Article

Cell

X Chromosome Reactivation Dynamics Reveal Stages of Reprogramming to Pluripotency

Vincent Pasque,[1,6] Jason Tchieu,[2,6] Rahul Karnik,[3] Molly Uyeda,[1] Anupama Sadhu Dimashkie,[1] Dana Case,[1] Bernadett Papp,[1] Giancarlo Bonora,[1] Sanjeet Patel,[1] Ritchie Ho,[1] Ryan Schmidt,[1] Robin McKee,[1] Takashi Sado,[4] Takashi Tada,[5] Alexander Meissner,[3] and Kathrin Plath[1,*]

[1]Department of Biological Chemistry, Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, Jonsson Comprehensive Cancer Center, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA

[2]Developmental Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA

[3]Department of Stem Cell and Regenerative Biology, Harvard University, Harvard Stem Cell Institute, Broad Institute of MIT and Harvard, Cambridge, MA 02138, USA

[4]Department of Advanced Bioscience, Graduate School of Agriculture, Kinki University, 3327-204 Nakamachi, Nara, 631-8505, Japan

[5]Department of Stem Cell Engineering, Stem Cell Research Center, Institute for Frontier Medical Sciences, Kyoto University, 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan

[6]Co-first author

*Correspondence: kplath@mednet.ucla.edu

http://dx.doi.org/10.1016/j.cell.2014.11.040

SUMMARY

Reprogramming to iPSCs resets the epigenome of somatic cells, including the reversal of X chromosome inactivation. We sought to gain insight into the steps underlying the reprogramming process by examining the means by which reprogramming leads to X chromosome reactivation (XCR). Analyzing single cells in situ, we found that hallmarks of the inactive X (Xi) change sequentially, providing a direct readout of reprogramming progression. Several epigenetic changes on the Xi occur in the inverse order of developmental X inactivation, whereas others are uncoupled from this sequence. Among the latter, DNA methylation has an extraordinary long persistence on the Xi during reprogramming, and, like *Xist* expression, is erased only after pluripotency genes are activated. Mechanistically, XCR requires both DNA demethylation and *Xist* silencing, ensuring that only cells undergoing faithful reprogramming initiate XCR. Our study defines the epigenetic state of multiple sequential reprogramming intermediates and establishes a paradigm for studying cell fate transitions during reprogramming.

INTRODUCTION

Understanding the mechanisms by which the identity of a cell is established and maintained is a key goal of contemporary biology. Somatic cells can be reprogrammed into induced pluripotent stem cells (iPSCs) through transcription factor expression (Takahashi and Yamanaka, 2006). This process entails profound changes in genome organization, histone modifications, DNA methylation, and gene expression (reviewed in Apostolou and Hochedlinger, 2013). Questions of outstanding interest are whether reprogramming proceeds through specific stages that can be defined based on epigenetic features and how and in what order the epigenetic features gradually acquired during differentiation are reversed during reprogramming. One approach to address these questions is to focus on events for which the sequence of epigenetic changes that occur during differentiation is well defined and to ask how it is reversed during reprogramming to iPSCs.

X chromosome inactivation (XCI) is induced upon differentiation of female mouse pluripotent cells and leads to the inactivation of one of the two X chromosomes (reviewed in Lee and Bartolomei, 2013; Barakat and Gribnau, 2010; Chow and Heard, 2009). The sequence of epigenetic events accompanying the silencing of the X chromosome during differentiation has been examined extensively (Chow and Heard, 2009). These events include an initiation phase characterized by the coating of the future inactive X chromosome (Xi) by the large noncoding RNA *Xist*, which creates a nuclear compartment devoid of RNA polymerase II (Chaumeil et al., 2006) and leads to transient recruitment of the Polycomb Repressive Complex 2 (PRC2) and the deposition of the repressive histone mark H3K27me3 by its catalytic subunit EZH2 (Plath et al., 2003; Silva et al., 2003), closely followed by gene silencing. Later in differentiation, these events are followed by incorporation of the repressive histone variant macroH2A1 and DNA methylation, stabilizing the silenced state (Gendrel et al., 2012; Mermoud et al., 1999). Thus, once established, the Xi is extraordinary stable and is only reversed in a process termed X chromosome reactivation (XCR), which, in embryos, is limited to the inner cell mass and to germ cells (Lee and Bartolomei, 2013). XCR results in erasure of Xi-heterochromatin marks, and, importantly, can also be induced experimentally by reprogramming of female mouse somatic cells to iPSCs and somatic cell nuclear transfer (SCNT) (Maherali et al., 2007; Eggan et al., 2000). It is known that XCR is a late event during reprogramming to iPSCs (Payer et al., 2013;

Stadtfeld et al., 2008), but the exact dynamics of XCR and how the epigenetic hallmarks of the Xi change in this process have remained unclear.

Most insight into the molecular events of reprogramming to iPSCs have been gained from gene expression studies of populations of cells undergoing reprogramming and of subpopulations isolated using cell surface markers (O'Malley et al., 2013; Golipour et al., 2012; Polo et al., 2012; Samavarchi-Tehrani et al., 2010; Stadtfeld et al., 2008; Mikkelsen et al., 2008; reviewed in Buganim et al., 2013). These studies indicated that reprogramming is a multistep process with two predominant "waves" of gene expression changes: an early wave marked by enhanced proliferation and a mesenchymal-to-epithelial transition (MET), characterized by *Cdh1* (E-cadherin) expression (Polo et al., 2012; Samavarchi-Tehrani et al., 2010; Li et al., 2010), and a late wave, characterized by reactivation of pluripotency genes such as *Nanog* (O'Malley et al., 2013; Buganim et al., 2012; Golipour et al., 2012; Polo et al., 2012). The variable latency and relatively low efficiency by which individual cells reprogram have also encouraged gene expression measurements at the single-cell level at various stages of reprogramming and in clonal late intermediates. These experiments have argued for a sequence of stochastic transcriptional changes early in reprogramming, where expression programs vary dramatically between individual cells, eventually leading to hierarchical activation of pluripotency genes during the final phase, which, however, may occur through multiple paths (Buganim et al., 2012; Polo et al., 2012; Parchem et al., 2014).

Despite these advances, further molecular insight into the reprogramming path and a continuous view of the molecular events and stages leading to pluripotency would benefit from alternative approaches. In situ temporal analyses that integrate the position of cells within their native reprogramming environment, as well as the level of proteins and chromatin marks and their subcellular localization, may be particularly useful. Given that reprogramming to iPSCs is associated with XCR, and in light of the detailed characterization of sequential steps of XCI during differentiation, the reprogramming process provides an unprecedented opportunity to study XCR. In turn, the Xi provides an exceptional possibility to characterize the dynamics of the reversal of epigenetic marks during reprogramming.

Here, we followed epigenetic changes on the Xi during reprogramming to iPSCs in individual cells using detailed, high-resolution in situ time course analyses to address the question of whether XCR and somatic cell reprogramming follow a precise sequence of epigenetic changes. Due to the sheer size of the X chromosome, this analysis can be done at the single-cell level using immunofluorescence and RNA FISH approaches, allowing for the identification of reprogramming stages that have been elusive in transcriptional and chromatin studies to date. Our work demonstrates that the epigenetic state of the Xi changes sequentially throughout reprogramming, along with global changes in chromatin character. To shed light on the mechanisms by which XCR takes place, we used genetically manipulated somatic cells and examined the role played by *Cdh1*, *Nanog*, *Xist*, *Tsix*, *Tet1*, *Tet2*, and DNA methylation. The highly reproducible sequence of epigenetic steps leading to XCR and induced pluripotency provides a simple readout of reprogramming progression and a basis for studying cell fate transitions during reprogramming.

RESULTS

Reprogramming Steps Defined by the Dynamics of Xi Chromatin Marks

To define epigenetic steps of XCR and reprogramming, we determined the dynamics of Xi hallmarks during the establishment of pluripotency in mouse embryonic fibroblasts (MEFs). We induced female MEFs to reprogram with retroviruses encoding *Oct4*, *Sox2*, and *Klf4* and analyzed single cells in their native reprogramming environment throughout detailed time courses every other day for 1–3 weeks using multicolor immunostaining (Figure 1A). This allowed us to assess the state of the Xi and of evolving global epigenetic states in any cell of the reprogramming cultures and to delineate the sequence of epigenetic events during reprogramming relative to other markers.

We first analyzed the dynamics of PRC2 on the Xi. EZH2 did not accumulate on the Xi within the first 6 days of reprogramming (Figures 1B and 1C, i). However, after CDH1 (E-cadherin) became expressed, which marks the MET (Li et al., 2010), and before the pluripotency factor NANOG was detectable, a strong nuclear EZH2 staining focus characteristic of Xi accumulation (Xi^{EZH2+}) arose in a small fraction of the cells (Figures 1B, 1C, ii, and 1D). The same result was obtained for SUZ12, another PRC2 subunit, and the PRC2-recruitment factor JARID2 (da Rocha et al., 2014) (Figure S1A available online). The Xi^{EZH2+} was restricted to CDH1+ cells (Figure 1E) and only occurred in a subset of CDH1+ cells, around 50% at day 10 of reprogramming. These findings show that Xi^{EZH2+} arises after an epithelial cell character is established during reprogramming, indicative of the existence of a reprogramming stage immediately downstream to MET that is more restrictive than CDH1 expression. In agreement with this, Xi^{EZH2+} was also present in known late reprogramming intermediates such as pre-iPSCs (Figures S1B–S1D) and was only detectable in reprogramming cultures cotransduced with viruses encoding *Oct4*, *Sox2*, and *Klf4* with or without *cMyc*, but not when fewer reprogramming factors were employed (Figure S1E), demonstrating that the PRC2 composition of the Xi only changes when reprogramming factor combinations able to induce pluripotency are used.

Notably, we also observed that the level of nuclear EZH2 (i.e., on autosomes) gradually increased during reprogramming in both male and female cells, which was initiated specifically in a subset of CDH1+ cells before Xi^{EZH2+} was induced (Figure 1C, progression from i to iv). However, ectopic EZH2 expression in female MEFs did not induce Xi^{EZH2+} (Figure S1F), indicating that the global EZH2 increase during reprogramming is not sufficient for Xi^{EZH2+}. These results reveal that, downstream of MET, PRC2 is gradually upregulated at the global level, irrespective of sex chromosome content, and additionally relocalizes to the Xi in female cells, providing a direct readout of reprogramming progression.

To uncover whether CDH1-positive (CDH1+)/Xi^{EZH2+} cells are intermediates on the path to the NANOG-positive (NANOG+) reprogramming stage, we determined the presence of Xi^{EZH2+} in the first cells that express the NANOG protein during

reprogramming. We found that NANOG activation initiated within a subset of Xi^{EZH2+} colonies, with nearly all NANOG+ cells that first appeared in reprogramming cultures carrying the Xi^{EZH2+} (Figures 1C, iii, and 1F). Later in reprogramming, and usually in large NANOG+ colonies, almost all NANOG+ cells lacked Xi^{EZH2+} (Figures 1C, iv, and 1F), and the absolute number of Xi^{EZH2+} cells and colonies decreased accordingly (data not shown). Moreover, NANOG+ cells were initially surrounded by NANOG-negative (NANOG−)/Xi^{EZH2+} cells (Figure S1G), which is consistent with the induction of NANOG occurring in a subset of CDH1+/Xi^{EZH2+} cells followed by removal of Xi^{EZH2+} within NANOG+ colonies.

H3K27me3, the downstream mark of PRC2, enriched on the Xi in NANOG− cells, was lost from the Xi exclusively within NANOG+ cells with kinetics slightly delayed compared to the loss of Xi^{EZH2+}, such that NANOG+ cells with $Xi^{H3K27me3+}$ but without Xi^{EZH2+} could be briefly detected (Figures 1G and 1H). These data suggest that the loss of $Xi^{H3K27me3+}$ is a consequence of the removal of EZH2 from the Xi. Thus, NANOG expression precedes both loss of Xi^{EZH2+} and $Xi^{H3K27me3+}$.

Taken together, our findings suggest that cells go through defined epigenetic steps as they progress toward pluripotency. Specifically, we reveal four steps by simply following PRC2, at the Xi-specific and global level, relative to CDH1 and NANOG. Downstream of MET, PRC2 proteins increase in overall levels and accumulate on the Xi in a subset of CDH1+ cells. Then, a subset of CDH1+/Xi^{EZH2+} cells reactivates NANOG, which precedes EZH2 and H3K27me3 removal from the Xi specifically in these NANOG+ cells (Figure 1I). Importantly, the reacquisition of Xi^{EZH2+} represents the inversed sequence of events of developmental XCI, where PRC2 accumulates on the Xi immediately after *Xist* RNA initially coats the X and disappears from the Xi later in differentiation (Plath et al., 2003; Silva et al., 2003), suggesting that Xi^{EZH2+} reflects the extent of reversal of the differentiated state established during reprogramming.

The histone variant macroH2A1 associates with the Xi late during developmental XCI and has been shown to act as a barrier to reprogramming (Pasque et al., 2012). We found that the $Xi^{macroH2A1+}$ of MEFs was maintained during reprogramming until Xi^{EZH2+} was lost (Figures S1H–S1K). Unexpectedly, the global level of macroH2A1 first increased from the somatic level before dropping again to the lower level of pluripotent cells in both female and male cells (Figure S1I; data not shown), indicating that an epigenetic mark associated with resistance to reprogramming is transiently induced during the reprogramming process (Figure S1O). Altogether, these results strengthen the conclusion that the Xi-specific and global epigenetic states of cells define multiple stages of reprogramming. Unlike Xi^{EZH2+}, the kinetics of $Xi^{macroH2A1+}$ loss during reprogramming does not represent the reversed sequence of developmental $Xi^{macroH2A1+}$ dynamics, suggesting that distinct mechanisms regulate the temporal Xi accumulation of different epigenetic marks during reprogramming.

Subpopulations with Increased Reprogramming Capacity Recapitulate Xi Events

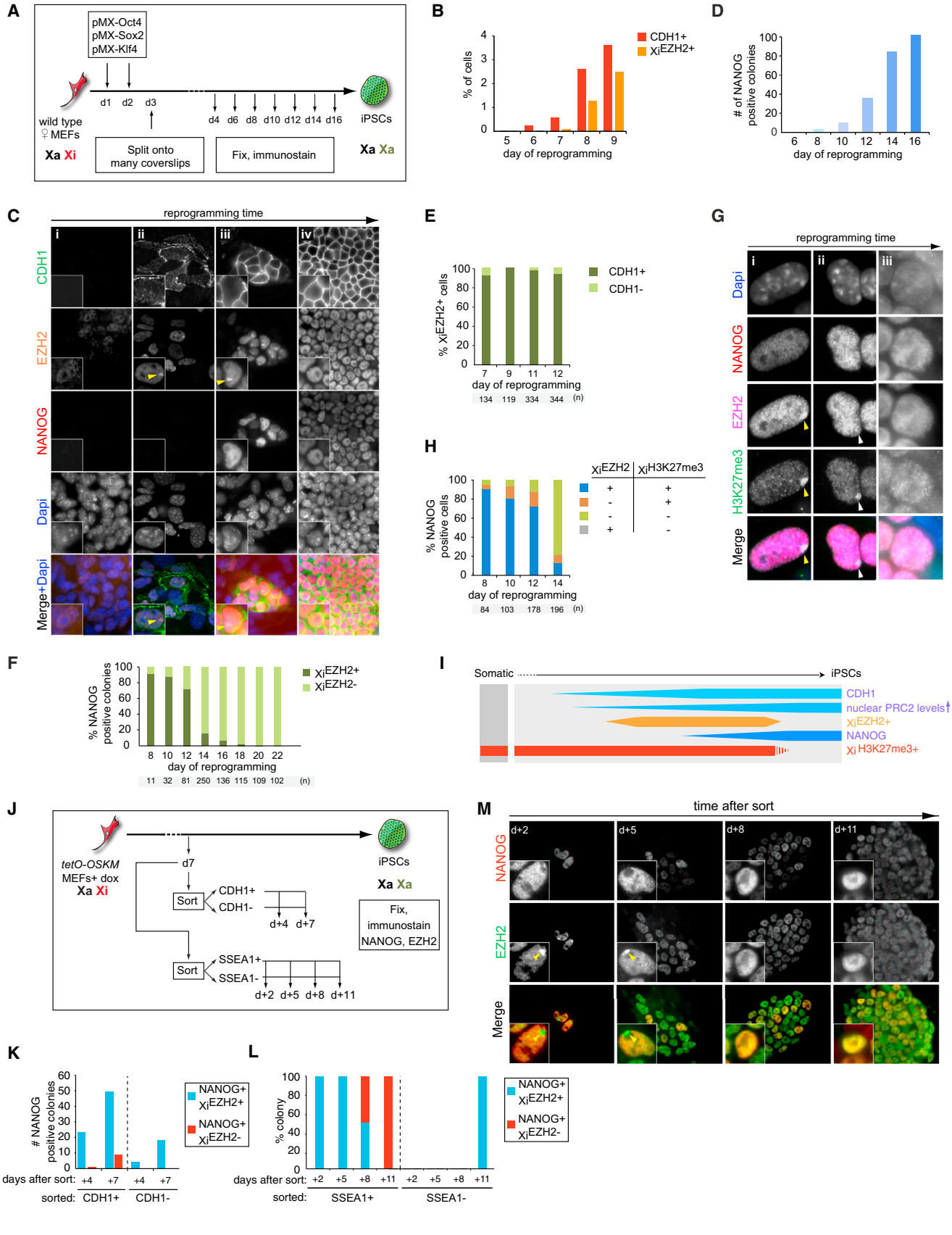
To test whether the steps identified based on fixed cultures represent dynamics of cells that would, if not fixed, continue along the path to pluripotency, we considered the use of pluripotency reporters such as *Oct4-GFP* or *Nanog-GFP*. However, we found that their activation occurred well after the endogenous NANOG protein was detectable, at a time when EZH2 is already removed from the Xi (Figures S1L–S1N), precluding their use for monitoring reprogramming events that occur when NANOG becomes initially expressed.

Instead, we asked whether NANOG+/Xi^{EZH2+} cells arise from CDH1+ cells by sorting CDH1+ and CDH1− cells at day 7 of reprogramming and assessing their ability to give rise to NANOG+/Xi^{EZH2+} cells after replating an equal number of both sorted cell populations (Figures 1J and S1P). We found that CDH1+-sorted cells preferentially gave rise to NANOG+/Xi^{EZH2+} colonies compared to NANOG+/Xi^{EZH2-} colonies (Figure 1K), supporting the conclusion that NANOG+/Xi^{EZH2+} cells originate from CDH1+ cells. Furthermore, in the time frame considered, replated CDH1−-sorted cells also proceeded to the NANOG+/Xi^{EZH2+} state but with delayed kinetics and reduced efficiency (Figure 1K), which is in agreement with the notion that cells reprogram with variable latencies (Hanna et al., 2009). We also performed sorting experiments employing SSEA1, a marker of a reprogramming intermediate arising within CDH1+ cells (Polo et al., 2012) (Figures 1J and S1Q). As expected, shortly after replating, NANOG+ colonies were detected specifically from the SSEA1+ population, and cells within these colonies were initially exclusively Xi^{EZH2+} (Figures 1L and 1M). Remarkably, as these NANOG+ colonies grew bigger over time, they completely lost Xi^{EZH2+} (Figures 1L and 1M). SSEA1− cells gave rise to NANOG+ cells later, and these were all first Xi^{EZH2+} (Figure 1L). Therefore, we conclude that the reprogramming steps defined based on our fixed time courses correctly capture the trajectory of cells moving toward pluripotency.

XCR Occurs after Loss of *Xist* RNA in NANOG+ Cells

To determine the dynamics of *Xist* RNA, the key regulator of developmental XCI, during reprogramming, we combined immunostainings with RNA FISH for *Xist*. Early in reprogramming, virtually all cells showed *Xist* RNA coating, detectable as a large "cloud" of RNA FISH signal (Figure 2A, i). At late reprogramming time points, *Xist* RNA was specifically absent from the Xi within NANOG+ colonies, whereas NANOG− cells still exhibited *Xist* RNA coating (Figure 2A, iii). However, the first NANOG+ cells to appear in culture were always Xi^{Xist+} (Figures 2A, ii, and 2B), indicating that *Xist* repression follows NANOG activation. Furthermore, we found that Xi^{EZH2+} in NANOG+ cells highly correlated with the presence of *Xist* RNA and that their loss occurred with similar dynamics (Figures S2A and S2B), which is consistent with *Xist*-dependent recruitment of PRC2 to the Xi (Plath et al., 2003).

Next, we used RNA FISH to examine when genes on the Xi reactivate during reprogramming (Figure S2C) and found that cells mostly displayed monoallelic expression of the X-linked genes *Mecp2*, *Atrx*, *Gpc4*, and *Rlim* when NANOG+ cells first appeared, indicative of maintenance of XCI in these cells (Figures 2C and S2D–S2F). Later in reprogramming, NANOG+ cells exhibited biallelic expression of these genes, a sign of XCR (Figures 2C and S2D–S2F). For all tested genes, reactivation occurred with delay relative to the loss of *Xist* RNA. We conclude that XCR is a very late event of reprogramming that occurs in a



(legend on next page)

coordinated fashion along the chromosome after *Xist* RNA coating has disappeared. Our results also suggest that XCR takes place independently in multiple cells of a given NANOG+ colony because all cells in NANOG+ colonies initially are Xi^{Xist+}/Xi^{EZH2+}, whereas, at later reprogramming time points, the cells in larger NANOG+ colonies are not (Figure S2A).

To confirm that XCR occurs only in NANOG+ cells, we performed two additional assays. First, we observed that the exclusion of RNA polymerase II from the Xi domain was maintained in all NANOG+ or NANOG− cells that carried $Xi^{H3K27me3+}$ (Figure 2D), which is consistent with the maintenance of silencing at these reprogramming stages. Second, we examined the expression of an Xi-linked GFP reporter (Maherali et al., 2007) and did not detect GFP reporter reactivation in NANOG− cells, whereas NANOG+ cells consistently expressed GFP at late reprogramming time points (Figure S2G). We conclude that, late in reprogramming, NANOG expression precedes the loss of *Xist* RNA, which coincides with loss of Xi^{EZH2+} and occurs before XCR (Figure 2E).

