מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Thesis for the degree
Doctor of Philosophy

עבודת גמר (תזה) לתואר
דוקטור לפילוסופיה

Submitted to the Scientific Council of the
Weizmann Institute of Science
Rehovot, Israel

מוגשת למועצה המדעית של
מכון ויצמן למדע
רחובות, ישראל

By
Ohad Gafni

מאת
אוהד גפני

התווית מסלולים ויישומים של המצב הנאיבי הרב תכליתי

Delineating routes and applications of naïve pluripotency

Advisor:
Dr. Jacob (Yaqub) Hanna -
M.D. Ph.D

מנחה:
ד"ר יעקוב חנא

Month and Year
August 2016

חודש ושנה עבריים
אב התשע"ו

List of Abbreviation

ICM	Inner Cell Mass
EBs	Embryonic Bodies
BAC	Bacterial artificial chromosome
MEFs	Mouse Embryonic Fibroblasts
SCNT	Somatic Cell Nuclear Transfer
iPSC	Induced Pluripotent Stem Cell
ESC	Embryonic Stem Cell
PSCs	Pluripotent Stem Cells, including ESCs and iPSCs
MSC	Multipotent stem cell
hESCs	Human Embryonic Stem Cells
mESCs	Mouse Embryonic Stem Cells
EpiSCs	Epiblast Stem Cells
OSKM	Oct4, Sox2, Klf4 and c-Myc
PGCs	Primordial Germ Cells
DOX	Doxycycline
KSR	Knock-out Serum Replacement
FBS	Fetal Bovine Serum
LIF	Leukemia inhibitory factor
2i	PD0325901 and CHIR99021 (2 inhibitors of the ERK1/2 and GSK3β)
2i/LIF	serum free medium with addition of Lif, Erki and Gsk3βi
PE	Proximal Enhancer of OCT4 gene
DE	Distal Enhancer of OCT4 gene
IVF	*In-Vitro* Fertilization
NHSM	Naïve Human Stem Cells Medium
TSS	Transcription start site
MHC	Major Histocompatibility Complex
HLA	Human Leukocyte Antigen
FDR	False Discovery Rate
RPKM	Reads Per Kilobase per Million reads
RQ	Relative Quantity
CRISPR	Clustered Regularly-Interspaced Short Palindromic Repeats

GFP	Green fluorescent protein
LSD	Lysine specific demethylase
PRC	Polycomb repressive complex
Utx	Ubiquitously transcribed tetratricopeptide repeat, X chromosome
XaXa	Both active X chromosomes
XaXi	One active X chromosome, one inactive X chromosome

Table of Contents

Abstract

Even though it has been almost 40 years since the first mammalian embryo was made by somatic cell nuclear transfer, fundamental routes and principals of somatic cell reprogramming into the naïve state are still elusive. Pluripotency can be induced in somatic cells by ectopic expression of a variety of transcription factors, classically Oct4, Sox2, Klf4 and c-Myc (abbreviated as OSKM)[1-4]. This process is accompanied by genome wide epigenetic Histone 3 Lysine 4 tri-methylation (H3K4me3) and Histone 3 Lysine 27 tri-methylation (H3K27me3) correlated with active and repressive transcriptional states, respectively. The questions of how do these critical changes get biochemically dictated and which epigenetic regulators drive the progression of cellular reprogramming require further investigation [5, 6].

My work is divided into two major parts; the first part revealed Utx H3K27me3 demethylase as a key component required for reprogramming. We have shown that although Utx deficient embryonic stem cells maintain pluripotent characteristics, somatic cells lacking Utx failed to be reprogramed back to a pluripotent state, mainly due to severe accumulation of H3K27me3 repressive marks on key pluripotency regulating genes.

