HHS Public Access

Author manuscript

Cell Stem Cell. Author manuscript; available in PMC 2019 September 06.

Published in final edited form as:

Cell Stem Cell. 2018 September 06; 23(3): 396–411.e8. doi:10.1016/j.stem.2018.08.001.

Dppa2/4 facilitate epigenetic remodeling during reprogramming to pluripotency

Charles Hernandez[1,2,3], **Zheng Wang**[1,2,3], **Bulat Ramazanov**[1,2,3], **Yin Tang**[1,2], **Sameet Mehta**[1,2,4], **Cheryl Dambrot**[1,2,3], **Yu-Wei Lee**[1,2,3], **Kaleab Tessema**[1,2,3], **Ishan Kumar**[1,2,3], **Michael Astudillo**[1,2,3], **Thomas A. Neubert**[6], **Shangqin Guo**[1,3,5], and **Natalia B. Ivanova**[1,2,3,#]

[1]Yale University, New Haven, CT, USA

[2]Department of Genetics; New York University School of Medicine, New York, NY, USA

[3]Yale Stem Cell Center; New York University School of Medicine, New York, NY, USA

[4]Yale Center for Genome Analysis; New York University School of Medicine, New York, NY, USA

[5]Department of Cell Biology; New York University School of Medicine, New York, NY, USA

[6]Kimmel Center for Biology and Medicine at the Skirball Institute and Department of Cell Biology, New York University School of Medicine, New York, NY, USA

SUMMARY

As somatic cells are converted into iPSCs, their chromatin is remodeled to a pluripotent configuration with unique euchromatin-to-heterochromatin ratio, DNA methylation patterns and enhancer/promoter status. The molecular machinery underlying this process is largely unknown. Here, we show that ESC-specific factors Dppa2 and Dppa4 play a key role in resetting the epigenome to a pluripotent state. They are induced in reprogramming intermediates, function as a heterodimer and are required for efficient reprogramming of mouse and human cells. When co-expressed with OKSM factors, Dppa2/4 yield reprogramming efficiencies that exceed 80% and accelerate reprogramming kinetics, generating iPSCs in two to four days. When bound to chromatin, Dppa2/4 initiate global chromatin decompaction via the DNA damage response pathway, contribute to down-regulation of somatic genes and activation of ESC enhancers, all of which enables an efficient transition to pluripotency. Our work provides critical insights into how the epigenome is remodeled during acquisition of pluripotency.

Correspondence should be addressed to: Natalia Ivanova (natalia.ivanova@yale.edu), Yale Stem Cell Center, 10 Amistad Street, New Haven, CT 06520, Phone: 203-785-5957, FAX: 203-785-4305.

[#]Lead contact and corresponding author

AUTHOR CONTRIBUTIONS

CH and NBI conceptualized the study, designed experiments and wrote the manuscript. CH, ZW and BR performed most experiments with the help from CD, IK, MA (tissue culture), KT (RT-qPCR), YWL (ChIP-Seq), TAN (mass spectrometry) and SG (MEF stocks). YT and SM analyzed ChIP-Seq and RNA-Seq data.

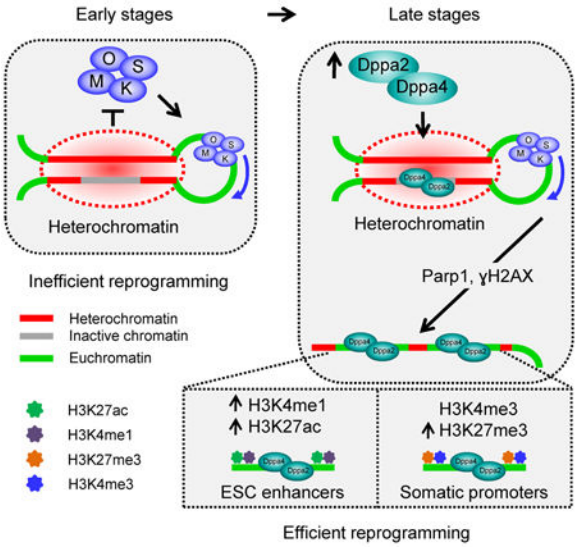
DECLARATION OF INTERESTS

The authors declare no competing interests.

In Brief

During cellular reprogramming the epigenome of a somatic cell is reset to a state compatible with pluripotency maintenance. The molecular machinery underlying this process remains poorly defined. Hernandez et al. identify chromatin-associated factors Dppa2 and Dppa4 as the key components mediating the reset of somatic chromatin to a pluripotent configuration.

Abstract



INTRODUCTION

Pluripotent stem cells (PSCs) can self-renew in culture while retaining the potential to form the full spectrum of cell lineages found in the body. Pluripotency can now be induced in fully differentiated somatic cells with four transcription factors: Oct4, Klf4, Sox2 and Myc (OKSM)(Takahashi and Yamanaka, 2006), yet the mechanistic understanding of the reprogramming process remains incomplete.

Reprogramming of mouse embryonic fibroblasts (MEFs) occurs over a period of 12-15 days and progresses through three phases. The initiation phase is characterized by a wave of transcriptional and epigenetic changes that result in the silencing of fibroblast-specific genes, an increase in proliferation rate, the mesenchymal-to-epithelial transition (MET), and changes in metabolism and cytoskeleton organization (Folmes et al., 2011; Li et al., 2010; Mathieu et al., 2014; Polo et al., 2012; Samavarchi-Tehrani et al., 2010). The maturation phase is marked by the gradual acquisition of early pluripotency markers such as SSEA1, Fbxo15 and Alpl, followed by a second wave of transcriptional and epigenetic remodeling that culminates in the activation of an endogenous pluripotency network capable of supporting transgene-independent growth (Golipour et al., 2012; Polo et al., 2012). During stabilization phase (day 12 and beyond) transgene-independent iPSCs reset DNA methylation profile, adjust telomere length and reactivate X chromosome in female iPSCs (Marion et al., 2009; Polo et al., 2012).

Reprogramming is inefficient even in cells with a uniformly high expression of reprogramming factors. The major rate-limiting event occurs at the end of the maturation phase. Indeed, while more than 90% of MEFs successfully convert into Thy1-SSEA1− and Thy1-SSEA1+ intermediates, only a small fraction of Thy1-SSEA1+ cells achieves stable pluripotency (Polo et al., 2012). Reprogramming efficiency can be enhanced via the modulation of specific pathways. Fast-cycling cells reprogram more efficiently (Guo et al., 2014) and the removal of cell-cycle checkpoints via inhibition of p53 or p21 increases the number of iPSC colonies (Hong et al., 2009; Kawamura et al., 2009; Utikal et al., 2009). Modulation of BMP4, TGF-β, and WNT pathways improves reprogramming efficiency through enhancement of MET (Li et al., 2010). Reprogramming efficiency can also be increased through activation of glycolysis or blockade of oxidative phosphorylation (Mathieu et al., 2014; Yoshida et al., 2009; Zhu et al., 2010). However, the aforementioned pathways primarily affect early stages of reprogramming and have only a modest effect.

In contrast, modulation of epigenetic pathways affects late reprogramming stages. The repressive heterochromatin mark H3K9me3 is enriched at pluripotency loci in somatic cells and has been shown to interfere with OKSM binding (Soufi et al., 2012). Depletion of this mark via knockdown of H3K9me3 methyltransferases Ehmt1/2 or Setdb1, or depletion of the H3K9me3 reader Cbx3, facilitates the transition from pre- to fully-reprogrammed iPSCs (Chen et al., 2013; Sridharan et al., 2013). Depletion of heterochromatic histone variant macroH2A which is detected at pluripotency loci in somatic cells, also results in more efficient reprogramming (Barrero et al., 2013). Knockdown of the histone chaperone CAF-1 significantly improves reprogramming efficiency and kinetics (Cheloufi et al., 2015). Collectively, these data support the long-held view that inefficient chromatin remodeling is the main bottleneck to reprogramming. However, critical chromatin-remodeling factors have not been identified.

Here, we show that chromatin-associated ESC-specific factors Dppa2 and Dppa4 function as key components of the chromatin remodeling network that governs the transition to pluripotency. These factors work as a heterodimer and both proteins must be present in order to efficiently bind and remodel chromatin. Dppa2/4 are required for successful reprogramming and—when co-expressed with the OKSM factors—greatly improve reprogramming efficiency and kinetics. Based on biochemical, bioinformatics, and functional data detailed below we propose a model for how Dppa2 and Dppa4 initiate epigenetic remodeling and facilitate a rapid and efficient transition to pluripotency.

RESULTS

Dppa2 and Dppa4 are required for reprogramming

Dppa2 and *Dppa4* are poorly characterized homologous genes expressed in pluripotent and germ cells (Maldonado-Saldivia et al., 2007). While both genes are dispensable for mESC maintenance (Madan et al., 2009; Nakamura et al., 2011), recent studies identified *Dppa2* among predictive markers of reprogramming whose expression labels rare cells that have a higher probability of becoming iPSCs (Buganim et al., 2012). To determine when *Dppa2* and *Dppa4* are expressed during reprogramming, we collected and analyzed RNA-seq data from reprogramming intermediates—as defined with a reprogrammable MEF model (Polo et

al., 2012; Stadtfeld et al., 2010). *Dppa2* and *Dppa4* were induced in Thy1-SSEA1+ pre-iPSCs and reached peak levels in Oct4-GFP+ iPSCs. At the protein level, both Dppa2 and Dppa4 were undetectable until the iPSC stage (Figure 1A-C). To determine whether Dppa2 and Dppa4 are required for reprogramming, shRNA pairs targeting both genes were co-transduced into the reprogrammable MEFs and cultures were maintained with doxycycline (Dox) for 12 days followed by 4 days of no Dox culture to select for fully reprogrammed iPSCs (Figure 1D, S1A, B). Although Dppa2/4 depletion in established ESCs/ iPSCs did not interfere with pluripotency maintenance (Figure S1C), Dppa2/4-KD reprogramming cultures yielded greatly reduced numbers of colonies, even when maintained in the presence of Dox for up to 20 days (Figure 1E). FACS analyses revealed that Dppa2/4-KD culture progression was arrested prior to the Thy1-SSEA1+ stage as evidenced by the accumulation of Thy1-SSEA1− cells and the lack of Thy1-SSEA1+ population (Figure 1F). To confirm these data, we derived Dppa2/4-KO iPSCs by CRISPR/Cas9 targeting (Figure S1E, F). These KO iPSCs, along with the parental iPSC line, were used to derive secondary (2^0) MEFs via blastocyst injections (Figure 1G). Similar to Dppa2/4-KD MEFs, 2^0 Dppa2/4-KO MEFs failed to reprogram. Transduction with single Dppa2 or Dppa4 transgenes did not rescue iPSC formation. However, reprogramming was fully restored when Dppa2 and Dppa4 were co-expressed from the same vector (Figure 1H, I, S1G, S2A). These data indicate that Dppa2 and Dppa4 complement each other during reprogramming and are required for transition to pluripotency.

Overexpression of Dppa2/4 accelerates the transition to pluripotency

We next tested whether Dppa2/4 could enhance reprogramming when added on top of OKSM factors. Dox-inducible vectors harboring single or double transgenes were transduced into the reprogrammable MEFs, sorted for transgene expression and cultured in the presence of Dox (Figure 2A, S2A, B). When overexpressed individually, Dppa2 and Dppa4 increased the number of colonies several fold. However, simultaneous overexpression of both Dppa2 and Dppa4 resulted in an additional gain in colony numbers, indicating cooperativity between the two factors (Figure 2B). In addition, the emergence of reprogramming intermediates was accelerated in overexpression cultures, with the most significant effect observed when both proteins were co-expressed together (hereafter referred as Dppa2/4^{OE}). SSEA1+ cells were present as early as day 2, while Oct4-GFP+ cells were detected at day 4 compared with day 8 and day 12 in cells transduced with empty vector (hereafter referred as Control^{OE}) (Figure 2C). While overexpression of Dppa2/4 was not sufficient to reprogram MEFs without OKSM factors (data not shown), double dosage of the factors in homozygous M2rtTA+/+; Col1a-OSKM+/+ MEFs further improved the kinetics of Dppa2/4^{OE} reprogramming (Figure 2C). Enhanced reprogramming was also observed with primary lentiviral transductions (Figure S2C).

To confirm that Dppa2/4^{OE}-iPSCs are functional, we evaluated the ability of iPSC clones to give rise to the germ layer derivatives *in vitro and in vivo*. Dppa2/4^{OE}-iPSCs exhibited correct pattern of marker expression when differentiated in embryoid bodies (Figure S2D-G) and gave rise to germline-competent chimeric mice via blastocyst injection and tetraploid complementation (Figure 2D, Tables S2, S3).

To test whether overexpression of DPPA2/4 could enhance human reprogramming we introduced the DPPA2/4^{OE} vector into iOKSM, TERT-immortalized 2^{0} human fibroblasts (Cacchiarelli et al., 2015) and quantified reprogramming after 38 days. A significant increase in AP+ colony numbers was observed in DPPA2/4^{OE} cultures (Figure 2E). Increase in colony numbers was also observed when rtTA and DPPA2/4^{OE} vector-transduced primary human BJ cells were reprogrammed with Sendai viral reprogramming factors (Figure 2F, G). As in the mouse model, DPPA2/4^{OE}-iPSCs exhibited the correct pattern of pluripotency markers and had the ability to give rise to endoderm- and neuroectoderm-like cells (Figures 2H, I). Taken together, these data identify Dppa2 and Dppa4 as regulators of cellular reprogramming.

Dppa2/4 boost reprogramming through all stages up to fully reprogrammed iPSCs.

To evaluate reprogramming kinetics in the presence of Dppa2/4, we monitored reprogramming cultures using live imaging. The cultures were set up as shown in Figure 2A and imaged for 14 days with 2-hour increments between frames. In Control^{OE} cultures, Oct4-GFP+ cells emerged by day 11 and Oct4-GFP+ colonies formed by day 14. In contrast, the first Oct4-GFP+ cells in the Dppa2/4^{OE} cultures were detectable 24 hours after plating. The majority of these cells gave rise to stable Oct4-GFP+ colonies by day 4. The emergence of Oct4-GFP+ colonies was accompanied by the silencing of the mCherry reporter, an additional indication of stable reprogramming (Figure 3A, B). To determine the minimal interval of time needed for the acquisition of stable pluripotency, Dppa2/4^{OE} OKSM-MEFs were induced with Dox, which was then withdrawn at different time points for the rest of the time course (Figure 3C, top). In agreement with the live imaging data, two days of transgene expression were sufficient to reprogram a fraction of cells in Dppa2/4^{OE} cultures, while four days were sufficient to reprogram the majority of cells (Figure 3C). To further confirm that early GFP+ colonies are not transient events but stably reprogrammed iPSCs, we removed Dox after 2 or 4 days of reprogramming and monitored the cultures for additional 3 days. GFP+ colonies observed at the time of Dox removal were also present after 3 additional days of no Dox culture, suggesting that early GFP+ colonies in Dppa2/4-OE cultures are indeed fully reprogrammed iPSCs (Figure S3A).