Reprogramming Reverses the Developmental Sequence of *Tsix* Expression

To establish the dynamics of activation of *Tsix* (transcribed antisense to *Xist)* RNA, a critical regulator of *Xist* during initiation of XCI (Lee and Bartolomei, 2013), we used strand-specific RNA FISH and found that *Tsix* was not expressed during the early stages of reprogramming (i.e., in NANOG− cells) or in the first NANOG+ cells that appear (Figure 2F). Within maturing NANOG+ cells, however, *Tsix* became first monoallelically expressed in cells still carrying Xi^{Xist+}. The monoallelic *Tsix* signal occurred specifically from the active X chromosome (Xa), as it never overlapped with Xi^{Xist+} (Figure 2G). *Tsix* activation on the Xi took place later, at the very tail end of Xi^{Xist} loss (Figures 2G and S2H). Together, these results show that reprogramming to pluripotency recapitulates the expression of *Tsix* in the reverse order from that of developmental XCI, where *Tsix* is first downregulated on the future Xi and then becomes repressed on the Xa (Lee and Lu, 1999) (Figure 2E).

Kinetics of XCR in Relation to Pluripotency Gene Activation

Given the stepwise changes of Xi hallmarks late in reprogramming, we aimed to determine the dynamics of these features in relation to the activation of pluripotency-associated factors. In agreement with the reported hierarchical activation of pluripotency factors late in reprogramming based on single-cell transcript analysis (Buganim et al., 2012), we observed the sequential induction of the pluripotency factors ESRRB, REX1, DPPA4, and PECAM1 at the single-cell level using multicolor immunostaining, which only occurred in NANOG+ cells (Figures S3A–S3F). In addition, silencing of the reprogramming factor-expressing retroviruses can be placed early in this hierarchy at around the time of ESRRB/REX1 activation (Figures S3G–S3I), which is consistent with the shift to endogenous pluripotency factor activation. Xi^{EZH2+} was lost after REX1 expression and just after DPPA4 activation (Figures 3A and 3B). PECAM1 expression, which is very late in the pluripotency factor hierarchy, marked cells that are devoid of Xi^{Xist+} and Xi^{EZH2+} (Figures 3C and 3D). Consistent with a delay of XCR relative to *Xist* RNA loss, XCR took place after DPPA4 activation, as small DPPA4+ colonies

Figure 1. Time Course Analysis of Xi^{EZH2+} during Reprograming to Pluripotency

(A) Diagram of reprogramming time course experiments. In all experiments, results for female cells are displayed except when stated otherwise, and the time points and number (n) of cells or colonies counted are given in each subfigure.

(B) Quantitation of the proportion of CDH1+ or Xi^{EZH2+} cells at indicated reprogramming time points. 100 cells in three randomly chosen microscopic fields were counted per time point.

(C) Multicolor immunostaining for CDH1 (green in merge), EZH2 (orange), and NANOG (red) at different stages of reprogramming. Dapi staining (blue) marks nuclei. Unlike MEFs (i), cells with elevated nuclear levels of EZH2 and Xi^{EZH2+} (arrowhead) are seen within CDH1+ cells during reprogramming starting around day 7 of reprogramming (ii). (iii) NANOG+ colonies are first marked by Xi^{EZH2+} and elevated EZH2 levels in the nucleus (image from day 9). (iv) Later, NANOG+ colonies become larger and are characterized by high nuclear EZH2 levels without Xi^{EZH2+} (image from day 14).

(D) Number of NANOG+ colonies throughout reprogramming (a colony is defined as four or more closely localized cells).

(E) Proportion of Xi^{EZH2+} cells with and without CDH1 expression during reprogramming.

(F) Proportion of NANOG+ colonies with or without Xi^{EZH2+} at indicated time points. All NANOG+ colonies present in the reprogramming cultures were counted up to day 14 and only a subset thereafter.

(G) Multicolor immunostaining for EZH2 (magenta in merge) and NANOG (red) in combination with H3K27me3 (green). The images depict various states of Xi^{EZH2+} and $Xi^{H3K27me3+}$ in NANOG+ cells quantified in (H). During reprogramming, (i) NANOG+/Xi^{EZH2+} cells are initially $Xi^{H3K27me3+}$ (ii) and, at a later time point, become $Xi^{EZH2-}/Xi^{H3K27me3+}$ for a very short time and subsequently become (iii) $Xi^{EZH2-}/Xi^{H3K27me3-}$. Yellow and white arrowheads indicate $Xi^{EZH2+}/Xi^{H3K27me3+}$ and $Xi^{EZH2-}/Xi^{H3K27me3+}$ patterns, respectively.

(H) Quantitation of the immunostaining experiment in (G), giving the proportion of NANOG+ cells with Xi^{EZH2+} or $Xi^{H3K27me3+}$ at indicated time points.

(I) Summary of Xi and global dynamics of PRC2 and H3K27me3 during reprogramming, relative to CHD1 and NANOG expression. Female-specific features are shown in orange/red, and those occurring in both female and male reprogramming are shown in blue. The width of the boxes represents the level of the epigenetic mark considered.

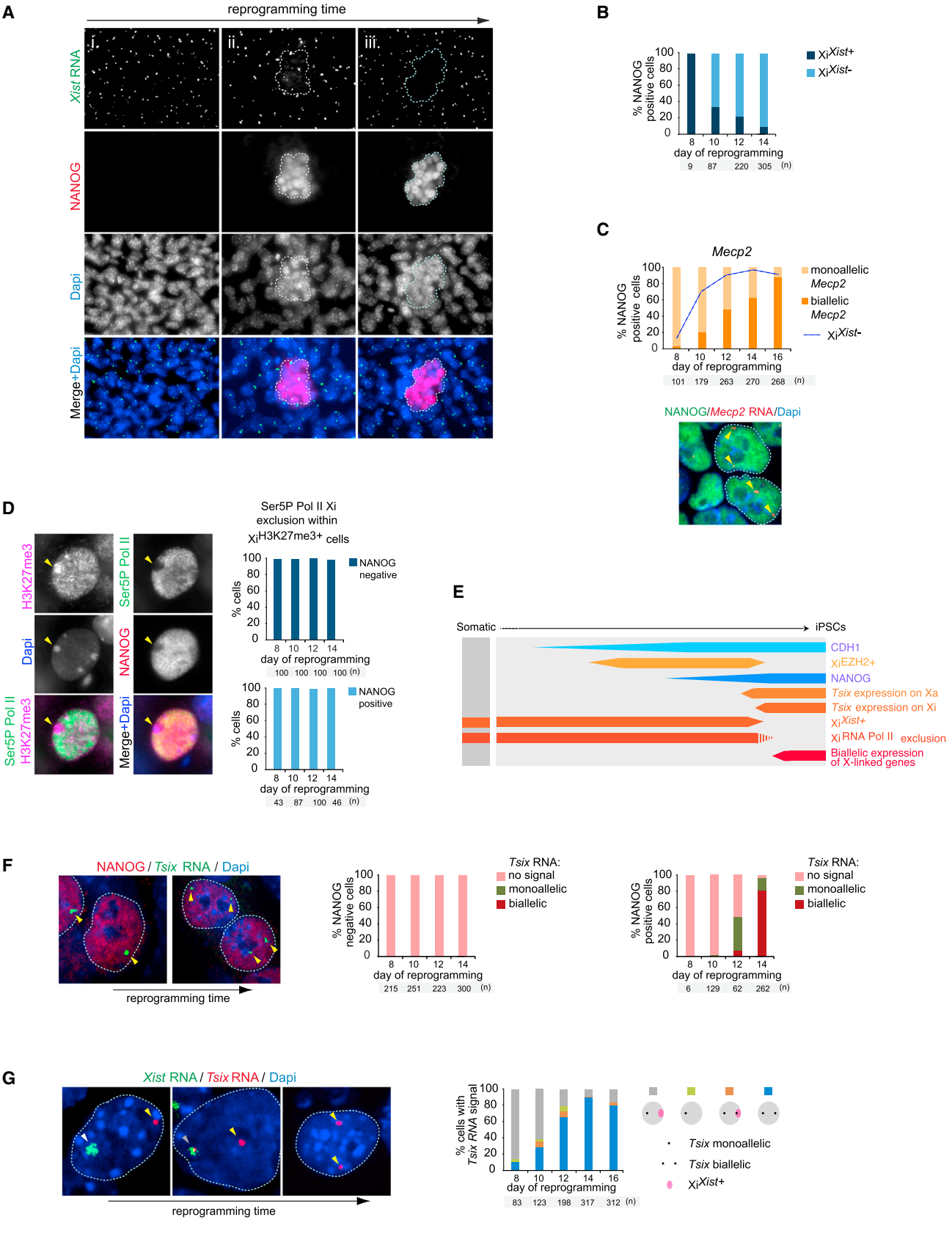
(J) Experimental design for the isolation and characterization of CDH1+/− or SSEA1+/− reprogramming subpopulations.

(K) Number of NANOG+ colonies with or without Xi^{EZH2+} in CDH1+ and CDH1− sorted cell populations isolated as shown in (J), at indicated days after replating.

(L) Proportion of NANOG+ colonies with or without Xi^{EZH2+} in SSEA1+ and SSEA1− sorted cell populations isolated as shown in (J), at indicated days after replating. n = 6 for each SSEA1+ time point, and n = 1 for the SSEA1− count at +d11. Colonies appearing in SSEA1+ replated cells become larger throughout this time course as shown in (M).

(M) Visualization of Xi^{EZH2+} changes in replated SSEA1+ reprogramming intermediates over time from the experiment shown in (J) and (L). Replated cells were immunostained for EZH2 (green in merge) and NANOG (red) at the indicated days. Note the increase in colony size and disappearance of Xi^{EZH2+} (yellow arrowhead) with time in culture.

See also Figure S1.



(legend on next page)

expressed the X-linked gene *Atrx* monoallelically and large colonies expressed the X-linked gene mostly biallelically (Figure 3E). Together, these results demonstrate the molecular timeline of Xi changes relative to hierarchical pluripotency gene activation (summarized in Figure 3F).

Sequential Xi States Are Conserved across Different Reprogramming Systems

To establish whether the sequential changes on the Xi during reprogramming are specific to the reprogramming system used, we reprogrammed MEFs carrying a single dox-inducible, polycistronic reprogramming cassette encoding *Oct4*, *Sox2*, *Klf4*, and *cMyc* in a defined locus instead of retroviral infection. In addition, we reprogrammed another starting cell type, mouse embryonic endoderm cells, and also used different culture conditions. The dynamics of Xi^{EZH2+} and sequential pluripotency gene activation were reproduced in each case (Figures S3J–S3N). We conclude that the epigenetic states identified represent fundamental changes inherent to reprogramming, applicable to multiple starting cell types and reprogramming systems.

CDH1 and NANOG Are Required, but Not Sufficient, for the Efficient Induction of Reprogramming Steps Leading to XCR

Our data indicate that CDH1 expression marks the cells that subsequently induce Xi^{EZH2+} and NANOG and that only those cells that activate NANOG are fated to induce *Xist* loss, *Tsix* activation on the Xi and Xa, pluripotency-associated factor activation, and XCR, suggesting that both CDH1 and NANOG are critical for this hierarchy of events. To address the role of CDH1 and NANOG for epigenetic changes taking place downstream of their expression, we performed both knockdown and overexpression experiments. We found that knockdown of *Cdh1* during reprogramming with shRNAs decreased the number of Xi^{EZH2+}, NANOG+, and DPPA4+ colonies (Figures 4A and S4A). In contrast, *Cdh1* overexpression did not promote any of the epigenetic events that normally take place downstream of CDH1 induction (Figures S4B–S4E). Depletion of *Nanog* transcripts during reprogramming using an inducible shRNA (Figures S4F and S4G) did not prevent CDH1 activation, global upregulation of EZH2, and Xi^{EZH2+} (Figures S4H and S4I), in agreement with its induction later during reprogramming (Silva et al., 2009). By contrast, the activation of *Tsix* on the Xa and Xi, as well as of pluripotency-associated transcription factors, was strongly reduced by *Nanog* depletion (Figures S4J–S4M). The lack of biallelic *Tsix* expression in the absence of *Nanog* also suggests that XCR was impaired without *Nanog*. Thus, NANOG orchestrates the efficient transition through the later molecular events, including XCR, although the requirement for *Nanog* can be bypassed (Carter et al., 2014; Schwarz et al., 2014). Overexpression of *Nanog* late in reprogramming promoted steps toward XCR as judged by the increased number of DPPA4+/Xi^{EZH2-} colonies, but not those before NANOG is normally induced (i.e., Xi^{EZH2+}) (Figures 4B, S4N, and S4O). However, most NANOG-overexpressing cells did not induce the subsequent reprogramming steps. Together, these results demonstrate that both *Cdh1* and *Nanog* are required, but not sufficient, for the induction of the epigenetic events leading to XCR.

The above finding raised the question of whether XCR represents a barrier to reprogramming. To test this, we obtained a large number of female and male MEFs preparations from four independent litters and measured the efficiency with which NANOG+, DPPA4+, or PECAM1+ colonies formed, without prior knowledge of the sex. This experiment revealed no difference in the reprogramming efficiency between male and female MEFs in KSR or FBS culture media and in the transition to different reprogramming stages (Figures 4C and 4D). Thus, even though the Xi represents the most extreme form of facultative heterochromatin, XCR does not limit reprogramming to induced pluripotency.

Requirement of *Xist* Silencing, but Not *Tsix* Expression, for XCR

To examine the molecular mechanism of XCR during reprogramming, we focused on the requirement for *Xist* and for *Tsix*, its negative regulator in pluripotent cells (Lee and Bartolomei, 2013). Despite *Tsix* becoming expressed on the Xi as *Xist* RNA disappears (Figure 2), deletion of *Tsix* did not alter the kinetics of *Xist* repression in NANOG+ cells (Figures 5A and 5B), indicating that *Tsix* does not negatively regulate *Xist* at the end of reprogramming. Conversely, to test whether repression of *Xist* RNA is required for XCR, we ectopically expressed *Xist* from the Xi during reprogramming (Figure 5C). Constitutive *Xist*

Figure 2. Kinetics of *Xist* and *Tsix* RNA and XCR during Reprogramming

(A) Representative images of *Xist* RNA (green in merge), NANOG (red), and Dapi (blue) from immunoFISH stainings at different time points of reprogramming reflecting different states of Xi^{Xist+} and NANOG expression as determined in (B). Dotted lines indicate the position of NANOG+ colonies across different channels. (i) day 8, (ii) day 10, and (iii) day 14. Each image represents a series of ten Z-sections merged onto a single plane.

(B) Proportion of NANOG+ cells with or without Xi^{Xist+} at different time points.

(C) ImmunoFISH analysis of NANOG expression and nascent transcripts of the X-linked gene *Mecp2* (seen as a strong pinpoint) during reprogramming. In the image, the biallelic *Mecp2* expression pattern is indicated (two arrowheads), and the proportion of NANOG+ cells with mono- or biallelic *Mecp2* expression is given below. The dotted line indicates the proportion of Xi^{Xist-} cells from the same time course.

(D) Images depict an immunostaining for H3K27me3, Ser5P polymerase II (Ser5P Pol II), and NANOG at day 12 of reprogramming. Quantification gives the proportion of Ser5P Pol II Xi-exclusion cells in NANOG− (top) or NANOG+ (bottom) cells that also display $Xi^{H3K27me3+}$ at indicated time points.

(E) Summary of reprogramming stages related to this figure, displayed as described in Figure 1I. Dashed lines indicate the window of time we narrowed down for the feature to occur or disappear.

(F) ImmunoFISH analysis as in (C), except for NANOG (red) and *Tsix* RNA (green). Proportion of NANOG− (left) and NANOG+ (right) cells, respectively, displaying monoallelic, biallelic, or no *Tsix* RNA FISH signal.

(G) RNA FISH analysis of the relationship between *Xist* and *Tsix* RNA in reprogramming. In the images, yellow arrowheads represent *Tsix* expression without *Xist* RNA present on the same X chromosome, and the white and gray arrowheads represent an *Xist* RNA cloud that does not or does overlap a *Tsix* signal, respectively. The quantification of cells with *Tsix* expression showing mono- or biallelic *Tsix* expression with and without Xi^{Xist+} is shown.

See also Figure S2.

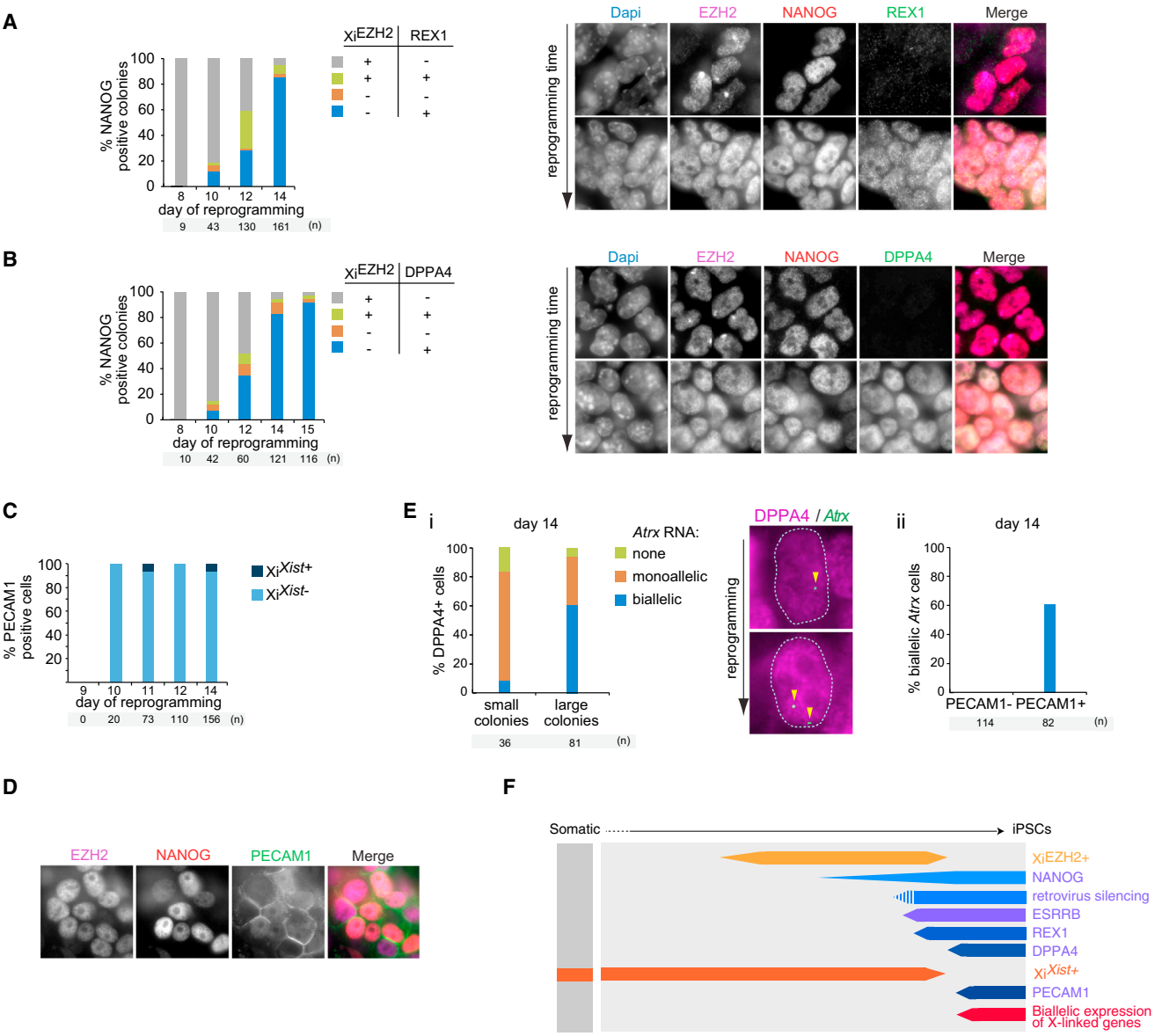


Figure 3. Xi Features in Relation to the Sequential Expression of Pluripotency Factors during Reprogramming

(A) Quantitation of an immunostaining analysis of NANOG, Xi^{EZH2+}, and REX1, presenting the proportion of NANOG+ colonies with Xi^{EZH2+} and/or REX1 expression at indicated time points. Right, representative immunostaining images for EZH2 (magenta in the merge), NANOG (red), and REX1 (green).

(B) As (A) for NANOG, Xi^{EZH2+}, and DPPA4.

(C) Quantitation of an immunoFISH analysis for PECAM1 and *Xist* RNA displaying the proportion of PECAM1+ cells with Xi^{Xist+}.

(D) Representative immunostaining image for EZH2 (magenta in the merge), NANOG (red), and PECAM1 (green), demonstrating the absence of Xi^{EZH2+} in PECAM1+ cells.

(E) (i) Quantitation of mono- and biallelic expression of the X-linked gene *Atrx* within cells of small (<12 cells) and large (>20 cells) DPPA4+ colonies and representative images from the DPPA4 (magenta) and *Atrx* (green) immunoFISH staining. Arrowheads indicate *Atrx* nascent transcription signals. (ii) Quantification of biallelic *Atrx* expression in PECAM1+/− cells.

(F) Summary of reprogramming stages identified in this figure as in Figure 2E.

See also Figure S3.

expression did not alter the efficiency by which ESRRB+ colonies appeared but resulted in a decrease in XCR within NANOG+ cells, as measured by the extent of biallelic *Atrx* expression (Figures 5D–5F). Thus, *Xist* silencing at the end of reprogramming is necessary for XCR.

