Stem Cell Reports
Article

OPEN ACCESS

X-Chromosome Dosage Modulates Multiple Molecular and Cellular Properties of Mouse Pluripotent Stem Cells Independently of Global DNA Methylation Levels

Juan Song,[1,*] Adrian Janiszewski,[1] Natalie De Geest,[1] Lotte Vanheer,[1] Irene Talon,[1] Mouna El Bakkali,[1] Taeho Oh,[1] and Vincent Pasque[1,*]

[1]KU Leuven – University of Leuven, Department of Development and Regeneration, Leuven Stem Cell Institute, Leuven Cancer Institute, Herestraat 49, 3000 Leuven, Belgium

*Correspondence: juan.song@kuleuven.be (J.S.), vincent.pasque@kuleuven.be (V.P.)

https://doi.org/10.1016/j.stemcr.2018.12.004

SUMMARY

Reprogramming female mouse somatic cells into induced pluripotent stem cells (iPSCs) leads to X-chromosome reactivation. The extent to which increased X-chromosome dosage (X-dosage) in female iPSCs compared with male iPSCs leads to differences in the properties of iPSCs is still unclear. We show that chromatin accessibility in mouse iPSCs is modulated by X-dosage. Specific sets of transcriptional regulator motifs are enriched in chromatin with increased accessibility in XX or XY iPSCs. The transcriptome, growth and pluripotency exit are also modulated by X-dosage in iPSCs. To understand how increased X-dosage modulates the properties of mouse pluripotent stem cells, we used heterozygous deletions of the X-linked gene *Dusp9*. We show that X-dosage regulates the transcriptome, open chromatin landscape, growth, and pluripotency exit largely independently of global DNA methylation. Our results provide insights into how gene dosage modulates the epigenetic and genetic mechanisms that regulate cell identity.

INTRODUCTION

Pluripotent stem cells (PSCs) are important for modeling development and diseases and for the design of future regenerative medicine approaches (Avior et al., 2016). A key question in the field is which mechanisms underlie the establishment and maintenance of pluripotency. Somatic cells can be reprogrammed into induced PSCs (iPSCs) by transcription factor (TF) overexpression (Takahashi and Yamanaka, 2006), and mouse embryonic stem cells (ESCs) can be derived directly from early embryos. Both cell types have the capacity to self-renew and maintain embryonic lineage differentiation potential in culture (Ying et al., 2008). It is of outstanding interest to understand which epigenetic and genetic mechanisms influence the molecular and functional properties of PSCs.

Several mammalian species including mice and human have adopted X-chromosome inactivation (XCI) as a means to compensate between the genetic imbalance of XX female and XY male cells. XCI is established during early embryogenesis following the expression of the long non-coding RNA *Xist*, and maintained in most somatic cells. Female cells undergo X-chromosome reactivation (XCR) in the mouse inner cell mass (ICM) resulting in two active X chromosomes (XaXa), a state maintained in female ESCs (Mak, 2004; Okamoto et al., 2004). XCR is also induced following somatic cell reprogramming to iPSCs, reviewed in Pasque and Plath (2015). XaXa is a hallmark of mouse naive pluripotency, the latter being characterized by unbiased embryonic lineage differentiation potential. Consequently, XX mouse ESCs have a higher dose of X-linked gene transcripts and hence an increased X-to-autosome gene expression ratio compared with XY cells. Increasing evidence suggests that the presence of XaXa can modulate the molecular and functional properties of mammalian PSCs (Bruck et al., 2013; Choi et al., 2017a, 2017b; Habibi et al., 2013; Hackett et al., 2013; Ooi et al., 2010; Ronen and Benvenisty, 2014; Schulz et al., 2014; Shirane et al., 2016; Yagi et al., 2017; Zvetkova et al., 2005). Work over the past decade showed that XX female ESCs (XX ESCs) exhibit global DNA hypomethylation affecting most genomic features including imprint control regions. More recent work showed that XX female iPSCs (XX iPSCs) also display global hypomethylation (Milagre et al., 2017; Pasque et al., 2018). Differences in global DNA methylation have been attributed to X-chromosome dosage (X-dosage), since female XO cells display male-like DNA methylation levels. Thus, mouse ESCs and iPSCs both show global DNA methylation differences due to X-dosage.

It was also discovered that XX ESCs show increased expression of several pluripotency-associated genes and display delayed pluripotency exit, suggesting that naive pluripotency features are promoted in XX ESCs (Schulz et al., 2014). Differences in transcription have also been attributed to X-dosage, since XO female ESCs (XO ESCs), or *Xist*-induced XCI, are associated with male-like pluripotency-associated gene expression and pluripotency exit. Despite the potential influence of X-dosage on iPSCs, X-dosage has been largely ignored in iPSC reprogramming studies so far, and it remains unclear whether X-dosage influences the molecular features of iPSCs beyond DNA methylation. Therefore, it is important to determine the potential influence of X-dosage on the molecular and cellular properties of iPSCs, which could influence mechanistic studies of reprogramming. A systematic comparison of transcriptional states, open chromatin landscapes, growth, and pluripotency exit in XX female and XY male mouse iPSCs has not yet been performed.

While several advances have been made, the molecular pathways by which XaXa modulate pluripotency remain incompletely understood (Schulz, 2017). At the mechanistic level, XaXa inhibit MAPK (mitogen-activated protein kinase) and GSK3 (glycogen synthase kinase 3) signaling (Bruck et al., 2013; Schulz et al., 2014), and global DNA hypomethylation has been attributed to reduced expression of DNMT3A and DNMT3B (Zvetkova et al., 2005), or DNMT3L (Ooi et al., 2010), or UHRF1 (Choi et al., 2017a; Milagre et al., 2017; Yagi et al., 2017) in XX ESCs/iPSCs. More recently, it was discovered that increased dosage of the X-linked MAPK inhibitor *Dusp9* (dual-specificity phosphatase 9) is in part responsible for inhibiting DNMT3A/B/L and global DNA methylation in XX ESCs (Choi et al., 2017a). The expression level of *Dusp9* is higher in XX ESCs than in XY ESCs, and overexpression of *Dusp9* in XY ESCs induced female-like global DNA hypomethylation and a female-like proteome. Conversely, heterozygous deletion of *Dusp9* in XX ESCs restored male-like global DNA methylation, suggesting that *Dusp9* is responsible for MAPK-mediated DNMT3A/B repression. However, whether *Dusp9* heterozygous deletion in XX ESCs has effects on the transcriptional regulatory network, open chromatin landscape, and pluripotency exit has not yet been explored. In addition, how and which X-linked genes modulate the pluripotency gene network of naive PSCs remains unclear. Furthermore, novel insights may be gained by identification of *cis*-regulatory elements that drive X-dosage-specific PSC states.

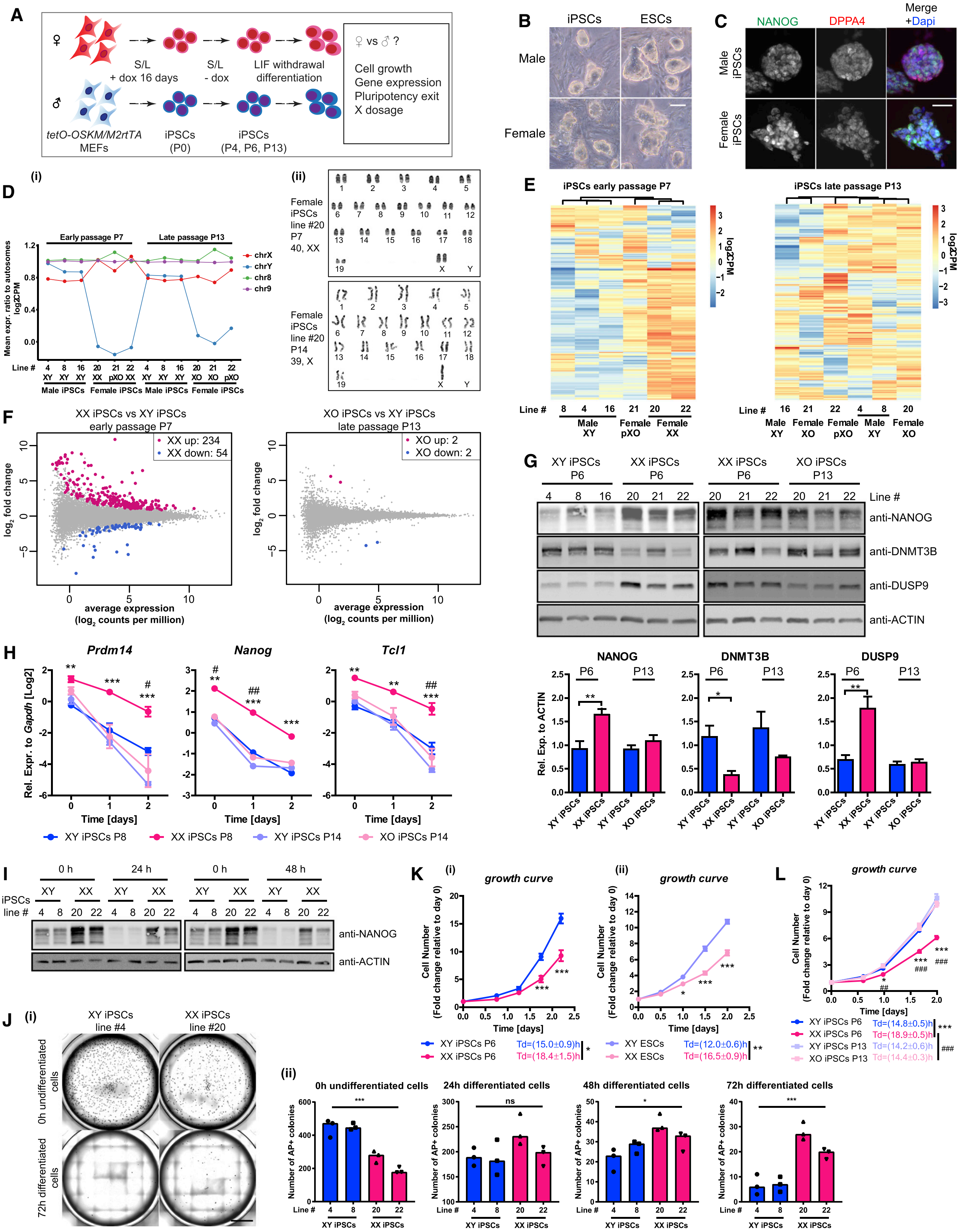
Here, to investigate the influence of X-dosage on iPSCs, we systematically compared multiple molecular and cellular properties of mouse XX and XY iPSCs at different passages. We found that X-dosage is associated with differences in chromatin accessibility, cell growth, the transcriptome, and pluripotency exit in early-passage iPSCs, which are subsequently resolved as a result of X-chromosome loss in female iPSCs upon prolonged culture. We further investigated the regulatory landscape of XX and XY iPSCs and ESCs. We found that thousands of chromatin regions differ in accessibility between XX and XY iPSCs. Motif discovery analysis identified that chromatin more accessible in XX iPSCs is enriched for binding sites of key pluripotency regulators including KLF/ZIC3/NANOG, suggesting stabilization of the naive pluripotency regulatory network via these regulators. By contrast, chromatin sites more accessible in XY iPSCs are enriched for activator protein 1 (AP-1) motifs, downstream effectors of signaling pathways including MAPK. We also show that XY iPSCs grow faster than XX iPSCs, irrespective of culture conditions. We further demonstrate that *Dusp9* heterozygous XX ESCs maintain female-like chromatin accessibility, growth, and delayed exit from pluripotency in the presence of male-like global DNA methylation. Altogether, our study uncovers X-dosage as a previously unrecognized modulator of chromatin accessibility and of growth in PSCs. Our results clarify the effects of X-dosage on the pluripotency transcriptome, revealing the uncoupling of DNA methylation from chromatin accessibility. This provides principles for using gene dosage in designing experiments to understand the epigenetic and genetic mechanisms regulating cell identity.

RESULTS

Differences in Transcriptional Landscapes and Pluripotency Exit Correlate with the Presence of XaXa in iPSCs

To explore the importance of X-dosage on the transcriptome and pluripotency exit of mouse iPSCs, we derived XX and XY iPSC lines. We used isogenic mouse embryonic fibroblasts (MEFs) carrying a tetO inducible transgene encoding the reprogramming factors *Oct4*, *Sox2*, *Klf4*, and *c-Myc* in the *Col1A* locus and the reverse tetracycline transactivator (M2rtTA) in the *Rosa26* locus (Figure 1A and Table S1) (Carey et al., 2010; Pasque et al., 2018). After 16 days of doxycycline (dox) treatment to induce reprogramming, 10 female and 11 male iPSC lines were expanded on feeders in the presence of serum and leukemia inhibitory factor (LIF) (S/L) in the absence of dox (Figure 1A), or adapted to dual ERK/GSK3 inhibition and LIF conditions (2i/L). This scheme allowed us to compare female and male iPSCs without the influence of differences in genetic background, reprogramming system, or derivation method. Both female and male iPSCs could be propagated over multiple passages while maintaining their morphology, indicative of self-renewal, and expressed pluripotency-associated factors NANOG and DPPA4 (Figures 1B, 1C, S1A, and S1B). As expected, the transcriptome of our iPSCs was similar to that of naive ESCs (Figure S1C). Thus, derivation of isogenic female and male iPSCs allowed us to systematically compare the transcriptome and epigenome of these cells.

First, we confirmed that XX iPSCs reactivated the inactive X chromosome, a hallmark of naive pluripotency (De Los Angeles et al., 2015), using RNA sequencing (RNA-seq) analysis (Figure 1D). These results were also in agreement with an independent single-cell level assay using RNA *in situ* hybridization for X-linked gene *Tsix* (Figure S1D; Pasque et al., 2014). XX ESCs are prone to lose



(legend on next page)

one of the two Xas upon extended *in vitro* culture (Choi et al., 2017b; Yagi et al., 2017; Zvetkova et al., 2005), and we recently showed that early-passage XX iPSCs are XaXa and become XO iPSCs upon passage (Pasque et al., 2018). To infer X-chromosome loss in our iPSC lines, we measured the average X-chromosome-to-autosome gene expression ratio using RNA-seq (Figure 1D). We found that early-passage XX iPSCs had increased X-dosage, in agreement with the XaXa state of female iPSCs (Maherali et al., 2007). However, female iPSCs at late passage showed reduced X-dosage, consistent with X-chromosome loss, and we termed these XO iPSCs. In addition, we found that X-chromosome loss in female iPSC lines displayed clonal variability. One early-passage and one late-passage female iPSC line showed partial X-dosage, consistent with partial X-chromosome loss, which we termed partial XO (pXO) iPSCs. In further support of our finding that XX iPSCs undergo X-chromosome loss rather than XCI, we designed a simple qPCR assay to determine the X/autosome genomic DNA ratio by measuring four X-linked genes (*Tfe3*, *Bcor*, *Pdha1*, and *Mid1*, located on either distal region on the X chromosome) and the autosomal gene *Gapdh*. We confirmed that late-passage iPSCs were XO (Figure S1E). Karyotype analyses corroborated these results (Figures 1Dii and S1F). These observations are consistent with XCR during reprogramming followed by X-chromosome loss in female iPSCs.

Using RNA-seq of XX, XY, and XO iPSCs grown in S/L, we asked whether the transcriptome of iPSCs is influenced by X-dosage. Unsupervised clustering of the top 200 most variable autosomal genes, or genes associated with stem cell maintenance, distinguished early-passage XX and XY iPSCs (Figures 1E and S1G). However, XO and XY late-passage iPSCs could not be distinguished, indicating that X-dosage rather than sex modulates the transcriptome of iPSCs. Furthermore, gene expression analysis identified 288 differentially expressed genes (DEGs) between XX and XY iPSCs, but only 4 DEGs between XO and XY iPSCs (Figure 1F, 1.5-fold, false discovery rate [FDR] = 0.05, Table S2). Using qRT-PCR, we found that in S/L, XX iPSC lines consistently expressed higher levels of pluripotency-associated genes *Prdm14*, *Nanog*, and *Tcl1* compared with XY iPSCs (Figure S1H). Western blot analysis showed that XX iPSCs had increased NANOG protein levels compared with XY and XO iPSCs (Figure 1G). These marked differences between XX and XY iPSCs are consistent with patterns observed in mouse ESCs (Choi et al., 2017a; Schulz et al., 2014) (Figures S1I and S1J; Table S2), and in agreement with the notion that iPSCs are molecularly equivalent to ESCs. Despite this, X-dosage has been largely

Figure 1. Two X chromosomes Modulate the Transcriptome, Cellular Growth, and Pluripotency Exit in Mouse iPSCs
(A) Scheme of female and male iPSCs derivation, characterization, and differentiation.
(B) Representative images of female and male iPSCs/ESCs grown on feeders in S/L. Scale bar, 50 μm.
(C) Immunofluorescence analysis for NANOG/DPPA4 in iPSCs grown in S/L. Representative images of all lines examined for NANOG (red), DPPA4 (green), and DAPI (blue, nuclei counterstaining) are shown. Scale bar, 50 μm.
(D) (i) Mean expression ratio to autosomes for sex chromosomes and chromosomes 8 and 9. The dosage of X- and Y-linked genes was used to infer XX, XY, XO, and partial XO (pXO) genotypes. (ii) Representative karyotype images of XX and XO iPSC lines grown in S/L.
(E) Unsupervised hierarchical clustering of top 200 most variable autosomal genes in XY, XX, pXO, and XO iPSCs. Early-passage iPSCs cluster by X-dosage, late-passage iPSCs do not.
(F) DEG analysis, identifying clear differences between XX and XY iPSCs, but not XO and XY iPSCs ($\log_2\text{fold} \geq \log_2 1.5$, FDR ≤ 0.05).
(G) Western blot analysis for NANOG, DNMT3B, and DUSP9 protein in iPSCs grown in S/L. Lower panel: quantification using actin as a loading control. Statistical significance was analyzed using unpaired two-tailed t test. P6, 6 XY versus 6 XX iPSC lines. P13, 3 XY versus 3 XO iPSC lines (n = 1).
(H) qRT-PCR analysis for pluripotency-associated gene expression during LIF withdrawal differentiation of both early- and late-passage iPSCs. Two-way repeated-measures ANOVA with Bonferroni post tests. P8, 3 XX versus 3 XY iPSC lines. P14, 3 XO versus 3 XY iPSC lines (n = 1).
(I) Western blot analysis for NANOG during pluripotency exit. The time after LIF withdrawal is indicated (n = 1).
(J) Two XX and 2 XY iPSC lines were subject to 0, 24, 48, and 72 hr of LIF withdrawal before replating 5,000 cells/well on feeders in 12-well plates. After 5 days in 2i/L, (i) representative images of alkaline phosphatase-positive (AP^+) colonies for replated XX and XY iPSCs are shown (scale bar, 5,000 μm) and (ii) the number of AP^+ colonies is indicated. Results are presented as averages (±SEM) of triplicates for each cell line (n = 1). One-way ANOVA with Tukey's multiple comparisons test.
(K) Growth curves and doubling times of XX and XY iPSCs (i) and ESCs (ii) in S/L condition. Cells were counted at the indicated time points and presented as fold changes relative to day 0. P6, 3 XY versus 3 XX iPSC lines (n = 1, left panel); 1 XY versus 1 XX ESC line (n = 3, right panel). Growth curve: two-way repeated-measures ANOVA with Bonferroni post tests. Doubling time (Td): unpaired two-tailed t test.
(L) As in (K) but for XY, XX, and XO iPSCs (three lines each, n = 1).
Growth curve: two-way repeated-measures ANOVA with Bonferroni post tests. Td: unpaired two-tailed t test.
P6 XX versus P6 XY iPSCs: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; P6 XX versus P13 XO iPSCs: $^{\#}p < 0.05$, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$.
See also Figure S1.

ignored in mechanistic iPSC reprogramming studies so far. Importantly, differences between XX and XY iPSCs cannot be attributed to differences in genetic background, since these differences were found when comparing cells of the same genetic background. Thus, reprogramming to iPSCs results in differences in the transcriptome of iPSCs, some of which can be attributed to differences in X-dosage.

Next, we investigated the extent to which X-dosage affects exit from pluripotency in iPSCs. We subjected XX, XY, and XO iPSCs to LIF withdrawal-mediated differentiation and measured the downregulation of pluripotency-associated genes by qRT-PCR (Figures 1A and 1H). Exit from pluripotency was delayed in XX iPSCs for *Prdm14*, *Nanog*, and *Tcl1*, but not XY and XO iPSCs (Figure 1H). We confirmed these results using an alternative differentiation protocol that mimics epiblast differentiation (Figures S1K and S1L) (Guo et al., 2009; Schulz et al., 2014), and also at the protein level (Figure 1I). These differences had functional consequences on pluripotency exit: replating an equal number of XX or XY cells before and after LIF withdrawal followed by 2i/L culture confirmed delayed pluripotency exit in XX cells (Figure 1J). Thus, XX iPSCs functionally exit pluripotency with delayed kinetics compared with XY and XO iPSCs. Altogether, these findings show that early-passage iPSCs display previously unrecognized X-dosage specific behavior in transcriptome, including pluripotency gene expression, and in pluripotency exit kinetics, consistent with X-dosage differences in ESCs (Schulz et al., 2014).

X-Dosage Modulates Cellular Growth in Mouse iPSCs and ESCs

To determine the effect of X-dosage on cell growth, we counted the number of XX and XY iPSCs over 2 days starting from the same amount of cells. We found that XX iPSC lines grew slower than XY iPSCs, with a doubling time (Td) extended by ~3.4 hr compared with XY iPSCs grown in S/L (Td XX iPSCs = 18.4.±1.5 hr versus Td XY iPSCs = 15.0 ± 0.9 hr) (Figure 1K). XX ESCs also grew slower than XY ESCs (Figure 1K). The delayed growth of XX iPSCs was attributed to the presence of XaXa, since XO iPSCs behaved like XY iPSCs (Figure 1L). The differences in growth of XX and XY iPSCs and ESCs did not depend on culture conditions because XX ESCs and iPSCs still grew slower than XY cells in 2i/LIF (Figure S1M). XX female mouse and human embryos show a delay in post-implantation development that has been attributed to the presence of two X chromosomes in female cells (Burgoyne et al., 1995). Our observations support the idea that the growth delay of XX female mammalian embryos is recapitulated *in vitro* in iPSC and ESC cultures, providing a platform to study this process.

To assess the effect of X-dosage on the cell cycle, we used EdU (5-ethynyl-2′-deoxyuridine) incorporation and propidium iodide staining in combination with flow cytometry to determine the distribution of cells over the different phases of the cell cycle. We found that the majority of both XX and XY iPSCs or ESCs reside in S phase, in line with the literature. The proportion of XX iPSCs and ESCs in S phase was larger than that of XY iPSCs and ESCs, whereas the number of XX iPSCs and ESCs in the G_1 phase was smaller than that of XY cells (Figure S1Ni). To further validate these results, we introduced a fluorescence ubiquitination cell-cycle indicator (FUCCI) into XX and XY ESCs. This system provides for direct fluorescent readout of ESCs in G_1 phase, G_1/S transition, or S/G_2/M phase (Sakaue-Sawano et al., 2008). This analysis confirmed, for XX ESCs, an increase in the proportion of cells in S phase, and a reduced proportion of XX cells in G_1 phase, compared with XY ESCs (Figure S1Nii).