To quantify the efficiency of reprogramming, single Dppa2/4^{OE} cells were sorted into wells of 96-well plates and emerging iPSC colonies counted at days 4, 8 and 12. While only 10% of Control^{OE} cells gave rise to Oct4-GFP+ iPSCs, the reprogramming efficiency of Dppa2/4^{OE} cells exceeded 80% (Figure 3D).

To identify reprogramming stages that respond to Dppa2/4 overexpression, rtTA+/+ MEFs were co-transduced with the constitutive OK-ires-SM (Sommer et al., 2012) and the Dox-inducible Dppa2/4^{OE} vectors. Dox was then added for different intervals of time (Figure S3B). Surprisingly, all experimental groups exhibited significant enhancement of reprogramming. However, the effect was more pronounced when Dppa4/2 were overexpressed during the late reprogramming stages (Figure S3B). To measure the responses of reprogramming intermediates to Dppa2/4 overexpression, we FACS-purified Thy1+SSEA1−, Thy1-SSEA1− and Thy1-SSEA1+ cells from day 9 reprogramming cultures, transduced them with Dppa2/4^{OE} and Control^{OE} vectors, and allowed to reprogram

for another six days. Dppa2/4^{OE} rescued the reprogramming of the refractory cells and significantly enhanced reprogramming of intermediates. Strikingly, the Dppa2/4^{OE} vector had no effect on colony formation when transduced into Oct4-GFP+ iPSCs (Figure 3E). Based on these data we concluded that Dppa2/4 function by removing reprogramming barriers that are present throughout all reprogramming stages up to fully reprogrammed iPSCs.

Dppa2/4 facilitates reprogramming via activation of DNA damage response pathway

Previous studies suggested that Dppa2/4 interact with DNA and/ or chromatin (Masaki et al., 2010; Siegel et al., 2009). We, therefore, hypothesized that Dppa2/4 may facilitate the remodeling of somatic chromatin, making it more permissive for the induction of pluripotency. We used confocal microscopy to examine global levels of several histone modifications known to be elevated in pluripotent cells compared to MEFs (Figure 4B) (Mattout et al., 2011). Assays were performed at day 2, 4 and 12 of reprogramming to compare the early and the late changes. Indeed, elevated nuclear levels of H3K4me3, H3K27ac and H3K27me3 were observed in Dppa2/4^{OE} cells at days 2 and 4 (Figure 4A). A significant increase was also observed for phosphorylated histone H2AX (ɣH2AX), a DNA damage-associated mark which is induced in late reprogramming intermediates (Gonzalez and Huangfu, 2016). While H3K9me3 levels exhibited a slight increase, the number of H3K9me3-positive heterochromatic foci was significantly reduced in Dppa2/4^{OE} cells compared to Control^{OE} cells. A similar reduction was observed when heterochromatic regions were quantified based on DAPI staining (Figure 4C, D). These data suggest that Dppa2/4 may exert its effect on reprogramming via remodeling of the chromatin.

Highly elevated levels of ɣH2AX and a positive correlation of ɣH2AX and Dppa2/4 binding profiles in Dppa2/4^{OE} reprogramming cells (Figures 5A) prompted us to examine a potential link between the DNA damage response (DDR) pathway and Dppa2/4-mediated chromatin remodeling. Previous studies demonstrated that during reprogramming ɣH2AX levels increase in the Thy1-SSEA1+ cells (Gonzalez and Huangfu, 2016), coincidental with the induction of *Dppa2/4* expression (Figure 1A, B). However, Dppa2/4^{OE}-MEFs that did not express reprogramming transgenes had low levels of ɣH2AX, suggesting that input from the OKSM factors is required synergistically with Dppa2/4 in order to elicit ɣH2AX response (Figure 5C). Notably, Dppa2/4^{OE}-MEFs also exhibited minimal—if any—loss of DAPI/ H3K9me3-positive heterochromatic foci compared to Dppa2/4^{OE} OKSM-MEF cultures (Figure 5D), also indicating that OKSM-mediated activation of the DDR pathway and/or ɣH2AX itself could be required for chromatin remodeling downstream of Dppa2/4. Indeed, when H2AX was depleted by shRNAs the loss of DAPI/ H3K9me3-positive foci in Dppa2/4^{OE} OKSM cultures was no longer detected (Figure 5B, E, F). Knockdown of H2AX also reduced the number of iPSC colonies in Dppa2/4^{OE} reprogramming cultures (Figure 5G). Chromatin remodeling and reprogramming efficiency were also negatively affected when Dppa2/4^{OE} reprogramming cultures were treated with NU7026, a specific inhibitor of protein kinase DNA-PK which mediates H2AX phosphorylation (Veuger et al., 2003). Interestingly, chemical inhibition of ATM, another canonical mediator of H2AX phosphorylation (Hickson et al., 2004), had a lesser effect on ɣH2AX levels and did not significantly affect Dppa2/4-mediated chromatin remodeling at the early time point (Figure

S4A-C). To better understand when DNA-PK and ATM are required during reprogramming, we quantified reprogramming efficiencies in Dppa2/4^{OE} and Control^{OE} cultures treated with DNA-PK or ATM inhibitors for specific intervals of time. Inhibition of DNA-PK had a strong negative effect on reprogramming efficiency when applied to early and intermediate stages, whereas inhibition of ATM had most effect when applied to intermediate stages (Figure S4C). This data suggest that the initial stages of reprogramming are dependent on DNA-PK, the intermediate stages utilize both DNA-PK and ATM whereas the late stages are less sensitive to either DNA-PK or ATM inhibition. The redundant usage of DNA-PK and ATM during intermediate stages could potentially explain why inhibition of DNA-PK or ATM alone does not significantly affect reprogramming in Control^{OE} cultures (Figure S4C).

We next tested whether overexpression of H2AX could mimic Dppa2/4^{OE} reprogramming phenotype. While we did observe an increase in γH2AX levels in H2AX^{OE} cultures, the reprogramming efficiency was lower than that in Control^{OE} cultures (Figure S4D-G). Thus, H2AX/ γH2AX deposition outside the Dppa2/4-bound domains may inhibit reprogramming. It has recently been shown, for example, that high levels of H2AX in ESCs reduce rate of cell proliferation through inhibition of rRNA transcription (Eleuteri et al., 2018). We also did not observe the loss of H3K9me3 foci in H2AX^{OE} cells, suggesting that γH2AX deposition is required but not sufficient for chromatin decompaction observed in the presence of Dppa2/4 (Figure S4G).

To determine which reprogramming factors are required for Dppa2/4-dependent γH2AX/ H3K9me3 response we overexpressed Dppa2/4 on top of individual factors, and their double and triple combinations and quantified γH2AX levels and H3K9me3 foci number by confocal microscopy. While none of the single factors or double factor combinations induced robust γH2AX deposition, a strong γH2AX response was observed with OSM cells, whereas OSK, OKM and SKM cells exhibited lower γH2AX levels and retained significantly more H3K9me3 foci (Figure S5A,B). Taken together, our data suggest that γH2AX deposition in Dppa2/4^{OE} reprogramming cells is mediated via a combined action of Oct4, Sox2 and Myc and is required for chromatin remodeling and enhanced reprogramming.

γH2AX deposition, chromatin remodeling and enhanced reprogramming in Dppa2/4-overexpressing cells is dependent on Parp1

Dppa2/4 are small proteins that lack enzymatic domains and thus could act as scaffolds for the recruitment of chromatin modifiers to chromatin. We sought to isolate Dppa2/4-interacting proteins from ESCs where these factors are naturally expressed. We found, however, that at least 800 mM NaCl was needed to extract Dppa2/4 from the nucleus. On the other hand, Dppa2 and Dppa4 were readily solubilized when nuclei were pre-treated with micrococcal nuclease (MNase). Dppa2/4-nucleosome complexes migrated in the 400-660 kDa range (Figure 6A, B). We purified Dppa2/4-bound complexes from ESCs expressing 3xFLAG-Dppa2 transgene using anti-FLAG affinity beads and identified co-purifying proteins by mass spectrometry. Dppa4 was recovered in these IPs in an equimolar ratio to Dppa2 (Figure 6C). Apart from the nucleosomal histones, several chromatin modifiers were identified, including H3K36me2 demethylase Kdm2a, the ISWI ATPase Smarca5, the ATP-

dependent DNA helicases Ruvbl1/2, the poly-ADP-ribose polymerase Parp1, the histone acetyltransferase Myst4 and the H3K4 methyltransferase Mll4 (Table S4). Select interactions were confirmed by western blotting (Figure 6D). Importantly, these Dppa2/4-associated chromatin modifiers were expressed in OKSM-MEFs, suggesting that they may function with Dppa2/4 during reprogramming (Figure 6E).

In view of our γH2AX data, we hypothesized that Parp1 and Smarca5 which have been reported to function in the DDR pathway (Smeenk et al., 2013; Strickfaden et al., 2016) may be required for the Dppa2/4-mediated chromatin remodeling and reprogramming. Indeed, shRNA depletion of either Parp1 or Smarca5 in Dppa2/4^{OE} cultures resulted in a significant decrease in iPSC numbers (Figure 6F, G, S5C). Depletion of Parp1 also reduced γH2AX levels and prevented the loss of DAPI/H3K9me3-positive foci similar to what was observed upon depletion of H2AX (Figures 6H, I) while depletion of Smarca5 did not affect γH2AX levels/ heterochromatin remodeling (Figure S5D, E).

Taken together, these data suggest that chromatin remodeling induced by overexpression of Dppa2/4 is mediated by the components of the DDR pathway and Parp1 which is induced in response to reprogramming factor expression.

Dppa2/4 binding correlates with epigenetic remodeling of promoter and enhancer regions

To evaluate the impact of Dppa2/4 overexpression on the transcriptome, we analyzed reprogramming cultures using RNA-seq. We sorted and profiled the canonical reprogramming populations from Dppa2/4^{OE} and Control^{OE} cultures at day 2, 4, 8, 11 and 14 along with the Thy+SSEA1+ cells which were observed at days 2 and 4 only in Dppa2/4^{OE} cultures (Figure S6A). Heterozygous OSKM+/−; rtTA+/−; Oct4-GFP+/− MEFs, which exhibit slower reprogramming kinetics, were used in RNA-Seq experiments to better resolve reprogramming intermediates.

Principal component analysis of RNA-Seq data and the FACS data suggested that over 50% of cells in Dppa2/4^{OE} cultures progressed via a distinct Thy1+SSEA1+ intermediate prior to conversion into Thy1-SSEA1+ cells (Figure 7A, S6A). To test whether this is indeed the case, the reprogramming capacity of Thy1+SSEA+ cells was compared against the canonical Thy1-SSEA1− and Thy1-SSEA1+ intermediates sorted from Dppa2/4^{OE} culture at the same time. We found that when cultured with Dox Thy1+SSEA1+ cells generated the same number of colonies as the canonical Thy1-SSEA1− intermediate and ~40% less than Thy1-SSEA+ cells (Figure S6B).

We next compared transcriptional profiles associated with the alternative and the canonical reprogramming trajectories. Hierarchical clustering of genes with significant changes in expression along each trajectory identified seven clusters of differential expression comprising a total of 2142 genes (Figure 7B, C). The alternative trajectory was characterized by faster down-regulation of MEF genes representing secreted and extracellular matrix proteins, focal adhesion and PI3K-Akt signaling (cluster i) and early induction of DNA damage response and mitochondrion-associated genes (clusters iii). Genes with functions in DNA replication, cell cycle and p53 signaling were transiently down-regulated in the canonical reprogramming intermediates but retained their expression in the alternative

trajectory (cluster iv), whereas lysosome and adherens junction associated genes transiently induced in the canonical intermediates were expressed at lower levels in the alternative trajectory (cluster vii). Unexpectedly, a cohort of early-induced iPSC genes exhibited delayed induction in the alternative trajectory (cluster vi) (Figure 7B, C). Additionally, 270 genes were induced in Dppa2/4^{OE} but not in Control^{OE} cultures (cluster v). Induction of the DDR pathway genes in the alternative reprogramming trajectory connects well with our γH2AX data. Therefore, other molecular pathways selectively up- or down-regulated in the alternative trajectory could also contribute to the enhanced reprogramming in Dppa2/4^{OE} cells.

We next mapped the binding of Dppa2 and Dppa4 in Dppa2/4^{OE}-expressing MEFs without OKSM factors (MEF-Dppa2/4^{OE}), in day 3 Dppa2/4^{OE}-expressing OKSM-MEFs undergoing reprogramming (R-Dppa2/4^{OE}) and in WT-ESCs using ChIP-Seq. We found that Dppa2 and Dppa4 exhibited extensive and highly correlated binding in these three cell types (Figure 7D, S6C, D, F). Dppa2/4-bound regions observed in MEF-Dppa2/4^{OE} showed significant overlap with and represented a subset of regions bound by Dppa2/4 in R-Dppa2/4^{OE} cells and ESCs. When expressed alone neither protein could bind chromatin efficiently (Figure 7D, E). No Dppa4 binding peaks were detected in MEF-Dppa4^{OE} cells and only 1,303 Dppa2 peaks were detected in MEF-Dppa2^{OE} cells whereas over 40,000 peaks were detected for either Dppa2 or Dppa4 in MEF-Dppa2/4^{OE} cells (Figure S6D). The motif search using Dppa2-only peak set returned 87 instances of (CGG)×8 sequence, suggesting that Dppa2 could be recruited to a subset of CpG islands independently of Dppa4 (Figure S6E). Dppa2 and Dppa4 binding correlated positively with gene density and was observed in both coding and non-coding regions of the genome with a bias toward introns and intergenic regions (Figure 7E, F). When correlated with the 35 chromatin trajectories that capture chromatin dynamics through canonical reprogramming stages (Chronis et al., 2017), Dppa2/4 binding was preferentially detected at active and Polycomb-repressed promoters (tr.E1 and E4). Strikingly, Dppa2/4 binding was also observed at ESC-specific enhancers (tr. E18) whereas binding to MEF or pre-iPSCs enhancers was minimal (Figure S6G).