The second part focused on defining growth conditions that facilitate the derivation of naïve human pluripotent stem cells (PSCs; refer to both ESCs and iPSCs). Mouse embryonic stem cells (mESCs) are isolated from the inner cell mass (ICM) of the blastocyst, and can be preserved *in vitro* in a naïve ICM-like configuration by providing exogenous stimulations, typically with leukemia inhibitory factor and small molecule inhibition of ERK1/2 and GSK3β signaling (termed 2i/LIF conditions). Upon withdrawal of 2i/LIF, naïve mESCs epigenetically drift towards a primed pluripotent state resembling that of the post-implantation epiblast. While human embryonic stem cells (hESCs) share several molecular features with naïve mESCs, they also share a variety of epigenetic properties with primed murine Epiblast stem cells (mEpiSCs). A combination of small inhibitors of signaling pathways and cytokines established in our laboratory retained growth characteristics, molecular circuits, chromatin landscape, and signaling pathways that are highly similar to mouse naïve ESCs, and drastically distinct from conventional primed hPSCs. The potential of naïve hiPSCs to faithfully contribute to chimeric animals is not defined yet. To address this issue, I engineered naïve human iPSC knock-in line that constitutively expresses GFP. These cells were injected into mouse E2.5 morulas, implanted into pseudo-pregnant female mice and their progeny was traced during embryonic development, using immuno-histological analysis and *in-toto* confocal live imaging.

תקציר

אף על פי שחלפו קרוב ל-40 שנים מאז שיצרו את עובר היונקים הראשון באמצעות העברת גרעין סומטי לביצית (SCNT), מסלולים בסיסיים ועיקריים בתהליך ההשריה של המצב הפלוריפוטנטי נותרו נסתרים. המצב הפלוריפוטנטי ניתן להשריה גם ע״י ביטוי חיצוני של פקטורי שעתוק, כאשר השילוב הקלאסי של פקטורים אלו כולל את Oct4, Klf4, Sox2 ו-c-Myc (OSKM). תהליך זה טומן בחובו גם שינויים אפיגנטים הכוללים את הוספה והורדה של קבוצות מתיל על שיירי הליזין של היסטון 3, H3K4me3 ו- H3K27me3, המזוהים עם שפעול ועיכוב של מצבי שעתוק, בהתאמה. בכדי להבין כיצד שינויים אלה מבוקרים, מווסתים ומי השחקנים המעורבים בתהליכים אלה, נדרש עוד מחקר רב.

עבודתי במהלך הדוקטורט מחולקת לשניים; בחלקה הראשון הראיתי את חשיבותו של Utx, אנזים המוריד קבוצות מתיל משייר ליזין 27 של היסטון 3 (H3K27me3), במהלך תהליך ההשריה של המצב הפלוריפוטנטי. חסרונו של Utx אמנם לא השפיע על תכונותיהם של תאים פלוריפוטנטים, אך תאים סומטים חסרי Utx לא הצליחו לעבור את תהליך ה-reprogramming בשל צבירה של הסמן מעכב השעתוק H3K27me3.

חלקה השני של עבודתי עוסק בהתאמת תנאי הגידול לתאי אדם פלוריפוטנטים במצב הנאיבי. תאי גזע נאיביים מעכבר ניתנים לבידוד מה-ICM של בלסטוציסטים. ניתן לגדל ולשמר תאים אלה לאורך זמן על ידי מתן תוספים חיצוניים למדיום הגידול, בד״כ LIF ומעכבים של ERK1/2 ו-GSK3β. כאשר תוספים אלה מוסרים, תאי הגזע מתמיינים למצב primed המזוהה בעובר לאחר ההשתלה ברחם. בעוד תאי גזע מאדם חולקים תכונות מסוימות עם תאי גזע נאיביים מעכבר, הם חולקים גם מגוון של סימנים אפיגנטים עם תאי גזע primed מעכבר. שילוב של מעכבי מסלולים תוך תאיים וציטוקינים במדיום שבוסס במעבדתנו הצליח לשמר מאפייני גידול, מסלולים מולקולרים, מצב כרומטין ומסלולים תוך תאיים הדומים למצב הנאיבי בעכבר ושונים בתכליתם מהמצב ה-primed באדם, המצב בו היה ניתן לגדל תאי גזע מאדם עד כה. הפוטנציאל של תאי גזם נאיביים מאדם לתרום לחיות כימריות טרם יושם, לכן ייצרתי שורת תאים המבטאת באופן קבוע GFP על ידי הכנסה של הגן ללוקוס ספציפי. תאים אלה הוזרקו למורולות של עכברים, והושתלו למחרת בקרן רחם של נקבות הנמצאות במצב של הריון מדומה. עוברים בשלבי התפתחות שונים נבדקו להמצאות GFP בשיטות היסטולוגיות שונות ובשיטת הדמיה *in-toto* במיקרוסקופ קונפוקלי בזמן אמת.