What might be the functional relevance of differences in cell growth between cells with one or two Xas? It has been suggested that the presence of two X chromosomes slows down development to ensure that cells progress through XCI (Schulz et al., 2014). We sought to test, *in vitro*, the hypothesis that reduced X-dosage provides a competitive growth advantage to cells that have undergone XCI. We mixed XX ESCs and GFP-labeled XY ESCs in different ratios and followed the proportion of labeled cells over time. We found that the increased cell growth of XY ESCs can provide a small advantage over an 8-day period (Figure S1O). Collectively, these observations support the idea that cell growth is decreased as a result of increased X-dosage in pluripotent cells *in vitro* and *in vivo*.

Influence of X-Dosage on Chromatin Accessibility Landscape in iPSCs

To assess how X-dosage differentially primes mouse PSCs for rapid exit from pluripotency and to identify additional candidate regulators, we set out to globally define the open chromatin landscape of XX, XY, and XO iPSCs. We employed an assay for transposase-accessible chromatin (omniATAC-seq) to profile genome-wide chromatin accessibility with high resolution (Corces et al., 2017). We generated ATAC-seq datasets from isogenic XX, XY, and XO iPSC lines (and XX/XY ESCs), allowing us to define the open chromatin regions and the enrichment for TF binding motifs associated with open chromatin landscapes (Figure 2 and Table S3). As expected, we observed open chromatin peaks at transcription start site proximal and distal genomic regions, suggesting enrichment in *cis*-regulatory sequences (Figure S2A). We also used the mean read count ratio to autosome to infer the XX, XO, and XY state of the cells (Figure S2B). We then compared autosomal chromatin accessibility globally, and found that X-dosage affects the chromatin accessibility landscape of iPSCs. Broadly, we observed a correlation between the number of Xas and

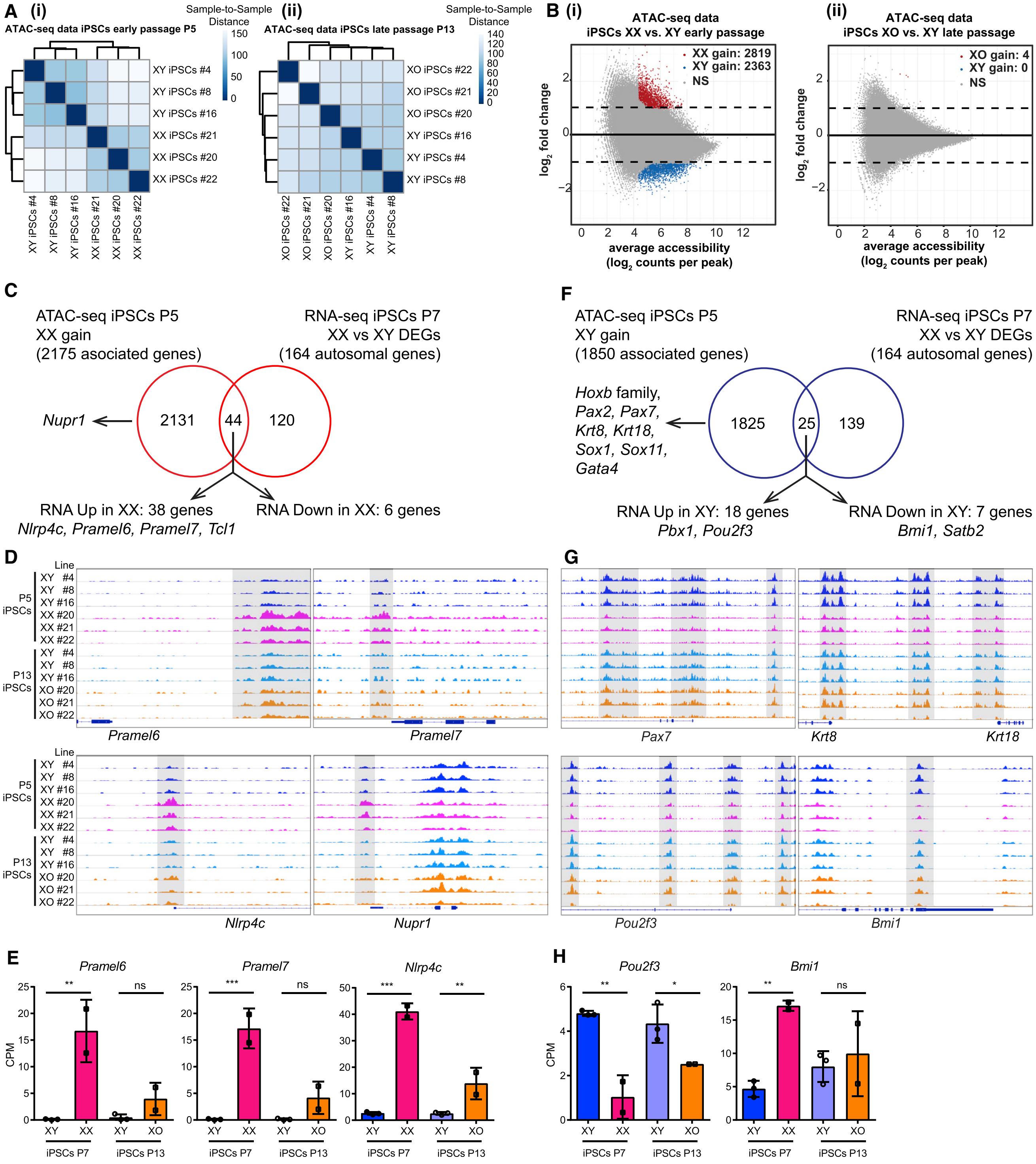


Figure 2. X-Dosage Modulates the Chromatin Regulatory Landscape of Mouse iPSCs

(A) ATAC-seq sample-to-sample distance heatmap showing the Euclidean distances (calculated from the rlog transformed counts, DESeq2) between iPSC samples. (i) XX versus XY iPSCs, chromatin landscapes cluster by X-dosage; (ii) XO versus XY iPSCs, the analysis cannot distinguish XO versus XY landscapes.

(legend continued on next page)

open chromatin landscapes (Figure 2A). We then assessed differential accessibility between XX and XY iPSCs, and between XO and XY iPSCs. We found that most open chromatin regions were shared between XX and XY iPSCs, suggesting that XX and XY iPSCs globally display similar open chromatin landscapes. However, thousands of chromatin regions showed increased accessibility in XX or in XY iPSCs, but not between XO and XY iPSCs (>2-fold, FDR < 0.05, Figure 2B). These results further support the idea that X-dosage influences chromatin accessibility in iPSCs. We identified 2,819 and 2,363 autosomal chromatin regions that are more open in XX iPSCs or in XY iPSCs, respectively (Figure 2B, defined as "XX gain" and "XY gain" regions, Table S4), which represent differences in chromatin accessibility driven by X-dosage. We also found a strong correlation between X-dosage and open chromatin landscapes in isogenic ESCs isolated from another genetic background (Figures S2C and S2D; Table S3). In summary, these results indicate that the chromatin landscape of XX and XY iPSCs is globally similar, but also contains differentially accessible chromatin at thousands of specific genomic regions, due to differences in X-dosage.

We next assessed the correlation between differentially open chromatin and gene expression. Broadly, we observed a weak correlation between changes in chromatin accessibility and changes in gene expression (Figures 2C–2H). Most differentially open chromatin regions did not associate with DEGs (2,131/2,175 genes for XX gain regions, 1,825/1,850 genes for XY gain regions, Figures 2C and 2F). Likewise, most DEGs did not associate with differentially accessible chromatin regions (120/164 and 139/164 DEGs were not associated with changes in chromatin accessibility in XX or in XY iPSCs, respectively, Figures 2C and 2F). Nevertheless, a small fraction of differentially open chromatin regions associated with DEGs (Figures 2C and 2F; Table S3). We identified 44 genes out of 164 autosomal DEGs that associated with chromatin regions more open in XX iPSCs. Most of these genes (86%, 38/44) were transcriptionally upregulated in XX iPSCs cells (Figure 2C). For example, there were chromatin regions more accessible in XX iPSCs that associated with pluripotency-associated genes *Pramel6* and *Pramel7*, both of which were upregulated in XX iPSCs but not in XO iPSCs (Figures 2C–2E and Table S3). Overexpression of *Pramel6* and *Pramel7* was found to oppose exit from pluripotency (Casanova et al., 2011) and *Pramel7* was shown to mediate ground-state pluripotency (Graf et al., 2017). We also observed increased accessibility in the vicinity of *Nlrp4c*, *Nupr1*, and *Tcl1* in XX iPSCs, but not XO iPSCs. These results indicate that the open chromatin landscape of iPSCs reflects specific cellular states, whereby XX-specific open chromatin could mediate stabilization of pluripotency in XX iPSCs and ESCs.

Chromatin regions more accessible in XY iPSCs associated with multiple genes involved in embryonic development and morphogenesis (several *Hoxb* genes, *Pax2*, *Pax7*, *Krt8*, *Krt18*, *Sox1*, *Sox11*, and *Gata4*, Table S3). Twenty-five genes associated with chromatin regions more open in XY iPSCs, 72% of which (18/25) were upregulated in XY iPSCs (Figure 2F). Examples of upregulated genes include *Pou2f3* and *Pbx1* (Figures 2G and 2H). Within the 25 DEGs associated with XY gain chromatin regions in iPSCs, 7 genes were downregulated in XY iPSCs (*Bmi1*). In summary, these findings indicate that chromatin more open in XY or in XX iPSCs is associated with several lineage specification/differentiation-related and pluripotency-associated genes, respectively.

Motif Analysis Reveals Potential Regulators of X-Dosage-Mediated Cell States

To identify TFs involved in modulating iPSCs as a result of differences in X-dosage, we searched for known TF motifs enriched in chromatin more open in XX or in XY iPSCs. Motif enrichment analysis of chromatin regions more open in XX iPSCs revealed a strong enrichment for the binding motif of TFs such as regulatory factor X (RFX)

(B) Differential chromatin accessibility analysis between XX (or XO) and XY iPSCs. $\log_2$ fold change (XX/XY or XO/XY) in reads per accessible region are plotted against the mean reads per ATAC-seq peak. Thousands of open chromatin regions are more open in XX iPSCs or in XY iPSCs (i), but not between XO and XY iPSCs (ii) ($\log_2$fold ≥ 1, FDR ≤ 0.05). These regions were defined as "XX gain" and "XY gain," respectively.

(C) Venn diagrams showing the overlap between genes nearest to the "XX gain" regions and the DEGs between XX and XY iPSCs (DEGs defined in Figure 1F).

(D) Integrated genome viewer track images of ATAC-seq signal for "XX gain" example regions in all iPSCs samples. Differentially open regions are shaded.

(E) Expression of *Pramel6*, *Pramel7*, and *Nlrp4c* in XX, XY, and XO iPSCs as assessed by RNA-seq. CPM (counts per million) values were plotted for each gene. One-way ANOVA with Sidak's multiple comparisons test: P7 2 XX versus 3 XY iPSC lines, P13 2 XO versus 3 XY iPSC lines; $^{**}p < 0.01$, $^{***}p < 0.001$.

(F) as in (C) for "XY gain" regions.

(G) As in (D) for "XY gain" regions.

(H) as in (E) for *Pou2f3* and *Bmi1*.

See also Figure S2.

(10%), KLF5 (48.14%), ZIC3 (26.32%), and NANOG (26.92%) (Figure 3A). RFX proteins encode TFs expressed in many tissues including brain and testes (Choksi et al., 2014). KLF5 and NANOG have been functionally implicated in ESC self-renewal (reviewed in Martello and Smith, 2014). ZIC3 is a pluripotency-associated factor required to maintain pluripotency (Lim et al., 2010). Interestingly, *Zic3* is located on the X chromosome, raising the possibility that ZIC3 dosage could drive X-linked driven stabilization of pluripotency in XX iPSCs (see below). In summary, several top TF motifs enriched in chromatin with increased accessibility in XX iPSCs belong to pluripotency-associated factors, suggesting that the identified pluripotency-associated TFs participate in stabilizing the pluripotency transcriptional regulatory network of XX iPSCs.

By contrast, pluripotency-associated TF motifs were less represented from the top motifs enriched in chromatin with increased accessibility in XY iPSCs, with a few exceptions (Figures 3B and 3C). Instead, within chromatin more open in XY iPSCs, motif enrichment analysis revealed binding motifs of multiple TFs of the AP-1 family such as JUN/AP-1 (39.27%), FOSL2 (43.76%), and ATF3 (51.08%) (Figures 3B and 3C). JUN/AP-1 is a transcriptional activator complex involved in regulating many processes (Shaulian, 2010) including cell growth and differentiation in response to a variety of stimuli including the MAPK pathway (Karin, 1995; Yang et al., 2012). FOSL2 is a member of the AP-1 complex (Shaulian, 2010). Collectively, these findings reveal that X-dosage modulates chromatin accessibility in iPSCs. As expected, we made similar observations in ESCs (Figures S2C–S2I). In addition, open chromatin regions that are common between XX and XY iPSCs still showed enrichment of pluripotency-related TFs (Figures S2I and S2J). We propose that the differential enrichment of TF binding sites in open chromatin regions modulated by X-dosage provides a molecular link between transcriptional regulators, stabilization of pluripotency in XX PSCs, and rapid exit from pluripotency in XY PSCs.

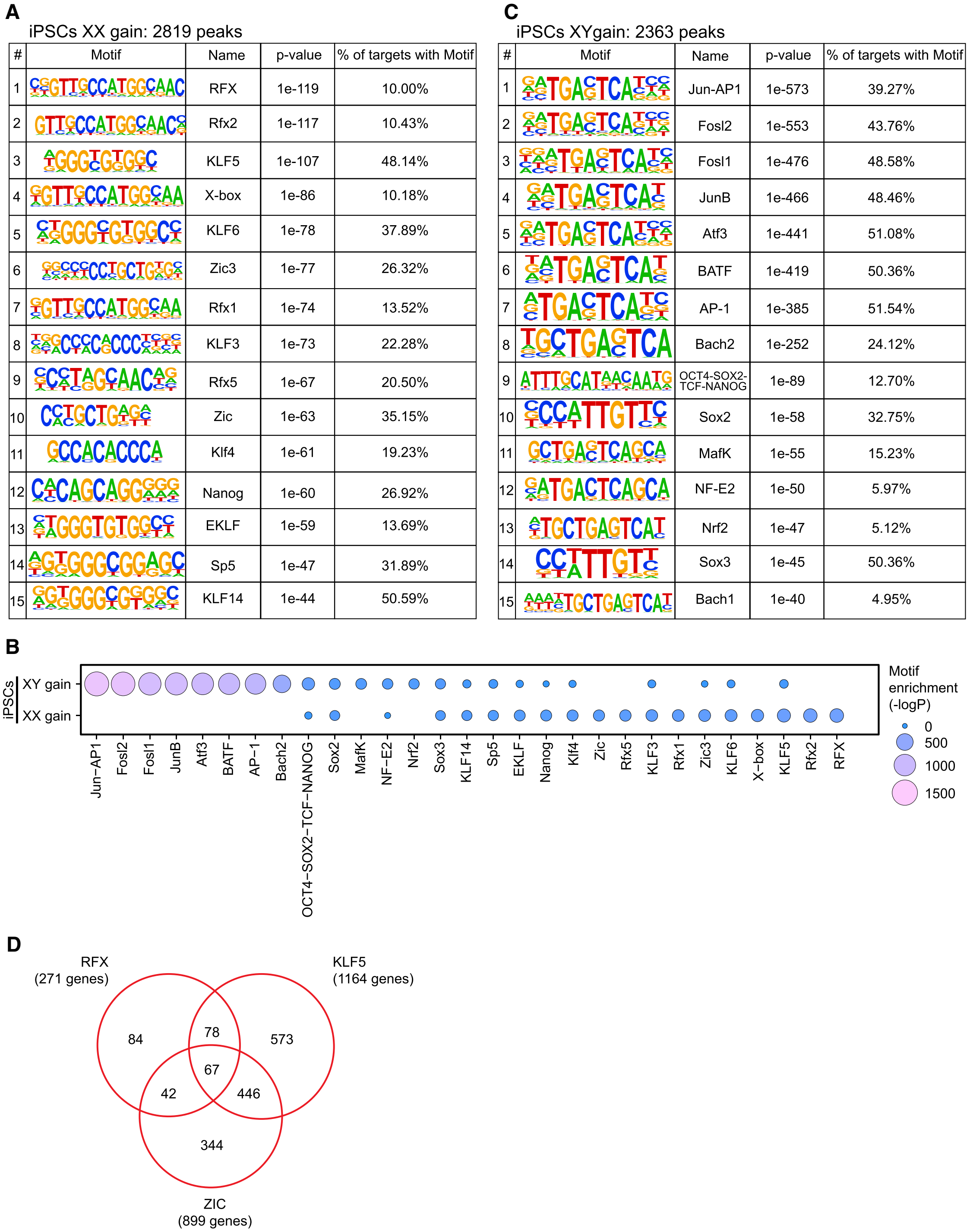
To identify the putative target genes, we searched for genes associated with open chromatin regions enriched for specific motifs, then determined the target genes shared for open chromatin containing more than one motif. In chromatin more open in XX iPSCs, we found that 67 genes were associated with binding motifs for all three TF motifs RFX, KLF, and ZIC (Figure 3D). Taken together, these analyses allowed the identification of TFs that regulate a large number of *cis*-regulatory regions, thereby improving our understanding on how X-dosage can drive two distinct PSC states.

Zic3 and *Tfe3* Dosage Do Not Explain X-Dosage Differences in Transcription and Pluripotency Exit

We sought to test whether X-linked pluripotency-associated genes with enriched motifs identified in chromatin more open in XX iPSCs stabilize pluripotency in XX PSCs. Our motif discovery analysis identified the X-linked gene *Zic3* within the top motifs enriched in chromatin more open in XX iPSCs and ESCs (Figures 3A and S2F). Western blot analysis showed that XX iPSCs and ESCs express higher ZIC3 protein than XY iPSCs and ESCs (Figure 4A). In addition, the increased *Zic3* transcript levels of XX iPSCs were restored to XY levels in XO iPSCs (Figure S3A). Moreover, *Zic3* was reported to prevent endodermal lineage specification and to act as a transcriptional activator of *Nanog* expression (Lim et al., 2010), further suggesting that it could have a role in stabilizing pluripotency in XX ESCs. To test the hypothesis that increased *Zic3* dosage stabilizes pluripotency in XX ESCs, we overexpressed *Zic3* in XY iPSCs and asked whether it induced XX-like features (Figure S3B). We achieved 3-fold overexpression of ZIC3 protein tagged in N- or C-terminal with hemagglutinin (HA) (Figure S3C). We then subjected the cells to LIF withdrawal. We found that overexpression of *Zic3* with an N-terminal HA tag, but not with the C-terminal HA tag, delayed pluripotency exit during LIF withdrawal differentiation (Figures S3B–S3E). These results suggested that increased *Zic3* dosage might be responsible for the pluripotency exit delay of XX PSCs. Using an independent approach, we generated *Zic3* heterozygous deletions in XX ESCs to reduce *Zic3* dosage, which is a more stringent method to test whether *Zic3* dosage stabilizes pluripotency in XX ESCs (Figures 4B–4D, S3F, and S3G). Two independent *Zic3*$^{+/-}$ XX ESC clones maintained XX-like expression of *Prdm14*, *Nanog*, and *Tcl1* and also maintained female-like delayed exit from pluripotency (Figure 4E). We performed similar experiments for another additional X-linked gene involved in pluripotency, *Tfe3*, and obtained similar results as for *Zic3* (Figures 4F–4H and S3H–S3J). These results support the idea that the dosage of *Zic3* and *Tfe3* does not explain the differences in pluripotency gene expression and exit from pluripotency between XX and XY ESCs.

Dkc1, *Otud6a*, *Fhl1*, *Zfp185*, and *Scml2* Dosage Do Not Explain X-Dosage-Specific Differences in Pluripotency Exit

We sought to find the X-linked regulators that drive stabilization of pluripotency in XX PSCs. We analyzed RNA-seq and published proteomics data of XX and XY ESCs (Choi et al., 2017a). We selected X-linked candidate factors with (1) increased expression in XX ESCs over XY or XO ESCs and ranked by expression ratio, (2) evidence that the genes are subject to XCI (Table S4), and (3) literature consistent with a possible role in stabilizing pluripotency. The selected candidate genes were *Dkc1*, *Otud6a*, *Fhl1*, *Zfp185*, and *Scml2*. We overexpressed their cDNAs in XY iPSCs (Figure S4). To test the effect of overexpression on



A

iPSCs XX gain: 2819 peaks

#	Motif	Name	p-value	% of targets with Motif
1	CGGTTGCCATGGCAAC	RFX	1e-119	10.00%
2	GTTGCCATGGCAACC	Rfx2	1e-117	10.43%
3	AGGGTGTGGC	KLF5	1e-107	48.14%
4	GGTTGCCATGGCAA	X-box	1e-86	10.18%
5	CTGGGCGTGGCC	KLF6	1e-78	37.89%
6	GGCCCCCTGCTGTGC	Zic3	1e-77	26.32%
7	GGTTGCCATGGCAA	Rfx1	1e-74	13.52%
8	TGGCCCCACCCCTCGC	KLF3	1e-73	22.28%
9	CCCTAGCAACAG	Rfx5	1e-67	20.50%
10	CCTGCTGAGA	Zic	1e-63	35.15%
11	GCCACACCCA	Klf4	1e-61	19.23%
12	CACAGCAGGGGG	Nanog	1e-60	26.92%
13	CTGGGTGTGGCC	EKLF	1e-59	13.69%
14	AGTGGGCGGAGC	Sp5	1e-47	31.89%
15	GGTGGGCGGGGC	KLF14	1e-44	50.59%

C

iPSCs XYgain: 2363 peaks

#	Motif	Name	p-value	% of targets with Motif
1	GATGACTCATCC	Jun-AP1	1e-573	39.27%
2	GATGAGTCATCC	Fosl2	1e-553	43.76%
3	GGATGAGTCATC	Fosl1	1e-476	48.58%
4	GATGACTCAT	JunB	1e-466	48.46%
5	GATGACTCATCC	Atf3	1e-441	51.08%
6	TATGAGTCAT	BATF	1e-419	50.36%
7	ATGACTCATC	AP-1	1e-385	51.54%
8	TGCTGAGTCA	Bach2	1e-252	24.12%
9	ATTTGCATAACAATG	OCT4-SOX2- TCF-NANOG	1e-89	12.70%
10	CCCATTGTTC	Sox2	1e-58	32.75%
11	GCTGAGTCAGCA	MafK	1e-55	15.23%
12	GATGACTCAGCA	NF-E2	1e-50	5.97%
13	ATGCTGAGTCAT	Nrf2	1e-47	5.12%
14	CCTTTGTT	Sox3	1e-45	50.36%
15	AAATTGCTGAGTCAT	Bach1	1e-40	4.95%

Figure 3. Identification of Candidate Regulators Mediating the Effects of X-Dosage on Open Chromatin

(A–C) TF motifs enriched in chromatin regions more open in XX or XY iPSCs.