Our confocal analyses detected elevated levels of several histone marks in Dppa2/4^{OE} cultures (Figure 4A). To determine whether these changes correlate with Dppa2/4 binding, we mapped genomic distributions of H3K4me3, H3K27me3, H3K27ac and H3K4me1 in MEFs, ESCs and R-Dppa2/4^{OE} and R-Control^{OE} cells at day 3 of reprogramming. We then compared the levels of these mark for Dppa2/4-bound elements in each trajectory in R-Dppa2/4^{OE} and R-Control^{OE} cells. Within the E1 trajectory Dppa2/4 binding was selectively detected at the elements that gained H3K27me3 in R-Control^{OE} cells and the median H3K27me3 levels of these elements were significantly higher in R-Dppa2/4^{OE} cells (Figure 7G-I, S7A), suggesting that these promoters are silenced more efficiently in the presence of Dppa2/4. Indeed, genes from the RNA expression clusters i and vi exhibited large overlap with Dppa2/4-bound E1 elements (Figure 7J, C). Dppa2/4-bound elements in the E2 and E3 trajectories (weak active and repressed promoters) exhibited higher H3K4me3 but lower H3K27me3 levels in Dppa2/4^{OE} cells compared to Control^{OE} cells. However, over 80% of these elements were located away from the annotated TSS regions and their associated genes did not overlap with any of the seven RNA clusters. The elements of the E4 trajectory

(Polycomb-repressed promoters) exhibited elevated H3K27me3 levels in Dppa2/4^{OE} cells and a small fraction of them (460 elements) gained H3K4me3 (Figure 7H, S7A). In agreement with their repressed state, genes located next to these E4 elements did not contribute to any of the RNA expression clusters. Lastly, the median levels of H3K4me1 and H3K27ac marks at Dppa2/4-bound elements within E18 trajectory (ESC-specific enhancers) were higher in R-Dppa2/4^{OE} cells compared to R-Control^{OE} cells, suggesting that a subset of ESC-specific enhancers is activated early in R-Dppa2/4^{OE} cells (Figure 7H, S7A).

To test whether the loss of H3K9me3/DAPI-dense heterochromatic foci in Dppa2/4^{OE} cultures (Figure 4C, D) occurs due to changes in H3K9me3 deposition, we mapped this mark in day 3 reprogramming cultures. Although analysis of MNase-digested material indicated that both total and K9-enriched chromatin from Dppa2/4^{OE} cells was less compacted than the chromatin from Control^{OE} cells (Figure S7B, C), the overall H3K9me3 levels/ deposition patterns were not significantly different between the two samples (Figure S7D). The majority of Dppa2/4 clusters mapped to H3K9me3-low regions. However, sparse binding was also observed within large H3K9me3-enriched domains. H3K9me3-high Dppa2/4 peaks were highly enriched for transposable elements of LINE and LTR families. Furthermore, a subset of these sites that gained H3K4me3 and H3K27me3 in R-Dppa2/4^{OE} cells was enriched for LINE transposons (Figure S7E, F).

Taken together, our data suggest that Dppa2/4 could enhance reprogramming via three distinct mechanisms. Dppa2/4 binding to E1 promoters could lead to a faster inactivation of somatic genes and delayed induction of early-induced iPSC genes, whereas binding to a subset of ESC-specific enhancers could facilitate their activation. Additionally, Dppa2/4 binding to transposable elements may contribute to chromatin remodeling of H3K9me3-enriched regions.

DISCUSSION

Inefficiency of epigenetic remodeling has long been recognized as the major bottleneck that limits reprogramming efficiency but remains poorly understood at the mechanistic level. Here, we identify Dppa2 and Dppa4 as key components mediating the reset of the epigenome during reprogramming.

We show that Dppa2 and Dppa4 work as a heterodimer – both proteins have to be expressed simultaneously in order to bind chromatin. Only about 3% of Dppa2/4 binding sites are bound by Dppa2 in Dppa2^{OE} MEFs. These Dppa2-only sites are enriched for CGG repeats and could represent a subset of CpG islands. Notably, Kdm2a, a H3K36me2 demethylase with CXXC DNA binding domain that recognizes unmethylated CpG islands, was recovered in Dppa2-IPs, suggesting that some genomic sites could utilize Kdm2a for Dppa2/4 recruitment.

Accelerated reprogramming in Dppa2/4^{OE} OKSM-MEFs is accompanied by highly elevated γH2AX levels. Increase in γH2AX levels normally occurs in late reprogramming intermediates and has been attributed to several potential causes, including oncogenic activities of reprogramming factors and global DNA demethylation that may produce DNA

intermediates processed via the DNA repair pathway (Gonzalez and Huangfu, 2016). However, in pluripotent cells high basal γH2AX levels are thought to be associated with global chromatin decompaction rather than pre-existing DNA damage (Banath et al., 2009) and have been suggested to play a role in the maintenance of the pluripotent state (Turinetto et al., 2012). Consistent with the latter view, we find that reducing γH2AX levels with shRNAs against H2AX or with DNA-PK inhibitor eliminates the beneficial effect of Dppa2/4 on chromatin remodeling and reprogramming efficiency. Our data, therefore, point to the essential role of DDR pathway in chromatin remodeling and prompt for detailed investigation into the potential mechanism.

We find that in ESCs Dppa2/4 associate with a set of chromatin remodelers many of which are also expressed in OKSM-MEFs and therefore may contribute to high reprogramming efficiency in Dppa2/4^{OE} cultures. Indeed, shRNA depletion of either Parp1 or Smarca5 eliminates the gain in reprogramming efficiency in Dppa2/4^{OE} cultures. Furthermore, Parp1 is required for Dppa2/4-mediated chromatin remodeling. Previous work demonstrated that Parp1 is required for efficient iPSC generation and Parp1 binding to the Nanog locus precedes Nanog activation during reprogramming (Doege et al., 2012). Our new data suggests a broader role for Parp1 in the context of Dppa2/4-mediated chromatin remodeling which is supported by prior observation that Parp1 may serve as the key component of DNA repair machinery responsible for chromatin decompaction (Strickfaden et al., 2016).

Our transcriptome analyses show that a large fraction of cells in Dppa2/4^{OE} cultures progresses via an alternative reprogramming trajectory which is characterized by faster down-regulation of somatic genes, many of which are directly bound by Dppa2/4 at the promoters and exhibit higher H3K27me3 levels in Dppa2/4^{OE} cells. In addition, a cohort of early-induced iPSC genes also acquire H3K27me3 marks at promoters and exhibit delayed induction in the alternative trajectory suggesting that suppression of these genes in reprogramming intermediates could be beneficial for reprogramming efficiency. Likewise, rapid induction of mitochondrion-associated genes in Dppa2/4^{OE} cultures indicates that modulation of mitochondrial activity could enhance reprogramming and fits well with previous work showing that activation of glycolysis or blockade of mitochondrial oxidative phosphorylation improves reprogramming efficiency (Zhu et al., 2010). In addition, Foxh1, a transcription factor transiently expressed in reprogramming intermediates of the alternative but not the canonical trajectory has been shown to facilitate reprogramming of human cells (Takahashi et al., 2014). We, therefore, anticipate that other members of this gene cohort could also improve reprogramming efficiency.

Apart from their role in reprogramming established in this study, Dppa2/4 are expressed during preimplantation development. Inactivation of either Dppa2 or Dppa4 causes developmental defects and lethality in both mice and frogs (Nakamura et al., 2011; Siegel et al., 2009). Furthermore, it has been suggested that both Dppa2 and Dppa4 could contribute to the development of tumors (Shabestarian et al., 2015). Thus, the mechanism presented here could provide a foundation for understanding the functions of the Dppa2/4 family in developmental and oncological contexts.

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Natalia Ivanova (natalia.ivanova@yale.edu), Department of Genetics, Yale Stem Cell Center, Yale University, New Haven, CT 06519.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The Institutional Animal Care and Use Committee of Yale University has approved our animal breeding and research protocols. Animals were used for isolation of cultures of primary cells from mice. Mouse embryonic fibroblasts (MEFs) harboring the M2rtTA transgene in the Rosa26 locus together Dox-inducible polycistronic cassette coding for OKSM in the *Col1A* locus (tetO-OKSM) and GFP in the Oct4 locus were used for mouse reprogramming experiments, ChIP-seq, and RNA-seq assays. Human TERT-immortalized secondary fibroblasts with inducible genome-integrated OSKM factors (hiF-T) were used for human reprogramming experiments. All cell lines are described in the Key Resources Table.

METHOD DETAILS

Plasmids and cloning

shRNA vectors: shRNAs were designed using the online siRNA design tool from Dharmacon (http://www.dharmacon.com/designcenter) with default parameters. shRNAs with potential off-target effects were filtered out using BLAST against mouse genome. Sense and antisense oligonucleotides of the shRNA duplex were synthesized and cloned into an in-house designed FUW-H1P-Hygro-mCherry vector. Multiple shRNAs were tested for each gene and shRNAs with knockdown efficiencies of seventy percent or more were used in the analyses. Sequence information for shRNAs used in this study is shown in Table S6.

sgRNA vectors: sgRNAs were designed using an online CRISPR design tool (http://crispr.mit.edu/). Sense and antisense oligonucleotides of the sgRNA duplex were synthesized and cloned into the pSpCas9n(BB)-2A-GFP vector. Sequence information for gRNAs used in this study is listed in Table S6.

Overexpression vectors: Dppa2 and Dppa4 cDNAs were PCR-amplified from ESCs and subcloned into an in-house FUW-pTRE-2A-mCherry or FUW-pTRE-2A-BFP vectors, in-frame with fluorescence reporters. Dppa2-2A-Dppa4 double transgene was generated by PCR and subcloned in the same vector backbones. Coding regions were verified by sequencing. DPPA2/4 overexpression vectors for reprogramming of human cells and FLAG-H2AX overexpression vector were constructed using the same strategy.

Cell culture

Lentivirus packaging and transduction: 15×10^6 HEK293FT cells were seeded onto gelatin-coated 15 cm tissue culture dish in DMEM media supplemented with 10% FBS, L-glutamine, penicillin-streptomycin, non-essential amino-acids, sodium pyruvate and β-mercaptoethanol. Cells were transfected the following day. For transfection, 9.5 μg pVSV-G,

18.9 μg pCMVdR8.2 and 7.5 μg of overexpression or shRNA vector were diluted in 2ml DMEM and vortexed. 67.5 μl of PEIMax (1mg/ml) was added to the plasmid mix, vortexed, incubated for 10 min at room temperature and added to cells. Next day media was replaced with fresh media and cells were maintained for 48 hours. Culture media was collected, filtered through 0.45 μm filter to remove debris. Virus was concentrated by ultracentrifugation at 50,000g for 1.5 hours, re-suspended in PBS, and stored at −80°C. For the lentiviral transduction, cells were plated on gelatin-coated dishes and allowed to attach overnight. Virus was added at MOI=5 to the maintenance media supplemented with Polybrene at a concentration 8 μg/ml. Cells were incubated with virus overnight, washed with PBS and fed with fresh medium.

MEF culture: Primary MEFs were isolated from 13.5-14dpc embryos. Prior to lentiviral transduction, cells were cultured in DMEM supplemented with L-glutamine, Na-pyruvate, β-mercaptoethanol, penicilin-streptomycin and 10% FBS for one or two passages.

mESC/iPSC culture: Cells were maintained on Mitomycin C treated MEF monolayers in DMEM supplemented with 15% ESC-grade FBS, L-glutamine, penicillin-streptomycin, non-essential amino-acids, Na- pyruvate, β-mercaptoethanol and ESGRO (10^3units/ml). Cells were split every two days.

CRISPR-mediated gene knock-out.—10^7 iPSCs were resuspended in 800 μL cold PBS, and mixed with 40 μg of CAS9-gRNA vector. Cells were electroporated using a Gene Pulser Xcell (Bio-Rad) at 250 V, 500 μF in a 0.4 cm Gene Pulser cuvette (Bio-Rad), and replated on irradiated CF1 MEFs. 2 days after electroporation, GFP+ cells single-cell were sorted and 10^4 cells were plated onto a 10cm dish with irradiated CF1 MEFs. Individual colonies were picked and genomic DNA was then purified using the KAPA Express Extract Kit (Kapa Biosystems). Mutations were validated by PCR and DNA sequencing. Protein expression was analyzed by western blot.

Reprogramming of MEFs: Reprogrammable MEFs were plated onto Mitomycin C treated MEF monolayers prepared from CD1 embryos. Reprogramming was performed in DMEM supplemented with L-glutamine, Na-pyruvate, β-mercaptoethanol, penicillin-streptomycin, 15% ESC-grade FBS and 10^3 units/ml of ESGRO. Doxycyclin (Dox) was added at concentration μg/ml. Cells were reprogrammed for 12 days and maintained in the same media without Dox for additional 4 days to select for fully reprogrammed iPSCs. To confirm pluripotent status, cells were immunostained with Nanog, Oct4 and SSEA1 antibodies.

Derivation of secondary MEFs: Dppa2/4-KO iPSCs (M2rtTA+/+; Col1a-OSKM+/+; Oct4-GFP+/+; PuroR) were injected into C57Bl/6 blastocysts, implanted into recipients and embryos collected 13.5-14 days after surgery. MEFs were isolated and cultured in MEF media supplemented with Puromycin (2.5 μg/ml) to establish pure Dppa2/4-KO MEFs for reprogramming.

In vivo assays: Blastocyst injections and tetraploid complementation assays were performed by the Yale Genome Editing Center according to IACUC-approved protocols using standard techniques.

Embryoid body (EB) differentiation: iPSCs were dissociated and plated onto tissue culture plates for 30 min to deplete feeder MEFs. Floating cells were collected and transferred into Petri dishes at concentration 0.75×10^6 cells in 5ml of media. Cultures were maintained in DMEM supplemented with L-glutamine, Na-pyruvate, β-mercaptoethanol, penicillin-streptomycin and 15% FBS prescreened for EB differentiation. Day 6 EBs were re-plated onto gelatin-coated plates, grown for 6 days, and fixed with 4% paraformaldehyde for marker analyses.

Live imaging: Reprogrammable MEFs were plated onto a 35mm Ibidi culture dish with a layer of mitotically inactivated MEFs prepared from CD1 embryos. Dishes were placed in Olympus VivaView LCV110U live imager where grid-format images were taken of bright field, GFP, and TRITC channels every 2 hours for 14 days. Images were processed in ImageJ to stitch individual fields and to enhance contrast.