Introduction

Background

Cell potency can be defined by three major stages: totipotency, pluripotency and multipotency. The first totipotent stem cell arises after the sperm fertilizes the oocyte to form the zygote. At this stage the cell undergoes several divisions until the morula stage when the first aberrant differentiation process takes place toward the blastocyst. Totipotency holds the greatest differentiation potential, these cells are able to produce all cell types in the organism including the extraembryonic tissues and the inner cell mass (ICM). The ICM is comprised of pluripotent embryonic stem cells (ESCs), which have the capacity to differentiate to any cell type within the embryonic three germ layers (endoderm, mesoderm and ectoderm). PSCs can be derived from the ICM at the naïve state, and from the epiblast in a slightly more differentiated state. ESCs have the ability to self-renew indefinitely and can be maintained *in-vitro*. Multipotent stem cells (MSCs) are progenitor cells with limited differentiation potential; their gene activation circuitry allows them to self-renew and to differentiate into multiple cell types from a certain lineage. Adult stem cells are considered MSCs that are present in specific niches in several adult tissues. For example, the intestinal epithelium is a rapid self-renewing tissue, Lgr5 positive cells are found in the intestinal crypts niche and are able to form all cells from intestinal epithelial lineages [7].

Embryonic and induced pluripotent stem cells

Derivation and culturing of embryonic stem cells from the mouse ICM was first possible in the 1980's with the use of a conditioned medium [8]. These mESCs were able to proliferate indefinitely, differentiate spontaneously when grown in suspension on non-adhesive plates to form embryonic bodies (EBs) and form teratomas constitute from all three germ layers by sub-coetaneous injection to immune deficient mice. Ultimately, Micro-injection of mESCs into host blastocysts gives rise to all embryonic cell types in the developing embryo, including germ cell lineages i.e. chimeric mouse [9]. In addition, mESCs retained the ICM epigenetic features, including open chromatin structure and lack of X chromosome inactivation in female cell lines. In tetraploid complementation assay, the two-cell embryo is subjected to an electric shock that causes fusion of the two nuclei to form one tetraploid nucleus, this embryo contributes solely to the extraembryonic tissues. mESCs injected into tetraploid blastocyst generate a whole embryo composed exclusively from the injected population, and develop to an adult animal. This assay provides conclusive evidence for mESCs pluripotency as these cells create an entire healthy and fertile animal.

Since then, mESCs were conventionally grown in conditioned media with fetal bovine serum (FBS) on mitotically inactive mouse embryonic fibroblasts (MEFs). Leukemia inhibitory factor (LIF) was discovered in 1988 as a critical mESCs stabilizing cytokine [10], LIF is an activator of the Stat3/Jak pathway that is activated in pluripotency. Human ESCs (hESCs) were first successfully derived from human blastocysts in 1998 [11], these hESCs formed teratomas comprised of the three germ layers, and EBs *in-vitro*. However, due to ethical limits, hESCs were never tested for their contribution to human embryos. hESCs conventional growth condition medium require FGF2 and TGFβ and not LIF dependent as mESCs.

Epigenetic reprogramming

The first evidence for reprogramming of a differentiated somatic cell nucleus to a pluripotent state was done by John Gurdon in 1962 [12]. In his research, John Gurdon replaced an oocyte nucleus with a terminally differentiated nucleus, this zygote-like cell maintained normal embryogenesis and developed to an adult animal. ESCs derived from blastocysts developed in this manner resembled ESCs derived from normal embryos. This discovery, later named somatic cell nuclear transfer (SCNT), had shown that oocytes have the potential to "reset" fully differentiated somatic cells nuclei.

SCNT is not commonly used today as it requires many non-fertilized oocytes, thus considered inefficient. Stabilization of the pluripotent state depends on critical transcriptional regulators, such as Oct4 (Pou5f1), Sox2 and Nanog [13-17]. In 2006, Shinya Yamanaka and Kazutoshi Takahashi performed a genetic screen of 24 transcription factors by ectopic induction on embryonic and adult fibroblasts. The screen results showed that direct reprogramming of somatic cells into a pluripotent state can be achieved by over expression of four transcription factors; Oct4, Sox2, Klf4 and c-Myc (OSKM) [3]. The induced pluripotent stem cells (iPSCs) were identical to ESCs in self-renewal and differentiation capabilities, and highly similar in gene expression profiles. The reprogramming process usually takes 1-4 weeks, and it completely erases the epigenetic landscape and transcription pattern. Nonetheless, these factors initiate poorly defined events, which eventually lead to the endogenous reactivation and self-maintenance of the pluripotent state. During iPSC reprogramming, the chromatin undergoes extensive changes and remodeling through yet to be fully identified mediators.