(D) Venn diagram showing the overlap between genes associated the ATAC-seq regions more open in XX iPSCs that contain a motif for KLF5, RFX, or ZIC. The number of genes associated with all three motifs is indicated.

See also Figure S2.

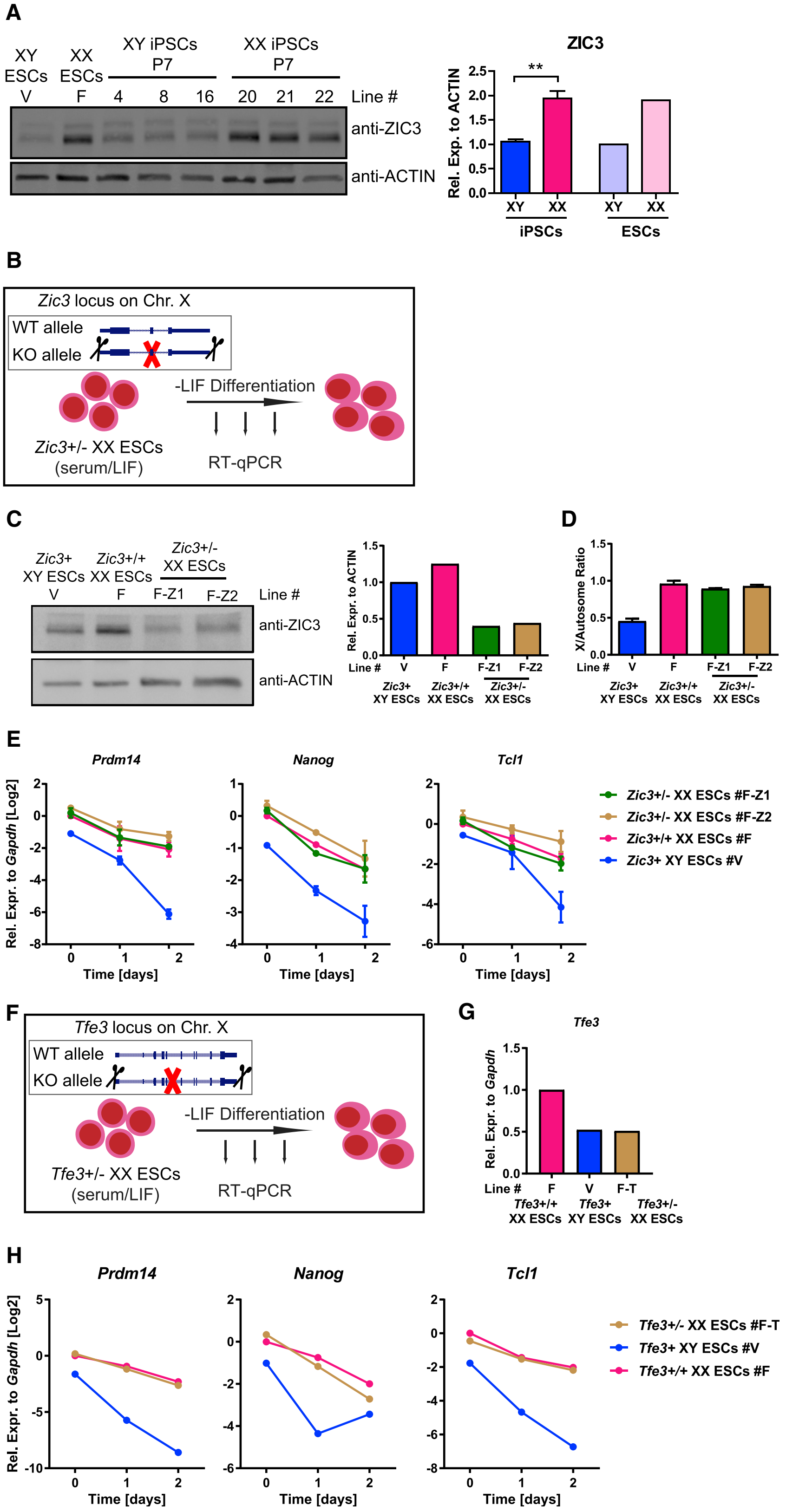


Figure 4. Effects of *Zic3/Tfe3* Heterozygous Deletion on Pluripotency Exit

(A) Western blot analysis for ZIC3 in iPSCs and ESCs grown in S/L. Right: quantification using actin as loading control. Three XY versus 3 XX iPSC lines, and 1 XY versus 1 XX ESC line. $**p < 0.01$, unpaired two-tailed t test. Data from a representative experiment from at least three independent experiments.

(B) Scheme of heterozygous *Zic3* deletion strategy in XX ESCs followed by LIF withdrawal.

(C) Western blot analysis for ZIC3 in two independent *Zic3*$^{+/-}$ ESC lines, *Zic3*$^{+/+}$ ESCs and XY ESCs all grown in S/L (n = 1). KO, knockout; WT, wild-type.

(D) qPCR analysis for X-chromosome DNA copy number. X copy numbers are presented as the average ratio of genomic DNA (gDNA) quantities for four X-linked genes (*Tfe3*, *Bcor*, *Pdha1*, and *Mid1*) to gDNA quantities for autosomal gene *Gapdh* (n = 2).

(E) qRT-PCR analysis for pluripotency-associated gene expression during LIF withdrawal in the two independent *Zic3*$^{+/-}$ XX ESC lines, the *Zic3*$^{+/+}$ XX parental ESC line, and an XY ESC line (n = 2).

(F) Scheme of heterozygous *Tfe3* deletion strategy in XX ESCs followed by LIF withdrawal.

(G) qRT-PCR analysis for *Tfe3* expression in the *Tfe3*$^{+/-}$ XX ESC line, the *Tfe3*$^{+/+}$ XX parental ESC line, and an XY ESC line (n = 1).

(H) As in (E) for *Tfe3*$^{+/-}$ ESCs.

Results are presented as averages (±SEM) of biological duplicates (n = 1). See also Figures S3 and S4.

pluripotency exit, we induced differentiation by LIF withdrawal and measured pluripotency gene expression at 24 and 48 hr. We found that overexpression of *Dkc1*, *Otud6a*, *Fhl1*, *Zfp185*, or *Scml2* was not sufficient to induce a delay in pluripotency exit (Figure S4). Collectively, these findings do not support a significant role for these X-linked pluripotency-associated genes in stabilizing pluripotency in XX ESCs.

Heterozygous *Dusp9* Deletion in XX ESCs Induces Male-like DNA Methylation yet Maintains a Female-like Transcriptome, Open Chromatin Landscape, Growth, and Delayed Pluripotency Exit

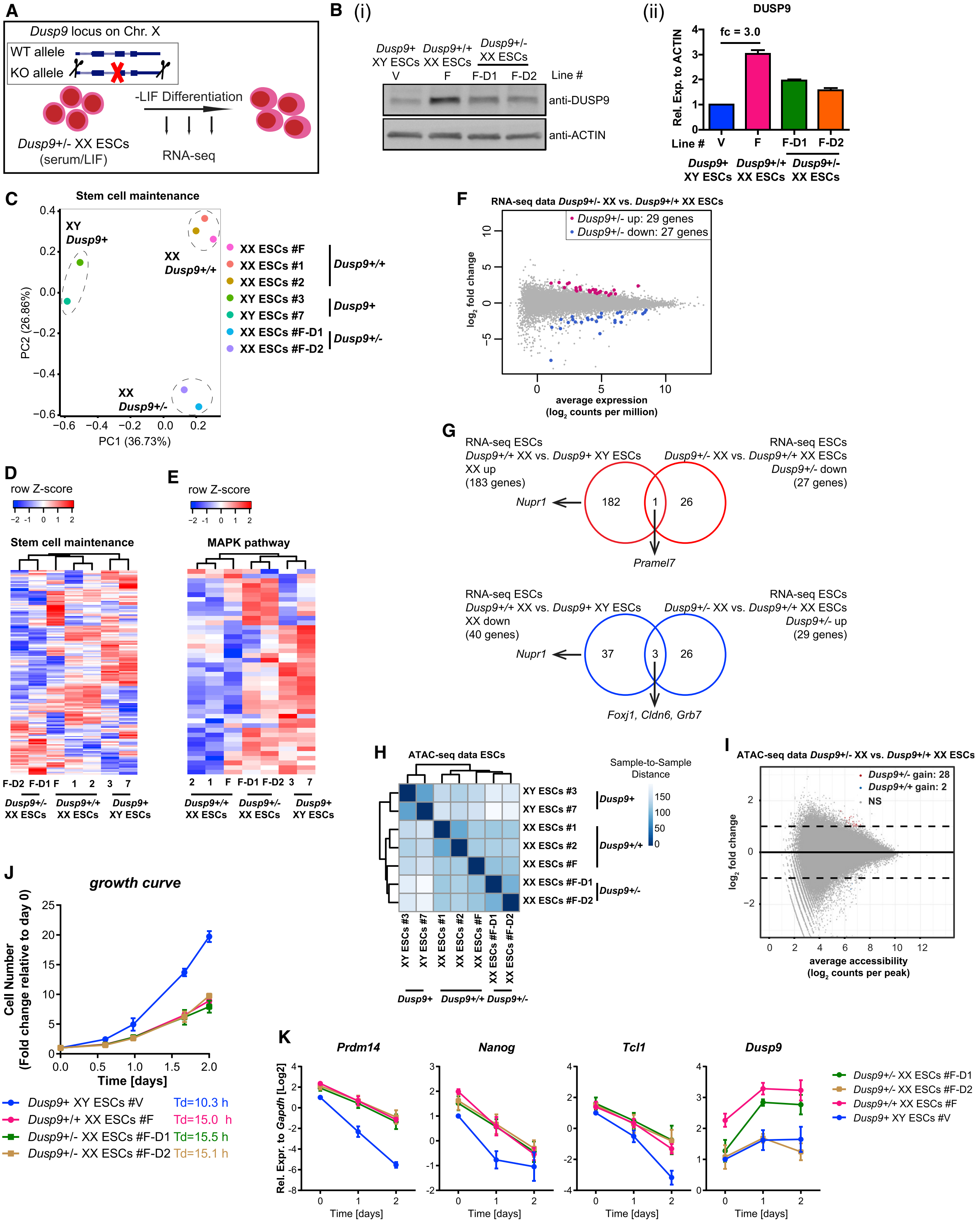
In an effort to understand the mechanisms by which X-dosage affects PSC properties, we generated *Dusp9* heterozygous deletions in XX ESCs, resulting in two independent $Dusp9^{+/-}$ XX ESC clones (Figures 5A, 5B, S5A, and S5B). To ensure the maintenance of two Xas in $Dusp9^{+/-}$ ESCs, we used polymorphic *Musculus/Castaneus* (Mus/Cas) ESCs, known to be less susceptible to X-chromosome loss (Choi et al., 2017a; Lee and Lu, 1999). We confirmed that $Dusp9^{+/-}$ ESCs maintained two Xas (Figures S5C and S5D). $Dusp9^{+/-}$ XX ESCs showed male-like global DNA methylation (Figure S5E), corroborating recent findings (Choi et al., 2017a). To determine whether $Dusp9^{+/-}$ XX ESCs with male-like DNA methylation acquire male-like transcription, we analyzed the expression of stem cell maintenance-related genes using RNA-seq in $Dusp9^{+/-}$ XX ESCs, $Dusp9^{+/+}$ XX ESCs, and $Dusp9^{+}$ XY ESCs, all sharing a Mus/Cas background to exclude potential strain-specific effects. Principal component analysis (PCA) placed $Dusp9^{+/-}$ XX ESCs away from both $Dusp9^{+/+}$ XX ESCs and $Dusp9^{+}$ XY ESCs, indicating that $Dusp9^{+/-}$ XX ESCs do not adopt a male-like transcriptional state (Figure 5C). We corroborated this finding using unsupervised clustering of stem cell maintenance-related gene expression, whereby $Dusp9^{+/-}$ XX ESCs clustered together with $Dusp9^{+/+}$ XX ESCs, and away from $Dusp9^{+}$ XY ESCs (Figure 5D). Unsupervised clustering analysis also showed the activation of most MAPK target genes in $Dusp9^{+/-}$ XX ESCs, in agreement with the function of *Dusp9* as an MAPK inhibitor (Figure 5E) (Li et al., 2012). Furthermore, we found more DEGs between $Dusp9^{+/+}$ XX ESCs and XY ESCs, and less DEGs between $Dusp9^{+/-}$ XX ESCs and $Dusp9^{+/+}$ XX ESCs, with little overlap between the two sets of DEGs (Figures 5F and 5G). The only exception was the *Pramel7* gene, the expression of which was reduced to XY levels in $Dusp9^{+/-}$ XX ESCs, indicating that transcription of *Pramel7* is influenced by *Dusp9* dosage (Table S4). Overall these results indicate that male-like DNA methylation can be induced in the absence of male-like transcription in $Dusp9^{+/-}$ XX ESCs. These experiments raise the possibility that distinct molecular features modulated by X-dosage in ESCs might be controlled by different regulators.

Next, we sought to determine whether chromatin accessibility is affected by reduced DUSP9 dosage in XX ESCs. ATAC-seq revealed that $Dusp9^{+/-}$ XX ESCs maintained a female-like open chromatin landscape (Figure 5H). We only observed very few differences in chromatin accessibility between $Dusp9^{+/-}$ XX ESCs and $Dusp9^{+/+}$ XX ESCs (Figure 5I), while the same analysis identified thousands of regions differentially accessible in XX and XY ESCs (Figure S2D). This finding indicates that the effects of X-dosage on chromatin accessibility can be largely dissociated from X-dosage effects on global DNA methylation levels.

We then measured the growth of $Dusp9^{+/-}$ XX ESCs, and found that the cells grew as slowly as their parental $Dusp9^{+/+}$ XX ESCs, both of which grew slower than XY ESCs (Figures 5J and 1K). Therefore, reducing *Dusp9* dosage and inducing global DNA methylation in XX ESCs is not sufficient to induce male-like cellular growth. We propose that X-dosage-specific growth and global DNA methylation are regulated by different pathways in mouse PSCs.

To study the effects of *Dusp9* heterozygous deletion in XX ESCs on pluripotency exit, we subjected $Dusp9^{+/-}$ XX ESCs, $Dusp9^{+/+}$ XX ESCs, and XY ESCs to LIF withdrawal differentiation for 48 hr followed by qRT-PCR analysis. The delay in pluripotency exit as judged by *Prdm14*, *Nanog*, and *Tcl1* expression was maintained in $Dusp9^{+/-}$ XX cells relative to $Dusp9^{+/+}$ XX cells (Figure 5K). In further support of the finding that $Dusp9^{+/-}$ XX ESCs maintain a delay in pluripotency exit, RNA-seq analysis showed that multiple pluripotency-associated genes behaved similarly in $Dusp9^{+/-}$ XX and $Dusp9^{+/+}$ XX cells undergoing differentiation (Figure S5F). Therefore, mechanistically, reducing *Dusp9* dosage is compatible with female-like pluripotency exit. We conclude that reducing the dosage of *Dusp9* in XX ESCs is not sufficient to induce a male-like transcriptome or accelerate pluripotency exit to a male-like state, despite changes in the expression level of multiple genes in the MAPK signaling pathway and despite male-like DNA methylation. In addition, *Dusp9* overexpression in XY ESCs did not induce a female-like delay in differentiation (Figures S5G–S5L) despite inducing female-like global DNA hypomethylation (Choi et al., 2017a).

Altogether, these results indicate that most changes in open chromatin, growth, and pluripotency exit as a result of differences in X-dosage are regulated independently of global DNA methylation in XX ESCs. Hence, mechanistically, heterozygous *Dusp9* deletion molecularly uncouples global DNA methylation from the open chromatin landscape, growth, and the pluripotency exit delay of XX ESCs. Importantly, this points toward the existence of other pathways and X-linked genes involved in mediating the effects of X-dosage in PSCs.



(legend on next page)

Large-Fragment Heterozygous Deletions

Two models emerged to explain delayed pluripotency exit in XX PSCs. In the first model, a single X-linked gene is responsible for delayed pluripotency exit. In the second model, multiple X-linked genes act together to delay pluripotency exit. To test these models, we generated a series of large-fragment (LF) heterozygous deletions of the X chromosome in XX ESCs, which were confirmed by genotyping and Sanger sequencing (Figure S6), then induced differentiation. There was a partial rescue of the pluripotency exit delay of XX ESCs in multiple, but not all, LF deletions (Figure 6 and Table S1). The partial rescue was gene specific and fragment specific. These results favor model 2, in which multiple X-linked genes participate in delayed pluripotency exit in XX PSCs.

DISCUSSION

Induction of naive pluripotency during reprogramming to iPSCs and during *in vivo* development in ICM leads to XCR in murine female cells (Maherali et al., 2007; Mak, 2004; Okamoto et al., 2004). The consequences of X-dosage imbalance between female (XX) and male (XY) cells on mouse iPSCs, and the regulatory mechanisms at the basis of distinct X-dosage-specific features in mouse PSCs, remain incompletely understood. In this study, we addressed these questions by analyzing the transcriptome, growth properties, chromatin accessibility landscape, and pluripotency exit of isogenic female and male iPSCs. We identified X-dosage as a factor influencing the molecular and cellular properties of iPSCs. By employing epigenomic analyses we found that X-dosage modulates open chromatin in iPSCs and ESCs. Furthermore, using genome editing we found that modulation of the transcriptome, open chromatin, cell growth, and pluripotency exit by X-dosage is largely independent of global DNA methylation. We provide evidence favoring a model in which multiple X-linked genes delay pluripotency exit.

One outcome of our study is that the number of Xas correlates with differences in the transcriptome and in pluripotency exit in iPSCs, in addition to differences in DNA methylation (Milagre et al., 2017; Pasque et al., 2018). Reprogramming somatic cells to iPSCs is an important system for studying erasure of epigenetic memory and pluripotency. Sex does not appear to influence the efficiency of iPSC generation, since we previously showed that female and male cells reprogram with similar efficiencies in this system (Pasque et al., 2014). However, we have now established that the presence of two Xas, as a result of reprogramming to pluripotency in female cells, is associated with the slower growth of XX iPSCs, an altered transcriptome including increased pluripotency-associated gene expression, and delayed pluripotency exit. These differences are caused by changes in X-dosage, since XO iPSCs revert to an XY iPSC phenotype. The influence of X-dosage on the growth of mouse PSCs, and on the transcriptome and pluripotency exit of iPSCs, is consistent with studies in mouse ESCs (Schulz, 2017; Schulz et al., 2014; Zvetkova et al., 2005) and human ESCs (Bruck et al., 2013; Ronen and Benvenisty, 2014), and in post-implantation mammalian embryos (Schulz, 2017). However, X-dosage is currently largely ignored in most reprogramming studies, in which neglecting sex-specific differences could have a negative

Figure 5. The Open Chromatin Landscape, Growth, and Pluripotency Exit Delay of XX ESCs Are Maintained in the Presence of Male-like Global DNA Methylation

(A) Scheme of *Dusp9* heterozygous deletion in XX ESCs followed by LIF withdrawal differentiation. KO, knockout; WT, wild-type.

(B) (i) Western blot analysis for DUSP9 in *Dusp9*$^{+/-}$ ESCs, *Dusp9*$^{+/+}$ ESCs, and XY ESCs grown in S/L. (ii) Quantification of DUSP9 levels using actin as a loading control (n = 2).

(C) PCA of stem cell maintenance genes (RNA-seq data) for the *Dusp9*$^{+/-}$, *Dusp9*$^{+/+}$, and XY ESCs grown in S/L conditions.

(D) Unsupervised hierarchical clustering of stem cell maintenance genes for the *Dusp9*$^{+/-}$, *Dusp9*$^{+/+}$, and XY ESCs.

(E) As in (D) for MAPK pathway-related genes (defined in Schulz et al., 2014).

(F) DEG analysis, identifying clear differences between *Dusp9*$^{+/+}$ XX and XY ESCs, but much less between *Dusp9*$^{+/-}$ XX and *Dusp9*$^{+/+}$ XX ESCs.

(G) Overlap between the DEGs of *Dusp9*$^{+/+}$ versus XY ESCs and *Dusp9*$^{+/-}$ versus *Dusp9*$^{+/+}$ ESCs for upregulated genes (up) and downregulated genes (down). *Dusp9* heterozygous deletion maintains a female-like transcriptome.

(H) ATAC-seq sample-to-sample distance heatmap in *Dusp9*$^{+/+}$, *Dusp9*$^{+/-}$, and XY ESCs. *Dusp9*$^{+/-}$ ESCs maintain a *Dusp9*$^{+/+}$-like open chromatin landscape.

(I) Differential chromatin accessibility analysis between *Dusp9* heterozygous mutant and wild-type XX ESCs. Log_2 fold change (mutant/wild-type) in reads per accessible region are plotted against the mean reads per ATAC-seq peak. *Dusp9* heterozygous mutants maintain a female-like open chromatin landscape.

(J) Growth curves and doubling times of *Dusp9*$^{+/-}$, *Dusp9*$^{+/+}$, and XY ESCs in S/L condition. Cells were counted at the indicated time points and presented as fold changes relative to day 0, averages (±SEM) of biological duplicates. Data from a representative experiment from at least 4 independent experiments.

(K) qRT-PCR for *Prdm14*, *Nanog*, *Tcl1*, and *Dusp9* expression before and after LIF withdrawal (n = 3).

See also Figure S5.