Reprogramming of secondary human fibroblasts: TERT-immortalized secondary fibroblasts with inducible genome-integrated OSKM factors (hiF-T) (Cacchiarelli et al., 2015). Prior to reprogramming cells were maintained in DMEM/F12 supplemented with L-glutamine, Na-pyruvate, β-mercaptoethanol, penicilin-streptomycin, 10% FBS bFGF at 8 ng/ml and puromycin at 0.5 μg/ml. Cells were transduced with DPPA2/4^{OE} and Control^{OE} vectors at MOI=5. Reprogramming was performed on irradiated MEF monolayers in DMEM/F12 supplemented with L-glutamine, Na-pyruvate, β-mercaptoethanol, penicillin-streptomycin and 20% KSR. Dox was added at concentration 2 μg/ml. bFGF was added at concentration 4 ng/ml. iPSCs were cultured in KSR-based media on irradiated MEF monolayers or in mTeSR1 media on Matrigel. Cultures were split every 5-6 days by adding 1 mg/ml Dispase for 5–10 min at 37ύC, dissociation, and replating onto fresh plate at the ration 1:4. To confirm pluripotent status, cells were immunostained with NANOG, Tra1-60 and SSEA4 antibodies.

Sendai virus reprogramming: Human BJ fibroblast reprogramming was performed with the Cytotune 2.0 kit (Invitrogen) according to the manufacturer's instructions. DPPA2/4^{OE} and Control^{OE} transduced cells were seeded at ~2×10^5 cells per well of a 6-well plate in MEF medium. Two days after, cells were transduced with Sendai viruses in MEF medium at MOI=5 for KOS and MOI=4 for c-MYC. After 24 hours, the medium was replaced with fresh MEF medium and changed every other day thereafter. On day 7, cells were trypsinized, counted and seeded onto a layer of irradiated MEF feeders in MEF medium. The next day, the cells are transferred to human iPSC culture media supplemented with Dox at 2μg/ml. Daily media change was performed thereafter. 26 days after transduction, cells were immunostained with NANOG, Tra 1-60 and SSEA4 antibodies.

In vitro hiPSCs differentiation: For the neuroectoderm differentiation, cells were treated with 15 μM SB431542, 0.25 μM PD0325901 and 1 μM dorsomorphin in neural induction media containing DMEM/F12, 1xN2, 1xB27, 0.1 mM β-mercaptoethanol, non-essential

amino acids, L-glutamine and penicillin-streptomycin for 7 days. Cells were fixed with paraformaldehyde and stained with PAX6 antibody. To induce endoderm differentiation, hESCs were treated with 100 ng/ml Activin A in advanced RPMI supplemented with L-glutamine and 0.2% FBS. Four days after, cells were fixed with paraformaldehyde and stained with the SOX17 antibody.

Assays

Alkaline phosphatase assay: Cells were fixed with 4% paraformaldehyde, washed with PBS and developed with Alkaline Phosphatase Detection Kit (Sigma) according to manufacturer's instructions.

Gamma H2AX inhibition ––—Inhibitors NU7026 (DNA-PK) and KU55933 (ATM) were used at a concentration of 10μM in DMSO. Inhibitors were added to reprogramming media, which was changed daily. Equal volume of DMSO was added to no inhibitor control culture.

Flow Cytometry: Cells were dissociated into single-cell suspension, blocked in PBS with 5% of newborn calf serum and incubated with fluorescently-labeled antibodies for 20 min on ice. Cells were washed, filtered and analyzed on BD LSRII instrument. Cell sorting was performed on BD Aria instrument. Cells were collected in media supplemented with 10% FBS, spun down and counted prior to plating into assays. Single cells were sorted directly into 96-well plates pre-plated with confluent mitotically inactive MEFs.

Immunofluorescence staining.—Cells were fixed in 4% paraformaldehyde. Slides were blocked in PBS containing 5% normal donkey serum (Millipore) and 0.3% Triton X-100 for 2 hours at room temperature, incubated with primary antibodies overnight at 4°C, washed, incubated with secondary antibodies for 2 hours at room temperature, washed, cover-slipped and imaged on the TCS SP5 AOBS Confocal Microscope (Leica). The list of antibodies used in this study is shown in Table S7.

Image quantification.—Images were taken on Leica TCS SP5 confocal and quantified using ImageJ software. Nuclei were marked using the oval tool and their integrated density measured using the measure tool. Dppa2/4 expressing cells were identified using Dppa2 or Dppa4 antibody and control cells were identified using mCherry fluorescence. Integrated Density parameter was taken for both Dapi and histone mark expression in both Dppa2/4 and Control MEFs. Background integrated density was subtracted from analysis by measuring regions without cells and Dapi was used as a normalizing control for all samples. Each sample had 100 cells measured, quantified, and plotted as dot plots. Cohen's effect size d and Student's t-test p-value are shown for each comparison and marked according the effect size as ns - very small (d>0.01), *small (d>0.2), **medium (d>0.5), ***large (d>0.8), ****very large (d>1.2) and *****huge (d>2.0).

Heterochromatic foci quantification.—Dapi/H3K9me3 foci were manually counted from Z-stacked confocal images and scored for total foci per cell. 100 cells were scored in Dppa2/4^{OE} and Control^{OE} cultures. Foci distribution was then plotted and Gaussian curves

were fitted over the data using Graphpad (Prism7) software to calculate the means of foci number ±SEM.

Stepwise salt extraction.—10×10^6 cells were resuspended in hypotonic buffer (10 mM HEPES, 10 mM KCl, 1.5 mM MgCl2, 0.34 M sucrose, 10% glycerol, 0.2 mM PMSF, 1 mM DTT, 1x proteinase inhibitor) on ice for 10 min. Nuclear pellets were treated with or without 2 U/ul MNase (NEB) at 37C for 5 min and then were sequentially incubated with increasing concentration of NaCl (50, 100, 200, 400 and 800 mM) on ice for 30 min. Nuclear extracts were separated from pellets by centrifugation and processed for WB.

Immunoblotting.—Cells were washed twice in ice-cold PBS and incubated in 1X RIPA buffer (Thermo Scientific) with protease and phosphatase inhibitor (Roche) for 5 min on ice. Lysates were pre-cleared by centrifugation and protein was quantified by BCA protein assay (Bio-Rad). 20 μg of protein lysate was resolved on SDS-PAGE gel and transferred onto Immobilon membrane (Millipore). Blots were blocked in 5% non-fat dry milk, incubated with primary antibodies overnight at 4°C, washed, incubated with secondary antibodies for 1 hour at room temperature, washed and developed with SuperSignal West Pico Chemiluminescent Substrate (Pierce). The list of antibodies used in this study is shown in Table S7.

Mass Spectrometry.—Co-IP samples were precipitated with TCA, resuspended in 8M urea, heated at 50°C for 5min, diluted tenfold with 12.5mM ammonium bicarbonate and digested with Trypsin. Samples were desalted on the C18 StageTip column. Eluates were collected, dried and dissolved in 0.1% formic acid. Samples were analyzed in a nano-HPLC coupled Q-Exactive mass spectromter (Thermo). Mass spectrometer was operated in data dependent mode using a top ten method. Full MS scans were acquired in the Orbitrap mass analyzer over a range of 300–1,650 m/z with resolution 70,000 (m/z 200). Tandem mass spectra were acquired in the Orbitrap mass analyzer with resolution 17,500 (m/z 200). Raw data were searched by Mascot Distiller with the following software settings: trypsin with one missed cleavage; oxidation (M), acetyl (N-term), deamidated (NQ) as variable modifications.

Co-IP.—10^8 cells were dissociated, pelleted and cytoplasm was removed by incubation in hypotonic buffer (10 mM HEPES, 10 mM KCl, 1.5 mM MgCl2, 0.34 M sucrose, 10% glycerol, 0.2 mM PMSF, 1 mM DTT). Nuclei were pelleted, washed and permeabilized with 0.1% n-Octyl-β-D-glucopyranoside. Nuclei were digested with MNase at concentration 2 U/ul at 37C for 5 min. Nucleosomes with associated proteins were extracted in IP buffer (10 mM HEPES, 150 mM KCl, 1.5 mM MgCl2, 0.34 M sucrose, 10% glycerol, 0.2 mM PMSF, 1 mM DTT), incubated with anti-FLAG magnetic beads overnight at 4°C, washed five times with IP buffer and eluted with FLAG peptide at 150 ng/ml. Immunoprecipitates were subjected to SDS-PAGE and probed with indicated antibodies or detected by silver staining according to the manufacturer's protocol.

RNA isolation.—Total RNA was isolated using Trizol, treated with DNAse I (Ambion) to remove genomic DNA, extracted with phenol:chloroform (1:1), precipitated and diluted in water.

Reverse Transcription and quantitative PCR (RT-qPCR): 1 μg of total RNA was reverse transcribed using the SuperScript II RT and random hexamers (Invitrogen). Twenty ng of diluted cDNA was used in each reaction. Reactions were run on the CFX96 instrument using iTaq Universal SYBR Green Supermix (Bio-Rad). Standard curves were generated for each primer set and were used to calculate Ct values with the expression threshold set to 100 RFU. To compare expression between samples expression values were scaled to *GAPDH*. Gene-specific primers used in PCR analyses are shown in Table S6.

RNA-seq analyses: RNA samples were processed by the Yale Stem Cell Center Genomics Core. Libraries were prepared following the standard Illumina protocol and sequenced on the HiSeq 2500 instrument (Illumina). Demultiplexed reads were trimmed with Trimmomatic (version 0.32) (parameters: ILLUMINACLIP: TruSeq3-SE.fa:2:30:10 LEADING: 3 TRAILING: 3 SLIDINGWINDOW: 4:5 MINLEN: 25). Trimmed reads were mapped onto the GRCm38/mm10 build of the mouse genome with STAR (version 2.4.2a). Mapped reads were indexed and sorted with Samtools (1.0). Coverage bigWig files for visualization were generated in R and scaled based on the number of uniquely mapping read counts to the ENSEMBL (version 89) annotation based on the TMM normalization in DESeq2 (1.16.1).

For transcript quantification, untrimmed reads were analyzed with Salmon (version 0.8.2) using transcripts defined by ENSEMBL (version 89) and imported into R with tximport (version 1.4.0). Principal component analysis and visualization was performed in R with functions PCA() (FactoMineR) and plot3d()(RGL).

Differentiation trajectories were defined based on the PCA clustering as follows. The canonical trajectories in Control^{OE} and Dppa2/4^{OE} cells (grey and dim purple lines in Figure 7A): MEFs (MEF sample), Thy1-SSEA1− (mean of day 8, day 11 and day 14 samples), Thy1-SSEA1+ (mean of day 8, day 11 and day 14 samples) and GFP+ (mean of day 11 and day 14 samples). For the alternative trajectory in Dppa2/4^{OE} cells (bright purple line) the values were assigned as follows: MEFs (MEF sample), Thy1+SSEA1+ (mean of day 2 and day 4 samples), Thy1-SSEA1+ (day 8 sample) and GFP+ (mean of day 8, day11 and day 14 samples). Genes that change in either direction along each reprogramming trajectory were identified using a Likelihood Ratio Test in DESeq2 (version 1.16.1) comparing models that did or did not have a cubic spline time term. Fold changes were determined as the $\log_2$ (Intermediate)-$\log_2$ (MEF). Prior to clustering, expression values for each gene were log-transformed and scaled to an average of 0 and standard deviation 1. Hierarchical clustering and visualizations were performed with Gene cluster (version 3.0) and TreeView (version 1.16r4). Trajectory profiles in each RNA expression cluster represent average of all genes in the cluster.

Gene Ontology analyses were performed using The Database for Annotation, Visualization and Integrated Discovery (DAVID) V6.8.

ChIP (Dppa2, Dppa4, H3K4me3, H3K4me1, H3K27me3 and H3K27ac).— 5-10×10^6 cells were cross-linked with 1% paraformaldehyde for 10 min and stopped by adding Glycine (to be 125 mM). The cross-linked material was washed with Buffer 1

(0.25% Triton, 10 mM EDTA, 0.5 mM EGTA, 10 mM Tris pH7.5) and Buffer 2 (200 mM NaCl, 10 mM EDTA, 0.5 mM EGTA, 10 mM Tris pH7.5) supplemented with protease inhibitor cocktail (Roche). Cell pellet was lysed in SDS lysis buffer (50 mM Tris pH 8.0, 10mM EDTA, 1% SDS, 50 mM PMSF) and sonicated to generate 200 to 600 bp fragments. Fragmented chromatin was immunoprecipitated with magnetic beads coupled with 5 μg of antibody. The beads were washed with low salt buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.1% SDS, 0.5% Deoxycholate, 1% NP40, 1 mM EDTA), high salt buffer (50 mM Tris pH 8.0, 500 mM NaCl, 0.1% SDS, 0.5% Deoxycholate, 1% NP40, 1 mM EDTA), LiCl wash buffer (50 mM Tris pH 8.0, 250 mM LiCl, 0.5% Deoxycholate, 1% NP40, 1 mM EDTA), Morohashi RIPA buffer (50mM Tris pH7.5, 150mM NaCl, 5mM EDTA, 0.5% NP40, 0.1% SDS), DOC/Triton Buffer (25mM Tris pH7.5, 150mM NaCl, 5mM EDTA, 1% Triton-X-100, 0.5% DOC) and Tris-EDTA buffer (10 mM Tris pH 8.0, 1 mM EDTA) for 10 min at 4 °C. Complexes were eluted in elution buffer (1% SDS, 0.1 M NaHCO3) for 2 h at 65 °C and reverse cross-linked for 6 h at 65°C. Reverse cross-linked DNA was purified by QIAquick PCR purification kit (Qiagen). The list of antibodies used in this study is shown in Table S7.

Native ChIP (H3K9me3): 10^7 cells were resuspended in buffer (10 mM Tris-HCl, pH 7.5, 2 mM MgCl2, 0.5 mM MgCl2) and permeabilized by adding 1% NP40. Chromatin was digested with 2 U/ml of MNase at 37 C for 5 min, and the reaction was quenched by 0.5 M EGTA. Chromatin was resuspended in N-ChIP buffer (2 mM EGTA, pH 8.0, 10 mM Tris pH7.4, 2 mM MgCl2, 0.5 mM PMSF and 1x protease inhibitor cocktail) and incubated on ice sequentially for 1 h with NaCl at 400 mM and 600 mM. Cellular debris was pelleted and the supernatant was recovered. Chromatin was pre-cleared with 20 ml of 1:1 protein A: G Dynabeads (Life Technologies) and immunoprecipitation was carried out with antibody–bead complexes overnight at 4 C. Bead-captured complexes were washed twice with N-ChIP buffer with 150 mM NaCl. Protein-DNA complexes were eluted in 200 μl of elution buffer (100 mM NaHCO3 and 1% SDS) O/N at 68 C. Eluted material was purified by phenol chloroform and processed for library construction.