Chromatin modifications

The eukaryotic chromatin is a dynamic structure influenced by transcription factor networks and transcriptional processes, which extensively harness chromatin modifiers and nucleosome remodelers. DNA and histone modification regulate gene expression patterns, and eventually impose cells identity [18]. Global epigenetic states must be tightly regulated during development to allow the proper transitions between cellular states. However, cell fates during development are neither restrictive nor irreversible.

The chromatin comprises nucleosomes, which are histone octamers that consist of two copies of each H2A, H2B, H3 and H4 proteins, wrapped with 146 bp of DNA. These nucleosome building blocks can be subjected to various post-translational modifications on their N-terminal tails such as phosphorylation, ubiquitination, acetylation and methylation. Each of the modification affects chromatin structure and different combinations of these modifications are the determinants of the chromatin accessibility that ultimately governs gene expression or repression [19]. For example, methylation of H3K4, H3K36 and H3K79 are related to transcriptional activation, while methylation of H3K9 and H3K27 are linked to transcriptional repression [9, 20].

Histone demethylases

There are three states mono-, di- and tri-methylation of each lysine residue, which is thought to differentially impose chromatin structure and transcription. A group of methyl transferases catalyze histone lysine methylation, a modification that was considered until a few years ago very stable and irreversible [21]. This notion was changed with the revelation of the amine oxidase Lysine Specific Demethylases 1 (LSD1) as an H3K4me2\me1 specific demethylase [22]. The Jumonji family of proteins is another family of demethylases, which shares the catalytic domain Jumonji C (JmjC). The JmjC catalytic domain is essential for the oxidative lysine demethylation reaction, which requires Fe(II) and α-ketoglutarate as co-factors, and by that differs from The LSD1 family in their ability to demethylate tri-methylated histone tails. The polycomb repressive complex 2 (PRC2) containing the (PcG) proteins Ezh2, Eed, Suz12 and the nucleosome binding protein RbpAp48 are responsible for the tri- and di-methylation of H3K27 [23]. The PRC2 proteins function is to maintain the repression of genes dictating cell fate and is crucial for normal development, differentiation and X chromosome inactivation in females [24-26].

H3K27 demethylases

Utx is an X chromosome encoded gene that together with Jmjd3 and Uty belongs to the JmjC-domain family [27, 28]. Utx and Jmjd3 appear to promote gene activation by mediating the demethylation of H3K27 tri- and di- methyl (me3/2) repressive chromatin marks, while Uty lacks catalytic activity and its function is still not clearly known [27, 29-33]. Interestingly, Utx was suggested not to play a role in pluripotency regulation, as Utx null mice displayed lethality only at embryonic stages E9.5-E10.5 as a result of defective cardiac development [27, 34]. The H3K27 demethylase activity is essential for normal development; it has been shown that inhibition of *Utx* expression prevented the activation of *HOX* genes. Moreover, inhibition of Utx orthologues in zebrafish also resulted in decreased expression of *Hox*, which resulted in defective posterior trunk [27, 31]. Although Utx and Jmjd3 share the specificity to H3K27 and have the same catalytic domain, they have different biological function. Over-expression of JMJD3, but not UTX, results in significant H3K27me3 demethylation *in vivo* [27, 31, 33]. This could imply that UTX needs to be part of a specific complex in order to be active whereas JMJD3 does not. Moreover, inhibition of UTX expression in HeLa cells results in a global increase in H3K27me3 levels, suggesting that UTX is involved in maintaining steady state H3K27me3 levels of a large number of genes [27, 33].

In the first project of my PhD work, I have elucidated the role of H3K27me3 demethylase Utx in the maintenance and re-establishment of pluripotency. I generated *Utx* knockout cell lines and examined their pluripotency properties both in mouse ESCs and during the reprogramming process *in-vitro*. In addition, by blastocyst injection of the *Utx* knockout cells, I obtained chimeric animals and inspected their germ cell development in the absence of Utx.