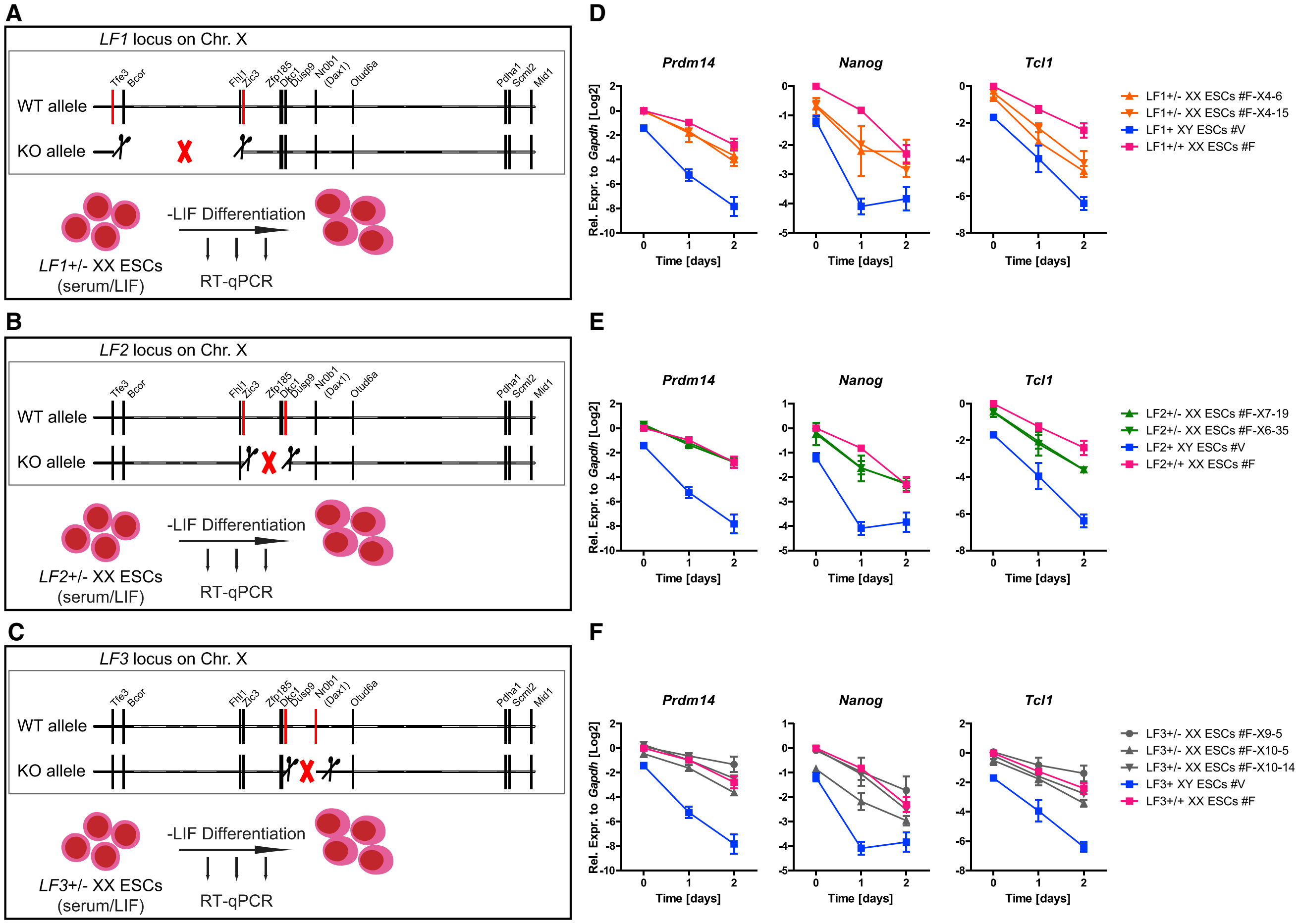


Figure 6. Multiple X-Linked Genes Modulate the Pluripotency Exit Delay of XX PSCs

(A–C) Scheme of large-fragment (LF) heterozygous deletions in XX ESCs followed by LIF withdrawal differentiation. (A) LF1, (B) LF2, (C) LF3. WT, wild-type.

(D–F) qRT-PCR for *Prdm14*, *Nanog*, and *Tcl1* expression before and after LIF withdrawal in 2–3 independent LF$^{+/-}$ XX ESC lines, the parental XX ESC line, and an XY ESC line (n = 2). (D) LF1, (E) LF2, (F) LF3.

See also Figure S6.

impact on interpretation of results. The notion that X-dosage influences the molecular and cellular properties of iPSCs is further supported by the loss of sex-specific differences concomitant with loss of one X chromosome in female iPSCs, in agreement with previous observations in ESCs (Choi et al., 2017a; Schulz et al., 2014; Zvetkova et al., 2005) and iPSCs (Pasque et al., 2018). The important point is that studies of reprogramming to iPSCs should consider the number of Xas as a modulator of the transcriptome, and cells of different sex should be studied separately, but also considered together.

The influence of X-dosage on the heterogeneity of ESCs also remained unclear. Our analysis of single-cell RNA-seq data for XX and XY ESCs in S/L and 2i/L (Chen et al., 2016) revealed that both XX and XY ESCs reside in a metastable state, with *Nanog*-high and *Nanog*-low cells (Figure S7). More XX ESCs express *Tcl1* than XY ESCs, in agreement with increased pluripotency-associated gene expression in XX ESCs (Schulz et al., 2014). Differences between the transcriptome of XX and XY ESCs persist in 2i/L, despite more homogeneous pluripotency-associated gene expression. Altogether, this analysis reveals that XX PSCs in S/L maintain a metastable state with a bias toward increased expression of specific pluripotency-associated genes, whereas transcriptomic differences between XX and XY PSCs persist in 2i/L.

Here we show that the presence of two Xas in iPSCs and ESCs is associated with delayed cellular growth. One possible interpretation is that the delayed growth of female post-implantation mammalian embryos (Burgoyne et al., 1995) is recapitulated in mouse PSCs. Since the growth differences of XX and XY/XO PSCs are maintained after dual

GSK3B and ERK inhibition, additional pathways are likely involved. One hypothesis is that there could be a competitive growth advantage of cells that have undergone XCI in the post-implantation mammalian embryo to select against remaining cells that may fail to undergo XCI, and maintain two Xas. Our *in vitro* experiment suggests that XCI could indeed provide a small growth advantage. However, this hypothesis remains to be tested *in vivo*.

To better understand what drives the features of the pluripotent regulatory network that are modulated by X-dosage in PSCs, we explored the open chromatin landscapes of female and male PSCs. While both female and male iPSCs/ESCs possessed globally similar open chromatin landscapes, thousands of chromatin regions were differentially accessible in XX and XY PSCs. These differentially accessible regions may underlie differences in the transcriptional regulatory network and functional properties of XX and XY PSCs. Decoding differentially accessible chromatin regions, we identified pluripotency-associated genes *Pramel6* and *Pramel7* with increased accessibility in XX iPSCs. *Pramel7* has been associated with naive pluripotency (Graf et al., 2017), and overexpression of *Pramel6* and *Pramel7* both compromise pluripotency exit (Casanova et al., 2011). We went further by identifying a catalog of *cis*-regulatory regions including promoters that are modulated by X-dosage in iPSCs and ESCs. These observations indicate that X-dosage can modulate chromatin accessibility in mouse PSCs.

Decoding differentially accessible chromatin allowed us to distinguish distinct sets of enriched TF binding motifs in XX and XY ESCs. Specifically, motifs for KLF5, ZIC3, and NANOG were enriched in chromatin more open in XX iPSCs, all of which have been implicated in pluripotency (reviewed in Martello and Smith, 2014). These results suggest that the stabilization of pluripotency in XX ESCs may be mediated by these core master regulators. In particular, *Zic3* is a known pluripotency factor (Lim et al., 2010), encoded on the X chromosome, and is not dosage compensated in XX PSCs. However, *Zic3* heterozygous deletion had no effect on stabilization of pluripotency. Although no TF chromatin immunoprecipitation sequencing (ChIP-seq) data are available for XX iPSCs to date, our *in silico* analyses identified a high-confidence set of direct putative KLF5, ZIC3, and NANOG targets in XX iPSCs, including known pluripotency-associated genes. Moreover, the specific putative regulatory region associated with *Pramel6*, which becomes more accessible in XX iPSCs, overlaps with ChIP-seq binding sites of OCT4, SOX2, and NANOG in XY ESCs (not shown), further suggesting that increased binding of master pluripotency regulators takes place at these more accessible regions in XX iPSCs. Our results raise the possibility that pluripotency is stabilized in XX iPSCs by binding of core pluripotency factors to a subset of regulatory elements whose accessibility is influenced by X-dosage.

In contrast to the XX state, chromatin more open in XY iPSCs identified AP-1 TFs as candidate regulators, which have not previously been implicated in X-dosage-specific regulation of pluripotency. JUN/AP-1 control many cellular processes including proliferation, apoptosis, and differentiation in response to a variety of stimuli including the MAPK pathway (reviewed in Shaulian and Karin, 2002). The role of AP-1 TFs in the context of X-dosage in iPSCs warrants future study.

A previous study showed that *Dusp9* modulates DNA hypomethylation and the proteome in XX female mouse ESCs (Choi et al., 2017a). However, the effects of reducing *Dusp9* dosage in XX ESCs on growth, transcription, and pluripotency exit were unknown. An important outcome of our analyses is that *Dusp9* heterozygous XX ESCs maintain a female-like open chromatin landscape, growth, and delayed pluripotency exit concomitant with male-like global DNA methylation levels. These results suggest that chromatin accessibility, growth, and delayed pluripotency exit can be regulated independently of global DNA methylation levels in mouse PSCs. This result was unexpected for two reasons. First, *Dusp9* overexpression in ESCs was reported to induce a female-like proteome, including activation of naive pluripotency marker PRDM14. Second, reducing the expression of DNMTs in male ESCs is associated, at least in part, with delayed pluripotency exit (Schulz et al., 2014). However, Choi et al. (2017a) reported that the ICM of female and male embryos shows comparable DNA methylation, despite delayed female development, suggesting that DNA hypomethylation and stabilization of pluripotency can be uncoupled both *in vivo* and *in vitro*. Our results therefore suggest that global DNA methylation levels are regulated, at least in part, by distinct X-linked genes, different from those regulating the open chromatin landscape and stabilization of pluripotency in PSCs (*Dusp9* for DNA methylation levels, other gene(s) for chromatin accessibility and delayed pluripotency exit and growth). Choi et al. (2017a) reported that *Dusp9* overexpression in male ESCs increases the expression of PRDM14, ROR2, and TFCP2L1. In our study, 3.5- to 3.7-fold overexpression of DUSP9 protein in male ESCs, achieving DUSP9 protein level comparable with that in XX ESCs, did not lead to an increase in *Prdm14* transcript level. This may be explained by differences in the level or method of *Dusp9* overexpression (inducible system versus piggyBac), or the assay used to judge expression of pluripotency markers (mass spectrometry versus qRT-PCR). Interestingly, *Pramel7* overexpression in ESCs was shown to induce DNA hypomethylation through the degradation of DNA methylation maintenance factor UHRF1 (Graf et al., 2017). At the same time, we found that reducing *Dusp9* dosage in XX

ESCs reduces *Pramel7* expression to male levels. The results suggest that *Pramel7* may be downstream of *Dusp9*, and may participate in the control of DNA methylation by X-dosage.

Our results do not support a model in which a single X-linked gene stabilizes pluripotency in XX PSCs. First, single heterozygous deletions of *Zic3*, *Tfe3*, or overexpression of *Zic3*, *Dkc1*, *Otud6a*, *Fhl1*, *Zfp185*, and *Scml2* had little effect. Second, distinct LF heterozygous deletions suggested that multiple X-linked genes participate in delaying pluripotency exit in XX PSCs. Therefore, identifying additional X-linked regulators that mediate the effects of X-dosage in PSCs requires future investigations. An interesting additional candidate is the recently identified X-linked transient octamer binding factor 1 (TOBF1) (Chakraborty et al., 2017), since it was shown to sustain pluripotency. It is also possible that other regulators of the Erk pathway are involved. A previous study in human ESCs reported that human primed PSCs with eroded XCI and increased expression of the MAPK/ERK downstream effector ELK-1 have decreased expression of TRA-1-60, a marker of the differentiated state (Bruck et al., 2013). However, human primed PSCs studies are likely not compatible with mouse naive PSC studies because cells reside in distinct pluripotent states.

To conclude, we revealed that global DNA methylation can be uncoupled from delayed pluripotency exit in XX mouse ESCs. Furthermore, our study shows that X-dosage-specific differences in cell growth, open chromatin landscape, transcription, and pluripotency exit in iPSCs correlate with the number of Xas. We also reveal a mechanism by which multiple genomic regions on the X chromosome are responsible for delaying pluripotency exit. Using information from the genome, the epigenome, and the transcriptome we gained insights into modulation of the open chromatin landscape and the transcriptional regulatory network of iPSCs by X-dosage. Furthermore, better understanding how X-dosage modulates pluripotency will have important implications for disease modeling and regenerative medicine.

EXPERIMENTAL PROCEDURES

Detailed information on cell line derivation, culture conditions, differentiation, plasmids, immunofluorescence, RNA fluorescence *in situ* hybridization, qPCR, qRT-PCR, and western blot, as well as clonal, cell growth, and cell-cycle assays are available in Supplemental Experimental Procedures.

Statistical Analysis

Statistical tests were performed using GraphPad Prism 5 (GraphPad Software). Unpaired two-tailed t test, one-way ANOVA with multiple comparisons test, or two-way repeated-measures ANOVA were used as indicated. The number of independent experiments performed is indicated (n). All data are presented as the mean ± SEM. p values of <0.05 were considered statistically significant.

RNA-Seq and Analysis

Total RNA was extracted from 2–3 independent cell lines for each cell type (Table S1) and used for construction of a single-end indexed, stranded poly(A) mRNA-seq library. DEG analysis and clustering analysis were performed in R.

ATAC-Seq and Analysis

OmniATAC-seq was performed as described by Corces et al. (2017) from 2–3 independent cell lines for each cell type (Table S1). Analysis of HiSeq4000 single-end reads used a pipeline from the Kundaje lab (version 0.3.3) (Lee et al., 2016). Differential chromatin accessibility analysis was performed using the DiffBind package. Motif analysis was performed using the HOMER package (v4.9.1).

ACCESSION NUMBERS

The GEO accession number for the RNA-seq and ATAC-seq data reported in this paper is GEO: GSE110215. The single-cell RNA-seq data (GEO: GSE74155) was from Chen et al. (2016), and V6.5 ESC and MEF RNA-seq (GEO: GSE90894) from Chronis et al. (2017).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and five tables and can be found with this article online at https://doi.org/10.1016/j.stemcr.2018.12.004.

AUTHOR CONTRIBUTIONS

Conception and design, J.S. and V.P.; Experiments, J.S., N.D.G., L.V., M.E.B., T.O., I.T., and V.P.; Analyses, J.S.; ATAC-seq and RNA-seq analyses, A.J. and J.S.; Writing, V.P. and J.S. with input from all authors; Supervision, V.P.

ACKNOWLEDGMENTS

We thank Stein Aerts, Konrad Hochedlinger, Frederic Lluis Vinas, Edda Schulz, and Edith Heard for discussions; Rudolf Jaenisch for providing mice; Qiaolin Deng for providing Mus/Cas ESCs; Rita Khoueiry and Michela Bartoccetti for help with derivation of Mus/Cas ESCs; Ye-Guang Chen for the *Dusp9* plasmids; Miguel Branco for the Scml2 plasmids; José Silva for the PiggyBac plasmid; Mitchell Guttman and Jesse Engreitz for the pZB-sg3 plasmid; Ali Brivanlou and Ariel Waisman for the FUCCI plasmids; Stein Aerts, Kristofer Davie, Liesbeth Minnoye, and Xinlong Luo for help with bioinformatics and ATAC-seq; and Kenjiro Shirane and Hiroyuki Sasaki for providing processed data. We thank Constantinos Chronis, Kian Koh, Kathrin Plath, and Edda Schulz for comments on an early version of the manuscript. We are grateful to the help provided by the KU Leuven FACS Core, Genomics Core, and Mouse Facility; Metabolomics Core, VIB/KU Leuven; and the Stem Cell Institute, Leuven. This work was supported by The Research Foundation Flanders (FWO) (Odysseus return grant

G0F7716N to V.P.), the KU Leuven Research Fund (BOFZAP starting grant StG/15/021BF to V.P., C1 grant C14/16/077 to V.P., and project financing), and FWO PhD fellowship to A.J. (1158318N).

Received: November 16, 2018
Revised: December 7, 2018
Accepted: December 10, 2018
Published: January 10, 2019

REFERENCES

Avior, Y., Sagi, I., and Benvenisty, N. (2016). Pluripotent stem cells in disease modelling and drug discovery. Nat. Rev. Mol. Cell Biol. *17*, 170–182.

Bruck, T., Yanuka, O., and Benvenisty, N. (2013). Human pluripotent stem cells with distinct X inactivation status show molecular and cellular differences controlled by the X-Linked ELK-1 gene. Cell Rep. *4*, 262–270.

Burgoyne, P.S., Thornhill, A.R., Boudrean, S.K., Darling, S.M., Bishop, C.E., and Evans, E.P. (1995). The genetic basis of XX-XY differences present before gonadal sex differentiation in the mouse. Philos. Trans. R. Soc. Lond. B Biol. Sci. *350*, 253–260, discussion 260–261.

Carey, B.W., Markoulaki, S., Beard, C., Hanna, J., and Jaenisch, R. (2010). Single-gene transgenic mouse strains for reprogramming adult somatic cells. Nat. Methods *7*, 56–59.

Casanova, E.A., Shakhova, O., Patel, S.S., Asner, I.N., Pelczar, P., Weber, F.A., Graf, U., Sommer, L., Bürki, K., and Cinelli, P. (2011). Pramel7 mediates LIF/STAT3-dependent self-renewal in embryonic stem cells. Stem Cells *29*, 474–485.

Chakraborty, D., Paszkowski-Rogacz, M., Berger, N., Ding, L., Mircetic, J., Fu, J., Iesmantavicius, V., Choudhary, C., Anastassiadis, K., Stewart, A.F., et al. (2017). lncRNA Panct1 maintains mouse embryonic stem cell identity by regulating TOBF1 recruitment to Oct-Sox sequences in early G1. Cell Rep. *21*, 3012–3021.

Chen, G., Schell, J.P., Benitez, J.A., Petropoulos, S., Yilmaz, M., Reinius, B., Alekseenko, Z., Shi, L., Hedlund, E., Lanner, F., et al. (2016). Single-cell analyses of X Chromosome inactivation dynamics and pluripotency during differentiation. Genome Res. *26*, 1342–1354.

Choi, J., Clement, K., Huebner, A.J., Webster, J., Rose, C.M., Brumbaugh, J., Walsh, R.M., Lee, S., Savol, A., Etchegaray, J.-P., et al. (2017a). DUSP9 modulates DNA hypomethylation in female mouse pluripotent stem cells. Cell Stem Cell *20*, 706–719.e7.

Choi, J., Huebner, A.J., Clement, K., Walsh, R.M., Savol, A., Lin, K., Gu, H., Di Stefano, B., Brumbaugh, J., Kim, S.-Y., et al. (2017b). Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. Nature *548*, 219–223.

Choksi, S.P., Lauter, G., Swoboda, P., and Roy, S. (2014). Switching on cilia: transcriptional networks regulating ciliogenesis. Development *141*, 1427–1441.

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

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat. Methods *14*, 959–962.

De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H.G., et al. (2015). Hallmarks of pluripotency. Nature *525*, 469–478.

Graf, U., Casanova, E.A., Wyck, S., Dalcher, D., Gatti, M., Vollenweider, E., Okoniewski, M.J., Weber, F.A., Patel, S.S., Schmid, M.W., et al. (2017). Pramel7 mediates ground-state pluripotency through proteasomal-epigenetic combined pathways. Nat. Cell Biol. *19*, 763–773.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development *136*, 1063–1069.

Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H.D., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N.C., et al. (2013). Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell *13*, 360–369.

Hackett, J.A., Dietmann, S., Murakami, K., Down, T.A., Leitch, H.G., and Azim Surani, M. (2013). Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency. Stem Cell Reports *1*, 518–531.

Karin, M. (1995). The regulation of AP-1 activity by mitogen-activated protein kinases. J. Biol. Chem. *270*, 16483–16486.

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

Lee, J., Christoforo, G., Christoforo, G., Foo, C.S., Probert, C., Kundaje, A., Boley, N., kohpangwei, Dacre, M., and Kim, D. (2016). kundajelab/atac_dnase_pipelines: 0.3.3. http://doi.org/10.5281/zenodo.211733.

Li, Z., Fei, T., Zhang, J., Zhu, G., Wang, L., Lu, D., Chi, X., Teng, Y., Hou, N., Yang, X., et al. (2012). BMP4 signaling acts via dual-specificity phosphatase 9 to control ERK activity in mouse embryonic stem cells. Cell Stem Cell *10*, 171–182.

Lim, L.S., Hong, F.H., Kunarso, G., and Stanton, L.W. (2010). The pluripotency regulator Zic3 is a direct activator of the Nanog promoter in ESCs. Stem Cells *28*, 1961–1969.

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.

Mak, W. (2004). Reactivation of the paternal X chromosome in early mouse embryos. Science *303*, 666–669.

Martello, G., and Smith, A. (2014). The nature of embryonic stem cells. Annu. Rev. Cell Dev. Biol. *30*, 647–675.

Milagre, I., Stubbs, T.M., King, M.R., Spindel, J., Santos, F., Krueger, F., Bachman, M., Segonds-Pichon, A., Balasubramanian, S., Andrews, S.R., et al. (2017). Gender differences in global but not targeted demethylation in iPSC reprogramming. Cell Rep. *18*, 1079–1089.

Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., and Heard, E. (2004). Epigenetic dynamics of imprinted X inactivation during early mouse development. Science *303*, 644–649.

Ooi, S.K., Wolf, D., Hartung, O., Agarwal, S., Daley, G.Q., Goff, S.P., and Bestor, T.H. (2010). Dynamic instability of genomic methylation patterns in pluripotent stem cells. Epigenetics Chromatin *3*, 17.

Pasque, V., and Plath, K. (2015). X chromosome reactivation in reprogramming and in development. Curr. Opin. Cell Biol. *37*, 75–83.

Pasque, V., Tchieu, J., Karnik, R., Uyeda, M., Sadhu Dimashkie, A., Case, D., Papp, B., Bonora, G., Patel, S., Ho, R., et al. (2014). X chromosome reactivation dynamics reveal stages of reprogramming to pluripotency. Cell *159*, 1681–1697.

Pasque, V., Karnik, R., Chronis, C., Petrella, P., Langerman, J., Bonora, G., Song, J., Vanheer, L., Sadhu Dimashkie, A., Meissner, A., et al. (2018). X chromosome dosage influences DNA methylation dynamics during reprogramming to mouse iPSCs. Stem Cell Rep. *10*, 1537–1550.

Ronen, D., and Benvenisty, N. (2014). Sex-dependent gene expression in human pluripotent stem cells. Cell Rep. *8*, 923–932.

Sakaue-Sawano, A., Kurokawa, H., Morimura, T., Hanyu, A., Hama, H., Osawa, H., Kashiwagi, S., Fukami, K., Miyata, T., Miyoshi, H., et al. (2008). Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. Cell *132*, 487–498.

Schulz, E.G. (2017). X-chromosome dosage as a modulator of pluripotency, signalling and differentiation? Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*. https://doi.org/10.1098/rstb.2016.0366.