ChIP-Seq analyses: ChIP samples were purified with MinElute PCR Purification kit (Qiagen) which yielded 20 μl ChIP DNA which was used immediately in a 23.3 μl combined end repair and A-tailing reaction using a KAPA Hyper Prep Kit (Kapa Biosciences) for 30 minutes at 20°C followed by 30 minutes at 65°C. 10 μl of ligase buffer, 3.7 μl of Y Adapters diluted 1:50 (Illumina 5179206) and 3.3 μl ligase (KAPA Hyper Prep Kit) were added and incubated at 20°C for 4 hours. Double-stranded DNA fragments were purified from this reaction using KAPA Pure Beads (Kapa Biosciences) and eluted in 22 μl 10 mM Tris pH8.0. The amplification cycle number was determined empirically by using 1 μl of the eluent in a 10 μl test reaction containing a 1:10,000 dilution of SYBR Green I (Invitrogen), 2x KAPA HotStart ReadyMix and custom barcoding primers. An appropriate number of cycles were used in a subsequent 50 μl reaction with the remaining eluent in the absence of SYBR Green. Libraries were size-selected on a 2% E-Gel EX agarose gel (Invitrogen) and fragments between 150 and 350 bp were extracted using a QIAEX II Gel Extraction Kit (Qiagen) performed at room temperature. Libraries were quantitated with a high sensitivity

Qubit dsDNA HS Assay Kit (Invitrogen) and pooled. Quality control and sequencing was performed by the Yale Center for Genome Analysis.

ChIP-Seq reads were mapped to the GRCm38/mm10 build of the mouse genome with Bowtie2 (version 2.2.9) using local mapping (--local). Only uniquely mapped reads were kept. Mapped reads were indexed and sorted with Samtools 1.5. Peaks in the ChIP-Seq datasets were called with MACS2 (version 2.1.1.20160309) using the -g 1.87e9 -f BAMPE (pair-end reads) or –f BAM (single-end reads) options.

For the heatmaps and metaplots in Figure 7D, we counted the number of reads in each ChIP experiment within 3 kb of each peak center. Unsupervised K-means clustering of these read counts was then performed with a fixed value of k=2, which was deemed optimal after preliminary analysis (clustering with k higher than 2 also resulted in two clusters). Red and green metaplots correspond to these two clusters. Visualizations were created with pheatmap(1.0.8), ggplot2(2.2.1) and custom R scripts.

To calculate correlation between Dppa2/4 binding and gene density (Figure 7E), anchor points were placed every 10 kb along individual chromosomes. Number of RefSeq promoters and number of Dppa2/4 peaks were counted within a 100 kb window around the anchor points and plotted.

To calculate the overlaps between Dppa2/4 binding and published chromatin trajectories during reprogramming, data from (Chronis et al., 2017) were converted into GRCm38/mm10 genomic coordinates using USCS LiftOver web tool with default parameters (minimum ratio of bases that must remap=0.95). Dppa2 binding peaks were scored as overlapping with chromatin trajectory elements if the overlap of genomic coordinates was at least 50% of the width of the smaller of two elements.

For chromatin mark quantification in Figure 7G, BEDTools (version 2.27.1) was used to calculate the signal in each Input and ChIP experiment within each trajectory element. Input subtracted cpkm counts were used for boxplots analyses.

For the correlation heatmap in Figure S6F, the Pearson correlation coefficients between pairs of experiments were calculated using ChIP-Seq signal in the genomic regions covered by the called peaks. The correlation values were clustered using Gene cluster (version 3.0) and visualized using TreeView (version 1. 16r4).

QUANTIFICATION AND STATISTICAL ANALYSIS

All assays were repeated at least twice. Statistical parameters for each experiment, including values of replicates and statistical significance, can be found in the figure legends.

DATA AND SOFTWARE AVAILABILITY

All software used in this study is freely or commercially available.

The ChIPseq and RNAseq data reported in this study have been deposited in the NCBI GEO database under the accession number GSE117173.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Tarjei Mikkelsen (Harvard University) for iOKSM-hF-TERT cells fibroblasts, Kathrin Plath (UCLA) for chromatin trajectory datasets and pre-iPSC lines, Michael Ziller and Alex Meissner (Harvard University) and Micah Gearhart (University of Minnesota) for help with bioinformatics analyses. We are also grateful to Timothy Nottoli and James McGrath (Yale Genome Editing Center) for help with blastocyst injections and tetraploid complementation assays, Mei Zhong (Yale Stem Cell Center Genomics Core), Kaya Bilguvar and Christopher Castaldi (Yale Center for Genome Analysis) for help with high-throughput sequencing and Diane Krause (Yale University) for the access to live imaging equipment. This work was supported by NIH (R01 GM105772 and GM107092 to NI and S10RR027990 to TAN) and the State of Connecticut (12-SBC-YALE-10 and 16-RMB-YALE-07 to NBI). CH was supported by the Yale CMB Training Grant and the Lo Family graduate fellowship, ZW was supported by the State of Connecticut (12-SCA-YALE-16).

REFERENCES

Banath JP, Banuelos CA, Klokov D, MacPhail SM, Lansdorp PM, and Olive PL (2009). Explanation for excessive DNA single-strand breaks and endogenous repair foci in pluripotent mouse embryonic stem cells. Experimental cell research 315, 1505–1520. [PubMed: 19154734]

Barrero MJ, Sese B, Kuebler B, Bilic J, Boue S, Marti M, and Izpisua Belmonte JC (2013). Macrohistone variants preserve cell identity by preventing the gain of H3K4me2 during reprogramming to pluripotency. Cell reports 3, 1005–1011. [PubMed: 23545500]

Bolger AM, Lohse M, and Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. [PubMed: 24695404]

Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, Klemm SL, van Oudenaarden A, and Jaenisch R (2012). Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell 150, 1209–1222. [PubMed: 22980981]

Cacchiarelli D, Trapnell C, Ziller MJ, Soumillon M, Cesana M, Karnik R, Donaghey J, Smith ZD, Ratanasirintrawoot S, Zhang X, et al. (2015). Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency. Cell 162, 412–424. [PubMed: 26186193]

Cheloufi S, Elling U, Hopfgartner B, Jung YL, Murn J, Ninova M, Hubmann M, Badeaux AI, Euong Ang C, Tenen D, et al. (2015). The histone chaperone CAF-1 safeguards somatic cell identity. Nature 528, 218–224. [PubMed: 26659182]

Chen J, Liu H, Liu J, Qi J, Wei B, Yang J, Liang H, Chen Y, Chen J, Wu Y, et al. (2013). H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. Nature genetics 45, 34–42. [PubMed: 23202127]

Chronis C, Fiziev P, Papp B, Butz S, Bonora G, Sabri S, Ernst J, and Plath K (2017). Cooperative Binding of Transcription Factors Orchestrates Reprogramming. Cell 168, 442–459 e420. [PubMed: 28111071]

de Hoon MJ, Imoto S, Nolan J, and Miyano S (2004). Open source clustering software. Bioinformatics 20, 1453–1454. [PubMed: 14871861]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Doege CA, Inoue K, Yamashita T, Rhee DB, Travis S, Fujita R, Guarnieri P, Bhagat G, Vanti WB, Shih A, et al. (2012). Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. Nature 488, 652–655. [PubMed: 22902501]

Eleuteri B, Aranda S, and Ernfors P (2018). NoRC Recruitment by H2A.X Deposition at rRNA Gene Promoter Limits Embryonic Stem Cell Proliferation. Cell reports 23, 1853–1866. [PubMed: 29742439]

Folmes CD, Nelson TJ, Martinez-Fernandez A, Arrell DK, Lindor JZ, Dzeja PP, Ikeda Y, Perez-Terzic C, and Terzic A (2011). Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. Cell metabolism 14, 264–271. [PubMed: 21803296]

Golipour A, David L, Liu Y, Jayakumaran G, Hirsch CL, Trcka D, and Wrana JL (2012). A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. Cell stem cell 11, 769–782. [PubMed: 23217423]

Gonzalez F, and Huangfu D (2016). Mechanisms underlying the formation of induced pluripotent stem cells. Wiley interdisciplinary reviews Developmental biology 5, 39–65. [PubMed: 26383234]

Guo S, Zi X, Schulz VP, Cheng J, Zhong M, Koochaki SH, Megyola CM, Pan X, Heydari K, Weissman SM, et al. (2014). Nonstochastic reprogramming from a privileged somatic cell state. Cell 156, 649–662. [PubMed: 24486105]

Hickson I, Zhao Y, Richardson CJ, Green SJ, Martin NM, Orr AI, Reaper PM, Jackson SP, Curtin NJ, and Smith GC (2004). Identification and characterization of a novel and specific inhibitor of the ataxia-telangiectasia mutated kinase ATM. Cancer research 64, 9152–9159. [PubMed: 15604286]

Hong H, Takahashi K, Ichisaka T, Aoi T, Kanagawa O, Nakagawa M, Okita K, and Yamanaka S (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature 460, 1132–1135. [PubMed: 19668191]

Huang da W, Sherman BT, and Lempicki RA (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols 4, 44–57. [PubMed: 19131956]

Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, DeCoste C, Schafer X, Lun Y, and Lemischka IR (2006). Dissecting self-renewal in stem cells with RNA interference. Nature 442, 533–538. [PubMed: 16767105]

Kawamura T, Suzuki J, Wang YV, Menendez S, Morera LB, Raya A, Wahl GM, and Izpisua Belmonte JC (2009). Linking the p53 tumour suppressor pathway to somatic cell reprogramming. Nature 460, 1140–1144. [PubMed: 19668186]

Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nature methods 9, 357–359. [PubMed: 22388286]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. [PubMed: 19505943]

Li R, Liang J, Ni S, Zhou T, Qing X, Li H, He W, Chen J, Li F, Zhuang Q, et al. (2010). A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell stem cell 7, 51–63. [PubMed: 20621050]

Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15, 550. [PubMed: 25516281]

Madan B, Madan V, Weber O, Tropel P, Blum C, Kieffer E, Viville S, and Fehling HJ (2009). The pluripotency-associated gene Dppa4 is dispensable for embryonic stem cell identity and germ cell development but essential for embryogenesis. Molecular and cellular biology 29, 3186–3203. [PubMed: 19332562]

Maldonado-Saldivia J, van den Bergen J, Krouskos M, Gilchrist M, Lee C, Li R, Sinclair AH, Surani MA, and Western PS (2007). Dppa2 and Dppa4 are closely linked SAP motif genes restricted to pluripotent cells and the germ line. Stem cells 25, 19–28. [PubMed: 16990585]

Marion RM, Strati K, Li H, Murga M, Blanco R, Ortega S, Fernandez-Capetillo O, Serrano M, and Blasco MA (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. Nature 460, 1149–1153. [PubMed: 19668189]

Masaki H, Nishida T, Sakasai R, and Teraoka H (2010). DPPA4 modulates chromatin structure via association with DNA and core histone H3 in mouse embryonic stem cells. Genes to cells : devoted to molecular & cellular mechanisms 15, 327–337. [PubMed: 20298437]

Mathieu J, Zhou W, Xing Y, Sperber H, Ferreccio A, Agoston Z, Kuppusamy KT, Moon RT, and Ruohola-Baker H (2014). Hypoxia-inducible factors have distinct and stage-specific roles during reprogramming of human cells to pluripotency. Cell stem cell 14, 592–605. [PubMed: 24656769]

Mattout A, Biran A, and Meshorer E (2011). Global epigenetic changes during somatic cell reprogramming to iPS cells. Journal of molecular cell biology 3, 341–350. [PubMed: 22044880]

Nakamura T, Nakagawa M, Ichisaka T, Shiota A, and Yamanaka S (2011). Essential roles of ECAT15-2/Dppa2 in functional lung development. Molecular and cellular biology 31, 4366–4378. [PubMed: 21896782]

Patro R, Duggal G, Love MI, Irizarry RA, and Kingsford C (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nature methods 14, 417–419. [PubMed: 28263959]

Polo JM, Anderssen E, Walsh RM, Schwarz BA, Nefzger CM, Lim SM, 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. [PubMed: 23260147]

Quinlan AR, and Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. [PubMed: 20110278]

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, and Mesirov JP (2011). Integrative genomics viewer. Nature biotechnology 29, 24–26.

Saldanha AJ (2004). Java Treeview--extensible visualization of microarray data. Bioinformatics 20, 3246–3248. [PubMed: 15180930]

Samavarchi-Tehrani P, Golipour A, David L, Sung HK, Beyer TA, Datti A, Woltjen K, Nagy A, and Wrana JL (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell stem cell 7, 64–77. [PubMed: 20621051]

Shabestarian H, Ghodsi M, Mallak AJ, Jafarian AH, Montazer M, and Forghanifard MM (2015). DPPA2 Protein Expression is Associated with Gastric Cancer Metastasis. Asian Pacific journal of cancer prevention : APJCP 16, 8461–8465. [PubMed: 26745102]

Siegel D, Schuff M, Oswald F, Cao Y, and Knochel W (2009). Functional dissection of XDppa2/4 structural domains in Xenopus development. Mechanisms of development 126, 974–989. [PubMed: 19772919]

Smeenk G, Wiegant WW, Marteijn JA, Luijsterburg MS, Sroczynski N, Costelloe T, Romeijn RJ, Pastink A, Mailand N, Vermeulen W, et al. (2013). Poly(ADP-ribosyl)ation links the chromatin remodeler SMARCA5/SNF2H to RNF168-dependent DNA damage signaling. Journal of cell science 126, 889–903. [PubMed: 23264744]

Sommer AG, Rozelle SS, Sullivan S, Mills JA, Park SM, Smith BW, Iyer AM, French DL, Kotton DN, Gadue P, et al. (2012). Generation of human induced pluripotent stem cells from peripheral blood using the STEMCCA lentiviral vector. Journal of visualized experiments : JoVE.

Soneson C, Love MI, and Robinson MD (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research 4, 1521. [PubMed: 26925227]

Soufi A, Donahue G, and Zaret KS (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell 151, 994–1004. [PubMed: 23159369]

Sridharan R, Gonzales-Cope M, Chronis C, Bonora G, McKee R, Huang C, Patel S, Lopez D, Mishra N, Pellegrini M, et al. (2013). Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1gamma in reprogramming to pluripotency. Nature cell biology 15, 872–882. [PubMed: 23748610]

Stadtfeld M, Maherali N, Borkent M, and Hochedlinger K (2010). A reprogrammable mouse strain from gene-targeted embryonic stem cells. Nature methods 7, 53–55. [PubMed: 20010832]

Strickfaden H, McDonald D, Kruhlak MJ, Haince JF, Th'ng JP, Rouleau M, Ishibashi T, Corry GN, Ausio J, Underhill DA, et al. (2016). Poly(ADP-ribosyl)ation-dependent Transient Chromatin Decondensation and Histone Displacement following Laser Microirradiation. The Journal of biological chemistry 291, 1789–1802. [PubMed: 26559976]

Takahashi K, Tanabe K, Ohnuki M, Narita M, Sasaki A, Yamamoto M, Nakamura M, Sutou K, Osafune K, and Yamanaka S (2014). Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. Nature communications 5, 3678.