To determine whether XCR depends solely on *Xist* repression, we asked whether *Xist* deletion leads to precocious activation of the Xi. Specifically, we deleted *Xist* early in the reprogramming process using female MEFs homozygous for a conditional (2lox) *Xist* allele (Csankovszki et al., 2001), which also carried a

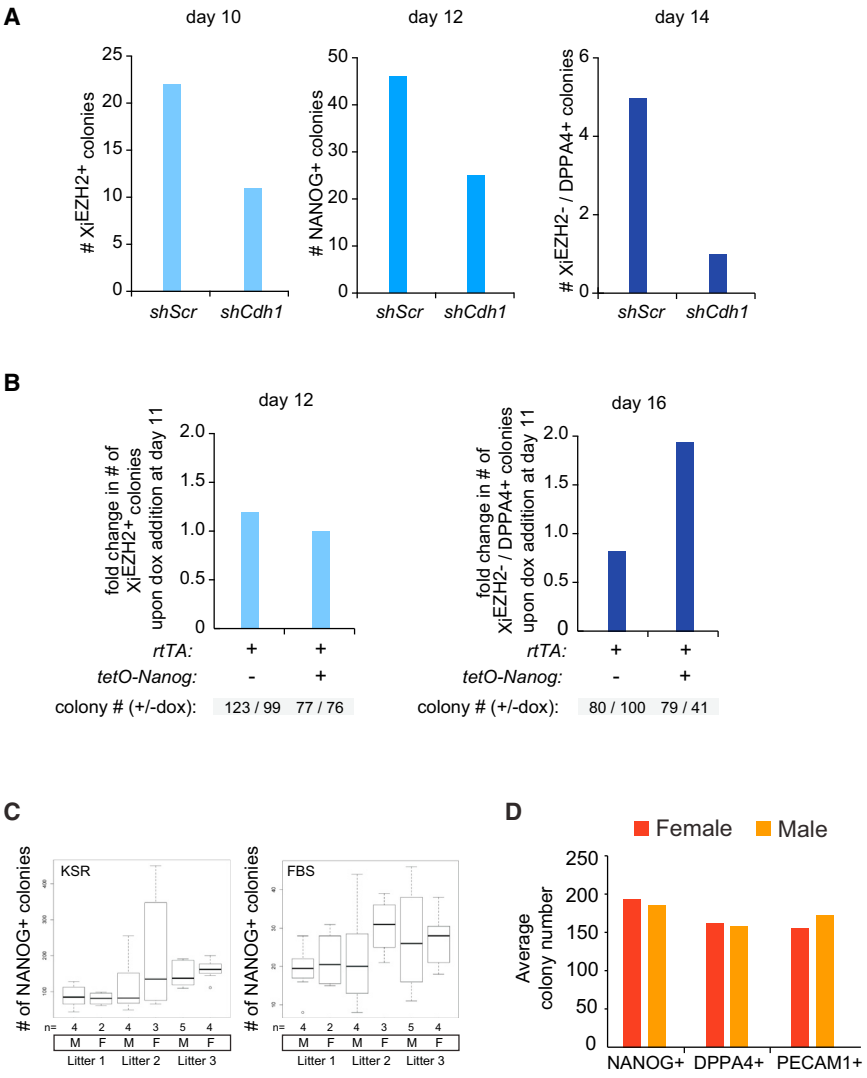


Figure 4. *Cdh1* and *Nanog* Modulate the Efficiency of Reprogramming and Dynamics of Xi Hallmarks, whereas XCR Does Not Represent a Reprogramming Barrier

(A) Number of Xi^{EZH2+}, NANOG+, and DPPA4+/Xi^{EZH2-} colonies obtained when *Cdh1* is knocked down by shRNAs (*shCdh1*) throughout reprogramming, compared to scrambled shRNA (*shScr*) reprogramming experiments.

(B) Reprogramming experiments with female MEFs carrying *rtTA* only or *rtTA* and the *tetO-Nanog* allele, and with and without dox addition at day 11. The number of Xi^{EZH2+} and DPPA4+/Xi^{EZH2-} colonies was determined at day 12 and 16 of reprogramming, respectively, and was plotted as fold change between dox relative to no dox treatment per cell line. Note that Xi^{EZH2+} counts are similar at day 12, whereas DPPA4+/Xi^{EZH2-} counts differ at day 14 when *Nanog* is overexpressed.

(C) Comparison of reprogramming efficiency between male (M) and female (F) MEFs. Box plots depict the number of NANOG+ colonies for reprogramming experiments in KSR and FBS media, respectively, with male and female MEFs, at day 14 and day 25 of reprogramming, respectively. MEFs are grouped by litter and the number of male and female MEF populations per litter is given (n). Whiskers demarcate the minimum and maximum of the data.

(D) Quantitation of different late reprogramming stages for MEFs isolated from seven female and seven male embryos, as judged by the number of NANOG+, DPPA4+, or PECAM1+ colonies at day 14.

See also Figure S4.

dox-inducible Cre recombinase (Figures 5G and 5H). *Xist* ablation had no effect on the efficiency with which NANOG+ colonies were generated (Figure 5I) and, surprisingly, did not alter XCR kinetics (Figure 5J). Therefore, *Xist* repression is necessary, but not sufficient, for XCR to occur, indicating the existence of other mechanisms that can maintain Xi silencing throughout reprogramming and even initially in NANOG+ cells.

High Persistence of Xi DNA Methylation during Reprogramming

We considered the possibility that DNA methylation could maintain the silent state of the Xi during reprogramming in the absence of *Xist*. DNA methylation at CpG islands of the Xi arises late in the sequence of epigenetic changes on the Xi during development (Gendrel et al., 2012). We examined the DNA methylation pattern of X-linked genes in SSEA1− and SSEA1+ subpopulations, isolated from day 9 reprogramming cultures, representing cell populations with different reprogramming capabilities (Figures 1L and 1M; Polo et al., 2012). Traditional bisulfite sequencing at promoters of the X-linked genes *Atrx* and *Rlim* demonstrated the presence of the hypermethylated Xi in female MEFs, as well as in SSEA1− and SSEA1+ subpopulations, but not in female ESCs (Figure 6A). In contrast, the *Nanog* promoter region, methylated at an intermediate level in MEFs and SSEA1− cells, displayed demethylation characteristic of pluripotent cells already in SSEA1+ cells. These findings suggested a differential persistence of the methylation mark between *Nanog* and Xi-linked genes.

To determine the DNA methylation status along the entire X chromosome, we employed reduced representative bisulfite sequencing (Meissner et al., 2008), which provides genome-scale single-base-resolution maps of DNA methylation. For this analysis, we additionally included early passage female iPSCs, as well as male MEFs and male ESCs for comparison. CpG islands on the Xa in male cells were hypomethylated to the same degree as those on autosomes in male or female cells (Figure 6B). By contrast, in female MEFs, CpG islands across the X chromosome showed an average of 20%–50% methylation, which is consistent with an Xi-specific methylation signature. This pattern was present in both SSEA1− and SSEA1+ subpopulations but absent in early-passage female iPSCs and female ESCs (Figure 6B). A similar result was obtained for CpG-island

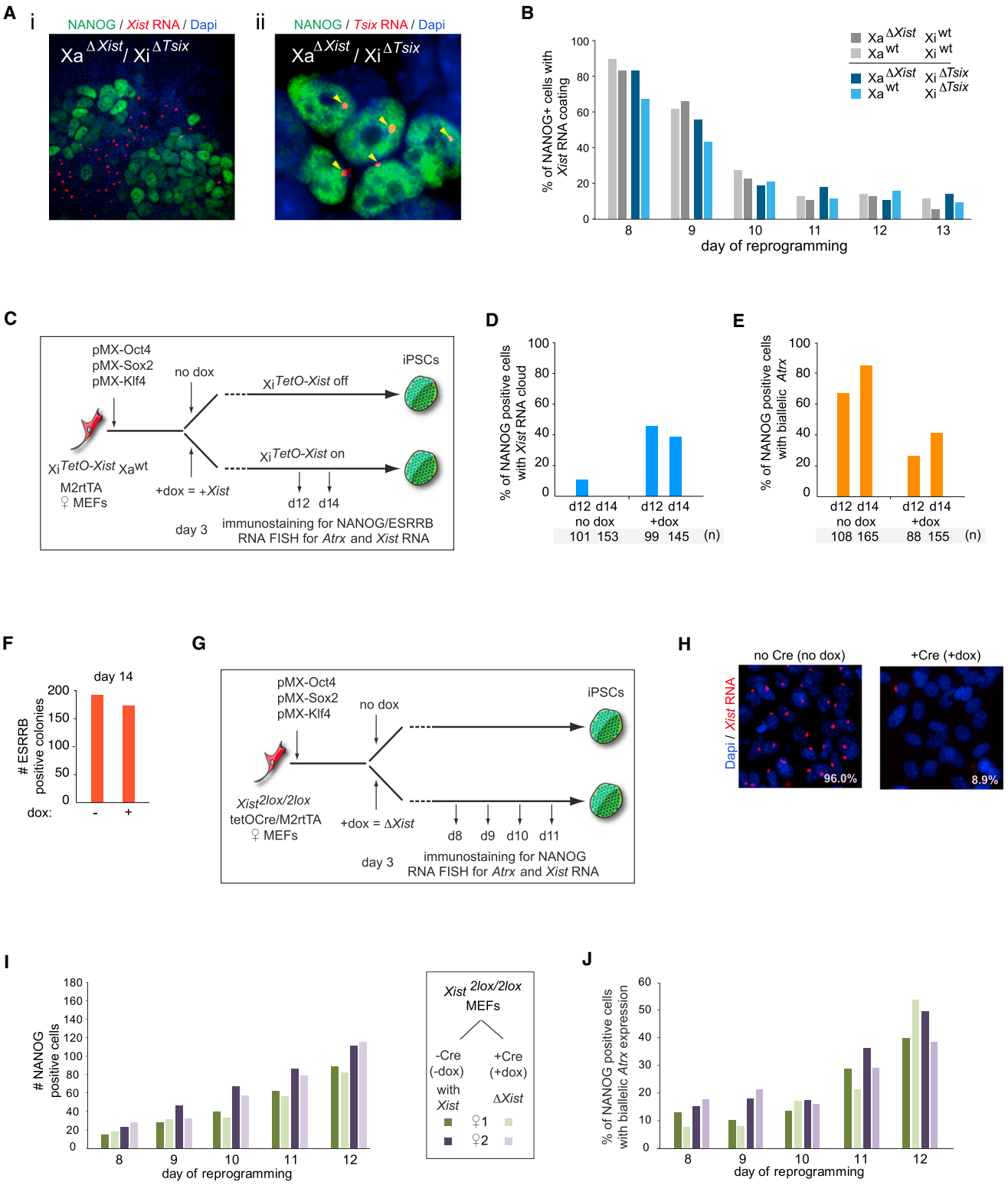


Figure 5. *Xist* Silencing Is Necessary, but Not Sufficient, for XCR

(A) (i) Representative image of an immunoFISH analysis for NANOG (green) and *Xist* RNA (red) at day 14 of reprogramming with MEFs carrying a deletion of *Tsix* on the Xi and of *Xist* on the Xa, illustrating that NANOG+ cells lose *Xist* RNA accumulation on the Xi even in the absence of *Tsix* on the Xi. (ii) iPSCs derived from the experiment in (i) were stained for NANOG (green) and *Tsix* RNA (red), confirming monoallelic expression of *Tsix* due to deletion on one X chromosome (arrowheads).

(legend continued on next page)

shores, high and low CpG-containing promoters (Figure S5A). These results indicate that DNA methylation established on the Xi late during differentiation (Gendrel et al., 2012) is preserved on the Xi until very late in reprogramming. The persistence of Xi-DNA methylation in reprogramming is *Xist* independent (Figures S5B and S5C), supporting the hypothesis that this Xi mark could maintain the silent state of the Xi until late in reprogramming, even when *Xist* is experimentally deleted. Because Xi-DNA methylation is not yet reversed when *Nanog* (Figure 6A) and many other ESC-specific enhancer elements have already become demethylated (V.P., R.K., C. Chronis, A.M., and K.P., unpublished data), we conclude that this Xi mark has a remarkable stability during reprogramming.

Xist RNA and DNA Methylation Both Maintain Xi Silencing Throughout Reprogramming

We determined whether XCR is mechanistically linked to the loss of both *Xist* and DNA methylation by deleting *Xist* and inhibiting *Dnmt1*, the maintenance DNA methyltransferase, during the late phase of reprogramming. The block of *Dnmt1* activity and loss of DNA methylation within Xi-linked CpG islands was confirmed (Figures S5D and S5E). In reprogramming cultures in which *Xist* on the Xi was experimentally deleted, 23% of NANOG+ cells displayed biallelic expression of the X-linked gene *Atrx* upon brief *Dnmt1* depletion, and this proportion was more than doubled when *Dnmt1* knockdown was combined with 5AzadC treatment to enhance DNA demethylation (Figures 6C and S5E). XCR was not detected at this time point in NANOG+ cells in control reprogramming cultures (Figure 6C). Importantly, we found that inhibition of DNA methylation only enhances XCR in NANOG+ cells in the absence of *Xist*, but not in its presence (Figure 6D). This finding also excludes the possibility that the acceleration of XCR upon inhibition of DNA methylation is simply due to faster overall reprogramming. We conclude that *Xist* RNA is able to maintain the Xi when DNA methylation is reduced and that DNA methylation is sufficient for Xi maintenance in the absence of *Xist*. Therefore, both DNA demethylation and *Xist* silencing are required for XCR late in reprogramming and occur downstream to the reactivation and demethylation of *Nanog* (Figure 6E).

Tet1, *Tet2*, and High Global 5hmC Levels Are Dispensable for XCR

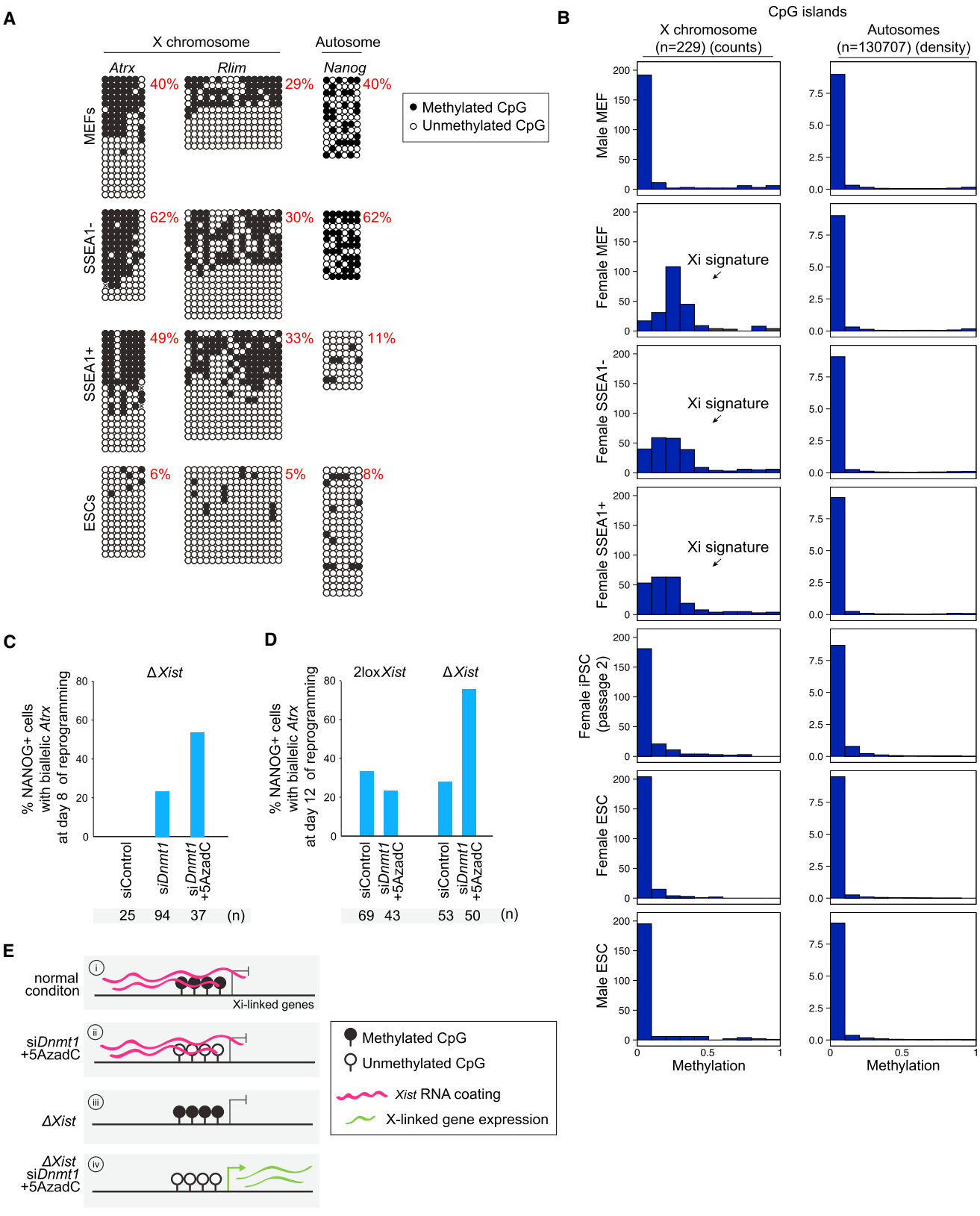
Given the implication of conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) in DNA demethylation processes (Wu and Zhang, 2014), we defined the Xi-specific and global dynamics of 5hmC during reprogramming. We found a striking increase in the global 5hmC level, specifically in those cells that globally upregulate EZH2 and gain Xi^{EZH2+} (Figures 7A–7C). Global upregulation of 5hmC also took place in male reprogramming cultures and in the absence of Vitamin C (Figure S6A), indicating that this epigenetic remodeling event is intrinsic to reprogramming across different culture conditions and sex chromosome content. Despite overall elevated 5hmC levels, this mark was depleted on the Xi in Xi^{EZH2+} cells (Figures 7D and 7E). Thus, during reprogramming, cells start off with low levels of 5hmC and EZH2 and then increase 5hmC and EZH2 downstream of MET, with PRC2 accumulating on the Xi and 5hmC remaining excluded from the Xi, all of which precedes the reactivation of pluripotency genes and transition to a pluripotent state with XCR devoid of Xi^{PRC2+} and 5hmC Xi exclusion.

Given the dynamics of 5hmC, we tested the requirement of *Tet1* and *Tet2* for XCR using female MEFs carrying *Tet1* knockout (*Tet1*$^{-/-}$) and *Tet2* conditional (*Tet2*$^{2lox/2lox}$) alleles, in which genetic deletion of *Tet2* could be induced by addition of Cre-expressing adenoviruses (AdCre) (Figure S6B). Strikingly, genetic ablation of both *Tet1* and *Tet2*, but not that of either *Tet1* or *Tet2* individually, prevented the global induction of 5hmC during reprogramming (Figures 7F, S6C, and S6D). Importantly, *Tet1*/*Tet2* double knockout and absence of global 5hmC did not affect the upregulation of nuclear EZH2 and occurrence of Xi^{EZH2+}, nor the efficiency with which NANOG+ colonies were obtained, nor the activation of the late pluripotency marker PECAM1 and XCR (Figures 7G, 7H, and S6E–S6G). Reprogramming experiments with ablation of either *Tet1* or *Tet2* resulted in similar results, and the resulting iPSCs contributed to chimeras and were effectively demethylated at *cis*-regulatory regions of the *Pou5f1* (*Oct4*) gene and Xi-linked promoters (Figures S6H–S6K and S7A–S7D). Additional shRNA-mediated depletion of *Tet3* transcripts in pre-iPSCs also carrying the *Tet1* and *Tet2* genetic deletion still enabled XCR (Figures S7E–S7K). We conclude that *Tet1* and *Tet2* and the global increase in 5hmC nuclear levels are dispensable for XCR and the transition through the reprogramming hierarchy that we have established.

DISCUSSION

A dramatic reorganization of the epigenome occurs during the reprogramming of somatic cells to iPSCs. Our findings demonstrate that changes in global and Xi-specific chromatin states, noncoding RNA expression, and pluripotency-associated factor

(B) Kinetics of *Xist* RNA loss in the absence of *Tsix* on the Xi. Proportion of NANOG+ cells with Xi^{Xist+} in reprogramming time courses performed with MEFs with (gray bars) and without (blue bars) *Tsix* on the Xi. Just like *Tsix* deletion, the additional deletion of *Xist* on the Xa (dark versus lighter bars) does not affect the kinetics of *Xist* RNA loss in NANOG+ cells.
(C) Diagram of *Xist* overexpression reprogramming experiments using MEFs in which the promoter of *Xist* on the Xi is replaced with a tet-inducible promoter.
(D) Proportion of NANOG+ cells with Xi^{Xist+} in reprogramming cultures described in (C) with and without ectopic *Xist* induction conditions (+/−dox), based on immunoFISH analysis.
(E) As (D), but for NANOG+ cells with biallelic *Atrx* expression.
(F) As (D), but for the number of ESRRB+ colonies.
(G) Diagram of the *Xist* deletion reprogramming experiments with female conditional *Xist* MEFs.
(H) *Xist* RNA FISH for MEFs described in (G) under control (−dox) and +dox conditions, the latter leading to *Xist* RNA loss in the majority of cells.
(I) Number of NANOG+ colonies at various time points of reprogramming for the experiment described in (G) under control (no dox/−Cre) and the *Xist* deletion (+dox/+Cre) conditions.
(J) As in (I), but quantitation of NANOG+ cells with biallelic *Atrx* expression based on immunoFISH analysis.



(legend on next page)

expression are highly reproducible and reveal the existence of a multitude of epigenetic steps that occur in a defined sequence throughout the reprogramming process (Figure 7I, i). For instance, focusing only on Xi^{EZH2+} dynamics relative to CDH1 and NANOG expression, transition through four steps can be defined: (1) CDH1+/Xi^{EZH2-}/NANOG-; (2) CDH1+/Xi^{EZH2+}/NANOG−; (3) CDH1+/Xi^{EZH2+}/NANOG+; and (4) CDH1+/Xi^{EZH2-}/NANOG+ (Figure 7I, ii). These stages are likely going to be generally applicable to female cells and not cell-type specific as the Xi enrichment of PRC2 is also expected to occur in epithelial cells at an intermediate step of reprogramming. The relocalization of EZH2 (PRC2) and its cofactor JARID2 to the Xi, along with global increases in PRC2, macroH2A1, and 5hmC downstream of MET and upstream of NANOG expression, indicate that major changes in chromatin structure take place in cells undergoing reprogramming, before pluripotency is reached.