Schulz, E.G., Meisig, J., Nakamura, T., Okamoto, I., Sieber, A., Picard, C., Borensztein, M., Saitou, M., Blüthgen, N., and Heard, E. (2014). The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. Cell Stem Cell *14*, 203–216.

Shaulian, E. (2010). AP-1—the Jun proteins: oncogenes or tumor suppressors in disguise? Cell. Signal. *22*, 894–899.

Shaulian, E., and Karin, M. (2002). AP-1 as a regulator of cell life and death. Nat. Cell Biol. *4*, E131–E136.

Shirane, K., Kurimoto, K., Yabuta, Y., Yamaji, M., Satoh, J., Ito, S., Watanabe, A., Hayashi, K., Saitou, M., and Sasaki, H. (2016). Global landscape and regulatory principles of DNA methylation reprogramming for germ cell specification by mouse pluripotent stem cells. Dev. Cell *39*, 87–103.

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.

Yagi, M., Kishigami, S., Tanaka, A., Semi, K., Mizutani, E., Wakayama, S., Wakayama, T., Yamamoto, T., and Yamada, Y. (2017). Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation. Nature *548*, 224–227.

Yang, S.-H., Kalkan, T., Morrisroe, C., Smith, A., and Sharrocks, A.D. (2012). A genome-wide RNAi screen reveals MAP kinase phosphatases as key ERK pathway regulators during embryonic stem cell differentiation. PLoS Genet. *8*, e1003112.

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–523.

Zvetkova, I., Apedaile, A., Ramsahoye, B., Mermoud, J.E., Crompton, L.A., John, R., Feil, R., and Brockdorff, N. (2005). Global hypomethylation of the genome in XX embryonic stem cells. Nat. Genet. *37*, 1274–1279.

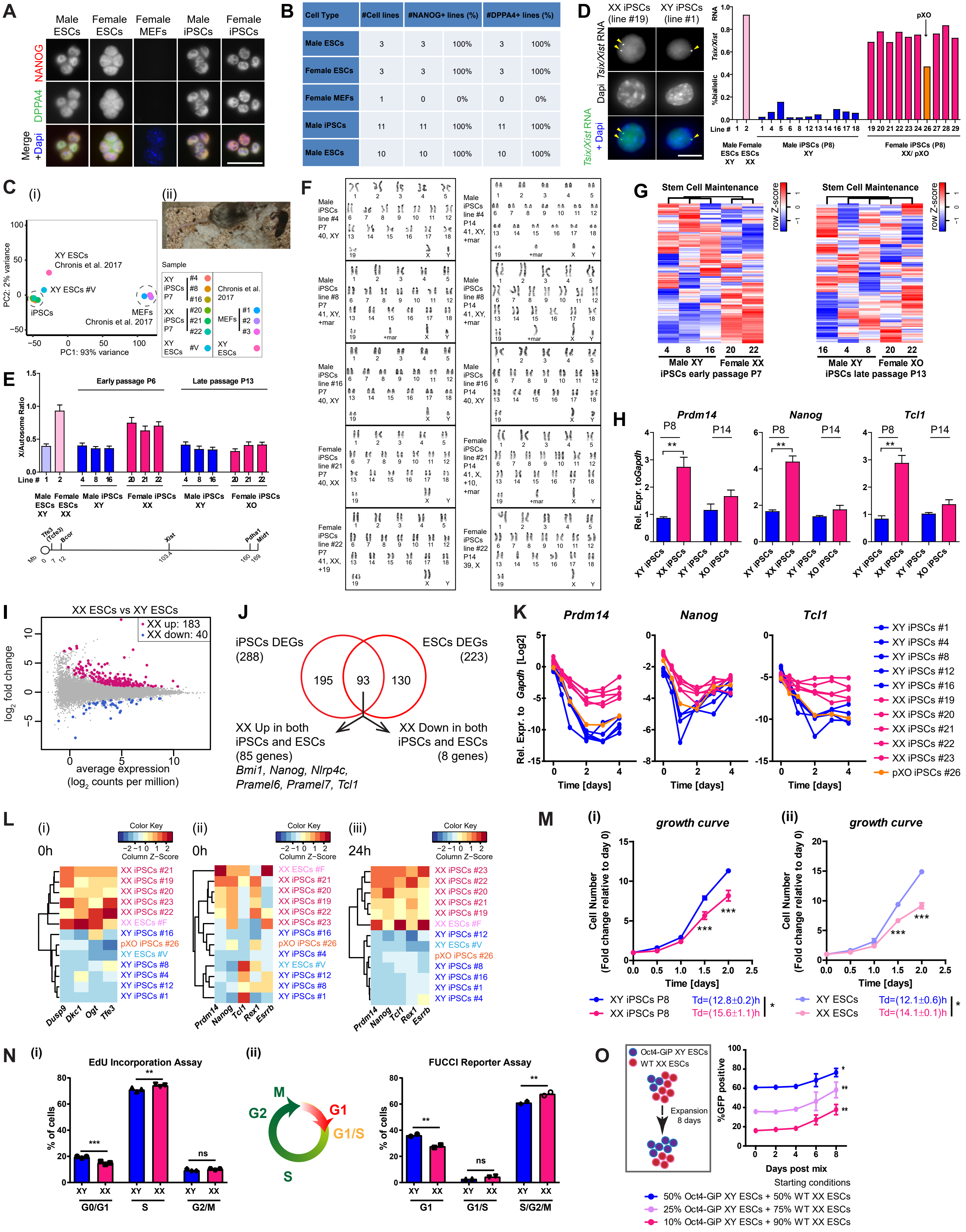
Stem Cell Reports, Volume *12*

Supplemental Information

X-Chromosome Dosage Modulates Multiple Molecular and Cellular Properties of Mouse Pluripotent Stem Cells Independently of Global DNA Methylation Levels

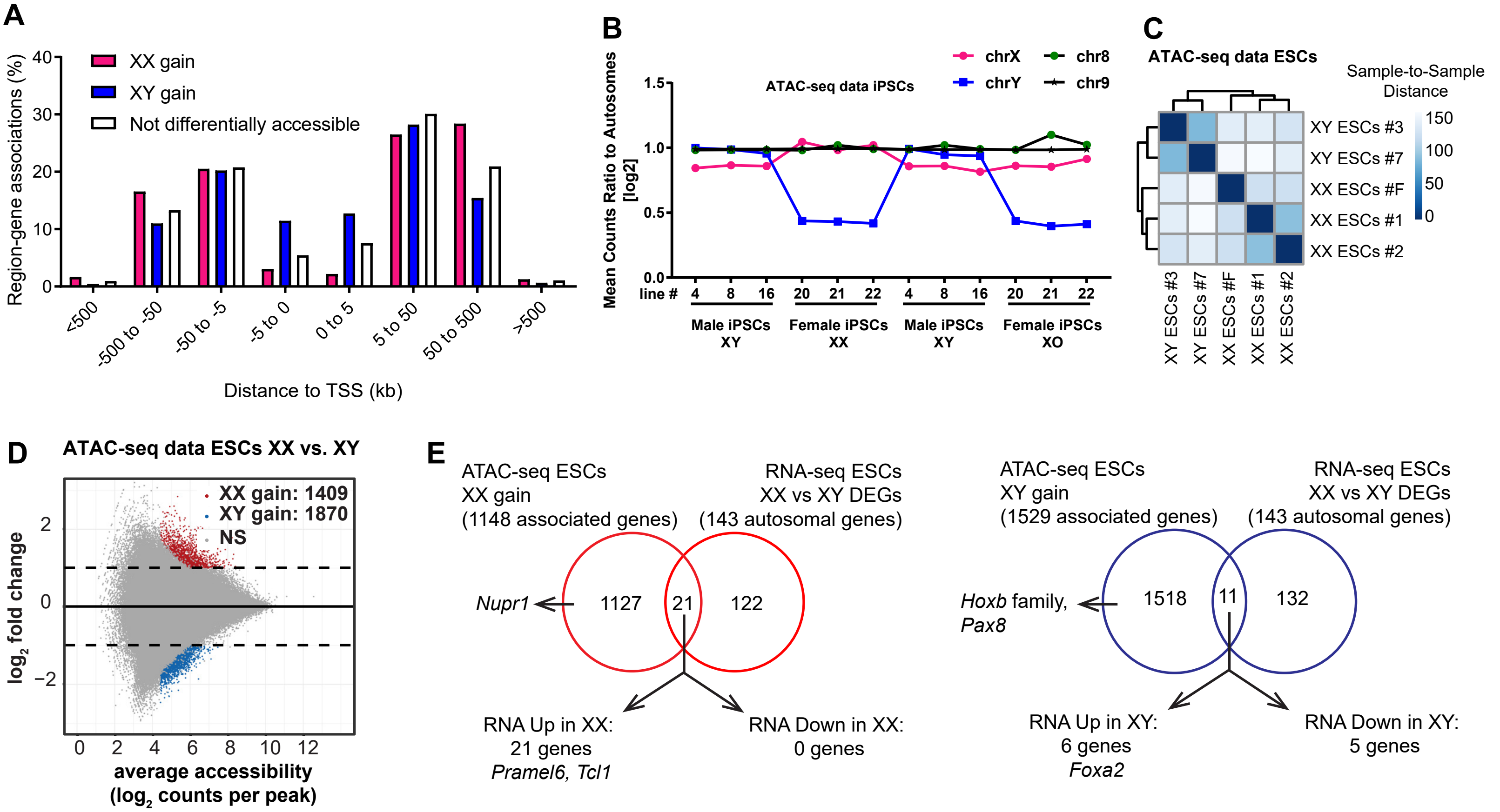
Juan Song, Adrian Janiszewski, Natalie De Geest, Lotte Vanheer, Irene Talon, Mouna El Bakkali, Taeho Oh, and Vincent Pasque

Figure S1



Cell Type	#Cell lines	#NANOG+ lines (%)		#DPPA4+ lines (%)	
Male ESCs	3	3	100%	3	100%
Female ESCs	3	3	100%	3	100%
Female MEFs	1	0	0%	0	0%
Male iPSCs	11	11	100%	11	100%
Male ESCs	10	10	100%	10	100%

Figure S2



F

ESCs XX gain: 1409 peaks

#	Motif	Name	p-value	% of targets with Motif
1		KLF5	1e-98	57.27%
2		KLF6	1e-68	44.78%
3		Sox2	1e-49	36.27%
4		Klf4	1e-44	22.43%
5		Sox3	1e-43	55.57%
6		Sox15	1e-43	39.89%
7		EKLF	1e-41	16.32%
8		Sox17	1e-41	30.16%
9		KLF3	1e-40	23.85%
10		Sox6	1e-39	50.75%
11		Rfx2	1e-39	8.73%
12		Sox10	1e-38	53.02%
13		RFX	1e-38	8.09%
14		KLF14	1e-36	56.35%
15		X-box	1e-32	9.01%
26		Zic3	1e-16	22.00%

G

ESCs XY gain: 1870 peaks

#	Motif	Name	p-value	% of targets with Motif
1		Fra1	1e-789	74.12%
2		Atf3	1e-783	78.29%
3		JunB	1e-763	72.83%
4		BATF	1e-746	16.03%
5		AP-1	1e-719	78.82%
6		Fosl2	1e-668	56.95%
7		Jun-AP1	1e-556	46.26%
8		Bach2	1e-246	27.86%
9		NF-E2	1e-57	7.70%
10		Pitx1:Ebox	1e-52	13.58%
11		RUNX-AML	1e-50	32.03%
12		Bach1	1e-49	6.90%
13		MafK	1e-48	16.10%
14		RUNX2	1e-47	37.86%

J

iPSCs XX and XY common regions: 282740 peaks

#	Motif	Name	p-value	% of targets with Motif
1		CTCF	1e-6495	11.02%
2		Fosl2	1e-5114	14.80%
3		BORIS	1e-5021	12.76%
4		Jun-AP1	1e-5012	11.80%
5		Fra1	1e-5010	20.54%
6		JunB	1e-4762	20.36%
7		Atf3	1e-4737	23.21%
8		AP-1	1e-4239	25.01%
9		BATF	1e-4191	22.86%
10		Bach2	1e-2741	8.79%
11		KLF5	1e-2351	39.80%
12		KLF3	1e-2119	18.46%
13		KLF6	1e-1866	32.00%
14		Sp5	1e-1835	30.36%

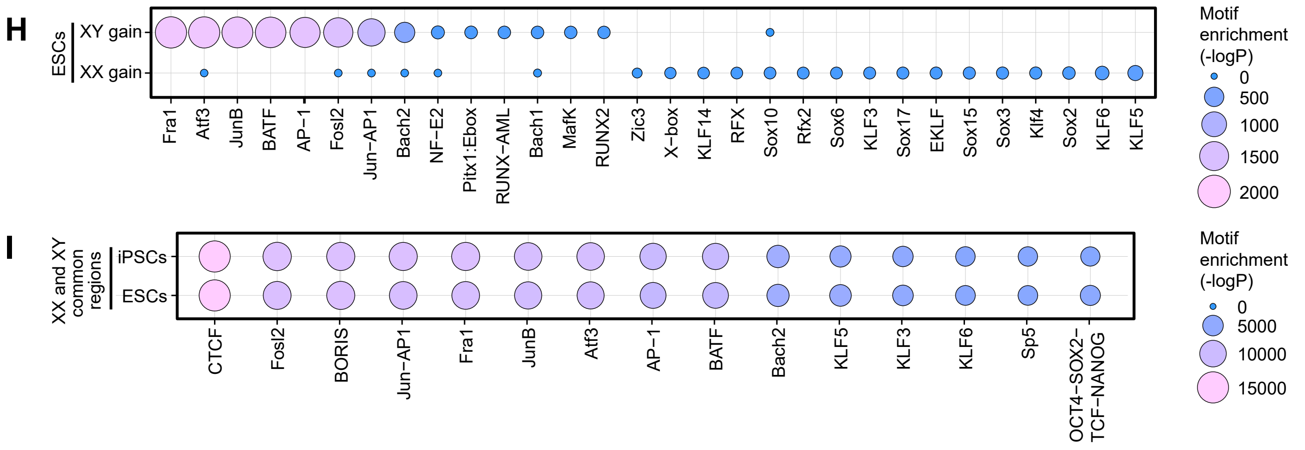


Figure S3

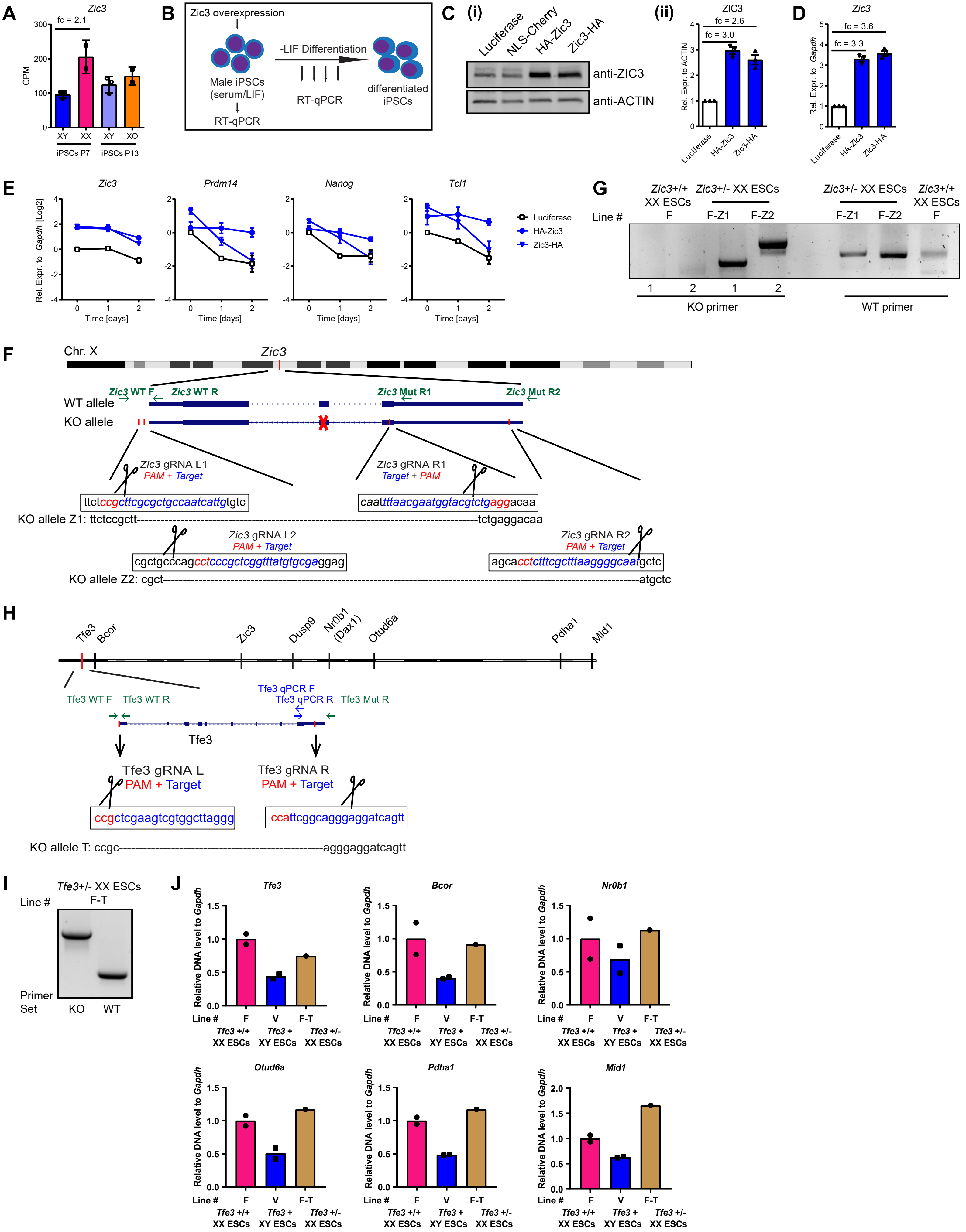

A
Zic3
fc = 2.1
CPM
300
200
100
0
XY
XX
XY
XO
iPSCs P7
iPSCs P13
B
Zic3 overexpression
Male iPSCs (serum/LIF)
RT-qPCR
-LIF Differentiation
RT-qPCR
differentiated iPSCs
C
(i)
Luciferase
NLS-Cherry
HA-Zic3
Zic3-HA
anti-ZIC3
anti-ACTIN
(ii)
ZIC3
fc = 2.6
fc = 3.0
Rel. Expr. to ACTIN
Luciferase
HA-Zic3
Zic3-HA
D
Zic3
fc = 3.6
fc = 3.3
Rel. Expr. to Gapdh
Luciferase
HA-Zic3
Zic3-HA
E
Zic3
Prdm14
Nanog
Tcl1
Rel. Expr. to Gapdh [Log2]
Time [days]
Luciferase
HA-Zic3
Zic3-HA
G
Zic3+/+ XX ESCs
Zic3+/- XX ESCs
Zic3+/- XX ESCs
Zic3+/+ XX ESCs
Line #
F
F-Z1
F-Z2
F-Z1
F-Z2
F
1
2
1
2
KO primer
WT primer
F
Chr. X
Zic3
Zic3 WT F
Zic3 WT R
Zic3 Mut R1
Zic3 Mut R2
WT allele
KO allele
Zic3 gRNA L1
PAM + Target
ttctccgcttcgcgctgccaatcattgtgtc
Zic3 gRNA R1
Target + PAM
caattttaacgaatggtacgtctgaggacaa
KO allele Z1: ttctccgctt--tctgaggacaa
Zic3 gRNA L2
PAM + Target
cgctgcccagcctcccgctcggtttatgtgcgaggag
Zic3 gRNA R2
PAM + Target
agcacctctttcgctttaaggggcaatgctc
KO allele Z2: cgct--atgctc
H
Tfe3
Bcor
Zic3
Dusp9
Nr0b1 (Dax1)
Otud6a
Pdha1
Mid1
Tfe3 WT F
Tfe3 WT R
Tfe3 qPCR F
Tfe3 qPCR R
Tfe3 Mut R
Tfe3
Tfe3 gRNA L
PAM + Target
ccgctcgaagtcgtggcttaggg
Tfe3 gRNA R
PAM + Target
ccattcggcagggaggatcagtt
KO allele T: ccgc--agggaggatcagtt
I
Tfe3+/- XX ESCs
Line #
F-T
Primer Set
KO
WT
J
Tfe3
Bcor
Nr0b1
Otud6a
Pdha1
Mid1
Relative DNA level to Gapdh
Line #
F
V
F-T
Tfe3 +/+ XX ESCs
Tfe3 + XY ESCs
Tfe3 +/- XX ESCs

