5 high GC buffer, and 0.5 μl Q5 polymerase (NEB). PCR was performed with the following program: 94°C for 4 min, 35 cycles of (94°C 30 s, 60°C 30 s, 72°C 45 s), and 72°C for 10 min. PCR products (~700 bp) for each sample were pooled, ethanol precipitated, resuspended, and gel purified using the QIAquick Gel Extraction Kit (Qiagen). The purified shRNA PCR products were used to (1) generate sublibraries for the next round of shRNA library screens and (2) generate sequencing libraries for Illumina sequencing.

For sublibrary generation, the purified PCR product was digested with NotI and MluI, and the ~400 -bp fragment that contains the shRNAs was gel purified. Separately, the pHAGE-Mir plasmid was also digested with NotI and MluI to recover the ~9 kb vector backbone. The purified shRNA fragment (25–50 ng) and the vector backbone (125–250 ng) were ligated in a 5 μl ligation reaction using NEB T4 ligase. A 1 μl ligation reaction was used to transform 20 μl Electromax competent cells DH10b (Life Technologies) with electroporation. One microliter of the transformation reaction was plated on one 10 cm LB-Amp (100 μg/ml) plate to estimate the total number of colonies, and the rest of the transformation reaction was plated on two 15 cm LB-carbenicillin (100 μg/ml) plates and grown overnight at 37°C. At least 100× coverage was needed to maintain the representation of the library (i.e., colony number = 100 × the number of shRNAs in the library). When necessary, the entire ligation reaction may be used for transformation in multiple electroporation reactions to increase the number of colonies. The next day, lawns formed on the two 15 cm plates were scraped off and cultured in 300 ml LB-carbenicillin (100 μg/ml) medium and grown at 30°C for 2–3 hr. The bacteria were collected and the cloned sublibrary DNA was extracted by the Genelute Maxiprep kit (Sigma).

For Illumina sequencing, the purified shRNA PCR product was used as template for another round of PCR: 500 ng purified shRNA PCR product, 200 μM dNTPs, 2 μM each PCR primer (p5 and p7), 1× Q5 reaction buffer, 1× Q5 high GC buffer, and 1 μl Q5 polymerase (NEB) in 100 μl PCR reaction buffer. PCR was performed with the following program: 94°C for 4 min, two cycles of (94°C 30 s, 50°C 20 s, 72°C 30 s), 20 cycles of (94°C 30 s, 60°C 20 s, 72°C 30 s), and 72°C for 10 min. PCR products (~120 bp) were gel purified using the QIAquick Gel Extraction Kit (Qiagen). Gel-purified products were submitted for Illumina sequencing on the Illumina MiSeq instrument, using a custom sequencing primer: mir30-EcoRI: 5′-TAGCCCCTTGAATTCCGAGGCAGTAGGCA.

PCR primers: p5-miSeq, 5′-ATGATACGGCGACCACCGAGATC TACACCTAAAGTAGCCCCTTGAATTC; p7-miSeq-1, 5′-CAAGCA GAAGACGGCATACGAGACGATAGTGAAGCCACAGATGTA; p7-miSeq-2, 5′-CAAGCAGAAGACGGCATACGAGACACTAGTGAAG CCACAGATGTA; p7-miSeq-3, 5′-CAAGCAGAAGACGGCATAC GAGACTATAGTGAAGCCACAGATGTA; and p7-miSeq-4, 5′-CAA GCAGAAGACGGCATACGAGACCTTAGTGAAGCCACAGATGTA. Different p7 primers were used for multiplexing purposes.

Single-end 51-bp reads were obtained using the Illumina HiSeq or MiSeq instrument. The reads are expected to have an initial 22 nucleotides that identify the shRNA, followed by a constant region that is the same for all shRNAs and a two-nucleotide barcode to identify the sample. Reads that contain perfect matches at the following six nucleotides were first extracted from the sequencing data: the two nucleotides adjacent to the initial 22 base sequence and the two nucleotides adjacent to the barcode on both sides. The shRNAs were then identified by requiring an exact match of the 22 nucleotides to the sequences in the shRNA library annotation file. The samples were identified by the two-nucleotide barcodes.

The total number of reads that were identified for each shRNA, sample, and round were counted. The counts were normalized to be directly comparable between samples and rounds by first dividing by the total number of counts for that sample and round, and then multiplying by the total number of shRNAs in the initial library. A pseudocount of 1 was added to each normalized count to downweight enrichment derived from low-read counts and to avoid division by zero in calculating fold changes.

The enrichment for each shRNA in each round was calculated as the log2 fold change of the Oct4-GFP$^+$ normalized counts over the maximum of the normalized counts of the controls (T0, No-Dox, and Oct4-GFP$^-$ cells). The cumulative enrichment for each shRNA in each round was calculated as the sum of the log2 fold changes for that round and all previous rounds. The overall enrichment of each shRNA was defined as the maximum of the cumulative enrichment scores among all rounds.