Compared to the establishment of Xi features during differentiation, we find that these have different propensities for reversal during reprogramming. Whereas Xi^{EZH2+} and the activation of *Tsix* from the Xa and Xi take place in an apparent reverse order of the developmental XCI program, macroH2A1 and DNA methylation, both associated with the differentiated state and resistance to reprogramming (Pasque et al., 2012; Mikkelsen et al., 2008), are reversed on the Xi only very late in reprogramming, despite the fact that they are established on the Xi late in differentiation. Similarly, the activation of Xi-linked genes during reprogramming only occurs after *Xist* RNA loss, even though *Xist* RNA coating precedes silencing of the X chromosome during differentiation. Thus, based on a subset of Xi hallmarks, reprogramming proceeds in a manner that would be expected for developmental reversal, indicating progressive dedifferentiation. However, based on another set of marks, cells undergoing reprogramming remain epigenetically distinct from those traversing differentiation. Thus, during reprogramming, certain epigenetic features follow the differentiation state of the cell, whereas others are uncoupled from this regulation.

During differentiation, *Xist* is required to initiate XCI, and its experimental silencing in the first days after the establishment of the Xi leads to immediate reactivation of the X chromosome (Wutz and Jaenisch, 2000). However, later in differentiation, *Xist* can be deleted from the Xi without dramatically affecting the stability of the silent chromosome, which, at this point, is thought to be maintained through the action of multiple repressive chromatin pathways (Csankovszki et al., 2001). In contrast to a recent study that used a different system to reduce *Xist* expression (Chen et al., 2014), we made the surprising observation that *Xist* ablation on the Xi does not alter the kinetics of XCR during reprogramming. Our result indicates that the extended, several-day-long window of *Xist* dependency of silencing seen during the initiation of XCI in differentiation (Wutz and Jaenisch, 2000) is not re-established during reprogramming. Silencing of the Xi in the absence of *Xist* remains stable until the very end of reprogramming because it is functionally maintained by DNA methylation, which has an extraordinarily high persistence on the Xi during reprogramming and is only erased after the pluripotency factor *Nanog* is already demethylated. Notably, the experimental interference with DNA methylation alone does not lead to precocious XCR, indicating that *Xist* RNA also actively contributes to the silencing of the Xi late in reprogramming. In agreement with this, we also discovered that forced *Xist* expression prevents XCR during reprogramming. Therefore, XCR requires loss of both *Xist* RNA and DNA methylation at the end of the reprogramming process. Because both events take place only late during hierarchical pluripotency-associated gene activation, these ensure that XCR only occurs in cells that establish faithful pluripotency (Figure 7I, iii). Accordingly, a block early in the pluripotency hierarchy blocks XCR (Figure S4). Notably, the pluripotency factor PRDM14 has been reported to be required for XCR during reprogramming (Payer et al., 2013), but whether *Prdm14* deletion blocks the reprogramming process at a stage prior to XCR needs to be resolved to understand its specific role in XCR.

The generation of 5hmC by Tet proteins has been suggested to play important roles during reprogramming to iPSCs and potentially mediates DNA demethylation through active and passive mechanisms (Hu et al., 2014; Wu and Zhang, 2014). Our findings reveal that *Tet1*, *Tet2*, and global 5hmC are dispensable for XCR. This raises the question of which DNA demethylation pathway, either active or passive, leads to XCR. We posit that loss of DNA methylation on the Xi during reprogramming likely occurs in a synchronous manner across the entire chromosome, requiring a mechanism that can act across a large number of CpG islands in a relatively short time frame. We expect that the characterization of the Xi DNA demethylation event will yield critical insights into mechanisms that control the final stages of reprogramming.

Notably, during reprogramming by SCNT, developmental defects are caused by misregulation of the XCI system, particularly due to ectopic *Xist* expression from the Xa (Inoue et al., 2010). By contrast, our data indicate that, during reprogramming to iPSCs,

Figure 6. Analysis of DNA Methylation on the X Chromosome during Reprogramming

(A) Bisulfite PCR analysis of the promoter regions of the X-linked genes *Atrx* and *Rlim* and of *Nanog* in female MEFs, ESCs, and day 9 SSEA1−/+ reprogramming intermediates. Black circles indicate methylated CpGs, and open circles indicate unmethylated CpGs. The proportion of methylated CpGs is given. For MEFs and SSEA1+/− cells, hemimethylation represents Xi methylation.

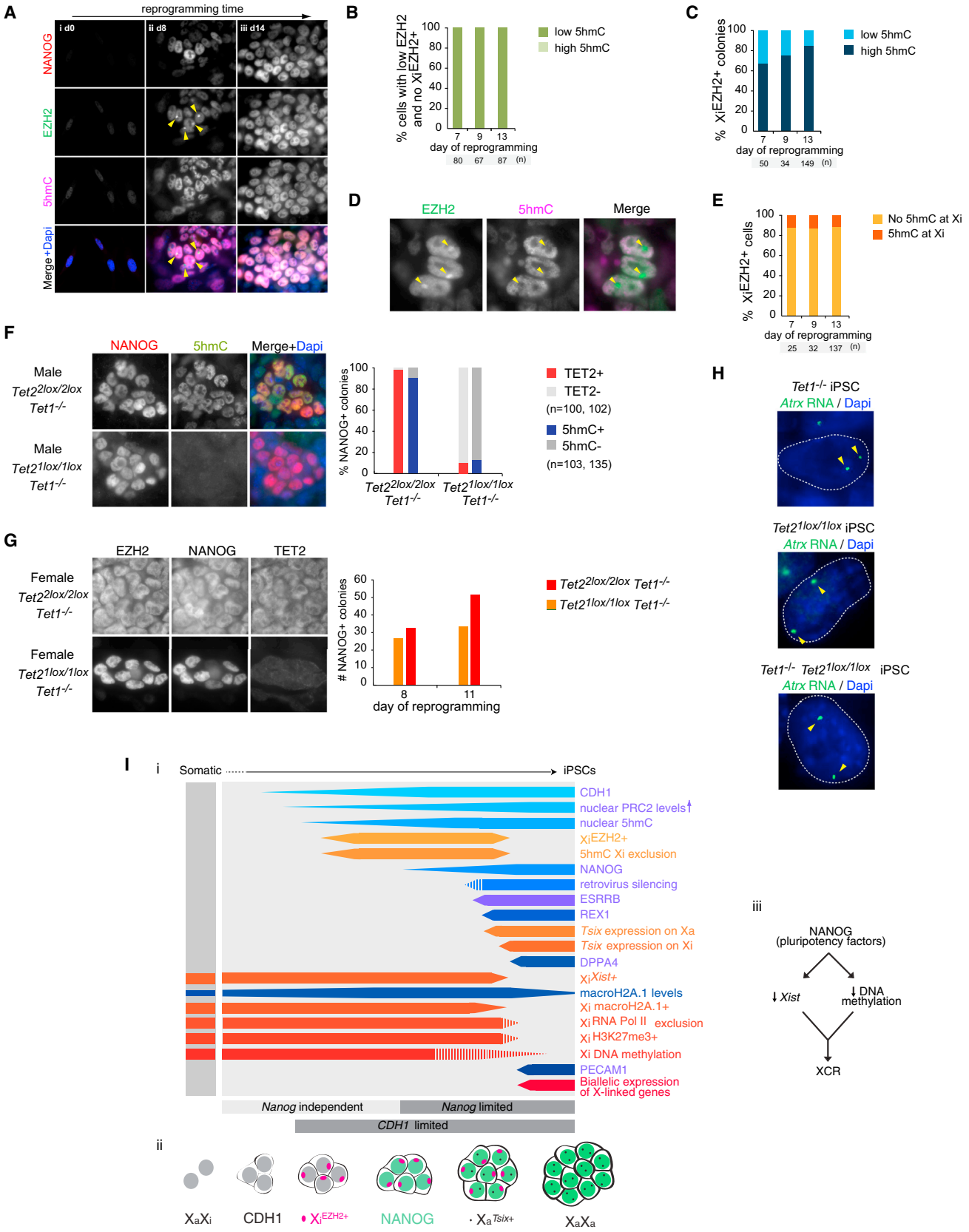
(B) Histograms showing the distribution of methylation levels across CpG islands on the X chromosome and autosomes in indicated cell types based on RRBS data (n, number of CpG islands). The arrow indicates the Xi-specific DNA methylation signature.

(C) Proportion of NANOG+ cells with biallelic *Atrx* expression based on immunoFISH analysis at day 8 of reprogramming with *2lox/2lox Xist* MEFs in which *Xist* was deleted by activation of the dox-inducible Cre-recombinase, and siControl, si*Dnmt1*, and si*Dnmt1* plus 5AzadC, respectively, the latter were added at day 5. All NANOG+ cells present in the culture were counted (n).

(D) Similar to (C), except that *Xist* deletion was performed only in half of the reprogramming culture, and siControl or si*Dnmt1*+5AzadC were applied on day 5 and day 8. At day 12, all NANOG+ cells present in the culture were assessed for biallelic *Atrx* expression.

(E) Summary of the role of *Xist* RNA and DNA methylation in the control of gene silencing on the Xi during reprogramming.

See also Figure S5.



(legend on next page)

the reactivation of the Xi (Figure 4) or ectopic XCI on the Xa (Figures 5B and S7L) does not seem to act as barriers to reprogramming, pointing to mechanistic differences between transcription factor- and oocyte-induced somatic cell reprogramming. Furthermore, in contrast to our findings in iPSC reprogramming, the activation of X-linked genes during XCR in preimplantation development occurs in the presence of *Xist* (Williams et al., 2011).

Importantly, our study defines many sequential reprogramming steps, extending previous reports based on gene expression studies that identified a limited number of reprogramming stages (Parchem et al., 2014; O'Malley et al., 2013; Buganim et al., 2012; Polo et al., 2012). We propose that the global epigenetic state of cells as they reprogram to iPSCs, and that of the Xi, is less variable than transcriptional states. However, our data do not exclude stochastic gene expression differences in cells with the same epigenetic state. One advantage of our analyses is that the stage of any cell in a reprogramming culture can be easily assessed, taking into account criteria such as colony growth and positional information of cells, as well as protein levels and subcellular localization. Notably, although most of our analyses focused on the female-specific XCR process, our work led to the identification of many reprogramming stages that are also applicable to male reprogramming (Figure 7I, i). For example, the global increase in EZH2 and 5hmC levels that occurs in both female and male cells was uncovered during our analysis of the localization of these marks on the Xi in female cells.

Our study provides an easily applicable platform for assaying the effects of interference with intrinsic and extrinsic factors on the stages of reprogramming and on the transitions between them. Additionally, we anticipate that the analysis of the transcriptome and other epigenetic features such as DNA methylation in the multiple reprogramming intermediates that we have identified will reveal insights into reprogramming. Another task ahead remains the continuous imaging of the transitions between the reprogramming steps identified here to quantitatively model the reprogramming process.

In conclusion, our comprehensive study yields insights into XCR and provides unprecedented details on the epigenetic dynamics of somatic cell reprogramming to induced pluripotency, establishing a valuable foundation exploitable for many applications, including staging of reprogramming cultures, isolation of intermediates, and to uncover mechanistically how cells transition toward pluripotency.

EXPERIMENTAL PROCEDURES

Reprogramming Experiments and Time Courses

Reprogramming was carried out using cells derived from reprogrammable mice or directly infected with retroviruses encoding *Oct4*, *Sox2*, and *Klf4*, as described in detail in the Extended Experimental Procedures. For time course analyses, reprogramming cultures on 22 × 22 mm gelatinized glass coverslips were fixed every other day, usually from day 6 to day 14, before carrying out immunostaining and RNA FISH analyses.

Flow Cytometry

Flow cytometry for SSEA1 and CDH1 was done starting from large reprogramming cultures using methods previously reported (Stadtfeld et al., 2008) with modifications described in the Extended Experimental Procedures.

Immunostaining and RNA FISH

Immunostainings and RNA FISH were carried out on 22 × 22 coverslips obtained from reprogramming cultures and as described previously (Maherali et al., 2007). Details are given in the Extended Experimental Procedures.

Bisulfite Analysis

Bisulfite-converted DNA was subjected to RRBS or analyzed by PCR as detailed in Table S1. Details are given in the Extended Experimental Procedures.

Data Analyses

See the Extended Experimental Procedures.

ACCESSION NUMBERS

The GEO accession number for the RRBS data reported in this paper is GSE58109.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2014.11.040.

Figure 7. *Tet1* and *Tet2* and Global 5hmC Are Dispensable for XCR

(A) Representative immunostaining images for different patterns of NANOG (red in merge), EZH2 (green), and 5hmC (magenta) arising at indicated days of reprogramming. Arrowheads indicate Xi^{EZH2+}.

(B) Proportion of cells with low nuclear EZH2 levels and no Xi^{EZH2+} that display either low or high nuclear levels of 5hmC at indicated time points.

(C) Proportion of Xi^{EZH2+} colonies that display either low or high 5hmC at indicated time points.

(D) Representative immunostaining image for EZH2 (green in merge) and 5hmC (magenta) in the Xi^{EZH2+} reprogramming intermediate. 5hmC Xi exclusion is indicated by arrowheads.

(E) Proportion of Xi^{EZH2+} cells that display 5hmC Xi exclusion (Xi^{5hmC-}) at indicated time points.

(F) Representative immunostaining images of male $Tet2^{2lox/2lox}Tet1^{-/-}$ reprogramming cultures infected with Ad5 (top) or AdCre (bottom) adenoviruses, stained for NANOG (red in merge) and 5hmC (green) at day 14 of reprogramming. AdCre induces *Tet2* deletion ($Tet2^{1lox/1lox}Tet1^{-/-}$). The graph gives the proportion of NANOG+ colonies positive for 5hmC and TET2, respectively, at day 14 based on immunostaining. The absence of the TET2 signal in NANOG+ cells confirms effective *Tet2* deletion. Loss of both *Tet1* and *Tet2* leads to loss of the 5hmC immunostaining signal (loss of global 5hmC).

(G) As in (F), except for female $Tet2^{2lox/2lox}Tet1^{-/-}$ and $Tet2^{1lox/1lox}Tet1^{-/-}$ reprogramming cultures, immunostained for EZH2, NANOG, and TET2. The number of NANOG+ colonies at indicated time points in these cultures is given in the graph.

(H) RNA FISH for *Atrx* nascent transcription on female $Tet1^{-/-}$, $Tet2^{1lox/1lox}$ and $Tet2^{1lox/1lox}Tet1^{-/-}$ iPSCs. Arrowheads indicate the biallelic *Atrx* signal.

(I) Stages of XCR and somatic cell reprogramming to induced pluripotency. Our view of the stages leading to XCR and the induction of pluripotency, shown as described in Figures 1I and 2E. Female-specific events are shown in orange/red, and those occurring in both female and male cells are shown in blue. With the exception of retroviral silencing in male reprogramming, all results presented are based on experimental evidence in both female and male reprogramming. See also Figures S6 and S7.

AUTHOR CONTRIBUTIONS

V.P., J.T., and K.P. designed experiments; V.P., J.T., R.K., M.U., A.S.D., D.C., B.P., S.P., R.M., and K.P. performed experiments; V.P., J.T., R.K., M.U., A.S.D., D.C., B.P., G.B., S.P., R.M., A.M., and K.P. analyzed data; R.H. helped with experiments; R.S., T.S., and T.T. generated reagents; and V.P. and K.P. wrote the manuscript with edits from J.T., R.K., B.P., and A.M.

ACKNOWLEDGMENTS

We are grateful to Drs. Nakamura, Xu, Eng, Berk, Silva, Pei, and Besser for providing reagents; Dr. Chronis for help with bioinformatics; Dr. Ch'ng and A. Sahakyan for help with microscopy and western blot analysis; F. Codrea and J. Scholes at the Broad Stem Cell Center FACS Core for help; and Dr. Lowry and members of the Plath lab for advice and critical reading of the manuscript. V.P., S.P., and A.S.D. are supported by CIRM Training Grants TG2-01169 and TB1-01183; J.T. is supported by a fellowship of the UCLA Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research; A.M. is a NYSCF Robertson Investigator and is supported by P01GM099117; K.P. is supported by the UCLA Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research and NIH P01 GM099134 and CIRM (RN1-00564); D.C. is supported by the NIH Ruth L Kirschstein National Research Service Award (GM007185); G.B. is supported by the Whitcome Pre-doctoral Training Program and a UCLA Dissertation Year Fellowship; M.U. is supported by R25GM055052; R.H. is supported by an NIH Training Grant (5T32AI060567-07) and the UCLA Graduate Division Dissertation Year Fellowship; and R.S. is supported by the UCLA Graduate Division Dissertation Year Fellowship.

Received: June 9, 2014
Revised: September 30, 2014
Accepted: November 12, 2014
Published: December 18, 2014

REFERENCES

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

Barakat, T.S., and Gribnau, J. (2010). X chromosome inactivation and embryonic stem cells. Adv. Exp. Med. Biol. *695*, 132–154.

Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, S.L., 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.

Buganim, Y., Faddah, D.A., and Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. Nat. Rev. Genet. *14*, 427–439.

Carter, A.C., Davis-Dusenbery, B.N., Koszka, K., Ichida, J.K., and Eggan, K. (2014). Nanog-independent reprogramming to iPSCs with canonical factors. Stem Cell Reports *2*, 119–126.

Chaumeil, J., Le Baccon, P., Wutz, A., and Heard, E. (2006). A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes Dev. *20*, 2223–2237.

Chen, Q., Gao, S., He, W., Kou, X., Zhao, Y., Wang, H., and Gao, S. (2014). Xist repression shows time-dependent effects on the reprogramming of female somatic cells to induced pluripotent stem cells. Stem Cells *32*, 2642–2656.

Chow, J., and Heard, E. (2009). X inactivation and the complexities of silencing a sex chromosome. Curr. Opin. Cell Biol. *21*, 359–366.

Csankovszki, G., Nagy, A., and Jaenisch, R. (2001). Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. J. Cell Biol. *153*, 773–784.

da Rocha, S.T., Boeva, V., Escamilla-Del-Arenal, M., Ancelin, K., Granier, C., Matias, N.R., Sanulli, S., Chow, J., Schulz, E., Picard, C., et al. (2014). Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. Mol. Cell *53*, 301–316.

Eggan, K., Akutsu, H., Hochedlinger, K., Rideout, W., 3rd, Yanagimachi, R., and Jaenisch, R. (2000). X-Chromosome inactivation in cloned mouse embryos. Science *290*, 1578–1581.

Gendrel, A.-V., Apedaile, A., Coker, H., Termanis, A., Zvetkova, I., Godwin, J., Tang, Y.A., Huntley, D., Montana, G., Taylor, S., et al. (2012). Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. Dev. Cell *23*, 265–279.

Golipour, A., David, L., Liu, Y., Jayakumaran, G., Hirsch, C.L., Trcka, D., and Wrana, J.L. (2012). A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. Cell Stem Cell *11*, 769–782.

Hanna, J., Saha, K., Pando, B., van Zon, J., Lengner, C.J., Creyghton, M.P., van Oudenaarden, A., and Jaenisch, R. (2009). Direct cell reprogramming is a stochastic process amenable to acceleration. Nature *462*, 595–601.

Hu, X., Zhang, L., Mao, S.-Q., Li, Z., Chen, J., Zhang, R.-R., Wu, H.-P., Gao, J., Guo, F., Liu, W., et al. (2014). Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. Cell Stem Cell *14*, 512–522.

Inoue, K., Kohda, T., Sugimoto, M., Sado, T., Ogonuki, N., Matoba, S., Shiura, H., Ikeda, R., Mochida, K., Fujii, T., et al. (2010). Impeding Xist expression from the active X chromosome improves mouse somatic cell nuclear transfer. Science *330*, 496–499.

Lee, J.T., and Bartolomei, M.S. (2013). X-inactivation, imprinting, and long noncoding RNAs in health and disease. Cell *152*, 1308–1323.

Lee, J.T., and Lu, N. (1999). Targeted mutagenesis of Tsix leads to nonrandom X inactivation. Cell *99*, 47–57.

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.

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

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766–770.

Mermoud, J.E., Costanzi, C., Pehrson, J.R., and Brockdorff, N. (1999). Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation. J. Cell Biol. *147*, 1399–1408.

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

O'Malley, J., Skylaki, S., Iwabuchi, K.A., Chantzoura, E., Ruetz, T., Johnsson, A., Tomlinson, S.R., Linnarsson, S., and Kaji, K. (2013). High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. Nature *499*, 88–91.

Parchem, R.J., Ye, J., Judson, R.L., LaRussa, M.F., Krishnakumar, R., Blelloch, A., Oldham, M.C., and Blelloch, R. (2014). Two miRNA clusters reveal alternative paths in late-stage reprogramming. Cell Stem Cell *14*, 617–631.

Pasque, V., Radzisheuskaya, A., Gillich, A., Halley-Stott, R.P., Panamarova, M., Zernicka-Goetz, M., Surani, M.A., and Silva, J.C.R. (2012). Histone variant macroH2A marks embryonic differentiation in vivo and acts as an epigenetic barrier to induced pluripotency. J. Cell Sci. *125*, 6094–6104.

Payer, B., Rosenberg, M., Yamaji, M., Yabuta, Y., Koyanagi-Aoi, M., Hayashi, K., Yamanaka, S., Saitou, M., and Lee, J.T. (2013). Tsix RNA and the germline factor, PRDM14, link X reactivation and stem cell reprogramming. Mol. Cell *52*, 805–818.

Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., de la Cruz, C.C., Otte, A.P., Panning, B., and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. Science *300*, 131–135.

Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J., et al. (2012). A

molecular roadmap of reprogramming somatic cells into iPS cells. Cell *151*, 1617–1632.

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

Schwarz, B.A., Bar-Nur, O., Silva, J.C.R., and Hochedlinger, K. (2014). Nanog is dispensable for the generation of induced pluripotent stem cells. Curr. Biol. *24*, 347–350.

Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T.B., Webster, Z., Peters, A.H.F.M., Jenuwein, T., Otte, A.P., and Brockdorff, N. (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. Dev. Cell *4*, 481–495.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. Cell *138*, 722–737.

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

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.

Williams, L.H., Kalantry, S., Starmer, J., and Magnuson, T. (2011). Transcription precedes loss of Xist coating and depletion of H3K27me3 during X-chromosome reprogramming in the mouse inner cell mass. Development *138*, 2049–2057.