Figure S4

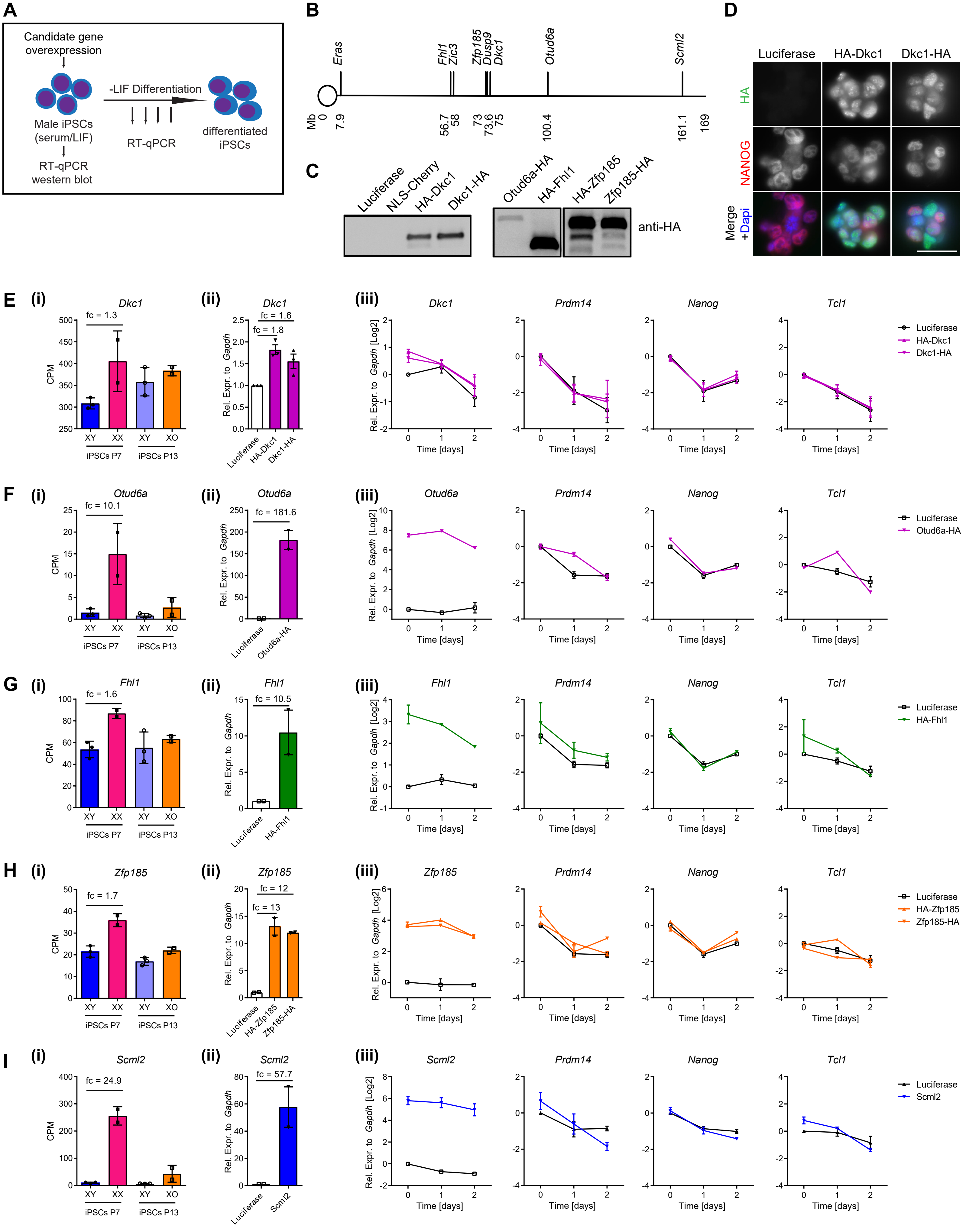

A
Candidate gene overexpression
Male iPSCs (serum/LIF)
RT-qPCR western blot
-LIF Differentiation
RT-qPCR
differentiated iPSCs
B
Eras
Fhl1
Zic3
Zfp185
Dusp9
Dkc1
Otud6a
Scml2
Mb
0
7.9
56.7
58
73
73.6
75
100.4
161.1
169
C
Luciferase
NLS-Cherry
HA-Dkc1
Dkc1-HA
Otud6a-HA
HA-Fhl1
HA-Zfp185
Zfp185-HA
anti-HA
D
Luciferase
HA-Dkc1
Dkc1-HA
HA
NANOG
Merge +Dapi
E (i)
Dkc1
fc = 1.3
CPM
XY
XX
XY
XO
iPSCs P7
iPSCs P13
(ii)
Dkc1
fc = 1.6
fc = 1.8
Rel. Expr. to Gapdh
Luciferase
HA-Dkc1
Dkc1-HA
(iii)
Dkc1
Prdm14
Nanog
Tcl1
Rel. Expr. to Gapdh [Log2]
Time [days]
Luciferase
HA-Dkc1
Dkc1-HA
F (i)
Otud6a
fc = 10.1
CPM
XY
XX
XY
XO
iPSCs P7
iPSCs P13
(ii)
Otud6a
fc = 181.6
Rel. Expr. to Gapdh
Luciferase
Otud6a-HA
(iii)
Otud6a
Prdm14
Nanog
Tcl1
Rel. Expr. to Gapdh [Log2]
Time [days]
Luciferase
Otud6a-HA
G (i)
Fhl1
fc = 1.6
CPM
XY
XX
XY
XO
iPSCs P7
iPSCs P13
(ii)
Fhl1
fc = 10.5
Rel. Expr. to Gapdh
Luciferase
HA-Fhl1
(iii)
Fhl1
Prdm14
Nanog
Tcl1
Rel. Expr. to Gapdh [Log2]
Time [days]
Luciferase
HA-Fhl1
H (i)
Zfp185
fc = 1.7
CPM
XY
XX
XY
XO
iPSCs P7
iPSCs P13
(ii)
Zfp185
fc = 12
fc = 13
Rel. Expr. to Gapdh
Luciferase
HA-Zfp185
Zfp185-HA
(iii)
Zfp185
Prdm14
Nanog
Tcl1
Rel. Expr. to Gapdh [Log2]
Time [days]
Luciferase
HA-Zfp185
Zfp185-HA
I (i)
Scml2
fc = 24.9
CPM
XY
XX
XY
XO
iPSCs P7
iPSCs P13
(ii)
Scml2
fc = 57.7
Rel. Expr. to Gapdh
Luciferase
Scml2
(iii)
Scml2
Prdm14
Nanog
Tcl1
Rel. Expr. to Gapdh [Log2]
Time [days]
Luciferase
Scml2

Figure S5

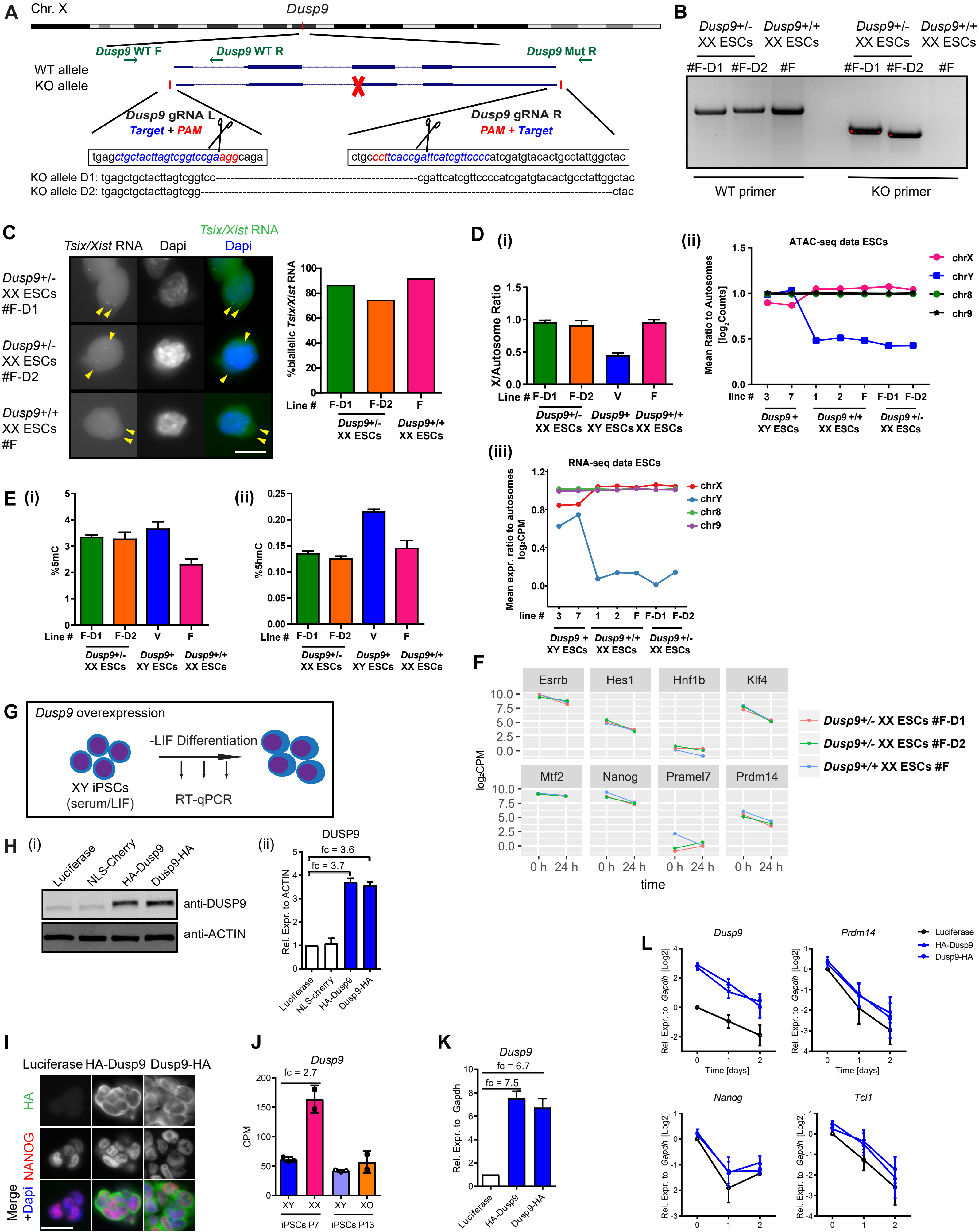

A
Chr. X
Dusp9
Dusp9 WT F
Dusp9 WT R
Dusp9 Mut R
WT allele
KO allele
Dusp9 gRNA L
Target + PAM
tgagctgctacttagtcggtccgaaggcaga
Dusp9 gRNA R
PAM + Target
ctgcccttcaccgattcatcgttccccatcgatgtacactgcctattggctac
KO allele D1: tgagctgctacttagtcggtcc---cgattcatcgttccccatcgatgtacactgcctattggctac
KO allele D2: tgagctgctacttagtcgg---ctac
B
Dusp9+/- XX ESCs
Dusp9+/+ XX ESCs
#F-D1
#F-D2
#F
WT primer
KO primer
C
Tsix/Xist RNA
Dapi
Tsix/Xist RNA Dapi
Dusp9+/- XX ESCs #F-D1
Dusp9+/- XX ESCs #F-D2
Dusp9+/+ XX ESCs #F
%biallelic Tsix/Xist RNA
0
20
40
60
80
100
Line #
F-D1
F-D2
F
Dusp9+/- XX ESCs
Dusp9+/+ XX ESCs
D
(i)
X/Autosome Ratio
0.0
0.5
1.0
1.5
Line #
F-D1
F-D2
V
F
Dusp9+/- XX ESCs
Dusp9+ XY ESCs
Dusp9+/+ XX ESCs
(ii)
ATAC-seq data ESCs
Mean Ratio to Autosomes [$\log_2$Counts]
0.0
0.5
1.0
1.5
chrX
chrY
chr8
chr9
line #
3
7
1
2
F
F-D1
F-D2
Dusp9+ XY ESCs
Dusp9+/+ XX ESCs
Dusp9+/- XX ESCs
(iii)
RNA-seq data ESCs
Mean expr. ratio to autosomes $\log_2$CPM
0.0
0.4
0.8
1.2
chrX
chrY
chr8
chr9
line #
3
7
1
2
F
F-D1
F-D2
Dusp9 + XY ESCs
Dusp9 +/+ XX ESCs
Dusp9 +/- XX ESCs
E
(i)
%5mC
0
1
2
3
4
5
Line #
F-D1
F-D2
V
F
Dusp9+/- XX ESCs
Dusp9+ XY ESCs
Dusp9+/+ XX ESCs
(ii)
%5hmC
0.00
0.05
0.10
0.15
0.20
0.25
Line #
F-D1
F-D2
V
F
Dusp9+/- XX ESCs
Dusp9+ XY ESCs
Dusp9+/+ XX ESCs
F
Esrrb
Hes1
Hnf1b
Klf4
Mtf2
Nanog
Pramel7
Prdm14
$\log_2$CPM
0.0
2.5
5.0
7.5
10.0
0 h
24 h
time
Dusp9+/- XX ESCs #F-D1
Dusp9+/- XX ESCs #F-D2
Dusp9+/+ XX ESCs #F
G
Dusp9 overexpression
-LIF Differentiation
XY iPSCs (serum/LIF)
RT-qPCR
H
(i)
Luciferase
NLS-Cherry
HA-Dusp9
Dusp9-HA
anti-DUSP9
anti-ACTIN
(ii)
DUSP9
fc = 3.6
fc = 3.7
Rel. Expr. to ACTIN
0
1
2
3
4
5
Luciferase
NLS-cherry
HA-Dusp9
Dusp9-HA
I
Luciferase
HA-Dusp9
Dusp9-HA
HA
NANOG
Merge +Dapi
J
Dusp9
fc = 2.7
CPM
0
50
100
150
200
XY
XX
XY
XO
iPSCs P7
iPSCs P13
K
Dusp9
fc = 6.7
fc = 7.5
Rel. Expr. to Gapdh
0
2
4
6
8
10
Luciferase
HA-Dusp9
Dusp9-HA
L
Dusp9
Prdm14
Nanog
Tcl1
Luciferase
HA-Dusp9
Dusp9-HA
Rel. Expr. to Gapdh [Log2]
Time [days]
0
1
2

Figure S6

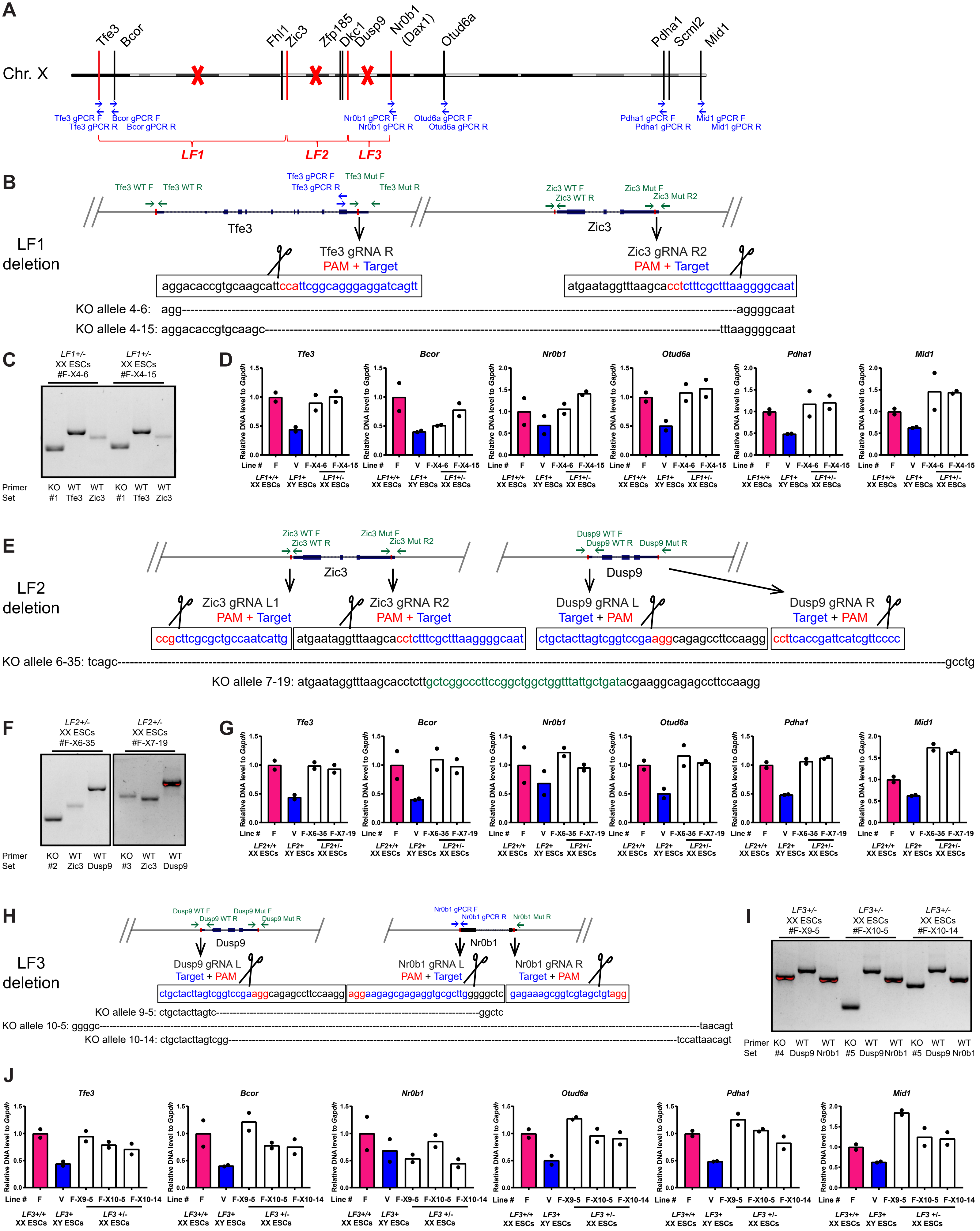
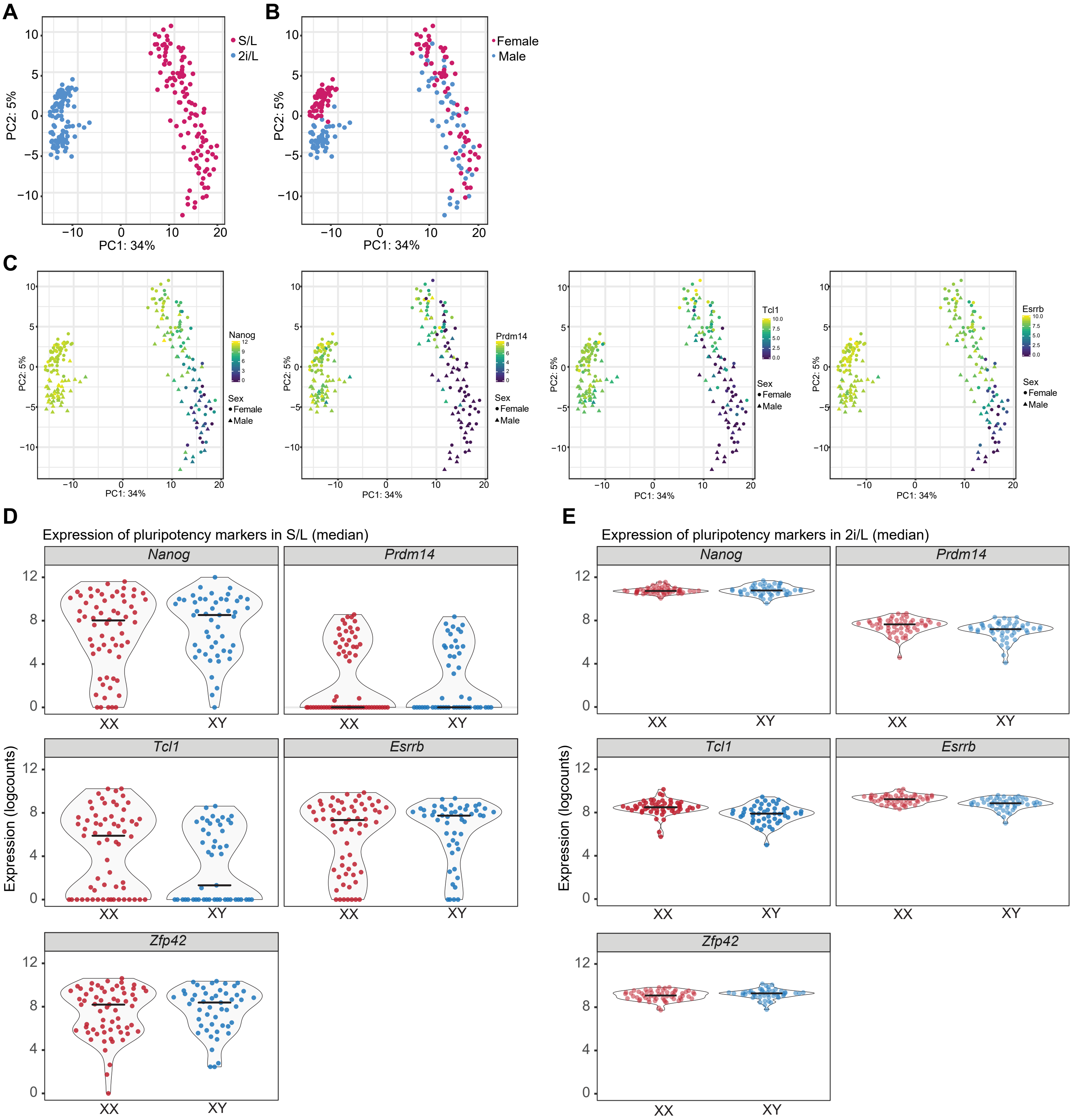

A
Chr. X
Tfe3
Bcor
Fhl1
Zic3
Zfp185
Dkc1
Dusp9
Nr0b1 (Dax1)
Otud6a
Pdha1
Scml2
Mid1
Tfe3 gPCR F
Tfe3 gPCR R
Bcor gPCR F
Bcor gPCR R
Nr0b1 gPCR F
Nr0b1 gPCR R
Otud6a gPCR F
Otud6a gPCR R
Pdha1 gPCR F
Pdha1 gPCR R
Mid1 gPCR F
Mid1 gPCR R
LF1
LF2
LF3
B
LF1 deletion
Tfe3 WT F
Tfe3 WT R
Tfe3 gPCR F
Tfe3 gPCR R
Tfe3 Mut F
Tfe3 Mut R
Tfe3
Zic3 WT F
Zic3 WT R
Zic3 Mut F
Zic3 Mut R2
Zic3
Tfe3 gRNA R
PAM + Target
aggacaccgtgcaagcattccattcggcagggaggatcagtt
Zic3 gRNA R2
PAM + Target
atgaataggtttaagcacctctttcgctttaaggggcaat
KO allele 4-6: agg--aggggcaat
KO allele 4-15: aggacaccgtgcaagc--tttaaggggcaat
C
LF1+/- XX ESCs #F-X4-6
LF1+/- XX ESCs #F-X4-15
Primer Set
KO #1
WT Tfe3
WT Zic3
KO #1
WT Tfe3
WT Zic3
D
Tfe3
Bcor
Nr0b1
Otud6a
Pdha1
Mid1
Relative DNA level to Gapdh
0.0
0.5
1.0
1.5
2.0
Line #
F
V
F-X4-6
F-X4-15
LF1+/+ XX ESCs
LF1+ XY ESCs
LF1+/- XX ESCs
E
LF2 deletion
Zic3 WT F
Zic3 WT R
Zic3 Mut F
Zic3 Mut R2
Zic3
Dusp9 WT F
Dusp9 WT R
Dusp9 Mut R
Dusp9
Zic3 gRNA L1
PAM + Target
ccgcttcgcgctgccaatcattg
Zic3 gRNA R2
PAM + Target
atgaataggtttaagcacctctttcgctttaaggggcaat
Dusp9 gRNA L
Target + PAM
ctgctacttagtcggtccgaaggcagagccttccaagg
Dusp9 gRNA R
Target + PAM
ccttcaccgattcatcgttcccc
KO allele 6-35: tcagc--gcctg
KO allele 7-19: atgaataggtttaagcacctcttgctcggcccttccggctggctggtttattgctgatacgaaggcagagccttccaagg
F
LF2+/- XX ESCs #F-X6-35
LF2+/- XX ESCs #F-X7-19
Primer Set
KO #2
WT Zic3
WT Dusp9
KO #3
WT Zic3
WT Dusp9
G
Tfe3
Bcor
Nr0b1
Otud6a
Pdha1
Mid1
Relative DNA level to Gapdh
0.0
0.5
1.0
1.5
2.0
Line #
F
V
F-X6-35
F-X7-19
LF2+/+ XX ESCs
LF2+ XY ESCs
LF2+/- XX ESCs
H
LF3 deletion
Dusp9 WT F
Dusp9 WT R
Dusp9 Mut F
Dusp9 Mut R
Dusp9
Nr0b1 gPCR F
Nr0b1 gPCR R
Nr0b1 Mut R
Nr0b1
Dusp9 gRNA L
Target + PAM
ctgctacttagtcggtccgaaggcagagccttccaagg
Nr0b1 gRNA L
PAM + Target
aggaagagcgagaggtgcgcttggggctc
Nr0b1 gRNA R
Target + PAM
gagaaagcggtcgtagctgtagg
KO allele 9-5: ctgctacttagtc---ggctc
KO allele 10-5: ggggc--taacagt
KO allele 10-14: ctgctacttagtcgg---tccattaacagt
I
LF3+/- XX ESCs #F-X9-5
LF3+/- XX ESCs #F-X10-5
LF3+/- XX ESCs #F-X10-14
Primer Set
KO #4
WT Dusp9
WT Nr0b1
KO #5
WT Dusp9
WT Nr0b1
KO #5
WT Dusp9
WT Nr0b1
J
Tfe3
Bcor
Nr0b1
Otud6a
Pdha1
Mid1
Relative DNA level to Gapdh
0.0
0.5
1.0
1.5
2.0
Line #
F
V
F-X9-5
F-X10-5
F-X10-14
LF3+/+ XX ESCs
LF3+ XY ESCs
LF3 +/- XX ESCs

Figure S7


A
PC1: 34%
PC2: 5%
S/L
2i/L
B
PC1: 34%
PC2: 5%
Female
Male
C
PC1: 34%
PC2: 5%
Nanog
Prdm14
Tcl1
Esrrb
Sex
Female
Male
D
Expression of pluripotency markers in S/L (median)
E
Expression of pluripotency markers in 2i/L (median)
Nanog
Prdm14
Tcl1
Esrrb
Zfp42
Expression (logcounts)
XX
XY

SUPPLEMENTAL FIGURES AND LEGENDS

Figure S1. X-dosage-specific differences in transcriptome, pluripotency exit and cell growth in mouse iPSCs. Related to Figure 1.