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. [PubMed: 16904174]

Turinetto V, Orlando L, Sanchez-Ripoll Y, Kumpfmueller B, Storm MP, Porcedda P, Minieri V, Saviozzi S, Accomasso L, Cibrario Rocchietti E, et al. (2012). High basal gammaH2AX levels sustain self-renewal of mouse embryonic and induced pluripotent stem cells. Stem cells 30, 1414–1423. [PubMed: 22628289]

Utikal J, Polo JM, Stadtfeld M, Maherali N, Kulalert W, Walsh RM, Khalil A, Rheinwald JG, and Hochedlinger K (2009). Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. Nature 460, 1145–1148. [PubMed: 19668190]

Veuger SJ, Curtin NJ, Richardson CJ, Smith GC, and Durkacz BW (2003). Radiosensitization and DNA repair inhibition by the combined use of novel inhibitors of DNA-dependent protein kinase and poly(ADP-ribose) polymerase-1. Cancer research 63, 6008–6015. [PubMed: 14522929]

Yoshida Y, Takahashi K, Okita K, Ichisaka T, and Yamanaka S (2009). Hypoxia enhances the generation of induced pluripotent stem cells. Cell stem cell 5, 237–241. [PubMed: 19716359]

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome biology 9, R137. [PubMed: 18798982]

Zhu S, Li W, Zhou H, Wei W, Ambasudhan R, Lin T, Kim J, Zhang K, and Ding S (2010). Reprogramming of human primary somatic cells by OCT4 and chemical compounds. Cell stem cell 7, 651–655. [PubMed: 21112560]

HIGHLIGHTS

- Depletion of Dppa2 and Dppa4 reduced reprogramming efficiency
- Overexpression of Dppa2 and Dppa4 improved reprogramming efficiency and kinetics
- Dppa2/4 overexpression induced γH2AX and chromatin decompaction via Parp1
- Depletion of Parp1 or H2AX abolished the effect of Dppa2/4 on chromatin remodeling

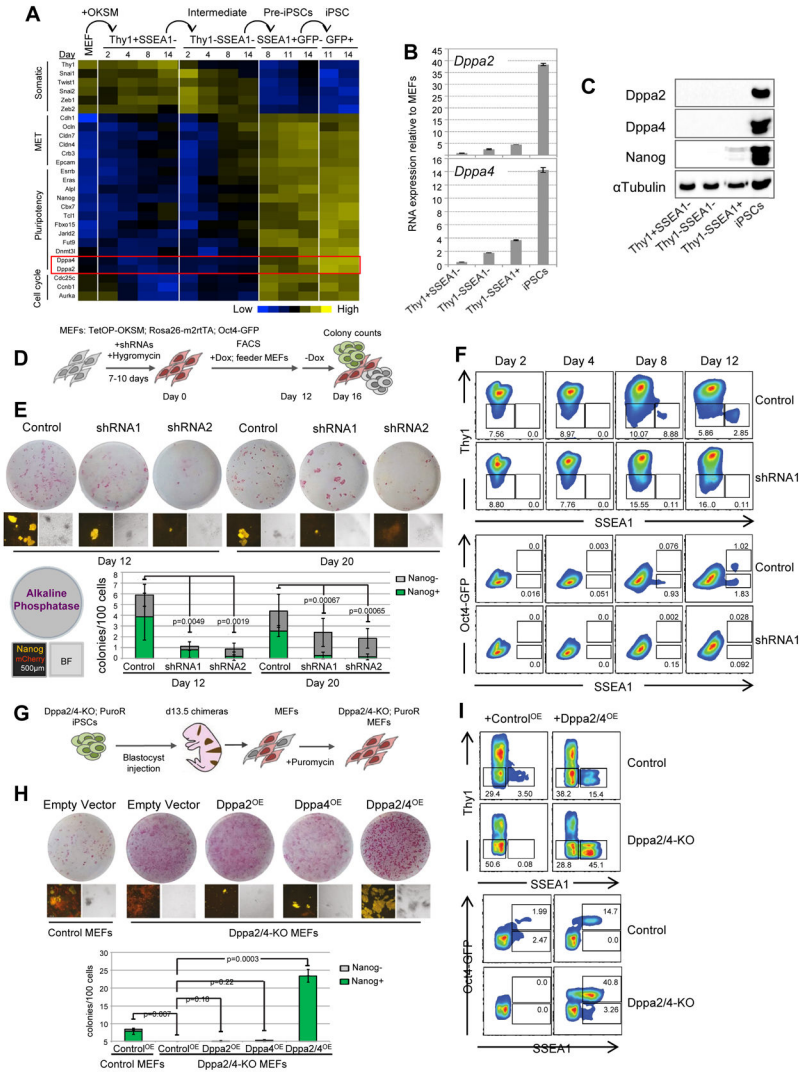


Figure 1. Dppa2 and Dppa4 are required for reprogramming.

(A) Expression of *Dppa2, Dppa4* and the signature genes in canonical reprogramming populations (Polo et al., 2012). Expression profile of each gene is scaled to the average value 0 and standard deviation 1 and color-coded from blue (low) to yellow (high).

(B) RNA expression of *Dppa2* and *Dppa4* in reprogramming intermediates and iPSCs. Expression was normalized to *Gapdh* and is shown relative to MEFs. Plotted are means ±SEM of technical triplicates.

(C) Protein levels of Dppa2 and Dppa4 in sorted day 9 reprogramming intermediates and iPSCs.

(D) Experimental setup used for knock-down experiments.

(E) Dppa2/4 knockdown and non-targeting control cultures were analyzed for alkaline phosphatase (AP) activity and Nanog immunofluorescence. Pairs of shRNAs were used to simultaneously knock-down both Dppa2 and Dppa4 prior to reprogramming. Plotted are mean colony counts ±SEM of biological duplicates.

(F) FACS analyses of knockdown cultures throughout reprogramming. Data is representative of three independent experiments.

(G) Derivation of 2^0 Dppa2/4-KO MEFs from Dppa2/4-KO and parental iPSCs.

(H) Dppa2/4-KO 2^0 MEFs were reconstituted with Dppa2, Dppa4 or Dppa2/4 expression vectors or empty vector, subjected to reprogramming and analyzed for alkaline phosphatase (AP) activity and Nanog immunofluorescence. Plotted are mean colony counts ±SEM of two biological replicates.

(I) FACS analyses of Dppa2/4-KO and Control cultures reconstituted with either Dppa2/4^{OE} or Control^{OE} vectors. Cultures were maintained with Dox for 12 days, followed by 4 days of culture without Dox. Data is representative of three independent experiments.

Homozygous OSKM+/+; rtTA+/+; Oct4-GFP+/+ cells were used in these experiments. p-values in E and H are calculated using Student's t-test. See also Figure S1.

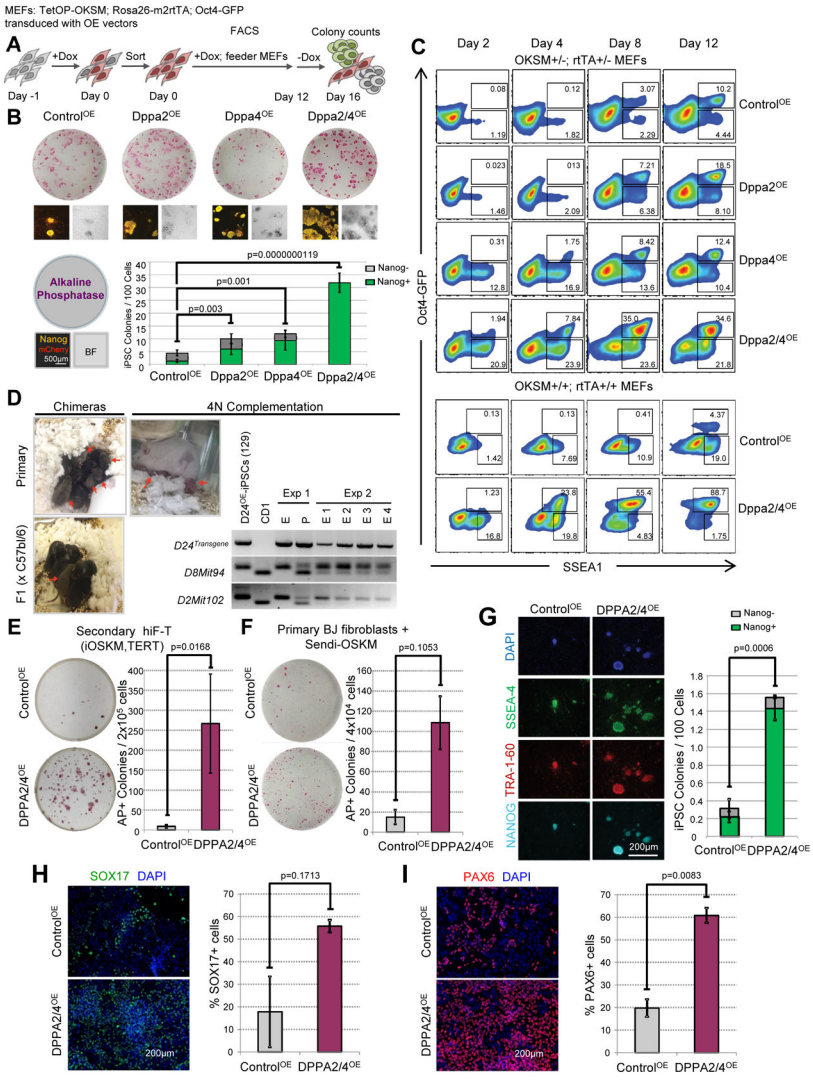


Figure 2. Overexpression of Dppa2 and Dppa4 improves reprogramming of mouse and human cells.

(A) Experimental strategy used for overexpression experiments.

(B) Single and double Dppa2/4 overexpression cultures alongside with empty vector control cultures were analyzed for AP activity and Nanog immunofluorescence. Plotted are mean colony counts ±SEM of two biological replicates.

(C) FACS analyses of overexpression cultures throughout reprogramming. Data is representative of three independent experiments.

(D) Dppa2/4^{OE}-iPSCs give rise to chimeric mice via blastocyst injections (left top). These chimeras were germline-competent (left bottom). Viable pups were also obtained via tetraploid complementation (right top and bottom). Three independently-derived clones were tested in chimera assay, one of which was used for tetraploid complementation assay.

(E) DPPA2/4 overexpression improves reprogramming of secondary human fibroblasts. iOKSM-hF-TERT cells (Cacchiarelli et al., 2015) transduced with lentiviral overexpression vectors were maintained with Dox for 30 days followed by 8 additional days without Dox. Plotted are mean AP+ colony counts ±SEM of two biological replicates.

(F) DPPA2/4 overexpression improves reprogramming of primary human cells. BJ cells were transduced with Sendai viral OSKM vectors and DPPA2/4^{OE} and rtTA vectors and reprogrammed for 30 days. Plotted are mean AP+ colony counts ±SEM of two biological replicates.

(G) Pluripotency markers were visualized by immunofluorescence in Sendai virus reprogrammed cultures. Plotted are mean colony counts ±SEM of two biological replicates.

(H, I) DPPA2/4^{OE}iPSCs give rise to SOX17+ endoderm and PAX6+ neuroectoderm. Plotted are mean percentages of marker positive cells ±SEM of two biological replicates.

p-values are calculated using Student's t-test. See also Figure S2, Tables S2, 3.

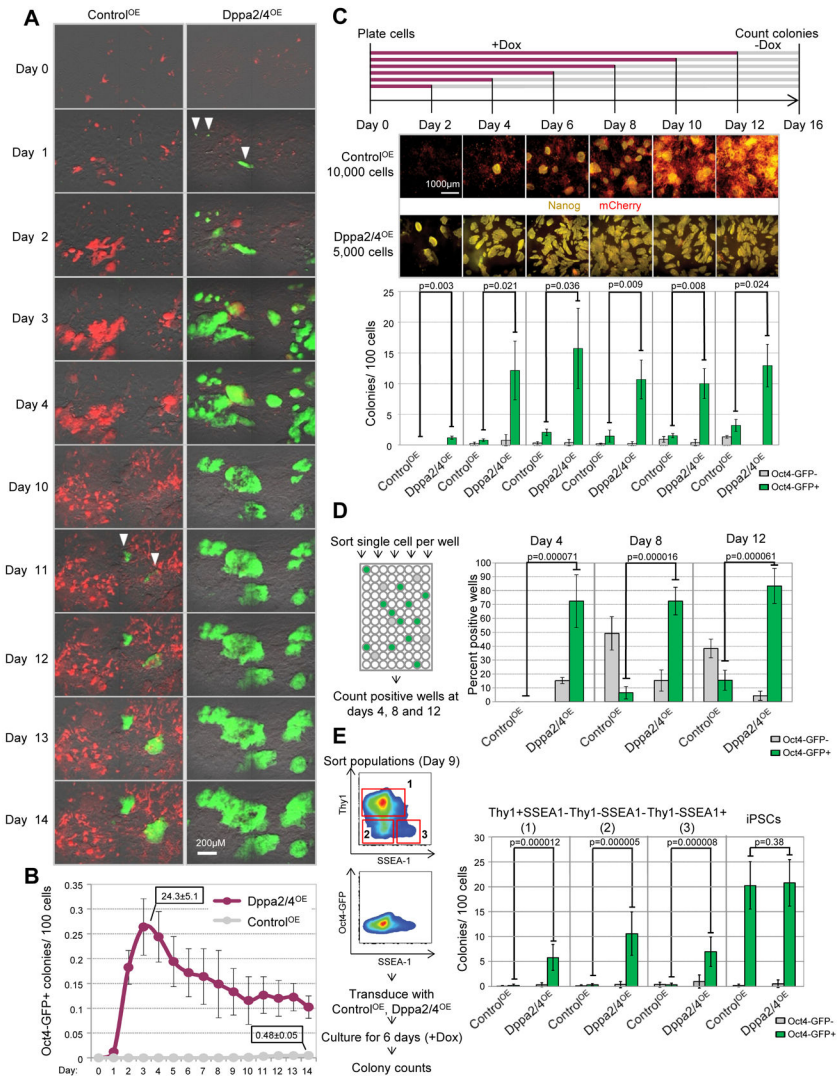


Figure 3. Dppa2/4 remove reprogramming barriers through all reprogramming stages up to iPSCs.