Naïve Vs. Primed pluripotent stem cells

Derivation of embryonic PSCs is feasible in two states. Naïve pluripotent ESCs are isolate from the ICM of the pre implantation embryo at the blastocyst stage. PSCs could also be derived from the Epiblast of post implantation embryos [35, 36]. While epiblast stem cells (EpiSCs) share several characteristics with ESCs and iPSCs such as teratoma formation, they differ in their ability to contribute efficiently to chimeric animals when injected to blastocysts. This inability could be caused by several reasons such as lower differentiation potential or expression of adhesion molecules in the advanced embryo. These adhesion molecules prevent EpiSCs and ICM cell to cell contact and integration. Similar to hESCs,

mEpiSCs are also dependent on FGF2 and Activin\Nodal signaling and not on LIF like mESCs and iPSCs.

Human is not the only species that could not yield ESCs in mouse conventional growth conditions, containing LIF and MEF coated plates. Monkey, pig and rat ESCs also could not be derived under these conditions. Based on this experience, the common assumption is that every species requires different growth conditions despite their identical origin. Furthermore, ESCs could not be derived in some mouse strains under conventional conditions. These facts emphasize the need for specific conditions for different species.

Murine naïve and primed pluripotency network

The quest for defined mPSCs growth condition revealed the inhibition importance of Erk and β-catenin signaling pathway. The use of serum free medium supplemented with Erk and Gsk3 small molecule inhibitors, termed 2i/LIF medium, prevents differentiation and allows proliferation even on plates without MEFs [35]. PSCs grown in 2i/LIF medium are more homogenous and show higher pluripotency gene expression due to intrinsic and extrinsic differentiation cues block. The exact effect of those pathways on pluripotency is yet to be revealed.

Inhibition of signaling pathways is crucial in order to maintain PSCs in an undifferentiated state, since PSCs in neutral conditions spontaneously differentiate. Several studies have shown that alternative 2i combinations can also stabilize PSCs; inhibition of atypical PKC (aPKCi Gö6983) in addition to MEK inhibitors and LIF enables the isolation of mESCs [37]. The use of medium containing Gsk3β and Src inhibitors resulted in murine germ line competent naïve ESCs [38].

Isolation of murine EpiSCs is done from E5.5-E7.5 embryos, their derivation and growth conditions are dependent on FGF2 and Activin signaling. EpiSCs expression levels of several differentiation genes are higher in EpiSCs, Otx2 ectoderm lineage marker and Brachyury mesoderm lineage marker are only two examples [39]. This fact, in addition to their low ability to engraft to the ICM, led researches to define EpiSCs as “primed” pluripotent stem cells, contrary to serum/LIF or 2i/LIF PSCs that exhibit more naïve traits [40]. Naïve and primed states differ in many characteristics; primed colony morphology is flat in comparison to naïve colony that have a more domed shape, primed PSCs are more heterogeneous within a cell line population, they show lower colonogenity of single cells [41], and higher expression of the surface molecule MHC class I [42]. Conversion between the two states is fairly simple, naïve PSCs can differentiate to prime state upon medium change after a few passage, and vice versa [43].

Stabilization of the pluripotency state is enabled by expression of transcription factors circuitry, where Oct4, Nanog, Klf4, Klf5, Sox2, Dppa3, Dppa5, Esrrb, Tfcp2l1 are only a few of them [44]. In contrast to naïve PSCs that express these genes in high level, primed cells sustain Oct4 and Sox2 expression levels, but express only low levels Nanog, Klfs, Dppas and Esrrb. This change in expression pattern between naïve and primed exhibit the rewiring of the pluripotency circuitry occurs in the *in-vivo* states pre and post-implantation embryos, respectively.

Murine naïve and primed PSCs epigenetic landscape

Global DNA methylation level in ESCs from the blastocyst ICM is extremely low. Naïve mPSCs exhibit similar level only in 2i/LIF conditions, much like somatic cells, while naïve mPSCs grown in serum/LIF show similar high levels of global DNA methylation [45].

In female mammals, X inactivation is required when exiting the naïve state of pluripotency for dosage compensation between the sexes. The mechanism of this process entails, except from *Xist* RNA expression, other processes including methylation to form H3K27me3 foci. During spermatogenesis, X and Y chromosomes are silenced through meiotic sex chromosome inactivation [46] and it has been suggested that the paternal X (Xp) arrives in a partially silenced state at fertilization. After fertilization, Xp undergoes first reactivation until the 2-cell stage. At the 4-cell stage, Xp gets inactivated until the embryo reaches the blastocyst stage, then the cells of the ICM undergo Xp chromosome second reactivation. At this state both X chromosomes are active and it is denoted as XaXa. When the ICM differentiates further to the epiblast primed state, the X chromosomes undergo random inactivation (XaXi) [43]. During germ cell differentiation, the X chromosomes undergo a third round of X reactivation. In reprogramming of somatic cells to naïve pluripotency, having two active X chromosomes is one of the key traits of this state.