Wu, H., and Zhang, Y. (2014). Reversing DNA methylation: mechanisms, genomics, and biological functions. Cell *156*, 45–68.

Wutz, A., and Jaenisch, R. (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Mol. Cell *5*, 695–705.

EXTENDED EXPERIMENTAL PROCEDURES

Cell Lines

Wild-Type MEFs

Wild-type MEFs were derived from E14.5 embryos each isolated separately to avoid mixing of male and female cells. The sex was determined for every embryo by H3K27me3 immunostaining (to detect $Xi^{H3K27me3+}$ in females, or its absence in males) and additional by genotyping for *Zfy* and *Ube1* (Table S1).

Stemcca MEFs and Endoderm Cells

Female or male MEFs containing a single polycistronic, tetO-inducible cassette carrying the four reprogramming factors *Oct4*, *Sox2*, *Klf4*, and *cMyc* in the *Col1A* locus and the reverse tet-transactivator M2rtTA in the R26 locus, were generated as described previously (referred to as Stemcca MEFs) (Sridharan et al., 2013; Stadtfeld et al., 2010). Endodermal cells were derived from mixed sexed E14.5 endoderm tissues plated on gelatin in DMEM medium containing 10% FBS penicillin/streptomycin, L-glutamine, β-mercaptoethanol, and nonessential amino acids.

Tet-Inducible Xist MEFs

Female MEFs heterozygous for a tet-inducible promoter in place of the endogenous *Xist* promoter on the Xi, also carrying the M2rtTA in the R26 locus (*Xist*: tetO/wt; R26: M2rtTA/wt), were obtained by crossing a male mouse of the genotype *Xist*: tetO; R26: M2rtTA/wt with a female mouse of the genotype *Xist*: tetO/wt; R26: M2rtTA/wt. During gestation, mice were kept on dox, embryos were derived at d14.5 of embryonic development, and upon genotyping, female MEFs were expanded in the absence of dox in the medium until the reprogramming experiment was started. RNA FISH with a strand-specific RNA probe targeting *Xist* demonstrated that virtually all cells in the MEF population showed an *Xist* RNA cloud upon dox addition for 24 hr, whereas less than 2% of the uninduced (-dox) cells displayed an *Xist* RNA cloud (number of cell counted = 101 and 109, respectively). Upon dox-addition, almost all cells with an *Xist* RNA cloud also displayed Xi-like accumulation of H3K27me3 under the *Xist* cloud (data not shown). Importantly, with and without dox, the majority of MEFs expressed the X-linked gene *Atrx* monoallelically as demonstrated by RNA FISH, and, in the +dox sample, the single *Atrx* pinpoint was never overlapping with the *Xist* RNA cloud (data not shown). Together, these data are most consistent with the interpretation that XCI in MEFs is skewed toward the X chromosome carrying the tet-inducible *Xist* promoter, enabling the efficient ectopic expression of *Xist* from the Xi.

Conditional Xist MEFs

To obtain female $Xi^{2loxXist/GFP}$ $Xa^{\Delta Xist\ (1lox)}$ fibroblasts, male mice carrying the conditional (2lox) *Xist* allele and an X-linked GFP transgene on the X (Hadjantonakis et al., 2001) were mated with female mice heterozygous for the *Xist* knockout (1lox) allele (Csankovszki et al., 1999). In addition, mice carrying both the tet-inducible Cre recombinase and M2rtTA in the R26 locus were bred with mice carrying the conditional *Xist* allele to obtain $Xist^{2lox/2lox}$ $R26^{tetOCre/M2rttA}$ fibroblasts (Pullirsch et al., 2010). *Xist* deletion was induced by addition of dox at day 1 or 3 of reprogramming.

Tsix KO MEFs

Female fibroblasts lacking *Tsix* expression on one of the two X chromosomes and with a wild-type Xa or an Xa lacking *Xist* expression were derived from a previously published mouse model in which *Tsix* transcription is prematurely truncated by a polyadenylation site (Sado et al., 2001), in combination with the *Xist* KO allele.

Tet1/Tet2 Knockout MEFs

Tet1 mutant and *Tet2* conditional mutant mice, respectively, (Dawlaty et al., 2011; Moran-Crusio et al., 2011) were obtained from Jax and MEFs were derived at d14.5. For combined deletion, $Tet1^{-/-}$ mice were crossed with $Tet2^{2lox/2lox}$ mice to obtain $Tet1^{-/-}$ $Tet2^{2lox/2lox}$ female embryos, from which MEFs were derived. Female cells were used unless stated otherwise.

shRNA Nanog MEFs

Female *Nanog* shRNA transgenic (NRi-Tg) were derived from mice previously published and genotyped as previously described (Yamaguchi et al., 2009). Nanog shRNA expression was induced by Tamoxifen treatment (5mM).

Nanog Reporter MEFs

For the *Nanog-H2B-cherry* reporter MEFs, the *Nanog* locus was modified in V6.5 ESCs by integration of a *H2B-Cherry-loxP-ires puro-loxP polyA* cassette at the *Nanog* start codon, using homologous recombination (Tan and Elowitz, 2014). Targeting was confirmed by Southern Blotting and proper *H2B-cherry* repression upon differentiation, as detected by microscopy. Chimeric mice were subsequently obtained by blastocyst injection at the UCLA transgenics core facility, and germline transmission was achieved, before MEFs were derived. The *Nanog-GFP-ires-puro* has been previously described (Maherali et al., 2007).

tetO-Nanog and TetO-Ezh2 MEFs

Male *tetO-Nanog* or *TetO-Ezh2* ESCs (V6.5 line) containing a *tetO*-inducible cassette carrying the *Nanog* or *Ezh2* cDNA in the *Col1A* locus and the reverse *tet*-transactivator *M2rtTA* in the R26 locus, were generated following previously described approaches (Beard et al., 2006). Chimeric mice were subsequently obtained by blastocyst injection of the targeted ESCs at the UCLA transgenics core facility, and germline transmission was achieved, before MEFs were derived. The resulting mice were crossed to obtain female MEFs homozygote for the *tetO-Nanog* and the *M2rtTA* alleles, or heterozygous for the *tetO-Ezh2* allele. 2 μg/ml doxycycline was used to induce *Nanog* or *Ezh2* expression.

pre-iPSCs

pre-iPSCs were isolated from reprogramming cultures in ESC medium (i.e., in the presence of FBS and without any KSR) of female MEFs of the genotype $Xi^{2lox Xist/GFP}$ $Xa^{1lox Xist}$, infected with pMX retroviruses encoding *Oct4*, *Sox2*, *Klf4* and *c-Myc*, as previously described (Sridharan et al., 2009). ESC-like colonies were picked at day 20 of reprogramming, clonally expanded, and screened for the absence of NANOG and GFP expression upon expansion. A clonal pre-iPSC line without NANOG and GFP expression, but with Xi^{EZH2+}, was then expanded for further use. *Xist* deletion was induced by infection of adenovirally encoded Cre-recombinase (a kind gift from Dr. Arnold J. Berk and Carol Eng (UCLA)).

Tet1 *Mutant,* $Tet2^{2lox/2lox}$ *Female Pre-iPSCs*

The generation of pre-iPSCs was achieved by infection of female $Tet1^{-/-}$; $Tet2^{2lox/2lox}$ MEFs with pMX retroviruses encoding *Oct4*, *Sox2*, *Klf4* and *c-Myc* in ESC medium (ie. in the presence of FBS and without any KSR). ESC-like colonies were picked, clonally expanded, and screened for the pre-iPSC state by immunostaining for NANOG and EZH2. Pre-iPSCs for further analysis were chosen based on the absence of NANOG expression and the presence of Xi^{EZH2+}.

iPSCs

iPSCs were picked from reprogramming cultures and expanded on male feeders in ESCs culture media with 15% FBS. In case of *Tet2* and *Tet1* knockout iPSCs, in addition to genotyping of reprogramming cultures, deletion was confirmed by immunostaining (for TET1, TET2 and 5hmC).

ESCs

V6.5 male ESCs and F1-2-1 female ESCs derived in the Jaenisch laboratory were used.

Cell Culture Methods and Reprogramming Approaches

ESCs, pre-iPSCs, and iPSCs were cultured on male irradiated MEFs in ESC media: KO DMEM (Invitrogen) containing 15% FBS, leukemia inhibiting factor (LIF), penicillin/streptomycin, L-glutamine, β-mercaptoethanol, and nonessential amino acids. MEFs were cultured in DMEM with the same components except for LIF and with 10% FBS.

Unless stated otherwise, reprogramming was usually induced using retroviral factor delivery. MEFs at passages 1–3 were infected at 50% confluency overnight with pooled viral supernatant generated by transfection of PlatE cells with individual pMX vectors encoding *Oct4*, *Sox2*, and *Klf4* as described previously (Sridharan et al., 2009). Typically, a second round of infection was performed the next day. To determine when retroviral silencing occurs, pMX-*Sox2* and pMX-*Klf4* were replaced by pMX-*FLAG-Sox2*, pMX-*FLAG-KLF4* retroviruses, which we first validated to be as active in reprogramming as the untagged versions. Cells were split 1:5 onto irradiated male feeders and gelatinized 22x22 mm coverslips at day three of reprogramming and ESC medium containing 15% KSR instead of FBS was added at day five (KSR medium), unless indicated otherwise. Together, we performed more than eighty independent reprogramming experiments in media containing KSR, and additional experiments with media supplemented with 15% fetal bovine serum (FBS), instead of KSR. We also performed experiments with additional ectopic *cMyc* expression (using the pMX *cMyc*) as indicated, for instance to derive pre-iPSCs. Due to the dramatically reduced number of partially reprogrammed cells in OSK versus OSKM reprogramming, we mostly focused on the analysis of reprogramming cultures transduced with *Oct4*, *Sox2*, and *Klf4*-encoding retroviruses, but also validated our results using a secondary system (polycistronic Stemcca MEFs). Notably, partially reprogrammed colonies in pMX-*Oct4*, *Sox2*, *Klf4*, *cMyc* –induced reprogramming cultures typically are stalled at the CDH1+/Xi^{EZH2+} state. Typically, in a representative reprogramming experiment, 48 to 192 coverslips were produced per experiment. While we examined thousands of coverslips for the present study, only a few time points are now sufficient to assay stages of reprogramming which should greatly facilitate further studies.

For Stemcca reprogramming, early passage female stemcca MEFs were induced to reprogram by addition of 2 μg/ul doxycycline to the culture media (ESC media up to day 5, at which point FBS was replaced with 15% KSR). Dox was replaced every two days. In case of experiments that involved FACS or some genotyping, no feeders were plated prior to FACS.

For inducible *Xist* reprogramming experiments and *Xist*-deletion reprogramming experiments using a tet-inducible Cre-recombinase, 2 μg/ul doxycycline were added only at day 1 or day 3 of reprogramming.

For *Tet1/2* single and double knockout reprogramming experiments, MEFs of the respective genotype were induced to reprogram using pMX retroviral vectors encoding *Oct4*, *Sox2*, and *Klf4* as described above. Deletion of *Tet2* in $Tet2^{2lox/2lox}$ MEFs or in $Tet1^{-/-}$ $Tet2^{2lox/2lox}$ MEFs was achieved by infecting cells for 1 hr with Cre- recombinase encoding- or control adenoviruses, and this within 6 hr of the end of infection with retroviruses encoding *Oct4*, *Sox2*, and *Klf4*. Adenovirus infections were done with Polybrene (4ug/ml) and in serum-free media. Successful deletion was assessed at day 4-5 of reprogramming by genotyping, and confirmed by immunostaining in reprogramming cultures or in isolated iPSCs lines.

General note regarding our reprogramming experiments: the day of reprogramming at which specific combinations of markers were first detected differed slightly across all experiments, which can be attributed to the particular fibroblast population used, passage number, and viral titer. However, trends were highly similar and reproducible, and we present results of representative experiments. Therefore, for each reprogramming time course, we are indicating the day of count and the number of cells/colonies assessed, but we would like to stress that the order of steps is highly reproducible, but due to the reasons given above, the exact time of a step can vary slightly from experiment to experiment and with the reprogramming system used.

For chemical treatments, 5AzadC (Sigma) was used at the low concentration of 0.3uM, and tamoxifen at 5mM (sh*Nanog* experiment).

For the *Tet1*$^{-/-}$; *Tet2*$^{2lox/2lox}$ pre-iPSC reprogramming experiment (with and without shRNA targeting *Tet3*), reprogramming was induced on coverslips with gelatin but without feeders by switching to KSR media containing 1% FBS, 3uM CHIR90021 and 1uM PD0325901.

Immunostaining, RNA FISH, and Image Analysis

Immunostainings

Cells were grown on gelatinized glass coverslips and washed with PBS, then fixed in PBS containing 4% paraformaldehyde for 10 min at room temperature, permeabilized with PBS containing 0.5% Triton X-100 for 5-10 min, and washed 2-3 times with PBS containing 0.2% Tween20 (PBST) for 5 min. Cells were then stored at 4°C, or immediately blocked for ~10 min with 10% goat or donkey serum in PBST and 0.1% fish skin gelatin. After incubation with primary antibodies overnight at 4°C in blocking solution, cells were washed 3 times 5 min with PBST and incubated with appropriate fluorophor-labeled secondary antibodies (Invitrogen) in blocking solution for 30 min in the dark, washed 3 times 5 min with PBST, stained with Dapi, and mounted in Aquapolymount (Polysciences Inc.) or ProLong Gold antifade reagent with Dapi (Invitrogen).

For 5hmC co-staining, a normal immunostaining was first performed according to the protocol described above, followed by another 10 min 4% paraformaldehyde 1x PBS fixation step, 2 washes with PBST, 10 min denaturation with 4N HCl, followed by 5 min 100mM TrisHCL pH 8.5 incubation. Cells were then fixed 10 min with 4% paraformaldehyde 1x PBS, washed 5 min in PBST, incubated with anti-5hmC antibody in blocking solution overnight at 4C, before proceeding with washes, secondary antibody incubation, washes and mounting as described above and adapted from a previously described protocol (Wossidlo et al., 2011).

RNA FISH

When combining immunostaining with RNA FISH, immunostaining was performed as described above with the addition of tRNA (Invitrogen), RNaseOUT (Invitrogen) and ribonucleoside-vanadyl complex to the blocking buffer in all steps. After immunostaining, cells were re-fixed for 10 min with PBS containing 4% paraformaldehyde, and serially dehydrated with 70%–100% ethanol. RNA FISH for X-linked genes was performed with double-strand DNA probes labeled with FITC or Cy3 generated with a Bioprime kit (Invitrogen) from a BAC containing *Atrx* (RP23-265D6), *Gpc4* (RP23-467J21), fosmids containing *Mecp2* (G135P69277C4 and G135P67639H10) or *Rlim* (w11-2704K12) and a plasmid containing full length *Pgk1* (pCAB17), following standard procedures. *Xist* and *Tsix* strand-specific RNA probes were made by in vitro transcription of T3-ligated PCR products of cDNA templates using Riboprobe system T3 (Promega) with ChromaTide Alexa 546-14-UTP (Invitrogen C11404) and ChromaTide Alexa 488-5-UTP (Invitrogen C11403) per manufacturer's instructions, as described previously (Maherali et al., 2007).

Image Analysis and Colony Counts

Images were usually taken with an Imager M1 microscope (Zeiss) and acquired with Axio Vision software or with a spinning disk confocal microscope (Intelligent Design) and acquired with Slidebook software. Epifluorescence images were analyzed, merged and quantified using ImageJ. Some of the images were taken as Z-stacks, projected and merged in ImageJ, as indicated in the figure legends. Images generated by confocal microscopy were manually analyzed upon viewing the images with Slidebook software.

For quantification, a colony was defined as positive when four or more closely localized or touching cells with clear nuclear staining for NANOG, ESRRB, REX1, DPPA4, or PECAM1 were detected within a reprogramming culture, unless otherwise stated. A NANOG+ colony was deemed Xi^{EZH2+} when 50% or more of the cells of that colony also showed Xi^{EZH2+}. When scored at the single cell level, the same conclusions were obtained (compare for instance Figure S2A and Figure 2B). A colony was scored NANOG+, ESRRB+, REX1+, or DPPA4+ when > 50% of the cells of a NANOG+, ESRRB+ or REX1+ colony showed ESRRB, REX1 or DPPA4 signal. For PECAM1, a NANOG+ or DPPA4+ colony was scored PECAM1+ when 4 or more cells showed clear PECAM1 expression. A colony was scored FLAG positive when > 50% of cells within a colony displayed FLAG expression. Cells expressing low EZH2 were defined as cells with endogenous EZH2 signal as that seen in MEFs.

Antibodies

Primary antibodies used for immunostaining were: NANOG (eBioscience 14-5761 clone eBioMLC-51, 1/200; and Abcam ab80892, 1/200), EZH2 (BD 612667, 1/200), SUZ12 (Cell Signaling 3737, 1/400), macroH2A1 (rabbit polyclonal, clone 89, generated in the Plath lab, 1/100), CDH1 (Abcam ab11512, 1/200), PECAM1/CD31 (BD 553370, 1/200), REX1 (Stemgent 09-0019, 1/200), DPPA4 (R&D AF3730, 1/200), H3K27me3 (active motif 39155, 1/200; Millipore 07-449, 1/200), ESRRB (R&D PP-H6705-00, 1/200), Ser5P Pol II (Millipore 05-623, clone CTD4H8, 1/1000), 5hmC (active motif 39791, 1/100), TET1 (Millipore 09-872, 1/200), TET2 (rabbit polyclonal, kind gift from Dr. Guo-Liang Xu, Shanghai Institute for Biological Sciences, 1/200), GFP (Nacalai Tesque, 04404-26, 1/500), JARID2 (Abcam ab48137, 1/800), and FLAG (Sigma F1804, F7425, 1/500). Primary antibodies used for Western analysis were: mouse anti-CDH1 (BD Biosciences 610182, 1/500), mouse anti-GAPDH (Fitzgerald 10R-G109a, 1/10000).

Plasmids

pMXs-*FLAG-Klf4* was obtained by cloning the full-length mouse *Klf4* cDNA fused to an amino terminal Flag tag using EcoRI/NotI In Fusion into a pMX backbone (see Table S1 for primers). pMX-*SOX2-FLAG* was obtained by cloning the full-length mouse *Sox2* cDNA fused to a C-terminal triple FLAG/triple His tag into pMX. Both constructs work efficiently in place of the untagged versions in reprogramming experiments. For CDH1 knockdown, we used a published plasmid, specifically pRetroSuper Ecad shRNA11 (Li et al., 2010), kindly provided by Dr. D. Pei, (Guangzhou Institutes of Biomedicine and Health) and pRetroSuper Scr shRNA (Pasque et al., 2011).

Full-length mouse *Cdh1* was cloned into pCR topo XL m*Cdh1* (See Table S1 for primer sequence). *Cdh1* was then subcloned EcoRI/NotI into pMX to give pMX-*Cdh1*.

Western Blot

For Western blot analysis of CDH1, cells were scrapped from plates (tryspin cleaves CDH1), pelleted before addition of combined RIPA buffer (50 mM Tris-HCL pH 8.0, 150 mM NaCl, 2 mM EDTA pH 8.0, 1% NP40, 0.1% sodium deoxycholate, 0.1% SDS, Protease inhibitor), and sonicated for 10 min. NuPAGE Buffer (Life technologies, NP007) and DTT (0.5 mM) were added to the 10 μl of sonicated cell extract. Samples were denatured 5 min at 98°C and incubated 5 min on ice. Entire mixtures were loaded onto a 4%–12% Bis-Tris gel (Life Technologies, NP0323BOX), electrophoresed, and transferred to nitrocellulose membranes. Membranes were blocked in TBS (LI-COR, 927-40003). Primary antibodies were incubated for 16h at 4°C. Primary antibodies used were: mouse anti-CDH1 (BD Biosciences 610182, 1/500), mouse anti-GAPDH (Fitzgerald 10R-G109a, 1/10000). After extensive TBS 0.2% Tween20 (TBS-T) washes, secondary antibody 800CW-conjugated goat anti-mouse (LI-COR, 1/10000) was incubated 30minutes at room temperature. After another round of extensive TBS-T washes, fluorescence was detected using the Odyssey® Infrared imaging system (LI-COR).

Bisulfite Sequencing for Candidate Regions

Feeder-depletion was carried out before gDNA was isolated for bisulfite PCR sequencing analyses. For single gene bisulfite sequencing, bisulfite treatment of DNA was performed with the EpiTect Bisulfite Kit (QIAGEN) according to manufacturer's instructions. Each 25ul PCR reaction included 1.25ul of 10uM gene specific forward primer, 1.25ul of 10uM of gene specific reverse primer, 0.625ul of 10mM dNTPs, 2.5ul of 10X Buffer with MgCl2 (QIAGEN), 1.5ul of 50mM MgCl2, 5ul of 5X Q solution (QIAGEN), 0.5ul HotStart Taq polymerase (QIAGEN), 2ul of extracted gDNA and 10.37ul of pico pure H_2O. Bisulfite specific primers were designed using MethPrimer. The primers for amplification of bisulfite converted DNA used are given in Table S1. Amplified products were run on a agarose gel, extracted, and purified using the Macherey nagal Nucleospin PCR purification kit (740609.50), cloned into the pJET-vector (Fermentas). Clones were picked for colony PCR using pJet forward and reverse primers (Fermentas), and subsequently sequenced. Sequenced products were aligned using Bioedit 7.1.3 with a ClustalW gap penalty of 100 and gap extension penalty of 10, and given CpG sites compared for each of the bisulfite converted sequences.