(A) Immunofluorescence analysis for NANOG and DPPA4 in iPSCs and ESCs grown in 2i/L. MEFs served as negative control. Representative images for NANOG (Red), DPPA4 (Green) and Dapi (Blue, nuclei counterstaining) are shown. Scale bar, 50 µm.
(B) Summary of ESC and iPSC lines used in this study as well as NANOG and DPPA4 protein expression analysis in 2i/L and in S/L.
(C) (i) PCA of our own RNA-seq data for XX and XY iPSCs, XY ESCs (V6.5), as well as published RNA-seq for XY ESCs (V6.5, (Chronis et al., 2017)) and somatic MEFs. (ii) Representative image of adult male chimeric mice (left) and germline transmitted pup (right) produced from XY ESCs (V6.5).
(D) RNA FISH analysis for *Tsix/Xist* expression in iPSCs grown in 2i/L. Representative images for *Tsix/Xist* RNA (Green) and Dapi (Blue, nuclei counterstaining) are shown. Yellow arrowheads point to *Tsix/Xist* transcriptional sites. Also see (Pasque et al., 2018). Scale bar, 10 µm. Right: Quantification of *Tsix/Xist* RNA FISH signal from Figure S1C and for all lines, plotted as the proportion of cells with biallelic *Tsix/Xist* signal (= number of cells with biallelic *Tsix/Xist* expression / number of cells with biallelic or monoallelic expression). The number of counted nuclei is > 50 per cell line.
(E) qPCR analysis for X-chromosome DNA copy number in both early passage and late passage iPSCs grown in S/L. X copy number are presented as the average ratio of gDNA quantities for four X-linked genes (*Tfe3, Bcor, Pdha1, and Mid1,* locations in X-chromosome shown in lower panel) to gDNA quantities for autosomal gene *Gapdh.*
(F) Karyotype analysis of iPSC lines grown in S/L. + mar denotes chromosomal fragment of unknown origin.
(G) Unsupervised hierarchical clustering of stem-cell maintenance gene expression in XX, XY and XO iPSCs.
(H) RT-qPCR analysis for pluripotency-associated gene expression in XX, XY and XO iPSCs grown in S/L. The expression values are represented as averages (±SEM) of XY, XX and XO iPSC lines (three different lines each) in early passage (P8) and late passage (P14), respectively. Statistical significance was analysed using the unpaired, two-tailed t-test (**p<0.01).
(I) DEG analysis, identifying clear differences between XX and XY ESCs.
(J) Venn diagrams showing the overlap between XX vs XY iPSCs DEGs and XX vs XY ESCs DEGs. Despite difference in genetic background between iPSCs and ESCs, 93 genes are differentially expressed between XX and XY cells both for iPSCs and ESCs.
(K) RT-qPCR for pluripotency-associated genes for individual XY (blue), XX (magenta) and pXO (orange, line 26) iPSC lines undergoing EpiLC differentiation.
(L) Unsupervised hierarchical clustering of XX, XY and XO iPSC and ESC lines based on expression of X-linked genes in undifferentiated state (i), and pluripotency-associated genes in undifferentiated state (ii) and 24h of LIF withdrawal differentiated state (iii).
(M) Growth curves and doubling times of XY and XX iPSCs (i) and ESCs (ii) in 2i/LIF condition. Cell counts were obtained at different time points, as indicated. Cell counts are shown as fold changes relative to 0h and represent the averages cell count (±SEM) over three cell lines each for XY and XX iPSCs (n=1, left panel) at early passage (P8) and over two lines each for two XY and XX ESCs (n=1, right panel). Growth curve: *p* value <0.001 (***) between XY and XX cell lines, by two-way repeated-measures ANOVA with Bonferroni posttests. Td: **p*<0.05, XY lines vs XX lines, by unpaired two-tailed t-test.
(N) (i) EdU incorporation in combination with DNA staining using PI shows a higher number of XX iPSCs residing in S phase, and a lower number of XX iPSCs residing in G1 phase when compared to XY iPSCs. Results show the averages (±SEM) over three XX female iPSC lines and three XY male iPSC lines (n=1). Significance was tested using one-way ANOVA with Sidak's multiple comparisons test (**p<0.01, ***p<0.001). (ii) FUCCI reporter expression shows the number of XX female ESCs activating the G1 phase reporter is lower than that of XY male ESCs, and the number of XX female ESCs activating the S/G2/M phase reporter is higher than that of XY male ESCs. Results show the averages (±SEM) over two XX female and two XY male FUCCI reporter ESC lines (n=1). Significance was tested using one-way ANOVA with Bonferroni's multiple comparisons test (**p<0.01, ***p<0.001).
(O) XY and XX ESCs competition assay. WT XX ESCs were mixed with GFP-labelled XY ESCs (Oct4-GiP) in different ratios (left panel, scheme of the experiment) and the proportion of GFP-labeled cells in the culture was

measured over time (n=2), as indicated. *p<0.05, **p <0.01, by two-way repeated-measures ANOVA with Dunnett's multiple comparisons test compared to day 0.

Figure S2. Influence of X-dosage on the chromatin regulatory landscape of mouse iPSCs and ESCs. Related to Figure 2 and Figure 3.

(A) Distance to closest TSSs of “Not differentially accessible”, “XX gain” and “XY gain” ATAC-seq regions in iPSCs.
(B) Mean read count ratio to autosomes showing which lines have increased/decreased ATAC-seq reads on the Y chromosome, X chromosome, chromosome 8 and chromosome 9. This analysis confirms the higher abundance of DNA sequence reads coming from the Y chromosome in XY lines, and from the X chromosome in XX lines.
(C) Sample-to-sample distance heatmap showing the Euclidean distances (calculated from the rlog transformed counts, DESeq2) between ESCs samples. Samples cluster by X-dosage (XX vs XY).
(D) Differential chromatin accessibility analysis between XX and XY ESCs. Log2 fold change (XX/XY) in reads per accessible region are plotted against the mean reads per ATAC-seq peak. Thousands of open chromatin regions that more open in XX ESCs or in XY ESCs (|log2fold|>=1, false discovery rate (FDR)<=0.05).
(E) Venn diagrams showing the overlap between genes nearest to the “XX gain” or "XY gain" regions defined in (D) and the DEGs between XX and XY ESCs (DEGs= $|\log_2 \text{fold}| >= \log_2 1.5$, FDR<=0.05).
(F) Transcription factor motifs enriched in chromatin regions more open in XX ESCs.
(G) Transcription factor motifs enriched in chromatin regions more open in XY ESCs.
(H) Overview of motif enrichment in XX gain and XY gain chromatin regions in ESCs.
(I) Overview of motif enrichment in open chromatin not differentially accessbilie between XX and XY cells in both ESCs and iPSCs.
(J) Transcription factor motifs enriched in open chromatin not differentially accessbilie between XX and XY iPSCs.

Figure S3. Effects of Zic3 overexpression and Zic3/ Tfe3 heterozygous deletion in XX ESCs on pluripotency exit. Related to Figure 4.

(A) Expression of *Zic3* in XY, XX and XO iPSCs showing 2.1 fold increased *Zic3* dosage in XX iPSCs over XY and XO iPSCs as assessed by RNA-seq.
(B) Scheme of *Zic3* overexpression in XY iPSCs, followed by characterization and LIF withdrawal differentiation.
(C) (i) Western blot analysis for ZIC3 and ACTIN in XY iPSCs after HA-Zic3 or Zic3-HA overexpression. (ii) Quantification using ACTIN as loading control. ZIC3 protein values are represented as averages (±SEM) of three independent experiments.
(D) RT-qPCR of *Zic3* in HA-tagged Zic3 overexpressing (3.3-3.6 fold) XY iPSCs. XY iPSCs overexpressing Luciferase served as negative control. Results are presented as averages (±SEM) of three independent experiments.
(E) RT-qPCR analysis for *Zic3*, *Prdm14*, *Nanog* and *Tcl1* expression during LIF withdrawal in XY iPSCs overexpressing HA-Zic3, Zic3-HA or Luciferase control. Results are presented as averages (±SEM) of three independent experiments (n=3).
(F) Scheme of heterozygous *Zic3* deletion strategy in XX ESCs. The sequences of the gRNAs used to deleted *Zic3* are shown in blue with PAM sequences in red. Two independent *Zic3*+/- XX ESC lines were derived. The sequences of the knockout (KO) alleles were obtained by Sanger sequencing. Red line shows the location of the gRNAs. Green arrows shows the locations of the primers for genotyping PCR.
(G) Genotyping of *Zic3* heterozygous deleted XX ESC lines for both WT and KO alleles. The parental *Zic3+/+* ESCs were used as positive control for the WT allele and as negative control for the KO allele.
(H) as in (F) but for heterozygous *Tfe3* deletion strategy in XX ESCs.
(I) Genotyping of *Tfe3* heterozygous deleted XX ESC lines for both WT and KO alleles.
(J) qPCR analysis for X-chromosome DNA copy number. X copy number are presented as the relative gDNA quantities for six X-linked genes (*Tfe3, Bcor, Nr0b1, Otud6a, Pdha1, and Mid1,* locations in X-chromosome shown in Figure S3H) to gDNA quantities for autosomal gene *Gapdh*.

Figure S4. Overexpression of X-linked candidates Dkc1, Otud6a, Fhl1, Zfp185 and Scml2 has no effect on the pluripotency exit kinetics of XY iPSCs. Related to Figure 4.

(A) Scheme for candidate X-linked gene overexpression in XY iPSCs, followed by LIF withdrawal.
(B) Map of the X-chromosome showing candidate X-linked genes.

(C) Western blot analysis for HA-tagged DKC1, OTUD6A, FHL1, ZFP185 using an anti-HA antibody. ACTIN was used as a loading control. Representative images are shown.
(D) Immunofluorescence analysis for NANOG and HA in stable XY iPSC lines overexpressing HA-tagged *Dkc1*. The stable XY iPSC line overexpressing Luciferase was used as a negative control. Representative images for NANOG (Red), HA (Green) and Dapi (Blue, nuclei counterstaining) are shown. Scale bar, 50 μm.
(E-I) (i) Expression of X-linked candidates *Dkc1, Otud6a, Fhl1, Zfp185 and Scml2* in XY, XX and XO iPSCs showing fold change (fc) in XX iPSCs over XY iPSCs as assessed by RNA-seq. **(ii)** RT-qPCR of X-linked candidates expression in XY iPSC lines with ectopic expression of HA-tagged *Dkc1* (E), HA-tagged *Otud6a* (F), HA-tagged *Fhl1* (G), HA-tagged *Zfp185* (H) or *Scml2* (I). XY iPSCs overexpressing Luciferase served as negative control. **(iii)** RT-qPCR for pluripotency-associated genes *Prdm14*, *Nanog* and *Tcl1*, and the respective X-linked candidate genes in XY iPSC lines stably overexpressing the respective HA-tagged X-linked candidates and subjected to LIF withdrawal. Results are presented as averages (±SEM) of three (E) or two (F-J) independent experiments, which are not statistically significant by two-way repeated-measures ANOVA with Bonferroni posttests.

Figure S5. Characterization of Dusp9 heterozygous mutant XX female ESCs. Related to Figure 5.

(A) Scheme of heterozygous *Dusp9* deletion in XX ESCs. The gRNAs sequences used to delete the *Dusp9* gene are shown in blue with PAM sequences in red. Two independent *Dusp9+/-* ESC lines were derived. The sequences of the KO alleles were obtained by DNA Sanger sequencing. Red line shows the location of the gRNAs. Green arrows shows the locations of the primers for genotyping PCR.
(B) Genotyping *Dusp9*+/- ESC lines for both the WT and the KO allele. The parental XX ESC line was used as a positive control for the WT allele PCR and as a negative control for the KO allele PCR.
(C) RNA FISH analysis for *Tsix/Xist* expression in the two independent *Dusp9+/-* ESC lines and their parental XX ESC line. Representative images for *Tsix/Xist* RNA (Green) and Dapi (Blue, nuclei counterstaining) are shown. Yellow arrowheads point to *Tsix/Xist* transcriptional sites. Scale bar, 10 μm. Right: Quantification of Figure S5C, plotted as the proportion of cells with biallelic *Tsix/Xist* signal (= number of cells with biallelic *Tsix/Xist* expression / number of cells with biallelic or monoallelic expression). Numbers of counted nuclei > 50 per cell line.
(D) (i) qPCR analysis for X-chromosome DNA copy number. X copy number are presented as the average ratio of gDNA quantities for four X-linked genes (Tfe3, Bcor, Pdha1, and Mid1) to gDNA quantities for autosomal gene Gapdh. Results are presented as averages (±SEM) of the same lines in two independent qPCR experiments (n=2). (ii) Mean read count ratio to autosomes showing which lines have increased/decreased ATAC-seq reads on the Y chromosome, X chromosome, chromosome 8 and chromosome 9. This analysis confirms the higher abundance of DNA sequence reads coming from the Y chromosome in XY lines, and from the X chromosome in XX lines. (iii) Mean expression ratio to autosomes for sex chromosomes and chromosomes 8 and 9.
(E) DNA methylation analysis of *Dusp9*+/- ESCs, *Dusp9*+/+ ESCs and XY ESCs by mass spectrometry. (i) 5mC, (ii) 5hmC.
(F) RNA-seq analysis of pluripotency associated gene expression during pluripotency exit in Dusp9+/- and Dusp9+/+ ESCs. The delay in pluripotency gene downregulation is maintained.
(G) Scheme of *Dusp9* overexpression in male iPSCs grown in S/L, followed by characterization and LIF withdrawal.
(H) (i) Western blot analysis for DUSP9 in XY iPSC lines with ectopic HA-tagged DUSP9 expression. ACTIN was used as a loading control. (ii) quantification using ACTIN as loading control. DUSP9 protein values are represented as averages (±SEM) of three independent experiments (n=3).
(I) Immunofluorescence analysis for NANOG and HA-tagged DUSP9 in stable male iPSCs with HA-tagged DUSP9 overexpression. The stable male iPSC line overexpressing Luciferase was used as a negative control. Representative images for NANOG (Red), HA (Green) and Dapi (Blue, nuclei counterstaining) are shown. Scale bar, 50 μm.
(J) Expression of *Dusp9* in XX, XY and XO iPSCs showing 2.7 fold increased *Dusp9* dosage in XX iPSCs over XY and XO iPSCs as assessed by RNA-seq.
(K) qRT-PCR for *Dusp9* in XY iPSCs overexpressing HA-Dusp9, Dusp9-HA and control Luciferase.
(L) qRT-PCR for *Dusp9* and pluripotency-associated genes *Prdm14*, *Nanog* and *Tcl1* in XY iPSCs overexpressing *Dusp9* following LIF withdrawal. The stable XY iPSC line overexpressing Luciferase served as a negative control. Results are presented as averages (±SEM) of three independent experiments.

Figure S6. Large fragment heterozygous deletion of X chromosome in XX ESCs. Related to Figure 6.

(A) Scheme of the three large fragment (LF) heterozygous deletions in XX female ESCs.

(B) Scheme of heterozygous LF1 deletion in XX ESCs. The gRNAs sequences are shown in blue with PAM sequences in red. Two independent *LF1*+/- ESC lines were derived. The sequences of the KO alleles were obtained by DNA Sanger sequencing. Red line shows the location of the gRNAs. Green arrows shows the locations of the primers for genotyping PCR used in (C).
(C) Genotyping *LF1*+/- XX ESC lines for both the WT and the KO allele.
(D) qPCR analysis for X-chromosome DNA copy number. X copy number are presented as the relative gDNA quantities for six X-linked genes (Tfe3, Bcor, Pdha1, and Mid1) to gDNA quantities for autosomal gene Gapdh. Results are presented as averages (±SEM) of the same lines in two independent qPCR experiments (n=2).
(E-G) As in (B-D) for heterozygous LF2 deletion in XX ESCs. Two independent *LF2*+/- ESC lines were derived and validated.
(H-J) As in (B-D) for heterozygous LF3 deletion in XX ESCs. Three independent *LF3*+/- ESC lines were derived and validated.

Figure S7. Single cell RNA-seq analysis of XX female and XY male ESCs in S/L and 2i/L.

(A) PCA analysis of XX female and male XY ESCs S/L and 2i/L single cell RNA-seq data from (Chen et al., 2016a). Cells grown in S/L are shown in magenta, those grown in 2i/L are shown in blue.
(B) Same analysis as in (A), showing XX female ESCs in magenta and XY male ESCs in blue.
(C) Expression of *Nanog*, *Prdm14*, *Esrrb* and *Tcl1* projected onto the PCA shown in (A) and (B).
(D) Expression (log read count) of pluripotency associated genes *Nanog*, *Prdm14*, *Tcl1*, *Esrrb* and *Zfp42* in single XX female and XY male ESCs grown in S/L. Violin plots indicate the distribution of single cells where each dot is a cell. Black lines indicate median gene expression.
(E) As in (D) for the same cells grown in 2i/L.

SUPPLEMENTAL TABLES

Table S1. Summary of cell lines used in this study.

Table S2. Differentially expressed genes between XX and XY iPSCs (ESCs). Related to Figure 1F and S1I respectively.

Table S3. Differentially accessible regions between XX and XY iPSCs (ESCs), and the association with differential gene expression. Related to Figure 2B/C/F, S2D/E.

Table S4 Expression level of X-linked candidate genes in XX and XY iPSCs and ESCs. Related to Figure 4A, 5, S3-5.

Table S5. Primer sequences.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Mice and reprogramming

MEFs were isolated from individual E14.5 mouse embryos obtained from a cross between wild type (WT) C57BL/6 and homozygous Rosa26:M2rtTA, TetO-OSKM mice (Carey et al., 2010). Individual embryos were genotyped for sex using *Ube1* as previously described (See S9 Table for primer sequence) (Pasque et al., 2014) using homemade Taq DNA Polymerase and grown in MEF medium [DMEM (Gibco, 41966-052) supplemented with 10% (v/v) fetal bovine serum (FBS, Gibco, 10270-106), 1% (v/v) penicillin/streptomycin (P/S, Gibco, 15140-122), 1% (v/v) GlutaMAX (Gibco, 35050-061), 1% (v/v) non-essential amino acids (NEAA, Gibco, 11140-050), and 0.8% (v/v) beta-mercaptoethanol (Sigma, M7522)]. Reprogramming was induced by doxycycline (final 2 µg/ml) in mouse ESC medium [KnockOut DMEM (Gibco, 10829-018) supplemented with 15% FBS, 1% (v/v) P/S, 1% (v/v) GlutaMAX, 1% (v/v) NEAA, 0.8% (v/v) beta-mercaptoethanol, and mouse LIF] in the presence of ascorbic acid (final 50 µg/ml). Individual colonies were picked at day 16 onto irradiated male feeders in ESC medium without doxycycline or ascorbic acid and expanded for three passages, eventually obtaining 10 female iPSC lines (lines 1, 4, 5, 6, 8, 12, 13, 14, 16, 17, 18) and 11 male iPSC lines (lines 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29) at passage (P) 4 (S1 Fig B). iPSC lines 1, 4, 8, 12, 14, 16, 20, 21, 22, 24, 26 and 28 were also used in another study (Pasque et al., 2018) (S1 Table). Mus/Cas ESCs were isolated from E3.5 embryos resulting from a cross between Cast/Eij males and C57B6/J females, as described (Czechanski et al., 2014). All animal work carried out in this study is covered by a project license approved by the KU Leuven Animal Ethics Committee.

Cell lines and culture

XY and XX Mus/Cas ESCs were newly derived in our lab and also obtained from the Deng laboratory (Chen et al., 2016a). XY ESCs (V6.5) and XX ESCs (F1-2-1) were obtained from the Plath laboratory. GFP-labelled (Oct4-GiP) XY ESCs were previously described (Ying et al., 2002). ESCs and iPSCs (male iPSC line 4, 8, 16; female iPSC line 20, 21, 22) were expanded on top of male WT feeders in mouse ESC medium (S/L condition), eventually early passage cells (iPSCs: P6-P8) and late passage cells (iPSCs: P13-P14) were used for further experiments. ESCs and iPSCs (male iPSC lines 1, 4, 8, 12, 16; female iPSC lines 19, 20, 21, 22, 23, 26) were adapted to 2i/LIF, where cells grown on feeders in S/L condition (iPSCs: P4) were switched to new tissue culture dishes precoated with gelatin (from porcine skin, 0.1% g/v final, Sigma, G2500) without feeders in 2i/LIF medium [N2B27 basal medium (Neurobasal medium (50% v/v final, Gibco, 21103-049) and DMEM/F-12 medium (50% final, Gibco, 11320-074) supplemented with L-Glutamine (1.25 mM final, Gibco, 25030081), NDiff Neuro2 supplement (1x final, Millipore, SCM012), B27 supplement (1x final, Gibco, 17504-044), 0.8% (v/v) beta mercapto-ethanol, and 1% (v/v) P/S) supplemented with 0.35% (g/v) Bovine Serum Albumin (BSA, Sigma, A7979), homemade mouse LIF, GSK3 inhibitor CHIR-99021 (3 µM final, Axon Medchem, Axon 1386) and MEK inhibitor PD0325901 (1 µM final, Axon Medchem, Axon 1408)] for four passages.