The heatmap for Figure 1G was plotted using the cumulative enrichment scores. Only shRNAs that have at least one read in the Oct4-GFP$^+$ sample in at least two rounds were used in the plot, resulting in a total of 23,853 shRNAs.

ACCESSION NUMBERS

Microarray expression data are available at the Gene Expression Omnibus database under accession number GEO: GSE76699.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2016.02.004.

AUTHOR CONTRIBUTIONS

M.B., B.D.B., N.M., S.J.E., G.H., and K.H. planned the experiments; M.B. and G.H. performed the serial enrichment screen, validated and characterized hits; B.D.B. conducted all bioinformatics analyses with assistance from Y.D. and D.C.F; B.L. performed sumoylation experiments; O.B.N. generated teratomas and helped with array analysis; J.B. generated chimeric mice; L.W. performed human reprogramming experiments; E.A. provided tetOP-OKS MEFs; S.C. shared unpublished information; M.B., B.D.B., G.H., and K.H. wrote the manuscript.

ACKNOWLEDGMENTS

We thank Laura Prickett, Kat Folz-Donahue, and Meredith Weglarz of the CRM Flow Cytometry core facility and all K.H. and G.H. lab members for advice and discussions. We are grateful to Dr. Peter Brown for providing the DOT1L inhibitor, and Dr. Jianping Jin for providing the SUMO2 entry vector. We thank Dr. Mehrnaz Ghazvini for help with human reprogramming experiments. We thank our animal facility for the care of mice; all mouse experiments have been approved by the Institutional Animal Care and Use Committee. M.B. was supported by a Boehringer Ingelheim Fonds predoctoral fellowship, O.B.N. by a MGH Tosteson fellowship, J.B. by the NIH (1F32HD078029-01A1) and S.C. by the PRCRP at the Department of Defense (CA 120212). This work was supported by grants from the NIH (GM44664 and AG011085) to S.J.E.; S.J.E. is an Investigator with the Howard Hughes Medical Institute. G.H. was supported by the Intramural Research Program at the National Institute of Environmental Health Sciences (Z01ES102745). Support to K.H. was from the Gerald and Darlene Jordan Endowment, the NIH (HD058013), and an Early Career Award from the Howard Hughes Medical Institute.

Received: July 2, 2015
Revised: February 4, 2016
Accepted: February 4, 2016
Published: March 3, 2016

REFERENCES

Apostolou, E., and Hochedlinger, K. (2013). Chromatin dynamics during cellular reprogramming. Nature *502*, 462–471.

Bar-Nur, O., Brumbaugh, J., Verheul, C., Apostolou, E., Pruteanu-Malinici, I., Walsh, R.M., Ramaswamy, S., and Hochedlinger, K. (2014). Small molecules facilitate rapid and synchronous iPSC generation. Nat. Methods *11*, 1170–1176.

Becker, J., Barysch, S.V., Karaca, S., Dittner, C., Hsiao, H.H., Berriel Diaz, M., Herzig, S., Urlaub, H., and Melchior, F. (2013). Detecting endogenous SUMO targets in mammalian cells and tissues. Nat. Struct. Mol. Biol. *20*, 525–531.

Cheloufi, S., Elling, U., Hopfgartner, B., Jung, Y.L., Murn, J., Ninova, M., Hubmann, M., Badeaux, A.I., Euong Ang, C., Tenen, D., et al. (2015). The histone chaperone CAF-1 safeguards somatic cell identity. Nature *528*, 218–224.

Dejosez, M., Ura, H., Brandt, V.L., and Zwaka, T.P. (2013). Safeguards for cell cooperation in mouse embryogenesis shown by genome-wide cheater screen. Science *341*, 1511–1514.

dos Santos, R.L., Tosti, L., Radzisheuskaya, A., Caballero, I.M., Kaji, K., Hendrich, B., and Silva, J.C. (2014). MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner. Cell Stem Cell *15*, 102–110.

Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S., et al. (2010). Vitamin C enhances the

generation of mouse and human induced pluripotent stem cells. Cell Stem Cell *6*, 71–79.

Flotho, A., and Melchior, F. (2013). Sumoylation: a regulatory protein modification in health and disease. Annu. Rev. Biochem. *82*, 357–385.

Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M., Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E., et al. (2001). Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. Nat. Genet. *27*, 31–39.

Krizhanovsky, V., and Lowe, S.W. (2009). Stem cells: the promises and perils of p53. Nature *460*, 1085–1086.