Reduced Representation Bisulfite Sequencing

Feeder-depletion was carried out before gDNA was isolated for bisulfite PCR sequencing analyses from indicated cell types. gDNA for reduced representation bisulfite sequencing (RRBS) was isolated using the DNeasy QIAGEN Blood and Tissue kit (Cat. 69504). RRBS was performed according to a previously published protocol (Boyle et al., 2012) starting from 15.000 cells for duplicate-sorted SSEA1- and SSEA1+ samples, unsorted female MEFs, and unsorted female ESCs (F1-2-1), and female wild-type iPSCs at passage 2. RRBS libraries were also prepared from > $75x10^4$ cells for male MEFs, male ESCs, female *Tet1*$^{-/-}$ iPSCs, female *Tet1*$^{+/-}$ iPSCs, female *Tet2*$^{2lox/2lox}$ iPSCs and female *Tet2*$^{1lox/1lox}$ iPSCs. We used two independent clones of *Tet1*$^{-/-}$ or *Tet1*$^{+/-}$ female iPSCs, *Tet2*$^{2lox/2lox}$ or *Tet2*$^{1lox/1lox}$ female iPSCs. Raw sequencing reads were mapped to the UCSD mm9 genome assembly with MAQ (Li et al., 2008) and methylation levels at each CpG were calculated as described previously (Gu et al., 2011). Promoters were defined as the regions 1kb upstream and 1kb downstream of transcription start sites annotated in RefSeq. LCPs and HCPs were defined as those promoters having a normalized CpG fraction (Saxonov et al., 2006) less than 0.5 and greater than 0.5, respectively. CpG island definitions were downloaded from UCSC and shores were defined as the 2kb regions flanking either side of CpG islands. The average methylation level for a genomic feature was defined as the mean (weighted by coverage) methylation of the CpGs within that region, with coverage capped at 25x per CpG.

siRNA Transfection

siRNAs targeting *Dnmt1* were obtained from Dharmacon (056796-09 and 056796-11). 20nM of siRNA were transfected into pre-iPSCs or reprogramming cultures with Lipofectamine RNAiMax (Invitrogen) as described by the manufacturer. Confirmation of knockdown efficiency of *Dnmt1* was determined by protein level. Primary antibodies were: DNMT1 (CosmoBio: BAM-70-201-EX) and GAPDH (Fitzgerald: 10R-1178).

Retroviral shRNA Knockdown of *Cdh1*

pRetro.puro vectors encoding shRNA against *Cdh1* (kind gift from Dr. D. Pei, Guangzhou Institutes of Biomedicine and Health, (Li et al., 2010) or *Scr* control (Pasque et al., 2011) were transfected into platE cells and the resulting viruses were used to infect female MEFs. Infected cells were selected with 2 μg/ml puromycin for two days starting 24h after infection. The resulting cells, normalized for cell number, where infected with retroviruses encoding *Oct4*, *Sox2 and Klf4* to initiate reprogramming, and split on male feeders at day 4 of reprogramming.

Lentiviral shRNA Knockdown of *Tet3*

MISSION Lentiviral transduction particles (SHCLNV) targeting *Tet3* (TRCN0000376843) and Mission pLKO.1-puro Non-Target shRNA Control Transduction Particles (SHC016V) were obtained from Sigma. Pre-iPSCs were infected with Polybrene in ESC media

for 8 hr. Cells were switched to 4 μg/ml puromycin starting 24h after infection, and selected for a week. The experiment is summarized in Figure S7E.

Flow Cytometry

For cell sorting, reprogramming cultures were collected by trypsin digestion followed by incubation with antibodies against CDH1 (ab11512) or SSEA-1 (R&D FAB2155P, Clone MC-480, lot LOY0410071) as well as isotype matched controls (BD 340764 and BD 555584) for 20-30 min at 4C in blocking solution (PBS, 0.5% BSA, 2mM EDTA). Dissociated cells were allowed to re-express CDH1 for 2 hr at 37C following trypsin digestion before primary antibody incubation started. Cells were washed twice in blocking solution and incubated for 20 min with secondary antibodies (Alexa488 anti-rat, Invitrogen A-11006), washed in blocking solution, passed through a 40um cell strainer, sorted on a FACSAria (BD Biosciences) at the UCLA FACS Core (with expert help from Felicia Codrea and Jessica Scholes), and replated on feeders in ESC medium with 15% FBS for 24hr then switched to 15% KSR medium. Analysis was performed in FlowJo.

Primers and Genotyping

Primers and genotyping conditions are reported in Table S1.

SUPPLEMENTAL REFERENCES

Beard, C., Hochedlinger, K., Plath, K., Wutz, A., and Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis *44*, 23–28.

Boyle, P., Clement, K., Gu, H., Smith, Z.D., Ziller, M., Fostel, J.L., Holmes, L., Meldrim, J., Kelley, F., Gnirke, A., and Meissner, A. (2012). Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. Genome Biol. *13*, R92.

Csankovszki, G., Panning, B., Bates, B., Pehrson, J.R., and Jaenisch, R. (1999). Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. Nat. Genet. *22*, 323–324.

Dawlaty, M.M., Ganz, K., Powell, B.E., Hu, Y.-C., Markoulaki, S., Cheng, A.W., Gao, Q., Kim, J., Choi, S.-W., Page, D.C., and Jaenisch, R. (2011). Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell *9*, 166–175.

Gu, H., Smith, Z.D., Bock, C., Boyle, P., Gnirke, A., and Meissner, A. (2011). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat. Protoc. *6*, 468–481.

Hadjantonakis, A.K., Cox, L.L., Tam, P.P., and Nagy, A. (2001). An X-linked GFP transgene reveals unexpected paternal X-chromosome activity in trophoblastic giant cells of the mouse placenta. Genesis *29*, 133–140.

Hoki, Y., Kimura, N., Kanbayashi, M., Amakawa, Y., Ohhata, T., Sasaki, H., and Sado, T. (2009). A proximal conserved repeat in the Xist gene is essential as a genomic element for X-inactivation in mouse. Development *136*, 139–146.

Imamura, M., Miura, K., Iwabuchi, K., Ichisaka, T., Nakagawa, M., Lee, J., Kanatsu-Shinohara, M., Shinohara, T., and Yamanaka, S. (2006). Transcriptional repression and DNA hypermethylation of a small set of ES cell marker genes in male germline stem cells. BMC Dev. Biol. *6*, 34.

Kang, L., Wang, J., Zhang, Y., Kou, Z., and Gao, S. (2009). iPS cells can support full-term development of tetraploid blastocyst-complemented embryos. Cell Stem Cell *5*, 135–138.

Li, H., Ruan, J., and Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. *18*, 1851–1858.

Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., Figueroa, M.E., Vasanthakumar, A., Patel, J., Zhao, X., et al. (2011). Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell *20*, 11–24.

Pasque, V., Gillich, A., Garrett, N., and Gurdon, J.B. (2011). Histone variant macroH2A confers resistance to nuclear reprogramming. EMBO J. *30*, 2373–2387.

Pullirsch, D., Härtel, R., Kishimoto, H., Leeb, M., Steiner, G., and Wutz, A. (2010). The Trithorax group protein Ash2l and Saf-A are recruited to the inactive X chromosome at the onset of stable X inactivation. Development *137*, 935–943.

Sado, T., Wang, Z., Sasaki, H., and Li, E. (2001). Regulation of imprinted X-chromosome inactivation in mice by Tsix. Development *128*, 1275–1286.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc. Natl. Acad. Sci. USA *103*, 1412–1417.

Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. Cell *136*, 364–377.

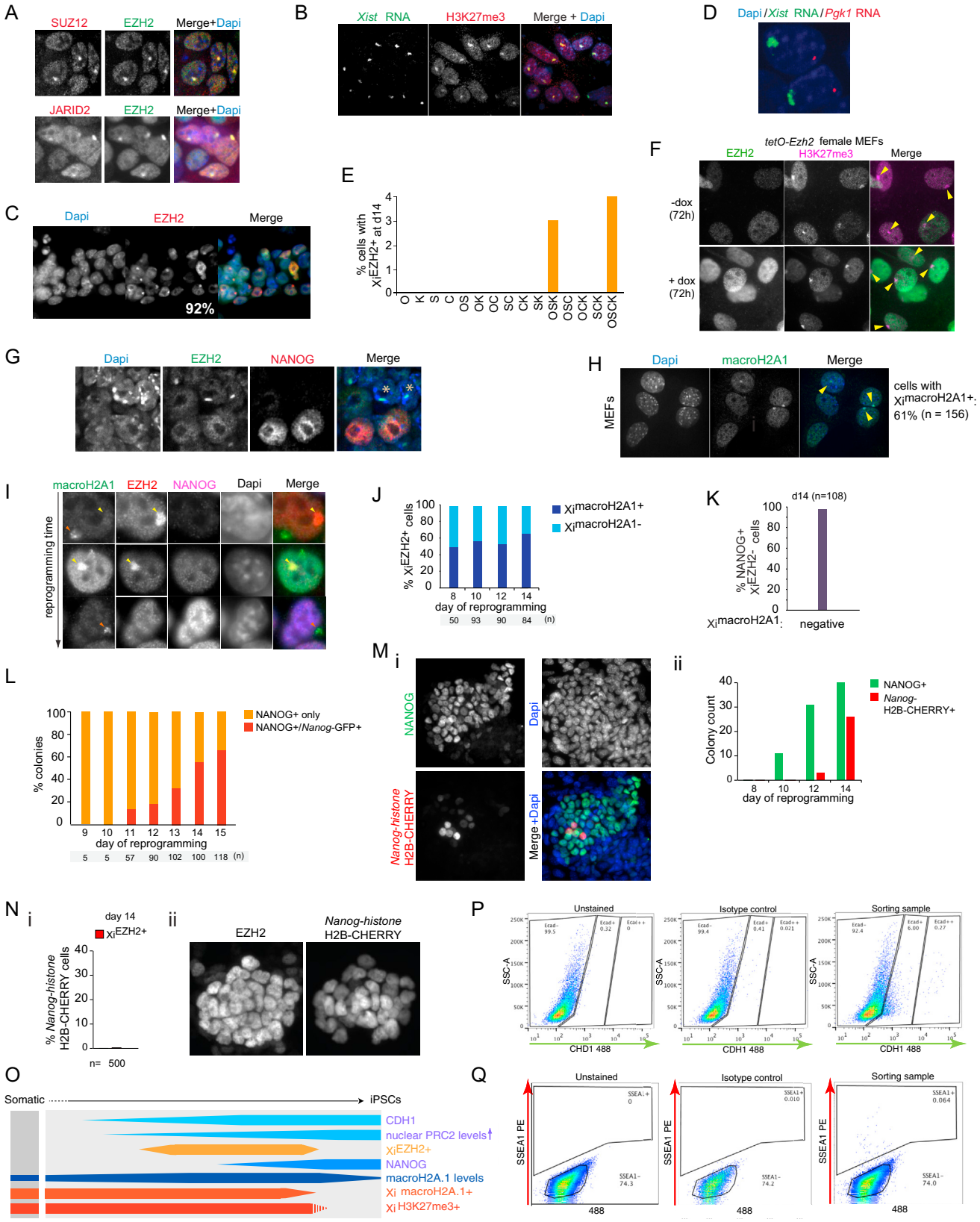
Sridharan, R., Gonzales-Cope, M., Chronis, C., Bonora, G., McKee, R., Huang, C., Patel, S., Lopez, D., Mishra, N., Pellegrini, M., et al. (2013). Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1γ in reprogramming to pluripotency. Nat. Cell Biol. *15*, 872–882.

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

Tan, F.E., and Elowitz, M.B. (2014). Brf1 posttranscriptionally regulates pluripotency and differentiation responses downstream of Erk MAP kinase. Proc. Natl. Acad. Sci. USA *111*, E1740–E1748.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat. Commun. *2*, 241.

Yamaguchi, S., Kurimoto, K., Yabuta, Y., Sasaki, H., Nakatsuji, N., Saitou, M., and Tada, T. (2009). Conditional knockdown of Nanog induces apoptotic cell death in mouse migrating primordial germ cells. Development *136*, 4011–4020.



(legend on next page)

Figure S1. Dynamics of PRC2, macroH2A1, and Pluripotency Reporter Expression during Reprogramming, Related to Figure 1

(A) Immunostaining for EZH2 (green in merge) and SUZ12 (red) at day 7 of reprogramming (top) or EZH2 (green in merge) and JARID2 (red) at day 8 of reprogramming (bottom). A representative field of cells shows the overlap, on the Xi, of SUZ12 and EZH2, or of JARID2 and EZH2. These proteins do not accumulate on the Xi in the starting female MEFs, but are induced to accumulate on the Xi during reprogramming.

(B) Characterization of female pre-iPSCs. Female pre-iPSCs, derived from *Oct4*, *Sox2*, *Klf4*, and *cMyc*-induced reprogramming cultures, carry an *Xist* RNA-coated X chromosome (green) with H3K27me3 enrichment under the *Xist* cloud (red). This pre-iPSC line carries the *Xist* 2lox allele and a *GFP* transgene on the Xi and a *Xist* knockout allele (1lox) on the Xa ($Xi^{2lox Xist/GFP}Xa^{1lox Xist}$), and was also used in Figures S5B–S5E.

(C) Immunostaining for EZH2 (red) in the same pre-iPSC line as in (B). The proportion of cells with Xi^{EZH2+} is given, 200 cells were counted.

(D) RNA FISH for nascent transcripts of the X-linked gene *Pgk1* (red) and *Xist* RNA (green) indicates that the *Xist* RNA-coated X chromosome in the pre-iPSCs described in (B) is inactive.

(E) Reprogramming factor-dependence of Xi^{EZH2+}. MEFs were infected with retroviruses carrying the indicated reprogramming factors and combinations thereof. The proportion of cells with Xi^{EZH2+} was quantified at day 14 post-infection. O = *Oct4*; K = *Klf4*; S = *Sox2*, C = *cMyc*.

(F) Overexpression of EZH2 in female MEFs does not induce accumulation of EZH2 on the Xi (Xi^{EZH2+}). Here, we used female MEFs carrying a tet-inducible *Ezh2* cDNA in the Col1A locus and the M2rtTA allele in the R26 locus. Immunostaining of female *tet-inducible-Ezh2* MEFs for EZH2 (green) and H3K27me3 (magenta) was performed 72h after dox induction. Representative images with and without dox indicate that a dramatic increase in nuclear EZH2 levels in MEFs is not sufficient for the generation of Xi^{EZH2+}. Arrowheads indicate the Xi.

(G) Immunostaining for EZH2 (green) and NANOG (red) at day 10 of reprogramming. In this image, NANOG+ cells lack Xi^{EZH2+}, while surrounding NANOG- cells still have a strong Xi^{EZH2+} (marked by asterisks). At this time point, there are more Xi^{EZH2+} cells in reprogramming cultures than NANOG+ cells. Thus, when NANOG+ colonies first lose Xi^{EZH2+}, they are typically surrounded by NANOG-negative cells with Xi^{EZH2+}. These data support the conclusion that cells with Xi^{EZH2+} progress into the NANOG+ state.

(H) Immunostaining for macroH2A1 in female MEFs demonstrating that macroH2A1 accumulates on the Xi in female MEFs. The proportion of cells with $Xi^{macroH2A1+}$ is indicated. n = number of counted nuclei.

(I) Immunostaining for macroH2A1 (green), EZH2 (red), and NANOG (magenta) during reprogramming. Three different patterns are depicted: (top) low nuclear macroH2A1 / $Xi^{macroH2A1+}$ / Xi^{EZH2+} / NANOG-; (middle) high nuclear macroH2A1 / $Xi^{macroH2A1+}$ / Xi^{EZH2+} / NANOG+; and (bottom) low nuclear macroH2A1 / $Xi^{macroH2A1-}$ / Xi^{EZH2-} / NANOG+), representing different, progressive reprogramming stages. Yellow arrowheads indicate $Xi^{macroH2A1+}$. Orange arrowheads indicate unspecific non-Xi staining of the centrosome using our anti-macroH2A1 antibody.

(J) Proportion of Xi^{EZH2+} cells that display $Xi^{macroH2A1+}$ during reprogramming at the indicated reprogramming time points. The number (n) of cells counted per time point are given.

(K) Proportion of NANOG+/Xi^{EZH2-} cells devoid of $Xi^{macroH2A+}$ at day 14 of reprogramming. The number of cells counted is indicated. We conclude that NANOG+ cells without Xi^{EZH2+} also lack macroH2A1 on the Xi.

(L) Kinetics of *Nanog-GFP* reporter expression relative to the detection of the NANOG protein during reprogramming. *Nanog-GFP* MEFs (heterozygous knockin, NGIP allele (Maherali et al., 2007), thus one *Nanog* allele expresses NANOG and the other the GFP reporter) were reprogrammed, immunostained for NANOG at various time points, and GFP and NANOG expression were quantified. The proportion of NANOG+ (immunostaining)/*Nanog*-GFP-negative and NANOG+/*Nanog*-GFP-(+) colonies at various reprogramming time points is given in the graph. The number of colonies counted is indicated. No NANOG-/*Nanog*-GFP-(+) cells were detected. We found that endogenous NANOG is clearly detected before transgene detection, indicating that the *Nanog-GFP* knock-in reporter is expressed with delayed kinetics compared to the NANOG protein from the endogenous locus during reprogramming of MEFs to iPSCs.

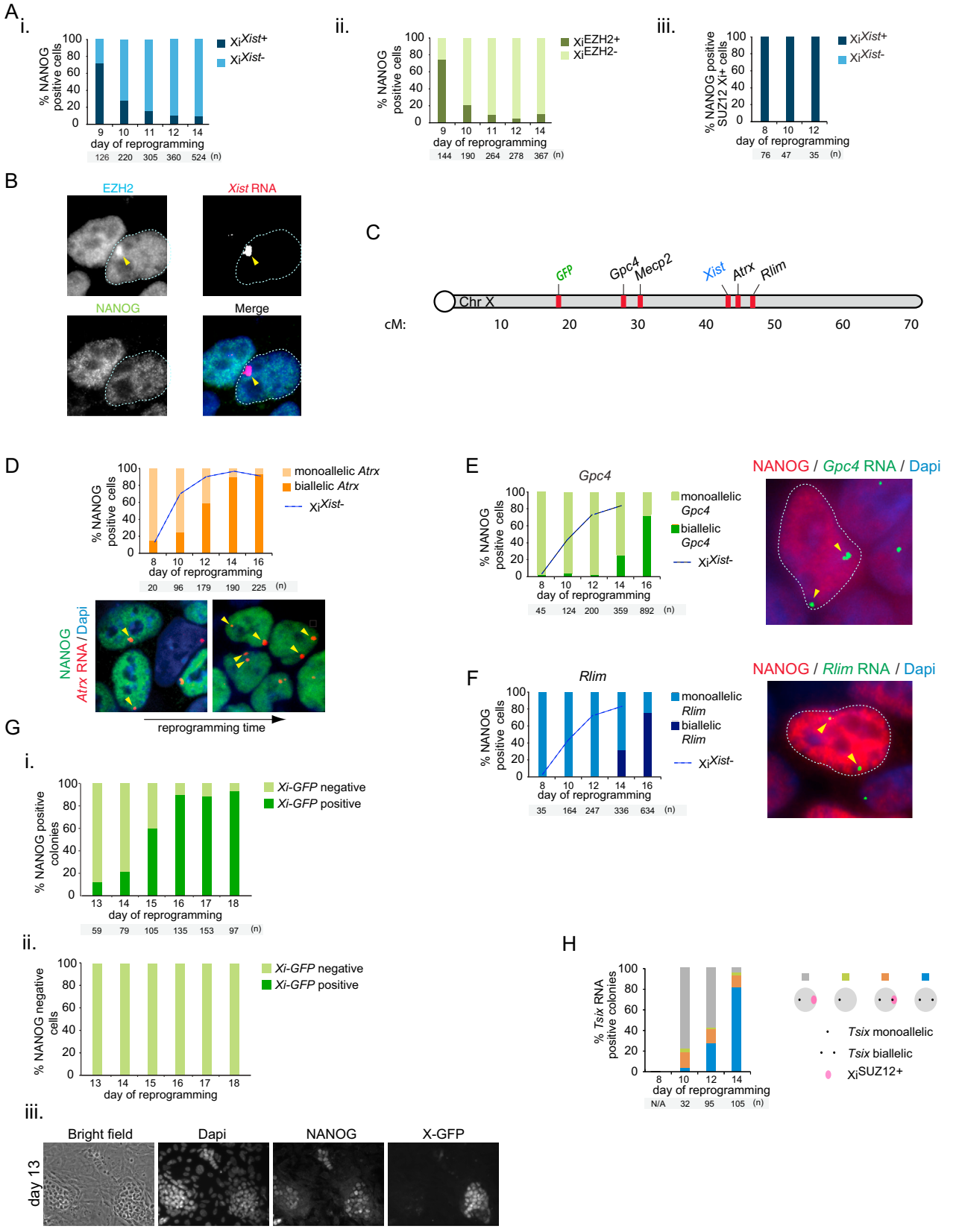
(M) Similar to (L), but with a different *Nanog*-reporter, demonstrating that the *Nanog-histoneH2B-mCherry* knock-in reporter is expressed with delayed kinetics compared to endogenous *Nanog* during reprogramming of MEFs to iPSCs. (i) Representative image of a NANOG+ colony (by immunostaining) displaying delayed reactivation of a *Nanog*-H2B-CHERRY reporter (knockin). (ii) Number of NANOG+ colonies (by immunostaining) and of colonies with *Nanog*-H2B-CHERRY expression. No NANOG-/*Nanog*-H2B-CHERRY-(+) cells were detected. Similar results were obtained with other reporters such as *Oct4-GFP* (data not shown).

(N) *Nanog*-reporter expression is only detectable when Xi^{EZH2+} is already lost from the Xi. (i) Proportion of *Nanog*-H2B-CHERRY-(+) cells with Xi^{EZH2+} at day 14 of reprogramming. (ii) Representative image of *Nanog*-H2B-CHERRY-positive cells immunostained for EZH2.

(O) Summary of reprogramming stages and dynamics of epigenetic changes defined in Figures 1 and S1. Female-specific features are shown in orange/red, those occurring in both female and male reprogramming in blue. The width of the boxes represents the level of the epigenetic mark considered.

(P) Isolation of CDH1+ reprogramming intermediates by FACS as described in Figure 1J. Flow cytometry analysis of day 7 reprogramming cultures for CDH1. Unstained, isotype control and sorting samples are shown.

(Q) Isolation of SSEA1+ reprogramming intermediates by FACS as described in Figure 1J. Flow cytometry analysis of day 7 reprogramming cultures for SSEA1. Unstained, isotype control and sorting samples are shown.