(A) Reprogramming kinetics in Dppa2/4^{OE} and Control^{OE} cultures monitored using live imaging. Shown are overlays of bright field, transgenic mCherry (red) and Oct4-GFP (green) channels for the initial (days 0-4) and late (days 10-14) stages of reprogramming. Arrows indicate the first GFP+ colonies that appear in each culture. Images were processed to enhance contrast.

(B) Reprogramming efficiency estimated form GFP+ colony counts in live imaging cultures. Plotted are mean colony counts ± SEM of two independent experiments.

(C) Two to four days of transgene expression is sufficient to reprogram cells in the presence of Dppa2/4. Cells were maintained with Dox for the intervals of time denoted by the purple line followed by Dox-free maintenance (grey line) for the rest of the 16 day time course. Plotted are mean colony counts ± SEM of two independent experiments.

(D) Reprogramming efficiency was determined in a single cell assay. Single mCherry positive cells (500 cells for each genotype) were sorted into wells of 96-well plates. 24 hours

after the sort mCherry positive wells were scored at 4, 8 and 12 days after plating. Plotted are mean well counts ± SEM of two independent experiments.

(E) Dppa2/4 overexpression restores reprogramming in refractory cells and accelerates reprogramming of intermediates. Day 9 reprogramming cultures were sorted, transduced with Dppa2/4^{OE} and Control^{OE} vectors, seeded onto a MEF monolayer and reprogrammed for 6 additional days. Plotted are mean colony counts ± SEM of two independent experiments.

All experiments were performed using homozygous OSKM+/+; rtTA+/+; Oct4-GFP+/+ cells. p-values are calculated using the Student's t-test.

See also Figure S3.

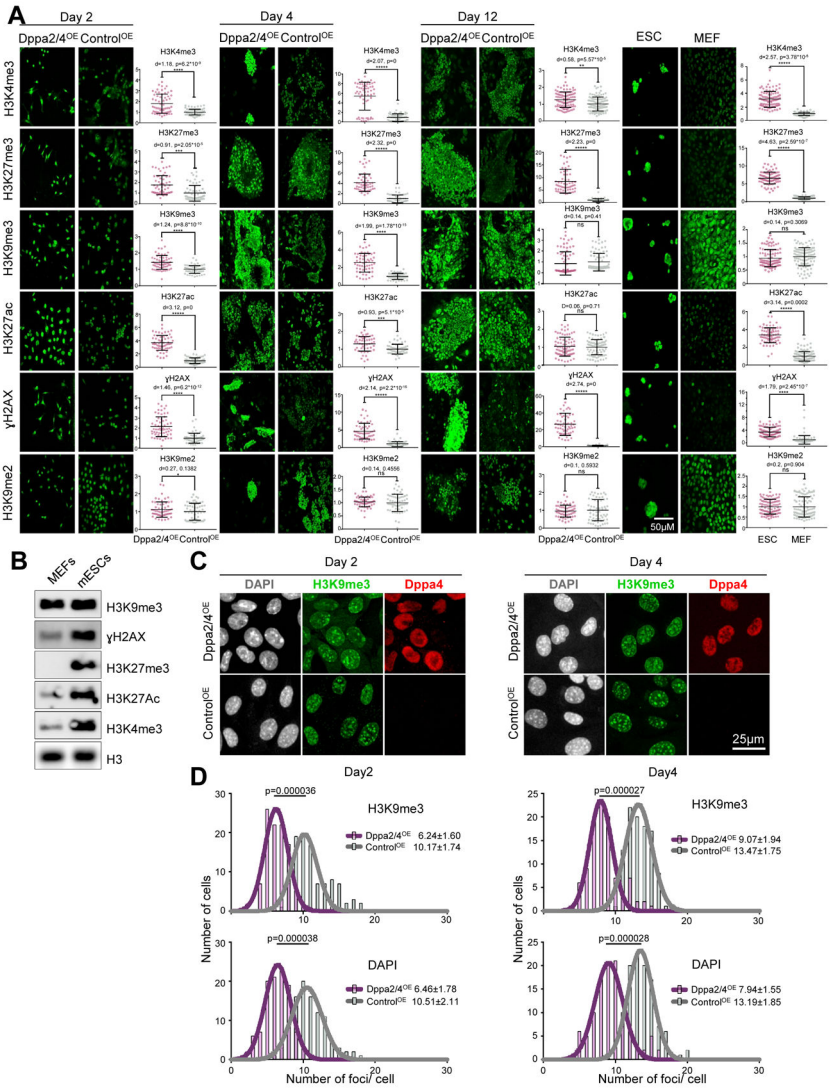


Figure 4. Dppa2/4 overexpression induces rapid changes in chromatin

(A) Overexpression of Dppa2/4 in OKSM-MEFs alters nuclear levels of H3K4me3, H3K27ac, H3K27me3, and γH2AX chromatin marks. Fluorescence intensities are plotted relative to the mean intensity in the Control^{OE} culture. Cohen's effect size d and Student's t-test p-value are shown for each comparison and marked according the effect size.

(B) Histone mark levels were quantified in MEFs and ESCs by western blotting. Histone H3 is shown as a loading control.

(C, D) Chromatin compaction was assessed by the number of DAPI/ H3K9me3 positive foci within the nucleus. Gaussian distribution curves were fitted over the data to calculate mean foci number ± SEM. p-value were calculated using Student's t-test.

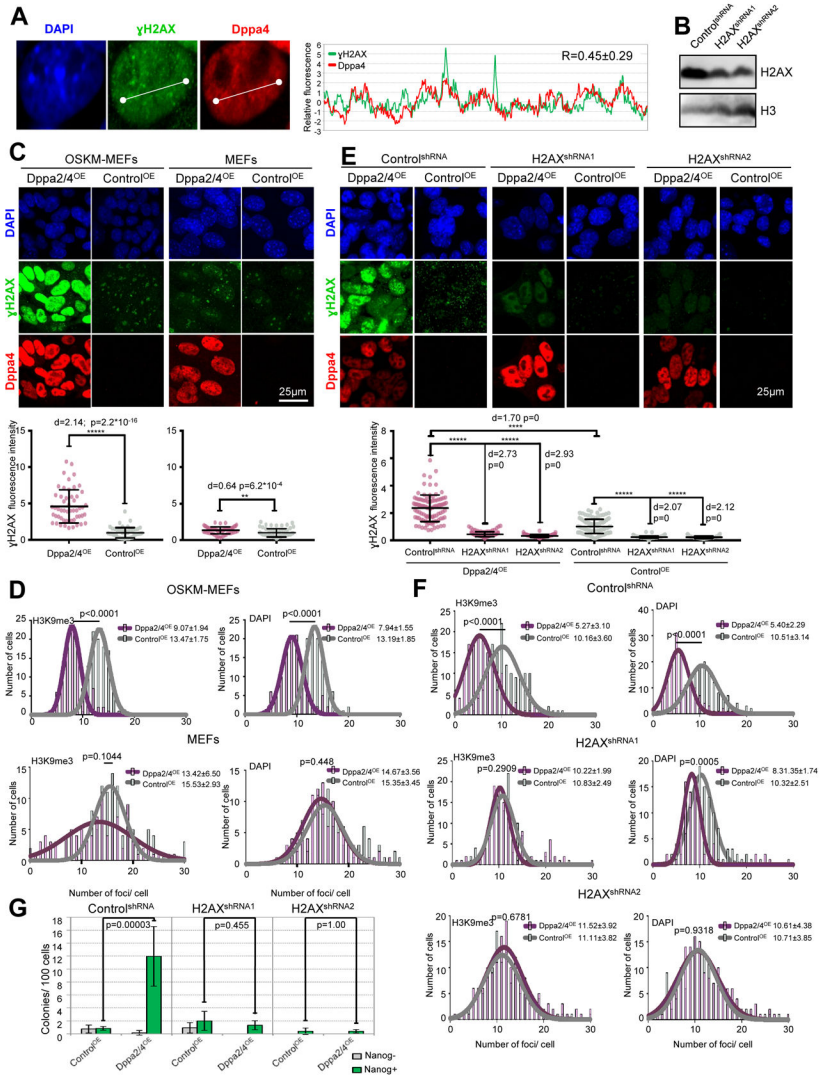


Figure 5. High γH2AX levels are required for Dppa2/4-induced chromatin remodeling and enhanced reprogramming.

(A) γH2AX and Dppa4 profiles within the nucleus are positively correlated. Intensity values along the indicated path were obtained using ImageJ and scaled to average 0 and standard deviation 1. Representative nucleus is shown.

(B) Lentiviral shRNAs deplete H2AX in MEFs.

(C) Overexpression of Dppa2/4 in OKSM-expressing but not in regular MEFs results in elevated γH2AX levels. Fluorescence intensities are plotted relative to the mean intensity in the Control^{OE} culture (bottom).

(D) Chromatin decompaction in response to Dppa2/4 overexpression requires input from the OKSM factors.

(E) Confocal imaging confirms significantly reduced γH2AX levels in day 2 shRNA-transduced reprogramming cultures.

(F) Chromatin decompaction in response to Dppa2/4 overexpression is not observed in Dppa2/4^{OE} reprogramming cultures depleted of γH2AX.

(G) Depletion of H2AX reduces reprogramming efficiency in Dppa2/4^{OE} cultures. Plotted are mean colony counts ±SEM of two independent experiments.
See also Figure S4.

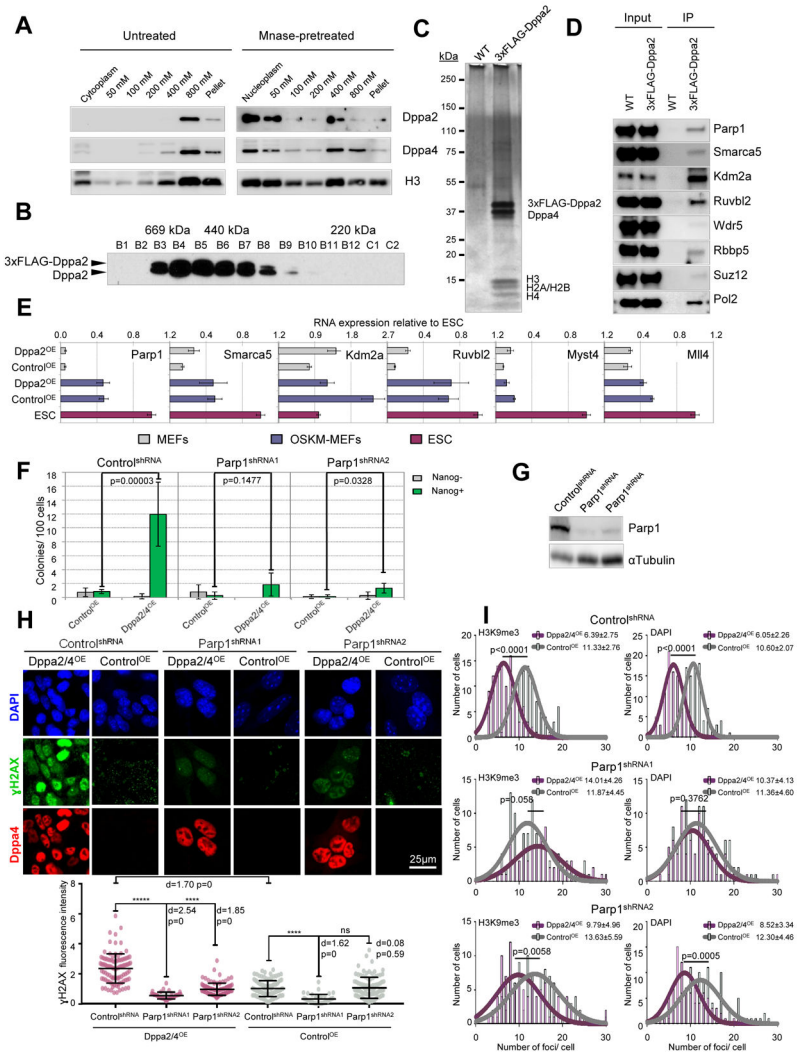


Figure 6. Dppa2/4-associated chromatin modifiers are required for chromatin remodeling and enhanced reprogramming.

(A) Dppa2 and Dppa4 are tightly bound to chromatin. Nuclei were isolated from ESCs and extracted sequentially with increasing concentration of NaCl. Nuclei pretreated with MNase prior to salt extraction release Dppa2/4-nucleosome complexes under low salt concentration.

(B) ESCs expressing 3xFLAG-Dppa2 were used to purify Dppa2-associated proteins. The fractionation pattern of the transgene was similar to that of endogenous Dppa2 protein.

(C) Dppa2-interacting proteins were visualized by silver staining. Core histones and Dppa4 were abundant in the IPs, multiple high molecular weight proteins were also detected and identified by mass spectrometry.

(D) Dppa2-associated proteins with functions in transcription and chromatin regulation were validated by western blotting.

(E) RNA expression of Dppa2-associated chromatin modifiers in ESCs, MEFs and OKSM-MEFs transduced with Dppa2/4^{OE} and Control^{OE} vectors and maintained with Dox for four days. Plotted are mean *GAPDH*-normalized expression levels ± SEM of technical triplicates.

(F) Depletion of Parp1 reduces reprogramming efficiency in Dppa2/4^{OE} cultures. Plotted are mean colony counts ± SEM of two independent experiments.
(G) shRNA vectors efficiently downregulate Parp1 in ESCs.
(H) Confocal imaging shows reduced γH2AX levels in day 2 Dppa2/4^{OE} cultures treated with Parp1 shRNAs.
(I) Chromatin decompaction is reduced in day 2 Dppa2/4^{OE} reprogramming cells treated with Parpl shRNAs.
See also Figure S5, Table S4.