Plasmids Constructs

The full-length mouse cDNAs of *Dusp9, Zic3, Dkc1, Otud6a, Fhl1, Zfp185*, and *Luciferase* (from pGL2-Basic Promage, E1641), *NLS-cherry* was cloned into pENTR vectors (Invitrogen, K240020) with either a C-terminal or a N-terminal HA tag, or no tag, and recombined into pPB-CAG-Dest-pA-pgk-bsd (PB-DEST-BSD) destination vectors. The PB-Scml2-BSD plasmid was obtained by recombining the pDONR221-Scml2 plasmid (Branco et al., 2016) into PB-DEST-BSD. Guide RNAs (gRNAs) were cloned into SapI digested pZB-sg3 (Fulco et al., 2016). All gRNAs sequences are included in S9 Table , S3 Fig F/H, S5 Fig A and S6 Fig B/E/H. All constructs were verified by DNA Sanger sequencing.

Generation of stable male iPSCs overexpressing X-linked candidate genes

Male iPSCs (line 4, P5, grown on feeders in S/L conditions) were feeder-depleted before seeding in six-well plates precoated with 0.1% gelatin in S/L medium at a density of 650,000 cells per well, which were co-transfected with 1 ug of PB expression constructs encoding candidate genes and 3 ug of pCAGP Base (Silva et al., 2009) using 10 µl Lipofectamine 2000 (Invitrogen, 11668027). Transfected cells were selected with 20 µg/mL blasticidin (Fisher BioReagents, BP2647100) supplemented to the medium for two days starting from 24h after transfection and maintained with 5 µg/mL blasticidin thereafter.

Generation of XX female ESC lines with heterozygous deletions of X-linked candidate genes

2000,000 female F1-2-1 ESCs (P19, grown on feeders in S/L condition) were resuspended in 1 ml of S/L medium and co-transfected with 2 ug of a plasmid expressing Cas9 under a CAG promoter and 1 ug of 2 plasmids (pZB-sg3 (Fulco et al., 2016)) containing gRNAs (S9 Table) using 10 µl Lipofectamine 2000 (Invitrogen, 11668027) (S3 Fig F/H, S5 Fig A and S6 Fig B/E/H) for one hour before plating on 4-drug resistant (DR4) feeders. Transfected cells were selected with 2 µg/mL puromycin (Fisher BioReagents, BP2647100) on DR4 feeders in ESC medium for two days starting from 24h after transfection, and expanded at low density on WT feeders in 10cm dishes. Individual colonies were picked onto WT feeders, expanded for another two passages and genotyped for both WT and mutant alleles (primers in Table S8). WT and mutant alleles were further verified by DNA Sanger sequencing.

Differentiation

To induce differentiation towards epiblast-like cells (EpiLCs), ESCs and iPSCs (male lines: 1, 4, 8, 12, 16; female lines: 19, 20, 21, 22, 23, 26; P8), which had been adapted to 2i/LIF conditions for 4 passages, were plated in N2B27 basal medium supplemented with 10 ng/ml Fibroblast Growth Factor-basic (Fgf2, Peprotech, 100-18C) and 20 ng/ml Activin A (ActA, Peprotech, 120-14E) on Fibronectin (5 ug/10 cm^2, Millipore, FC010-5MG)-coated tissue culture plates at a cell density of $8*10^4$ cells/cm^2 for four days, during which medium was refreshed daily and cells were harvested at different time points (0h, 12h, 1 day, 2 days, 3 days and 4 days), as previously described (Schulz et al., 2014). ESCs (WT female and male ESCs, *Dusp9+/-* ESCs, and *Zic3+/-* ESCs) and iPSCs (male iPSC lines 4, 8, 16; female iPSC lines 20, 21, 22; both early and late passages) grown in S/L condition were differentiated in the absence of feeders by LIF withdrawal (similar as mouse ESC medium but with 10% FBS and without LIF) at a cell density of $4*10^4$ cells/cm^2 for two days, during which medium was refreshed daily and cells were harvested at different time points (0h, 24h and 48h), as previously described (Schulz et al., 2014). Likewise, male iPSC lines overexpressing X-linked genes were differentiated by LIF withdrawal with 5 µg/mL blasticidin in the absence of feeders.

Clonal assays

iPSCs were subjected to LIF withdrawal, and 5000 cells were sorted onto feeders in S/L in each well of a 12-well plate in triplicate at 0h, 24h, 48h, 72h of LIF withdrawal. The next day cultures were switched to 2i/L. Alkaline Phosphatase (AP) staining was carried out 5 days after replating, using the VECTOR Red Alkaline Phosphatase (Red AP) Substrate Kit (VECTOR, SK-5100). Imaging was carried out using an Nikon Eclipse Ti2 Microscope equipped with an Nikon DS-Qi2 camera. AP positive colonies in each well of the 12-well plates were counted using NIS Element Auto Measurement.

Cell growth assay

ESCs and iPSCs were plated in 24-well plates at a cell density of $4*10^4$ cells/cm^2 for two days, during which medium was refreshed daily and cells were counted at different time points (0h, 12h, 24h, 36h and 48h). The cell numbers are presented as fold changes relative to cell numbers at 0h.

FUCCI cell-cycle reporter assay

We generated XY and XX Mus/Cas ESCs expressing the FUCCI fluorescent reporters together with an H2B nuclear marker by co-transfecting of the WT XY and XX ESCs with PB expression constructs including PB-mCherry-hCdt1-BSD, PB-mVenus-hGeminin-PURO and PB-mCerulean-H2B-NEO (Waisman et al., 2017) and pCAGP Base (Silva et al., 2009) using Lipofectamine 2000. Transfected cells were selected with 20 µg/mL blasticidin, 2 µg/mL puromycin and 100 µg/mL G418 supplemented to the medium for two days starting from 24h after transfection and maintained with 5 µg/mL blasticidin, 1 µg/mL puromycin and 50 ug/ml G418 thereafter. The BD FACSMelody cell sorter were used to analyze the FUCCI ESCs at different phases of the cell cycle.

EdU incorporation assay

ESCs and iPSCs were pulse-labeled with the Click-iT EdU Alexa Fluor 647 Flow Cytometry Assay Kit (Invitrogen, C10424) according to the manufacturer’s instructions. Briefly, cells were incubated with 10µM 5-ethynyl-2′-deoxyuridine (EdU) for 45 min at 37°C. Then, cells were detached from plates with 0.05% Trypsin-EDTA (Gibco, 25300054), washed with PBS/ 2% BSA and aliquoted into one million cells per tube. Cells were fixed with 4% PFA for 20 minutes, washed with PBS/ 2% BSA and followed by 20 minutes permeabilization with PBS/ 0.5% Triton X-100. Cells were further incubated with the staining cocktail for 10 minutes at room temperature in the dark to reveal EdU incorporation. After twice washes with PBS/ 2% BSA, cells were stained with 3µM PI (Invitrogen, P1304MP) for 15 minutes at room temperature and analyzed using the BD FACSCanto II HTS flow cytometer.

Immunofluorescence

Immunofluorescence analyses were carried out largely as described previously (Pasque et al., 2014), using the following primary antibodies: NANOG (eBioscience, 14-5761 clone eBioMLC-51, 1/200; and Abcam, ab80892, 1/ 200), DPPA4 (R&D, AF3730, 1/200), HA (Cell Signaling Technology, 2367S, 1/100), DUSP9 (Abcam, ab167080, 1/100). Images were acquired using an ApoTome Zeiss Microscope equipped with an AxioCam MRc5 camera. ESC and iPSC lines were defined as NANOG+ or DPPA4+ when >50% cells showed NANOG or DPPA4 staining signal.

RNA FISH

RNA Fluorescence In Situ Hybridization (RNA FISH) analyses were carried out mostly as described previously using double stranded directly labelled DNA probe for *Tsix/Xist (Pasque et al., 2014)*. Images were acquired using an ApoTome Zeiss Microscope equipped with an AxioCam MRc5 camera. Single-cell resolution analysis of *Tsix/Xist* biallelic expression in iPSCs and ESCs was determined by calculating the ratio of cells with biallelic *Tsix/Xist* expression to the cells with monoallelic or biallelic *Tsix/Xist* expression.

Genomic DNA extraction and qPCR

Genomic DNA (gDNA) was extracted from feeder-depleted ESCs and iPSCs using the PureLink Genomic DNA Kit (Invitrogen, K1820) and qPCR was performed using the Platinium SYBR Green qPCR SuperMix-UDG kit (Invitrogen, 11733046) on a ABI ViiA7 real-time PCR system (Applied Biosystems), following the manufacturer's protocol. Primers against four X-linked genes (*Tfe3, Bcor, Pdha1, and Mid1*) covering the two distal parts of the mouse X-chromosome are listed in S9 Table (Fig 1 G). The standard curve was derived from serial dilutions of gDNA from XY ESCs (V6.5). All qPCR assays used had an efficiency above 95%. Relative quantities of each gene were measured as arbitrary units from comparison to the standard curve. The ratio of X-chromosome to autosome (X/Autosome Ratio) in DNA level was presented as the average ratio of the X-linked gene quantity (*Tfe3, Bcor, Pdha1* and *Mid1*) to the autosomal gene quantity (*Gapdh*), in other words X/Autosome Ratio = (*Tfe3/Gapdh* + *Bcor/Gapdh* + *Pdha1/Gapdh* + *Mid1/Gapdh*)/4 .

RT-qPCR

Total RNA was extracted using the RNeasy Mini Kit (Qiagen, 74106) or TRIzol (Invitrogen, 15596026). cDNA synthesis was performed using the SuperScript III First-Strand Synthesis SuperMix kit (Invitrogen, 11752-050) and RT-qPCR was performed using the Platinium SYBR Green qPCR SuperMix-UDG kit (Invitrogen, 11733046) and on the ABI ViiA7 real-time PCR system, following the manufacturer's protocol. Primers used are listed in S9 Table. The standard curve was derived from serial dilutions of cDNA. All assays used had an efficiency above 95%. Relative quantities of each transcript were calculated as arbitrary units from comparison to the standard curve. Relative expression level of the target transcript was presented as the ratio of the target transcript quantity to the housekeeping transcript (Gapdh) quantity. Logarithm values (base 2) of relative expression levels were used for assessment of the gene expression kinetics during differentiation. The relative gene expression levels of five pluripotency-associated genes (*Prdm14, Nanog, Tcl1, Rex1* and *Esrrb*) from iPSCs (male lines: 1, 4, 8, 12, 16; female lines: 19, 20, 21, 22, 23, 26) and ESCs (V6.5 male ESCs and F1-2-1 female ESCs) at 0h and 24h of EpiLC differentiation were used for unsupervised clustering comparison, which was performed in R with heatmap.2 function in package "gplots".

RNA sequencing

Total RNA was isolated from two independent female *Dusp9*+/- ESC lines, *Dusp9+/+* XX and XY ESCs in both the undifferentiated state and the differentiated state after 24 hours of LIF withdrawal using TRIzol following the manufacturer's protocol. 4 µg of total RNA was used for construction of stranded poly(A) mRNA-Seq library with the KAPA stranded mRNA Library prep kit (KAPA Biosystems, KK8421). Library concentrations were quantified with the Qubit dsDNA HS (High Sensitivity) Assay Kit (Invitrogen, Q32854), and equimolar amounts were pooled for single-end sequencing on an Illumina HiSeq 4000 instrument (Illumina) to yield ~20 million (range 16-23 million) 36bp or 51bp long reads per sample.

Differential gene expression analysis

Reads from all datasets (*Dusp9+/-* ESCs, *Dusp9+/+* ESCs and XY ESCs) were aligned to mouse reference genome GRCm38/mm10 using STAR (v2.5.3a) with default parameters followed by conversion to BAM format sorted by coordinate. The mapping efficiencies of the datasets were >69% of uniquely mapped reads. Subsequently, the featureCounts function from the R Bioconductor package "Rsubread" was used to assign mapped reads to genomic features. For downstream analyses, only the genes with CPM value (count-per-million) higher than 0.5 in at least two libraries were retained. The resulting read count matrix (S8 Table) was used as the input for PCA with the top 500 most variable genes. Differential gene expression analysis was performed using the edgeR quasi-likelihood pipeline in R (Chen et al., 2016b). Obtained p-values were corrected for multiple testing with the Benjamini-Hochberg method to control the FDR. DEGs were defined on the basis of both FDR < 0.05 and fold difference ≧ 1.5. Venn diagrams were generated using an online tool as previously described (Heberle et al., 2015). Heatmaps were created using unsupervised hierarchical clustering of both 200 most variable genes (or stem cell maintenance related genes (GO:0019827), MAPK pathway related genes) and the different samples and generated in R using the heatmap.2 function of the package "gplots".

Omni-ATAC-seq

Assay for transposase accessible chromatin (ATAC) followed by sequencing was performed using the Omni-ATAC protocol (Corces et al., 2017). Briefly, iPSCs and ESCs were expanded on top of male WT feeders in mouse ESC medium (S/L condition). After feeder-depletion, 50,000 viable cells were pelleted at 500 RCF at 4°C for 5 min in a fixed angle centrifuge, and then the cells were gently washed once with 50 µl of cold PBS. Next, the cell pellets were resuspended in 50 µl of ATAC-lysis buffer (10mM Tris HCl pH7.4, 10mM NaCl, 3mM MgCl2, 0.1% Tween-20, 0.1% NP40, and 0.01% Digitonin) and incubated on ice for 3 min. Wash out lysis with 1 ml of cold ATAC-lysis buffer containing 0.1% Tween-20 but No NP40 or digitonin and invert tube 3 times to mix. Nuclei were pelleted at 500 RCF for 10 min at 4°C in a fixed angle centrifuge. After discarding all supernatant, nuclei were resuspended in 50 µl of transposition mixture (25 ul 2x TD buffer, 2.5 ul transposase (100 nM final), 16.5 ul PBS, 0.5 ul 1% digitonin, 0.5 ul 10% Tween-20, and 5 ul H2O) (Nextera DNA Sample Preparation Kit, Illumina, FC-121-1030). The reaction was performed at 37°C for 30 minutes in a thermomixer with 1000 RPM mixing. The transposed DNA was purified using a Zymo DNA Clean and Concentrator-5 Kit (D4014). DNA libraries were PCR amplified using NEBNext High-Fidelity 2x PCR Master Mix (Bioke, M0541), and size selected for 200 to 800 bp using homemade Serapure beads (Rohland and Reich, 2012). Library concentrations were quantified with the KAPA Library Quantification Kit (KK4854), and equimolar amounts were pooled for single-end sequencing on an Illumina HiSeq 4000 instrument (Illumina) to yield ~50 million (range 34-90 million) 51bp long reads per sample.

Differential chromatin accessibility analysis

Single-end ATAC-seq raw data were analyzed using the ATAC-seq pipeline from the Kundaje lab (Version 0.3.3) (Lee et al., 2016). Briefly, the raw reads were first trimmed using cutadapt (version 1.9.1) to remove adaptor sequence at the 3′ end. The trimmed reads were aligned to reference genome (mm10) using Bowtie2 (v2.2.6) using the '--local' parameter. Single-end reads that aligned to the genome with mapping quality ≥30 were kept as usable reads (reads aligned to the mitochondrial genome were removed) using SAMtools (v1.2). PCR duplicates were removed using Picard's MarkDuplicates (Picard v1.126). Open chromatin regions (peak regions) were called using MACS2 (v2.1.0) using the '-g 1.87e9 -p 0.01 --nomodel --shift -75 --extsize 150 -B --SPMR --keep-dup all --call-summits' parameter (Zhang et al., 2008). The differential chromatin accessibility analysis and related plots were performed using the DiffBind package with 'DESeq2, log2fold=1, FDR<=0.05' parameter (Stark and Brown, 2011). GO analysis for Biological Process terms was performed using GREAT (v3.0.0) analysis (McLean et al., 2010) with the mm10 reference genome, where each region was assigned to the single nearest gene within 1000 kb maximum distance to the gene's TSS.

Motif Discovery Analysis

Known motif search was performed using program of findMotifsGenome.pl in the HOMER package (v4.9.1) with 'mm10 -size -250,250 -S 15 -len 6,8,10,12,16' parameters (Heinz et al., 2010). Incidences of specific motif was examined by the program of annotate-Peaks.pl in the HOMER package with size parameter ''-size 500''.

Western blots

Cells were detached from plates with 0.25% Trypsin-EDTA (Gibco, 25200056), pelleted before addition of RIPA lysis buffer (Sigma, R0278-50ML) supplemented with 1% (v/v) Protease inhibitor cocktail (Sigma, P8340-1ml) and 1% (v/v) Phosphatase inhibitor Cocktail 3 (Sigma, P0044-1ML), and lysed on ice for 30 min. The lysates were spun for 10 min at 13000 rpm. The protein concentration was determined with BCA protein assay kit (Pierce, 23225). Each sample with 15 µg of total protein was denatured in 1x LDS Sample buffer (Life Technologies, NP0007) with 100 mM DTT for 5 min at 98°C. The cell lysates were loaded onto a 4%–15% mini-Protean TGX gel (Bio-Rad, 456-1083), electrophoresed, and transferred to nitrocellulose membranes (VWR,10600002). Membranes were blocked in PBS 0.1% (v/v) Tween-20 and 5% (g/v) blotting reagent (Bio-Rad, 1706404) and incubated with the following primary antibodies overnight at 4°C: rabbit anti-NANOG (Abcam, ab80892, 1/1000), rabbit anti-DUSP9 (Abcam, ab167080, 1/500), mouse anti-DKC1 (Santa Cruz, sc-365731, 1/250), mouse anti-HA (Cell Signaling Technology (CST), 2367S, 1/1000), sheep anti-ZIC3 (R&D Systems, AF5310, 1/250) and mouse anti-ACTIN (Abcam, ab3280, 1/5000). After extensive PBS 0.1% Tween-20 (PBS-T) washes, membranes were incubated with a secondary HRP-conjugated goat anti-mouse IgG antibody (Bio-Rad, 1706516, 1/5000) or goat anti-rabbit IgG antibody (Bio-Rad, 1706515 1/5000) for 30 minutes at room temperature. After another round of extensive PBS-T washes, protein expression was visualized using the ECL chemiluminescence reagent (Perkin-Elmer, NEL103001EA) and LAS-3000 imaging system (Fuji). Data were analyzed with ImageJ.

Single Cell RNA-seq analysis

Single cell RNA-seq data of XX female and XY male ESCs in S/L and 2i/L from published data set (Chen et al., 2016a) was re-aligned to N-masked mouse reference genome mm10 using hisat 2 (2.0.5) with disabled soft-clipping. Alignment was followed by conversion to BAM files using SAMtools 1.4.1, aligned reads were then summarized using featureCounts v1.5.2. Quality controls and downstream analyses were performed with the use of scater and SingleCellExperiment packages (Aaron Lun, Davide Risso, 2017; McCarthy et al., 2017). Cells displaying total counts lower than 500,000 reads and less than 9000 genes detected were discarded from the analysis.

SUPPLEMENTAL REFERENCES

Aaron Lun, Davide Risso (2017). SingleCellExperiment: S4 Classes for Single Cell Data. R package version 1.2.0 (Bioconductor).

Branco, M.R., King, M., Perez-Garcia, V., Bogutz, A.B., Caley, M., Fineberg, E., Lefebvre, L., Cook, S.J., Dean, W., Hemberger, M., et al. (2016). Maternal DNA Methylation Regulates Early Trophoblast Development. Dev. Cell *36*, 152–163.

Carey, B.W., Markoulaki, S., Beard, C., Hanna, J., and Jaenisch, R. (2010). Single-gene transgenic mouse strains for reprogramming adult somatic cells. Nat. Methods *7*, 56–59.

Chen, G., Schell, J.P., Benitez, J.A., Petropoulos, S., Yilmaz, M., Reinius, B., Alekseenko, Z., Shi, L., Hedlund, E., Lanner, F., et al. (2016a). Single-cell analyses of X Chromosome inactivation dynamics and pluripotency during differentiation. Genome Res. *26*, 1342–1354.

Chen, Y., Lun, A.T.L., and Smyth, G.K. (2016b). From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. F1000Res. *5*, 1438.

Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J., and Plath, K. (2017). Cooperative Binding of Transcription Factors Orchestrates Reprogramming. Cell *168*, 442–459.e20.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat. Methods *14*, 959–962.

Czechanski, A., Byers, C., Greenstein, I., Schrode, N., Donahue, L.R., Hadjantonakis, A.-K., and Reinholdt, L.G. (2014). Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains. Nat. Protoc. *9*, 559–574.

Fulco, C.P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S.R., Perez, E.M., Kane, M., Cleary, B., Lander, E.S., and Engreitz, J.M. (2016). Systematic mapping of functional enhancer-promoter connections with CRISPR interference. Science *354*, 769–773.

Heberle, H., Meirelles, G.V., da Silva, F.R., Telles, G.P., and Minghim, R. (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics *16*, 169.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Lee, J., Christoforo, G., Christoforo, G., Foo, C.S., Probert, C., Kundaje, A., Boley, N., kohpangwei, Dacre, M., and Kim, D. (2016). kundajelab/atac_dnase_pipelines: 0.3.3.

McCarthy, D.J., Campbell, K.R., Lun, A.T.L., and Wills, Q.F. (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics *33*, 1179–1186.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Pasque, V., Tchieu, J., Karnik, R., Uyeda, M., Sadhu Dimashkie, A., Case, D., Papp, B., Bonora, G., Patel, S., Ho, R., et al. (2014). X chromosome reactivation dynamics reveal stages of reprogramming to pluripotency. Cell *159*, 1681–1697.

Pasque, V., Karnik, R., Chronis, C., Petrella, P., Langerman, J., Bonora, G., Song, J., Vanheer, L., Sadhu Dimashkie, A., Meissner, A., et al. (2018). X Chromosome Dosage Influences DNA Methylation Dynamics during Reprogramming to Mouse iPSCs. Stem Cell Reports *10*, 1537–1550.

Rohland, N., and Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. *22*, 939–946.

Schulz, E.G., Meisig, J., Nakamura, T., Okamoto, I., Sieber, A., Picard, C., Borensztein, M., Saitou, M., Blüthgen, N., and Heard, E. (2014). The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. Cell Stem Cell *14*, 203–216.

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.

Stark, R., and Brown, G. (2011). DiffBind: differential binding analysis of ChIP-Seq peak data. R Package Version.

Waisman, A., Vazquez Echegaray, C., Solari, C., Cosentino, M.S., Martyn, I., Deglincerti, A., Ozair, M.Z., Ruzo, A., Barañao, L., Miriuka, S., et al. (2017). Inhibition of Cell Division and DNA Replication Impair Mouse-Naïve Pluripotency Exit. J. Mol. Biol. *429*, 2802–2815.

Ying, Q.-L., Nichols, J., Evans, E.P., and Smith, A.G. (2002). Changing potency by spontaneous fusion. Nature *416*, 545–548.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.