(legend on next page)

Figure S2. Kinetics of *Xist* Repression and XCR within NANOG+ Cells, Related to Figure 2

(A) Quantification of the relationship between NANOG/*Xist* RNA/EZH2 or SUZ12 at the single cell level. (i) The graph depicts the proportion of NANOG+ cells with or without Xi^{Xist+} at the indicated reprogramming time points and the number of cells counted per time point is given. NANOG+ cells are characterized by Xi^{Xist+} when they first appear in reprogramming cultures, but later lose this mark as colonies grow further. (ii) NANOG+ cells are characterized by Xi^{EZH2+} when they first appear in reprogramming cultures, but later lose this mark. The graph depicts the proportion of NANOG+ cells with or without Xi^{EZH2+} at the indicated reprogramming time points and the number of cells counted per time point is given. Compare this count on a cell basis with counts on a colony basis as shown in Figure 1F. Results in (i) and (ii) are from the same reprogramming experiment and indicate that Xi^{Xist+} and Xi^{EZH2+} in NANOG+ cells have with similar kinetics. (iii) SUZ12 is only enriched on the Xi in NANOG+ cells when Xi^{Xist+} is present. Strand-specific *Xist* RNA FISH in combination with immunostaining for NANOG and SUZ12 at the indicated reprogramming time points. The graph shows the proportion of NANOG+/Xi^{SUZ12+} cells with and without the *Xist* RNA cloud. The number of cells counted is indicated.

(B) NANOG+/Xi^{EZH2+} cells typically also have an overlapping Xi^{Xist+} signal. Immunostaining for EZH2 (blue) and NANOG (green) in combination with strand specific RNA FISH for *Xist* RNA (red). In this image, one NANOG+ cell (dashed line) is positive for Xi^{Xist+} and Xi^{EZH2+}, while the other cell has lost both of these Xi marks.

(C) Schematic of the locations of the *X*-linked *GFP*-reporter and the endogenous genes *Gpc4*, *Mecp2*, *Xist*, *Atrx* and *Rlim* on the mouse X chromosome. Units are Centimorgan (cM).

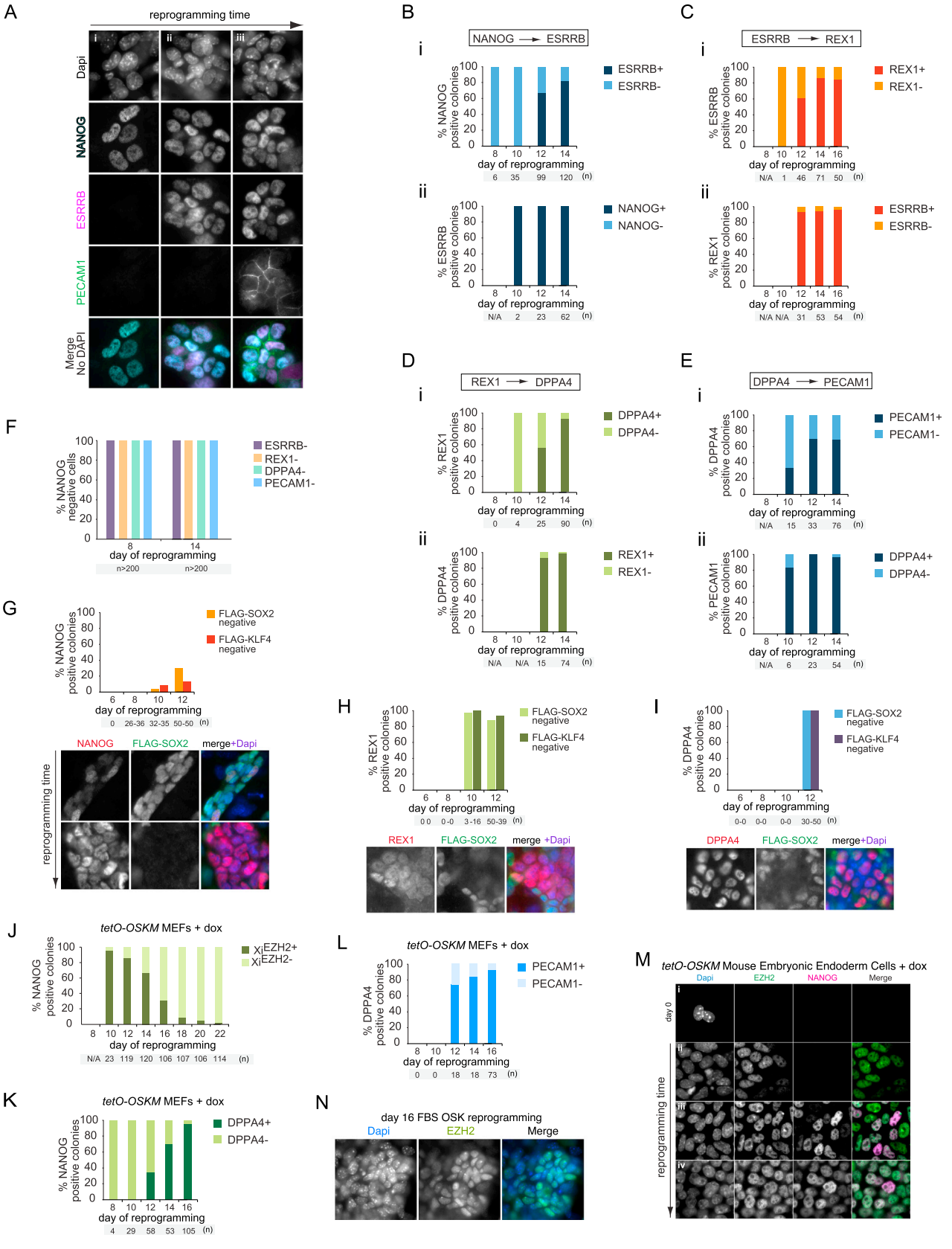
(D) Reactivation kinetics of the X-linked gene *Atrx* in NANOG+ cells. A reprogramming time course was analyzed for NANOG and *Atrx* expression by immunostaining for NANOG and RNA FISH for *Atrx*, detecting the sites of nascent transcription. The graph presents the proportion of NANOG+ cells with the mono- and bi-allelic expression of *Atrx* at the indicated time points. The number of cells counted per time point is given. The dotted line indicates the proportion of Xi^{Xist-} cells from the same reprogramming time course, which is higher than the proportion of NANOG+ with biallelic *Atrx* expression. Thus, as for the X-linked gene *Mecp2*, it appears that *Atrx* reactivates only after *Xist* RNA coating is lost. All NANOG+ cells present in the reprogramming cultures were counted up to day 10 and only a subset thereafter. In the images, monoallelic *Atrx* expression (one arrowhead) and biallelic *Atrx* expression (two arrowheads) patterns are indicated.

(E) As in (D) except that the relationship between NANOG and *Gpc4* nascent transcripts was analyzed. These data, including those for *Xist* RNA, are from a different reprogramming time course than that shown in (D).

(F) As in (E), except that the relationship between NANOG and *Rlim* nascent transcripts was analyzed.

(G) Reactivation of the *GFP* reporter located on the Xi during reprogramming takes place only long after NANOG is detectable by immunostaining. (i) Female *Xi-GFP* MEFs ($Xi^{X\text{-}GFP}$/$Xa^{Xist\ KO}$) were induced to reprogram and the resulting cultures were fixed at the indicated time points. The proportion of NANOG+ colonies without (Xi-GFP-negative) and with (Xi-GFP-positive) GFP expression is given, based on immunostaining for NANOG and detection of GFP fluorescence. (ii) For the same experiments described in (i), the proportion of NANOG- cells either Xi-GFP-negative or Xi-GFP-positive is given. No NANOG-(-)/Xi-GFP-positive colonies were detected, consistent with our conclusion that XCR is specifically induced within NANOG+ cells. (iii) The immunostaining image depicting representative NANOG+ colonies in which Xi-GFP-activation has or has not occurred.

(H) Proportion of *Tsix*-expressing cells with mono- or biallelic *Tsix* expression and with or without Xi^{SUZ12+} at different reprogramming time points.



(legend on next page)

Figure S3. XCR Dynamics in Relation to the Sequential Expression of Pluripotency-Associated Factors and Retroviral Silencing, as well as Conservation of Reprogramming Steps in Different Reprogramming Systems and Starting Cell Types, Related to Figure 3

(A) Detection of step-wise expression of pluripotency factors based on immunostaining. Immunostaining for NANOG (cyan in merge), ESRRB (magenta) and PECAM1 (green) at progressive time points during reprogramming, reflecting representative images of reprogramming stages identified through the quantifications in (B-E). (i) day 10, (ii and iii) day 14 of reprogramming.

(B) Proportion of (i) NANOG+ colonies with and without ESRRB expression and of (ii) ESRRB+ colonies with and without NANOG expression.

(C) as (B) for the relationship between ESRRB and REX1.

(D) as (B) for REX1 and DPPA4.

(E) as (B) for DPPA4 and PECAM1.

(F) Proportion of NANOG- cells also negative for ESRRB, REX1, DPPA4 or PECAM1 at days 8 and 14 of reprogramming. We conclude that NANOG- cells are not positive for any of these pluripotency factors in this assay.

(G) Retroviral silencing during reprogramming. Reprogramming was induced by retroviral expression of *Oct4*, *Klf4*, and *Flag-Sox2* (combination is referred to as FLAG-SOX2), or of *Oct4*, *Sox2*, and *Flag-Klf4* (combination is referred to as FLAG-KLF4). The graph gives the proportion of NANOG+ colonies lacking expression of FLAG-SOX2 or FLAG-KLF4, as detected by NANOG/FLAG co-immunostaining. Representative images reflecting the NANOG+/FLAG-SOX2+ and NANOG+/FLAG-SOX2- states are given. These data indicate that retroviral silencing is induced only after NANOG becomes expressed.

(H) As in (G), but the proportion of REX1+ colonies lacking expression of FLAG-SOX2 or FLAG-KLF4 is shown. These data indicate that retroviral silencing takes place before REX1 becomes expressed. In the images, note that cells surrounding the REX1+ colony are still FLAG-positive.

(I) As in (G), but for the proportion of DPPA4+ colonies lacking FLAG-SOX2 or FLAG-KLF4 expression. These data indicate that retroviral silencing takes place before DPPA4 becomes expressed.

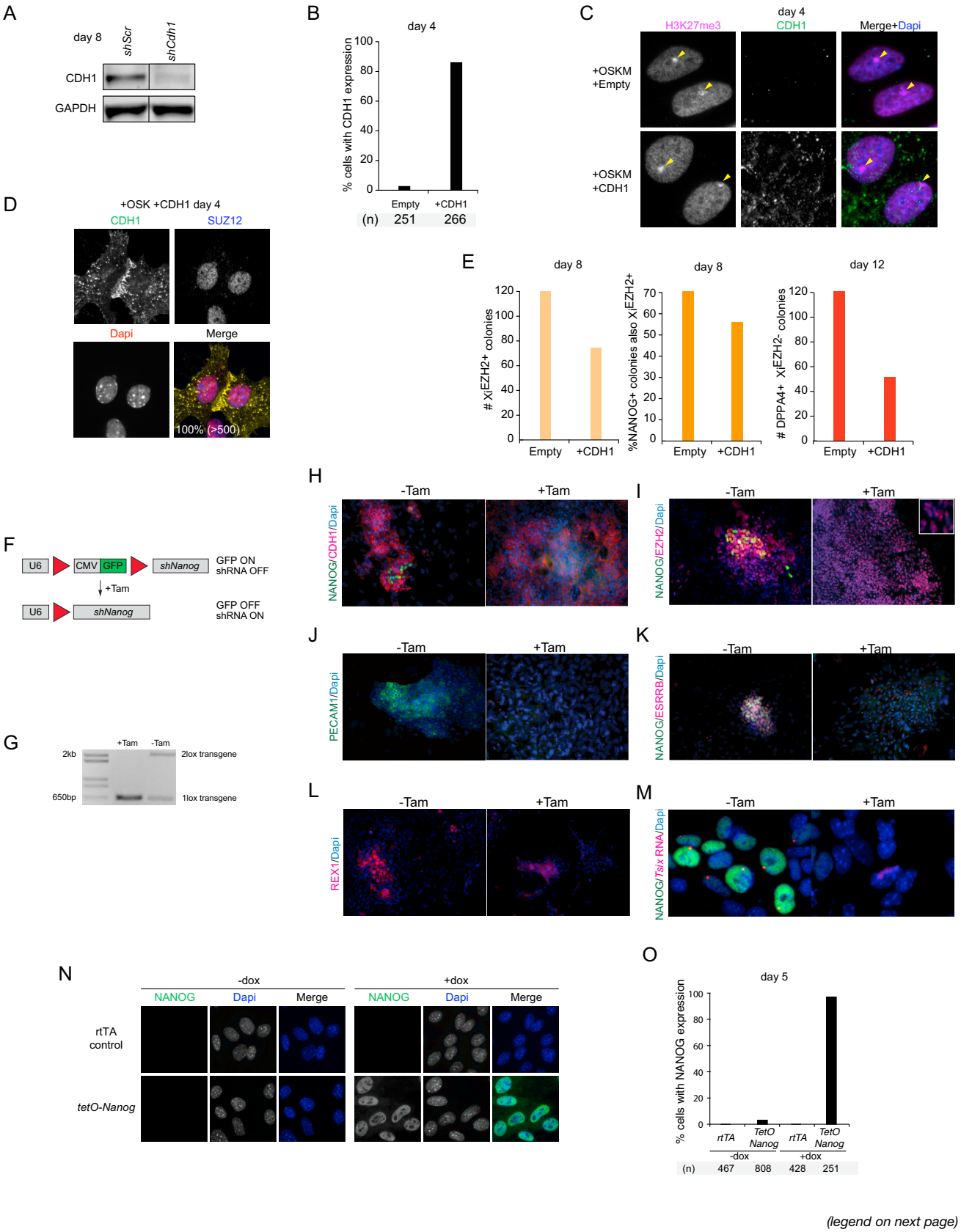
(J) Xi^{EZH2+} dynamics during OSKM reprogramming of MEFs. NANOG+ colonies are also characterized by Xi^{EZH2+} when they first appear during tetO-OSKM (stemcca) MEF reprogramming (see Extended Experimental Procedures). The graph depicts the proportion of NANOG+ colonies with or without Xi^{EZH2+} at the indicated time points and the number of colonies counted per time point is given. All NANOG+ colonies present in the reprogramming cultures were counted up to day 14 and only a subset thereafter. The transition from the NANOG+/ Xi^{EZH2+} to the NANOG+/ Xi^{EZH2-} state is similar to that of retroviral OSK reprogramming shown in Figure 1F.

(K) NANOG to DPPA4 transition during OSKM reprogramming. Proportion of NANOG+ colonies negative or positive for DPPA4 at various time points of tetO-OSKM MEF reprogramming, based on immunostaining. Again, this transition is similar to that of retroviral OSK reprogramming.

(L) DPPA4 to PECAM1 transition during OSKM reprogramming. Proportion of DDPA4+ colonies positive or negative for PECAM1 during tetO-OSKM MEF reprogramming. This transition is similar to that of retroviral OSK reprogramming shown in Figure S3E.

(M) Xi^{EZH2+} dynamics during OSKM reprogramming of endoderm cells. NANOG+ cells are characterized by Xi^{EZH2+} when they first appear in tetO-OSKM reprogramming of mouse embryonic endoderm cells. Images represent different patterns observed progressively in reprogramming cultures. (i) MEFs (low EZH2 levels); (ii) Xi^{EZH2+}/NANOG-; (iii) Xi^{EZH2+}/NANOG+; (iv) Xi^{EZH2-}/NANOG+.

(N) Immunostaining for EZH2 at day 16 of female MEF OSK reprogramming in FBS media (mouse ESC media), indicating that the Xi^{EZH2+} is also established in FBS-reprogramming cultures.



(legend on next page)

Figure S4. Modulation of Reprogramming Stages, Including XCR by CDH1 and NANOG, Related to Figure 4

(A) Reprogramming experiment with *Cdh1* depletion by shRNAs. Western blot analysis of CDH1 and GAPDH at day 8 of reprogramming induced by OSK under *shScr* control or *shCdh1* conditions.

(B) Reprogramming experiment with CDH1 overexpression. Female OSKM (stemcca) MEFs infected with pMX-empty or pMX-CDH1 retroviruses and induced to reprogram with dox. The proportion of cells in the reprogramming culture with a positive CDH1 immunostaining signal was determined at day 4 of reprogramming.

(C) Ectopic CDH1 expression during reprogramming induced by OSKM (stemcca) is not sufficient for *Xist* repression or loss of $Xi^{H3K27me3+}$ at day 4 of reprogramming. The proportion of cells with $Xi^{H3K27me3+}$ was unchanged.

(D) CDH1 overexpression during reprogramming, induced by retroviral OSK expression, is not sufficient for formation of Xi^{SUZ12+} at day 4 of reprogramming. The same result was obtained for OSKM (stemcca) reprogramming.

(E) Overexpression of CDH1 using pMX-*Cdh1*, but not control pMX vector, slightly delayed the appearance of colonies at several epigenetic stages of reprogramming toward XCR. We conclude that ectopic expression of CDH1 is not sufficient to enhance the formation of Xi^{EZH2+} colonies, the transition to the NANOG+/ Xi^{EZH2+} state or to the DPPA4+/ Xi^{EZH2-} state.

(F) Testing the requirement of *Nanog* for the transition through the reprogramming stages. Female MEFs carrying a shRNA construct targeting *Nanog* and a Cre-ER fusion were obtained from mice described previously (Yamaguchi et al., 2009). The drawing displays features of the pSico-based *shNanog* transgene. Addition of Tamoxifen (Tam) to the cell culture media allows for Cre-mediated excision of the GFP cassette, which activates expression of the shRNA, and thereby leads to *Nanog* depletion.

(G) MEFs from (F) were induced to reprogram retrovirally with *Oct4*, *Sox2*, and *Klf4*, split at day three post-infection, and treated with 5mM Tam (+Tam) or vehicle control (-Tam) for 24 hr. Genomic PCR demonstrates the excision of the loxP-flanked GFP cassette after Tam treatment. Some excision (presence of the 1lox band) had already occurred in the absence of Tam, likely due to leakiness of Cre-ER activity.

(H) -Tam and +Tam reprogramming cultures from (G) were immunostained for NANOG (green) and CDH1 (red) on day 14 of reprogramming. As expected, +Tam cultures largely lacked NANOG+ colonies, while a large number of NANOG+ colonies was found in -Tam reprogramming cultures. A few NANOG+ colonies were found in +Tam conditions, but these always expressed the GFP transgene, indicating that the hairpin was not activated in these colonies. Regardless, CDH1 expression was very similar between -Tam and +Tam cultures, in agreement with the notion that CDH1 expression clearly occurs before NANOG induction in reprogramming.

(I) As in (H), except stained for NANOG (green) and EZH2 (red) within the same reprogramming experiment. The inset gives a higher magnification view to visualize Xi^{EZH2+}. Xi^{EZH2+} was found in -Tam and +Tam reprogramming cultures to a similar extent, in agreement with our finding that Xi^{EZH2+} occurs earlier than NANOG expression during reprogramming. Thus, *Nanog* is not required for efficient transition toward the Xi^{EZH2+} state.

(J) As in (H), except that the same reprogramming experiment was stained for PECAM1 (green). No PECAM1 expression was found in the +Tam culture, except in those rare colonies that were GFP+. By contrast, -Tam cultures were characterized with many positive colonies.

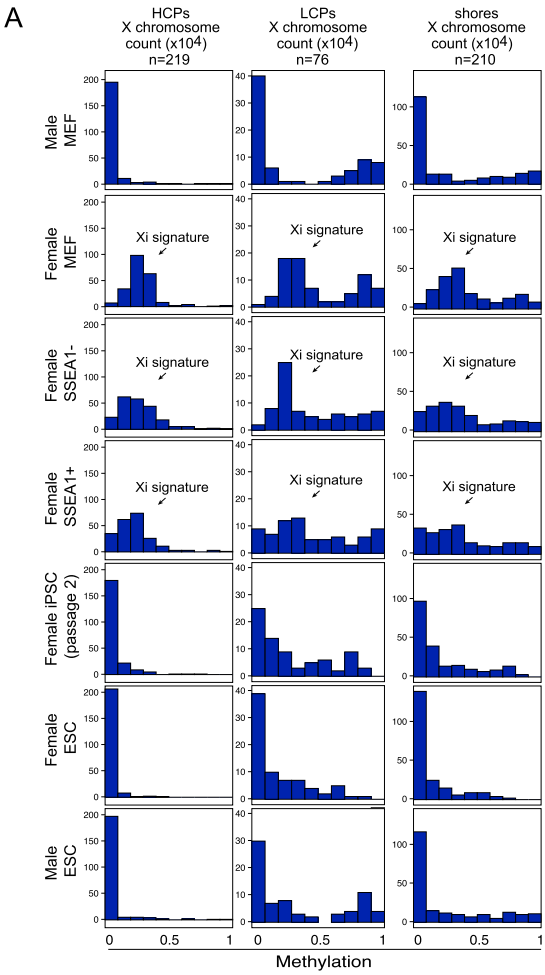
(K) As in (H), except that the same reprogramming experiment was stained for NANOG (green) and ESRRB (red). ESRRB is largely undetectable in +Tam cultures, except in those rare colonies that were GFP+. By contrast, -Tam cultures were characterized with many positive colonies.

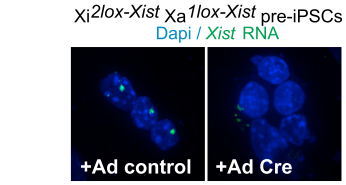
(L) As in (H), except that the same reprogramming experiment was stained for REX1 (red). REX1 is largely undetectable in +Tam cultures, except in those rare colonies that were GFP+. By contrast, -Tam cultures were characterized with many positive colonies.

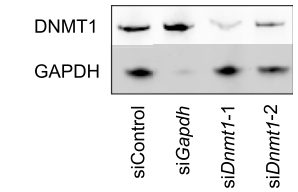
(M) As in (H), except that the same reprogramming experiment was immunostained for NANOG (green) in combination with strand-specific RNA FISH for *Tsix* RNA (red). *Tsix* expression was not induced in the absence of NANOG. This result indicated that *Tsix* is not activated on the Xi or the Xa when *Nanog* levels are reduced. We conclude from these findings that the activation of *Nanog* acts as a rate-limiting step for sequential activation of pluripotency-associated factors and XCR.