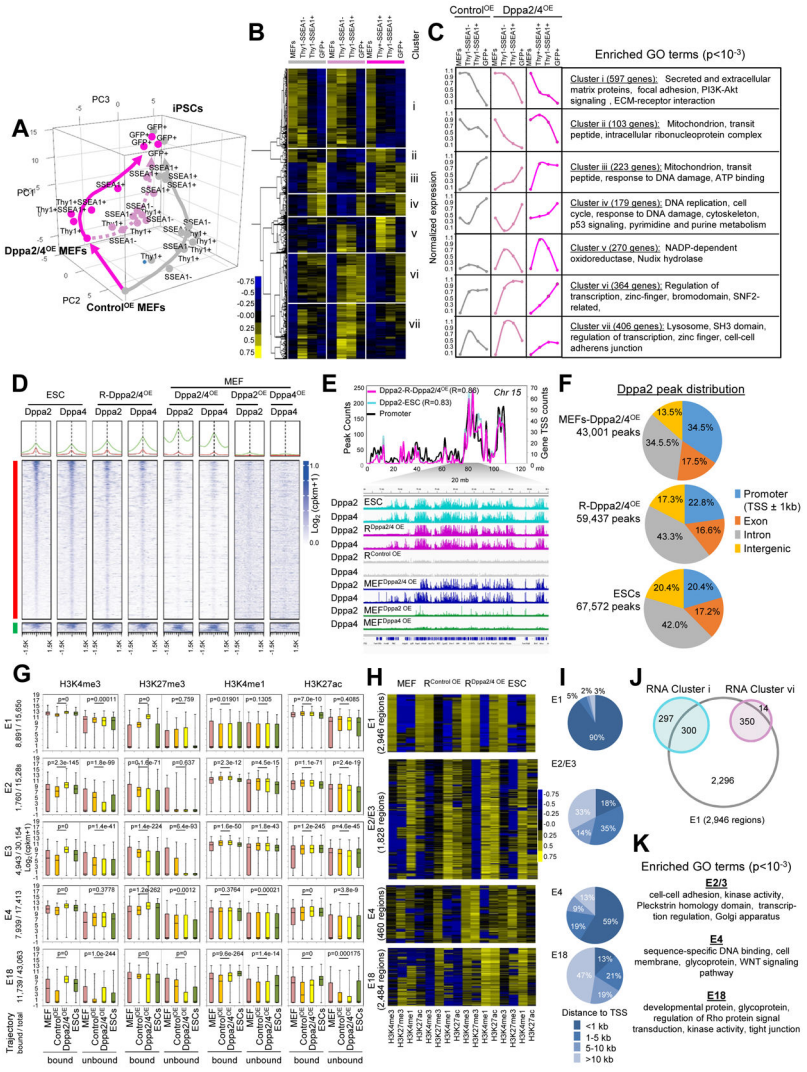


Figure 7. Dppa2/4 overexpression directs cells into an alternative reprogramming trajectory and contributes to epigenetic remodeling of promoter and enhancer regions

(A) Dppa2/4^{OE} cultures reprogram via an alternative trajectory. PCA analysis was performed on a set of genes with dynamic expression during reprogramming. The abbreviations are as follows: Thy1+: Thy1+SSEA1−Oct4-GFP−; Thy1+SSEA1+: Thy1+SSEA1+Oct4-GFP−; SSEA1−: Thy1−SSEA1−GFP−; SSEA1+: Thy1−SSEA1+Oct4-GFP−; GFP+: Thy1−SSEA1+Oct4-GFP+.

(B) Seven clusters of differential expression highlight the key differences between the alternative and the canonical reprogramming trajectories.

(C) Aggregate expression profiles and enriched GO terms for seven RNA expression clusters.

(D) Dppa2 and Dppa4 binding in Dppa2/4^{OE} MEFs and reprogramming cells is similar to that in ESCs. Shown are metaplots (top) and heatmaps (bottom) of Log_2-transformed ChIP-Seq read counts within a 3 kb window centered on the peak summit.

(E) Dppa2/4 binding correlates positively with gene density. Chromosome 15 is shown as an example, the genome-wide correlations were 0.50 (Dppa2-ESC) and 0.52 (Dppa2-R-Dppa2/4^{OE}).

(F) Distributions of Dppa2 binding peaks according to their genomic location in Dppa2/4^{OE} MEFs and reprogramming cells, and ESCs.

(G) Distributions of H3K4me3, H3K27me3, H3K4me1 and H3K27ac levels for Dppa2/4-bound and unbound elements within four chromatin trajectories as defined in (Chronis et al., 2017). E1-E4 are different classes of promoter elements, E18 are ESC-specific enhancer elements. Boxes denote second and third distribution quartiles, whiskers show the first and the fourth quartiles, the median is marked by horizontal line. p-values are from the Student's t-test.

(H) Heatmaps of chromatin marks in MEFs, R-Control^{OE}, R-Dppa2/4^{OE} and ESCs for representative Dppa2/4-bound elements from each trajectory. The Log_2(cpm+1) values were centered and normalized for each region across all cell types and clustered.

(I) Distributions of Dppa2/4-bound elements in each cluster according to distance to the nearest TSS.

(J) Dppa2/4-bound E1 genes with gain in H3K27me3 overlap with RNA expression clusters i and vi.

(K) Enriched gene ontologies for Dppa2/4-bound elements in trajectories E2, E3, E4 and E18. Elements were assigned to the nearest gene.

See also Figure S6, 7, Table S5.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rat monoclonal anti-CD90.2 (Thy-1.2) eFluor 450	eBioscience	Cat#48-0902-82; RRID:AB 1272200
Mouse monoclonal anti- SSEA-1- APC-Vio770	Miltenyi Biotec	Cat#130-104-939; RRID:AB 2653516
Rabbit polyclonal anti-Nanog	Abcam	Cat#ab80892; RRID:AB 2150114
Mouse monoclonal anti-Dppa2	MIllipore	Cat# MAB4356; RRID:AB 1977389
Goat polyclonal anti-Dppa4	R&D Systems	Cat#AF3730; RRID: AB 2094166
Mouse monoclonal anti-a-Tubulin	Sigma	Cat#T6074; RRID: AB 477582
Rabbit monoclonal anti-ParpI	Thermo Fisher	Cat#MA5-15031; RRID:AB 11000218
Rabbit polyclonal anti-Smarca5	Abcam	Cat# ab3749; RRID:AB 2191856
Rabbit polyclonal anti-Kdm2a	Abcam	Cat# ab31739; RRID:AB 775817
Mouse monoclonal anti-Wdr5	Abcam	Cat#ab56919; RRID:AB 946146
Rabbit polyclonal anti-Rbbp5	Abcam	Cat# ab154755; RRID:AB 2728762
Rabbit polyclonal anti-Ruvb2	Abcam	Cat# ab36569; RRID: AB 2301439
Rabbit polyclonal anti-Suz12	Cell Signaling Tecnology	Cat#3737; RRID: AB 2196850
Mouse monocolonal anti-Sma	Fisher Scientific	Cat#01-671-636; RRID:AB 1087372
Mouse monocolonal anti-Pol2	Covance	Cat#8WG16; RRID: AB 10013665
Mouse monoclonal anti-Gfap	Millipore	Cat# MAB360; RRID: AB 11212597
Rabbit monoclonal anti-NANOG	Cell Signaling Technology	Cat#4903; RRID: AB 10559205
Mouse monoclonal anti-Afp	R&D Systems	Cat# MAB1368; RRID:AB 357658
Mouse monoclonal anti-SSEA4	Stemcell Technologies	Cat#60062; RRID: AB 2721031
Mouse monoclonal anti-PAX6	BD Biosciences	Cat#561664; RRID: AB 10895587
Mouse monoclonal anti-TRA 1-60	Bd Cell Analysis	Cat# BDB560884; RRID: AB 10562572
Goat polyclonal anti-SOX17	R&D Systems	Cat#AF1924; RRID: AB 355060
Rabbit polyclonal anti-H3	Abcam	Cat#AB1791; RRID: AB 302613
Rabbit polyclonal anti-H3K4me3	Millipore	Cat#07-473; RRID: AB 1977252

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Rabbit polyclonal anti-H3K4me1	Abcam	Cat#ab8895; RRID: AB 306847
Rabbit polyclonal anti-H3K27me3	Cell Signaling Technology	Cat#9733; RRID: AB 2616029
Rabbit polyclonal anti-H3K27ac	Active Motif	Cat#AM39135; RRID: AB 2614979
Rabbit polyclonal anti-H3K9me3	Abcam	Cat#ab8898; RRID: AB 306848
Rabbit polyclonal anti-H3K9me2	Abcam	Cat#ab1220; RRID: AB 449854
Rabbit polyclonal anti-H2AX	Active Motif	Cat#AM39689; RRID: AB 2728764
Rabbit polyclonal anti-γH2AX	Millipore	Cat#05-636; RRID: AB_309864
Mouse monoclonal anti-FLAG	Sigma	Cat#F1804; RRID: AB 262044
Chemicals, Peptides, and Recombinant Proteins		
Matrigel	BD Biosciences	Cat#354277
mTeSR media	Stem Cell Technologies	Cat#85850
Dispase	Stem Cell Technologies	Cat#07913
Doxycyclin	Stem Cell Technologies	Cat#72742
Trizol	Invitrogen	Cat#15596018
SB43152	R&D Systems	Cat#1614
PD0325901	EMD Millipore	Cat#444968
Dorsomorphin	Stemgent	Cat# 04-0024
DMEM/F12 media	Life Technologies	Cat#11330032
DMEM media	Life Technologies	Cat#11965092
N-2	Invitrogen	Cat#17502048
B27	Invitrogen	Cat#17504044
ESGRO	Millipore	Cat#ESG1107
p-mercaptoethanol	Sigma Aldrich	Cat#M7522
Non-essential amino acids	Life Technologies	Cat#11140050
L-Glutamine	Life Technologies	Cat#25030081
Penicillin-streptomycin	Life Technologies	Cat#15140122
bFGF	Millipore	Cat#GF003AF-MG
Na-pyruvate	Life Technologies	Cat#11360070
KSR	Technologies	Cat#10828028
PEI MAX	Polysciences, Inc.	Cat#24765
Activin A	R&D	Cat#338-AC
bFGF	Millipore	Cat#GF003AF-MG
Accutase	Stem Cell Technologies	Cat#07920
Paraformaldehyde	EMS	Cat#15710
ROCK inhibitor	Abcam	Cat#ab120129

REAGENT or RESOURCE	SOURCE	IDENTIFIER
NU7026	Tocris	Cat#2828
KU55933	Tocris	Cat# 3544
DMSO	Sigma	Cat#D2650
DNAse I	Roche	Cat#04716728001
Micrococcal nuclease	NEB	Cat#M0247S
Polybrene	Sigma	Cat# 107689
Geneticin	Life Technologies	Cat#10131035
Puromycin	Life Technologies	Cat#A1113802
Normal donkey serum	Millipore	Cat#566460
iTaq Universal SYBR Green Supermix	BioRad	Cat#1725120
SYBR Green I Nucleic Acid Gel Stain	Invitrogen	Cat#S7563
SuperScript II Reverse Transcriptase	Invitrogen	Cat#18064014
cOmplete Protease Inhibitor Cocktail	Roche	Cat#11873580001
Dynabeads Protein A	Invitrogen	Cat#10001D
Anti-FLAG magnetic beads	MBL	Cat#M185-11
KAPA Pure Beads	Kapa Biosystems	Cat#KK8000
PhosSTOP protease inhibitor	Roche	Cat#4906845001
Octyl β-D-glucopyranoside	Sigma	Cat#O8001
Critical Commercial Assays		
Cytotune iPS 2.0 Kit	Invitrogen	Cat#A16517
KAPA Express Extract Kit	Kapa Biosystems	Cat#KK7151
Illumina TotalPrep RNA Amplification Kit	Ambion	Cat#AMIL1791
QIAquick PCR Purification Kit	Qiagen	Cat#28104
MinElute PCR Purification Kit	Qiagen	Cat#28004
KAPA Hyper Prep Kit	Kapa Biosystems	Cat #KK8502
Qubit dsDNA HS Assay Kit	Invitrogen	Cat#Q32854
SilverQuest Silver Staining Kit	Invitrogen	Cat#LC6070
Supersignal West Pico Chemiluminescent Substrate	Thermo Scientific	Cat#PI34080
Experimental Models: Cell Lines		
CF1 MEF	MTI-GlobalStem	Cat#GSC-6001G
HEK293FT	Invitrogen	Cat#R70007
Mouse embryonic fibroblasts isolated from *Collal-OSKM*+/+; *Rosa26-rtTA*+/+; *Oct4-GFP*+/+ Mice	(Guo et al., 2014)	N/A
human TERT-immortalized secondary fibroblasts with inducible genome-integrated OSKM factors (hiF-T)	(Cacchiarelli et al., 2015)	N/A
BJ	ATCC	Cat#CRL2522
Oligonucleotides		
shRNAs, see Table S6	This paper	N/A
gRNAs, see Table S6	This paper	N/A
Primers used for RT-qPCR analyses, see Table S6	This paper	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant DNA		
pCMV-VSV-G	Weinberg lab, MIT	Addgene#8454
pCMV-dR8.2	Weinberg lab, MIT	Addgene#8455
FUW-H1P-Hygro-mCherry	(Ivanova et al., 2006)	N/A
pSpCas9n(BB)-2A-GFP	Zhang lab, Broad Institute	Addgene #48140
FUW-pEF1a-IRES-Puro	(Ivanova et al., 2006)	N/A
FUW-pTRE-2A-mCherry	This study	N/A
FUW-pTRE-2A-BFP	This study	N/A
TetO-FUW-Sox2	Jaenisch lab, MIT	Addgene #20326
TetO-FUW-cMYC	Jaenisch lab, MIT	Addgene #20324
TetO-FUW-Oct4	Jaenisch lab, MIT	Addgene #20323
TetO-FUW-Klf4	Jaenisch lab, MIT	Addgene #20322
STEMCCA	(Sommer et al., 2012)	N/A
Data Resources		
Raw and analyzed RNA-Seq data	This paper	GSE117173
Raw and analyzed ChIP-Seq data	This paper	GSE117173
STACKED_35_segments	(Chronis et al., 2017)	Plath lab
Software and Algorithms		
ImageJ	NIH	https://imagej.net/
Prism 7	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
Trimmomatic	(Bolger et al., 2014)	http://www.usadellab.org/
STAR	(Dobin et al., 2013)	https://www.encodeproject.org/software/star/
DESeq2 v1.16.1	(Love et al., 2014)	v1.16.1
Salmon	(Patro et al., 2017)	0.8.2
tximport	(Soneson et al., 2015)	v1.4.0
Gene Cluster v3.0	(de Hoon et al., 2004)	http://bonsai.hgc.jp/~mdehoon/software/cluster/
TreeView v1.16r4	(Saldanha, 2004)	https://sourceforge.net/projects/jtreeview/files/
DAVID 6.8	(Huang da et al., 2009)	https://david.ncifcrf.gov/
Bowtie2	(Langmead and Salzberg, 2012)	v2.2.9
MACS2	(Zhang et al., 2008)	v2.1.1.20160309
Samtools	(Li et al., 2009)	v1.5
USCS LiftOver		http://genome.ucsc.edu/cgi-bin/hgLiftOver
BEDTools	(Quinlan and Hall, 2010)	2.27.1
Integrative Genomics Viewer	(Robinson et al., 2011)	http://software.broadinstitute.org/software/igv/