Meerbrey, K.L., Hu, G., Kessler, J.D., Roarty, K., Li, M.Z., Fang, J.E., Herschkowitz, J.I., Burrows, A.E., Ciccia, A., Sun, T., et al. (2011). The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. Proc. Natl. Acad. Sci. USA *108*, 3665–3670.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49–55.

Neyret-Kahn, H., Benhamed, M., Ye, T., Le Gras, S., Cossec, J.C., Lapaquette, P., Bischof, O., Ouspenskaia, M., Dasso, M., Seeler, J., et al. (2013). Sumoylation at chromatin governs coordinated repression of a transcriptional program essential for cell growth and proliferation. Genome Res. *23*, 1563–1579.

Onder, T.T., Kara, N., Cherry, A., Sinha, A.U., Zhu, N., Bernt, K.M., Cahan, P., Marcarci, B.O., Unternaehrer, J., Gupta, P.B., et al. (2012). Chromatin-modifying enzymes as modulators of reprogramming. Nature *483*, 598–602.

Poleshko, A., Kossenkov, A.V., Shalginskikh, N., Pecherskaya, A., Einarson, M.B., Marie Skalka, A., and Katz, R.A. (2014). Human factors and pathways essential for mediating epigenetic gene silencing. Epigenetics *9*, 1280–1289.

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.

Qin, H., Diaz, A., Blouin, L., Lebbink, R.J., Patena, W., Tanbun, P., LeProust, E.M., McManus, M.T., Song, J.S., and Ramalho-Santos, M. (2014). Systematic identification of barriers to human iPSC generation. Cell *158*, 449–461.

Rais, Y., Zviran, A., Geula, S., Gafni, O., Chomsky, E., Viukov, S., Mansour, A.A., Caspi, I., Krupalnik, V., Zerbib, M., et al. (2013). Deterministic direct reprogramming of somatic cells to pluripotency. Nature *502*, 65–70.

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.K., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., and Wrana, J.L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell *7*, 64–77.

Schlabach, M.R., Luo, J., Solimini, N.L., Hu, G., Xu, Q., Li, M.Z., Zhao, Z., Smogorzewska, A., Sowa, M.E., Ang, X.L., et al. (2008). Cancer proliferation gene discovery through functional genomics. Science *319*, 620–624.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol. *6*, e253.

Sommer, A.G., Rozelle, S.S., Sullivan, S., Mills, J.A., Park, S.M., Smith, B.W., Iyer, A.M., French, D.L., Kotton, D.N., Gadue, P., et al. (2012). Generation of human induced pluripotent stem cells from peripheral blood using the STEMCCA lentiviral vector. J. Vis. Exp. *68*, 4327.

Stadtfeld, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell *2*, 230–240.

Stadtfeld, M., Maherali, N., Borkent, M., and Hochedlinger, K. (2010). A reprogrammable mouse strain from gene-targeted embryonic stem cells. Nat. Methods *7*, 53–55.

Tahmasebi, S., Ghorbani, M., Savage, P., Yan, K., Gocevski, G., Xiao, L., You, L., and Yang, X.J. (2013). Sumoylation of Kruppel-like factor 4 inhibits pluripotency induction but promotes adipocyte differentiation. J. Biol. Chem. *288*, 12791–12804.

Tahmasebi, S., Ghorbani, M., Savage, P., Gocevski, G., and Yang, X.J. (2014). The SUMO conjugating enzyme Ubc9 is required for inducing and maintaining stem cell pluripotency. Stem Cells *32*, 1012–1020.

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.

Wang, T., Chen, K., Zeng, X., Yang, J., Wu, Y., Shi, X., Qin, B., Zeng, L., Esteban, M.A., Pan, G., et al. (2011). The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. Cell Stem Cell *9*, 575–587.

Wang, L., Wansleeben, C., Zhao, S., Miao, P., Paschen, W., and Yang, W. (2014). SUMO2 is essential while SUMO3 is dispensable for mouse embryonic development. EMBO Rep. *15*, 878–885.

Wu, Y., Guo, Z., Wu, H., Wang, X., Yang, L., Shi, X., Du, J., Tang, B., Li, W., Yang, L., et al. (2012). SUMOylation represses Nanog expression via modulating transcription factors Oct4 and Sox2. PLoS One *7*, e39606.

Yang, C.S., Chang, K.Y., and Rana, T.M. (2014). Genome-wide functional analysis reveals factors needed at the transition steps of induced reprogramming. Cell Rep. *8*, 327–337.

Yang, B.X., El Farran, C.A., Guo, H.C., Yu, T., Fang, H.T., Wang, H.F., Schlesinger, S., Seah, Y.F., Goh, G.Y., Neo, S.P., et al. (2015). Systematic identification of factors for provirus silencing in embryonic stem cells. Cell *163*, 230–245.