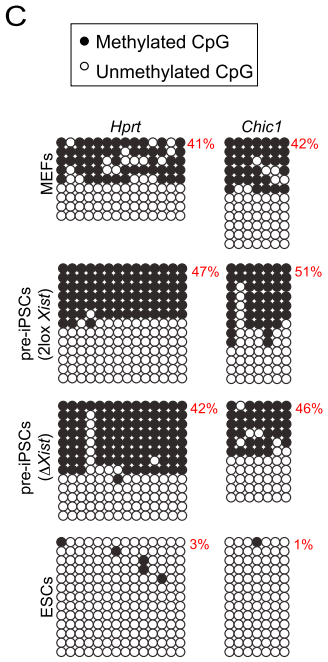
(N) Validation of our NANOG overexpression system. Female MEFs carrying a tet-inducible NANOG overexpression transgene and control MEFs, carrying only the reverse transactivator (rtTA), were induced to reprogram by retroviral infection with OSK and in with 2 μg/ul dox, as indicated. Immunostaining at day 5 of reprogramming indicated NANOG overexpression in nearly all induced cells.

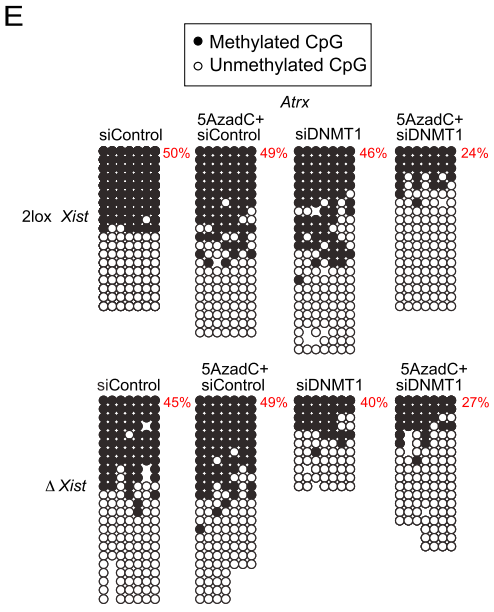
(O) Quantitation of the proportion of cells with NANOG expression from the experiment described in (N).











(legend on next page)

Figure S5. DNA Methylation and *Xist* Act as Barriers to XCR during Reprogramming, Related to Figure 6

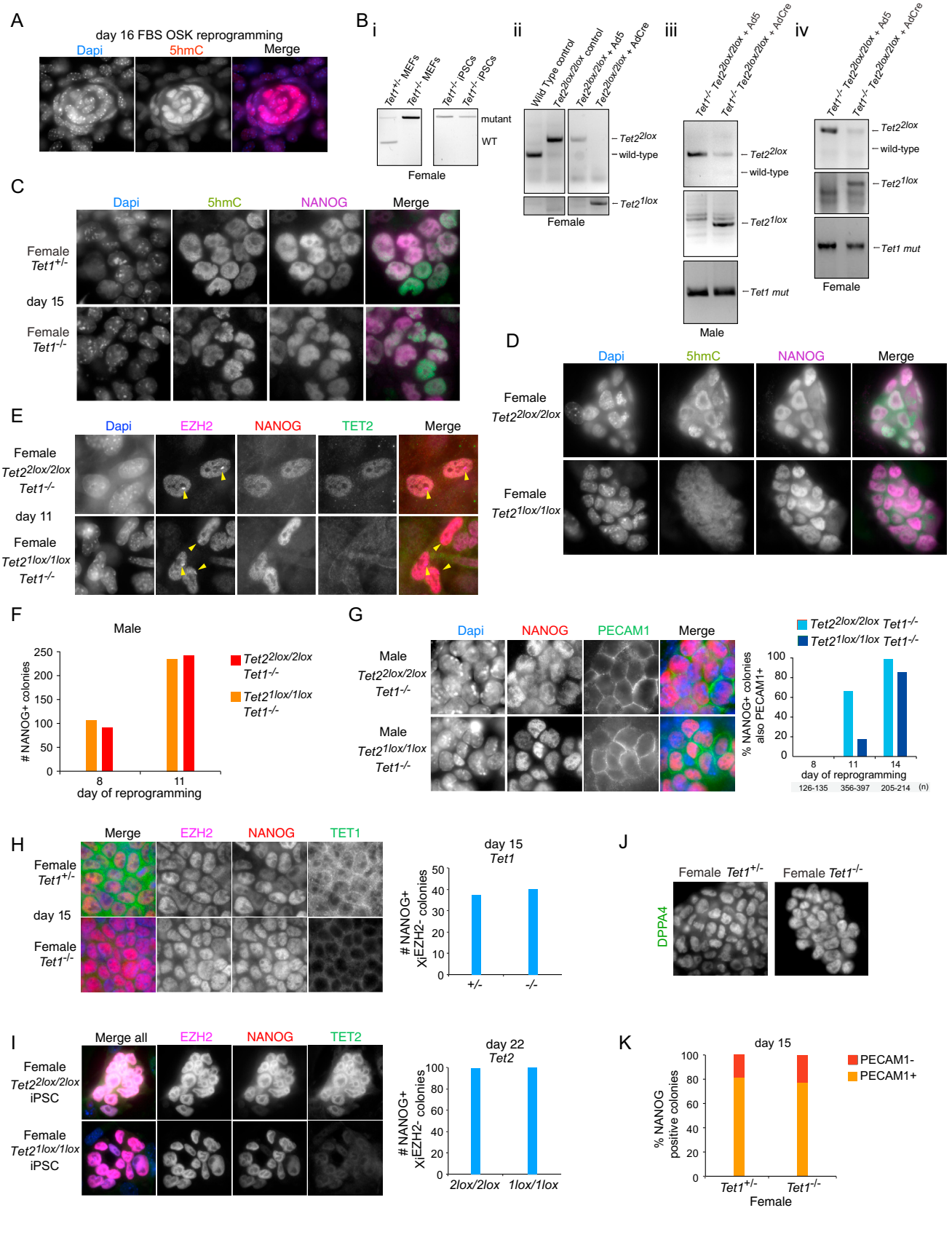
(A) Histograms showing the distribution of methylation levels across High CpG promoters (HCPs), Low CpG Promoters (LCPs), and CpG island shores of the X chromosome for the indicated cell types, based on RRBS data (n, number of HCPs, LCPs, and shores considered). The arrow indicates the Xi-specific DNA methylation signature.

(B) An adenovirus encoding Cre-recombinase (AdCre) was employed to delete the *Xist* gene on the Xi in our $Xi^{2loxXist/GFP}Xa^{1loxXist}$ pre-iPSC line described in Figures S1B–S1D. Deletion of *Xist* was confirmed by RNA FISH for *Xist*. In a control experiment, pre-iPSCs were treated with an empty adenovirus (Ad control) without changes in the normal *Xist* RNA pattern.

(C) Bisulfite PCR analysis of the DNA methylation state of the promoter regions of the two X-linked genes, *Hprt* and *Chic1*, in our female pre-iPSC line before and after *Xist* deletion from the Xi and in control female ESCs and female MEFs. Black circles indicate methylated CpGs, open circles unmethylated CpGs. We conclude that DNA methylation on the Xi is maintained in pre-iPSCs even in the absence of *Xist*.

(D) Validation of efficient knockdown of *Dnmt1* by siRNAs. Western blot for DNMT1 in $Xi^{2loxXist/GFP}Xa^{1loxXist}$ pre-iPSCs treated with two different siRNAs targeting *Dnmt1* or with control siRNAs. GAPDH was used as a loading control as well as a positive control for the knockdown approach (si*Gapdh*). siControl refers to knockdown with a scrambled sequence. Two siRNAs (si*Dnmt1*-1 and si*Dnmt1*-2) that target different sequences of *Dnmt1* depleted the DNMT1 protein efficiently. si*Dnmt1*-1 was used for the experiments in Figure 6.

(E) Confirmation of DNA demethylation efficiency following knockdown of *Dnmt1*, 5AzadC treatment, or combined treatment, in the presence or absence of *Xist*. Bisulfite PCR analysis of the promoter of the X-linked gene *Atrx* in female pre-iPSCs after 3 days of the indicated treatments. Combination of low 5AzadC concentration with *Dnmt1* knockdown is most effective in reducing DNA methylation.



(legend on next page)

Figure S6. *Tet1* and *Tet2* Are Dispensable for the Induction of Pluripotency-Associated Factors and Loss of Xi^{EZH2+}, Related to Figure 7

(A) Immunostaining for 5hmC at day 16 of female MEF reprogramming in FBS culture media, demonstrating that the increase in global 5hmC levels occurs also under these culture conditions.

(B) (i) Genotyping of $Tet1^{+/-}$ and $Tet1^{-/-}$ MEFs and resulting, picked and expanded iPSCs, which were used for bisulfite sequencing and chimera formation. (ii) Genotyping of $Tet2^{2lox/2lox}$ female MEF reprogramming cultures infected with control Ad5 or AdCre viruses (maintaining the $Tet2^{2lox/2lox}$ genotype or establishing the $Tet2^{1lox/1lox}$ genotype). DNA for genotyping was obtained at day 4 of reprogramming before addition of feeders. (iii-iv) Genotyping of (iii) male or (iv) female $Tet1^{-/-}Tet2^{2lox/2lox}$ MEF reprogramming cultures infected with control Ad5 or AdCre viruses at day 1 of reprogramming and collected for genotyping at day 4.

(C) Immunostaining of female $Tet1^{+/-}$ and $Tet1^{-/-}$ day 15 reprogramming cultures for NANOG (magenta in merge) and 5hmC (green). We conclude that global upregulation of 5hmC within NANOG+ cells still takes place in the absence of catalytically active TET1.

(D) Immunostaining for NANOG (magenta in merge) and 5hmC (green) of female $Tet2^{2lox/2lox}$ and $Tet2^{1lox/1lox}$ early passage iPSCs. The nuclear level of 5hmC are reduced, but not abolished within NANOG+ cells in the absence of TET2.

(E) Immunostaining of $Tet1^{-/-}Tet2^{2lox/2lox}$ and $Tet1^{-/-}Tet2^{1lox/1lox}$ female day 11 reprogramming cultures for EZH2 (magenta in merge), NANOG (red), and TET2 (green). Xi^{EZH2+} still takes place in NANOG+ cells in the absence of *Tet1/2* activity. Arrowheads indicate the Xi.

(F) Male $Tet2^{2lox/2lox}Tet1^{-/-}$ and $Tet2^{1lox/1lox}Tet1^{-/-}$ reprogramming cultures were immunostained for NANOG. The number of NANOG+ colonies at indicated time points in these cultures is given.

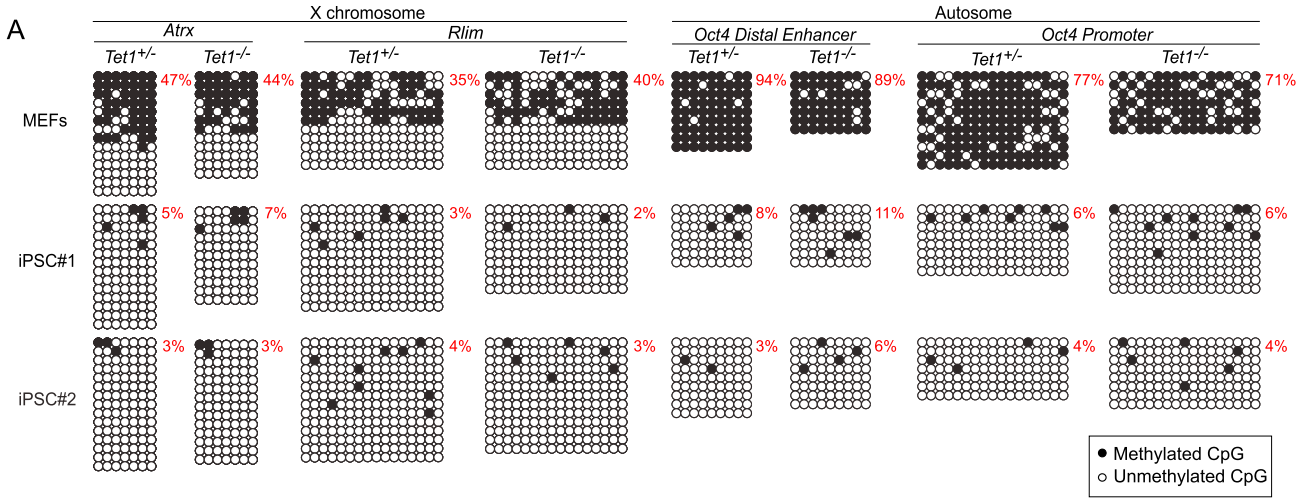
(G) Male $Tet2^{2lox/2lox}Tet1^{-/-}$ and $Tet2^{1lox/1lox}Tet1^{-/-}$ day 14 reprogramming cultures were immunostained for NANOG (red in merge) and PECAM1 (green). The proportion of NANOG+ colonies that are also PECAM1+ at day 8, 11 and 14 is plotted.

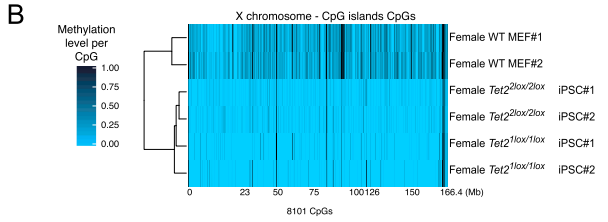
(H) Immunostaining of $Tet1^{+/-}$ and $Tet1^{-/-}$ reprogramming cultures at day 15 for EZH2 (magenta in merge), NANOG (red), and TET1 (green). Graph: efficiency of $Tet1^{+/-}$ and $Tet1^{-/-}$ reprogramming as judged by the number of NANOG+ colonies with no Xi^{EZH2+}.

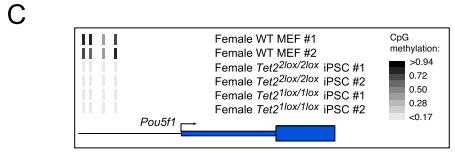
(I) Immunostaining for EZH2 (magenta in merge), NANOG (red), and TET2 (green) on female $Tet2^{2lox/2lox}$ and $Tet2^{1lox/1lox}$ iPSCs obtained from $Tet2^{2lox/2lox}$ female MEFs infected with control Ad5 or AdCre viruses. The graph shows the proportion of NANOG+ colonies without Xi^{EZH2+} in $Tet2^{2lox/2lox}$ and $Tet2^{1lox/1lox}$ iPSC cultures.

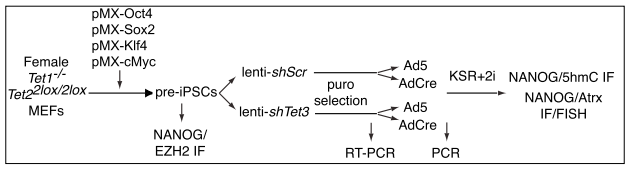
(J) Immunostaining of female $Tet1^{+/-}$ and $Tet1^{-/-}$ iPSCs for DPPA4, indicating that *Tet1* is not required for *Dppa4* induction.

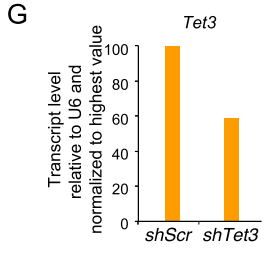
(K) Quantitation of the proportion of NANOG+ colonies with and without PECAM1 expression at day 15 of $Tet1^{+/-}$ and $Tet1^{-/-}$ MEF reprogramming.

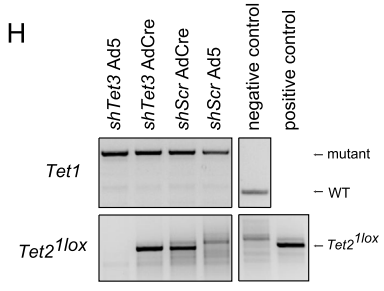


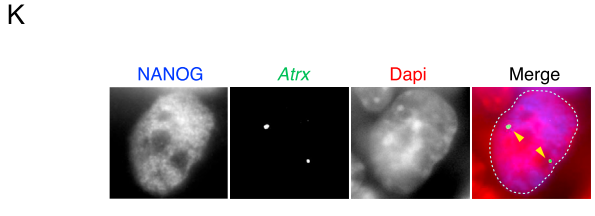


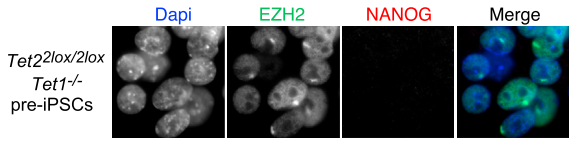


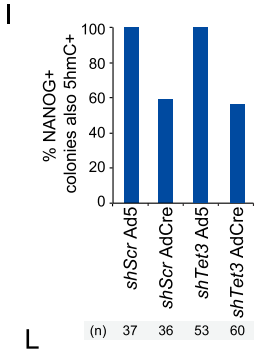


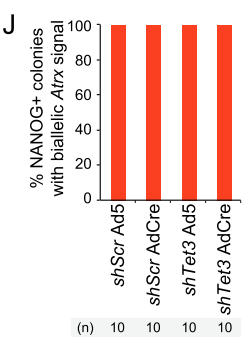


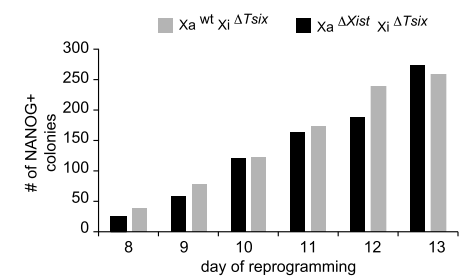












(legend on next page)

Figure S7. *Tet1* and *Tet2* Are Dispensable for DNA Demethylation of the Xi and of *Oct4* Regulatory Sites, Related to Figure 7

(A) Bisulfite PCR analysis of the DNA methylation state of the promoter regions of the X-linked genes *Atrx* and *Rlim*, and of the *Oct4* distal enhancer and promoter in $Tet1^{+/-}$ or $Tet1^{-/-}$ females MEFs and in resulting female iPSC lines (two different lines where analyzed). Black circles indicate methylated CpGs, open circles unmethylated CpGs, and the proportion of methylated CpGs is given. *Atrx* is methylated on the Xi in both $Tet1^{+/-}$ or $Tet1^{-/-}$ MEFs but fully demethylated in both $Tet1^{+/-}$ or $Tet1^{-/-}$ iPSC lines. *Tet1* is dispensable for DNA demethylation of *Oct4* regulatory regions during reprogramming.

(B) Hierarchical clustering of DNA methylation values of CpGs within X-linked CpG islands from two female MEF lines, two female $Tet2^{2lox/2lox}$ iPSC lines, and two female $Tet2^{1lox/1lox}$ iPSC lines. The number of CpGs shown is indicated. X-linked CpG islands are hemimethylated in wild-type MEFs (darker color), and become demethylated in both $Tet2^{2lox/2lox}$ and $Tet2^{1lox/1lox}$ iPSC lines, suggesting that *Tet2* is also dispensable for DNA demethylation of the Xi.

(C) Examples of CpG DNA methylation levels from the RRBS analysis in female MEFs, female $Tet2^{2lox/2lox}$ iPSC lines and female $Tet2^{1lox/1lox}$ iPSC lines for *Pou5f1 (Oct4)* regulatory regions. These data demonstrate DNA demethylation of these regions in the absence of *Tet2*.

(D) Images of chimera obtained from female $Tet1^{-/-}$ or $Tet2^{1lox/1lox}$ iPSCs obtained from reprogramming female MEFs depleted of *Tet1* and *Tet2*, respectively.

(E) Schematic of the experiment testing whether XCR can occur upon reprogramming of pre-iPSCs when, in addition to the absence of TET1/2 activity, *Tet3* is depleted by shRNA-mediated knockdown.

(F) $Tet1^{-/-}$ $Tet2^{2lox/2lox}$ female pre-iPSCs were obtained from MEFs of the same genotype and stained for EZH2 (green) and NANOG (red), demonstrating the lack of NANOG expression and presence of Xi^{EZH2} in these pre-iPSCs.

(G) *Tet3* transcript levels in lenti-*shScr* and lenti-*shTet3*-infected $Tet1^{-/-}$ $Tet2^{2lox/2lox}$ *female* pre-iPSCs selected for with puromycin for a week after lentiviral infection. *Tet3* transcript levels are reduced in these pre-iPSCs.

(H) Genotyping for the *Tet1* mutation and the deleted *Tet2* allele ($Tet2^{1lox}$) on genomic DNA from lenti-*shScr* and lenti-*shTet3* expressing $Tet1^{-/-}$ $Tet2^{2lox/2lox}$ female pre-iPSCs infected with Ad5 or AdCre viruses.

(I) Proportion of NANOG+ colonies displaying 5hmC signal upon reprogramming of lenti-*shScr* and lenti-*shTet3* $Tet1^{-/-}$ $Tet2^{2lox/2lox}$ female pre-iPSCs infected with Ad5 or AdCre viruses. The number of colonies counted is indicated. Note that the reprogramming efficiencies were not quantified here.

(J) Proportion of NANOG+ colonies displaying biallelic *Atrx* signal (XCR) after reprogramming of lenti-*shScr* or lenti-*shTet3* $Tet1^{-/-}$ $Tet2^{2lox/2lox}$ female pre-iPSCs infected with Ad5 or AdCre viruses. All colonies displayed cells with biallelic expression of *Atrx*.

(K) Immunostaining for NANOG (blue) combined with RNA FISH for *Atrx* (green) in reprogramming cultures obtained from $Tet1^{-/-}$ $Tet2^{2lox/2lox}$ *female* pre-iPSCs. The image shown is a representative example of a NANOG+ cell with biallelic *Atrx* expression (arrowheads), indicative of XCR under these conditions.

(L) Number of NANOG+ colonies at various reprogramming time points for reprogramming experiments where the active X chromosome (Xa) carried either wild-type *Xist* or a deletion of *Xist*, combined with deletion of *Tsix* on the Xi.