General low level of H3K27me3 repressive mark is another characteristic of the naïve state. This repressive modification is mainly found in somatic cells developmentally regulated genes, maintaining them at inactive or silent mode [47]. In primed PSCs, both H3K27me3 repressive mark and H3K4me3 active mark are present on many developmental genes. This bivalent state allows genes to remain silent but ready for rapid activation upon differentiation.

Oct4 is considered as the master regulator of pluripotency, its enhancer has two regulatory elements, the proximal enhancer (PE) and the distal enhancer (DE). Even though the expression levels of Oct4 are similar in naïve and primed PSCs, in the ICM and in naïve ESCs, both enhancers regulate Oct4 expression, while DE is more dominant. However, In the

post implantation embryo and primed PSCs, DE enhancer activity is reduced and the expression regulation is more dominant via PE [40, 48].

In general, mouse naïve PSCs resemble the blastocyst ICM by displaying a more undifferentiated state, while primed PSCs are similar to *in-vivo* EpiSCs. These states can be defined by differences at regulation mechanisms, including signaling pathways, epigenetic and expression patterns.

Human pluripotent stem cells

As mentioned previously, different species most likely require specific conditions to derive and to maintain PSCs. Even some strains in mice and rat could not be derived in conventional serum/LIF conditions, but only under inhibition of Gsk3 and Erk in 2i/LIF medium [49, 50].

Although human ESCs share several molecular features with naive mouse ESCs [11], they also share a variety of epigenetic properties with EpiSCs [42, 51]. These include predominant use of the proximal enhancer element to maintain OCT4 expression, pronounced tendency for X chromosome inactivation in most female human ESCs, increase in DNA methylation and prominent deposition of H3K27me3 and bivalent domain acquisition on lineage regulatory genes [42]. Human ESCs similar to primed PSCs are FGF2 and TGFβ growth dependent. Applying 2i/LIF conditions on hPSCs with or without FGF2 and TGFβ induce rapid differentiation. Much like mouse primed PSCs, hPSCs also express basal levels of key pluripotency genes OCT4, SOX2 and NANOG, and very low levels if at all of KLF4, KLF5, TFCP2L1 and DPPAs. Moreover, hPSCs have flat colony morphology and have low level single cell colonogenity [40].

The similarity between hPSCs to mouse primed cells is much higher than mouse naïve cells grown in 2i/LIF is an accepted notion [40, 52, 53]. Evidently, hPSCs derived from blastocysts could not be stabilized at the naïve state in known growth conditions. However, hPSCs grown in 2i/LIF with constitutive expression of pluripotency transcription factors OCT4 & KLF4 or KLF2 & KLF4, resembled human ICM and naïve mPSCs in both transcriptional and epigenetic parameters [42]. Furthermore, upon withdrawal of these factors ectopic expression, these cells differentiated to a prime like state.

The feasibility of establishing human state naive pluripotency *in-vitro* using different combinations of inhibitors and cytokines, with equivalent molecular and functional features to those characterized in mouse ESCs remains to be defined [40].

In the second project of my PhD, I helped establishing stable naïve growth conditions for hPSCs by applying a screen of inhibitors and cytokines combinations that

were previously described to have a role in pluripotency. Characterization of hPSCs grown in this medium showed molecular, epigenetic and transcriptional pattern similar to human and mouse ICM. Next, I designed and generated naïve hiPSCs expressing constitutive EGFP, and established a human PSCs injection protocol to mouse morulas that yielded human-mouse interspecies embryonic chimeras.

Results

Part I: The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming

The work described in this chapter was published under the same title in Nature, 2012 [27].

In order to identify inhibitors of EpiSC reversion to naive pluripotency, we established primed EpiSCs [35, 36] carrying a Nanog–GFP reporter and doxycycline-inducible OSKM alleles, and applied piggyBac (PB)-based gene-trap screening [54] (Fig. 1a–c). Several clonal populations analyzed carried a disruptive PB integration in the second intron of the *Utx* gene and failed to generate Nanog–GFP [3] naive pluripotent cells adequately (Fig. 1d). Functional rescue by overexpression of *Utx* induced rapid Nanog–GFP reactivation.

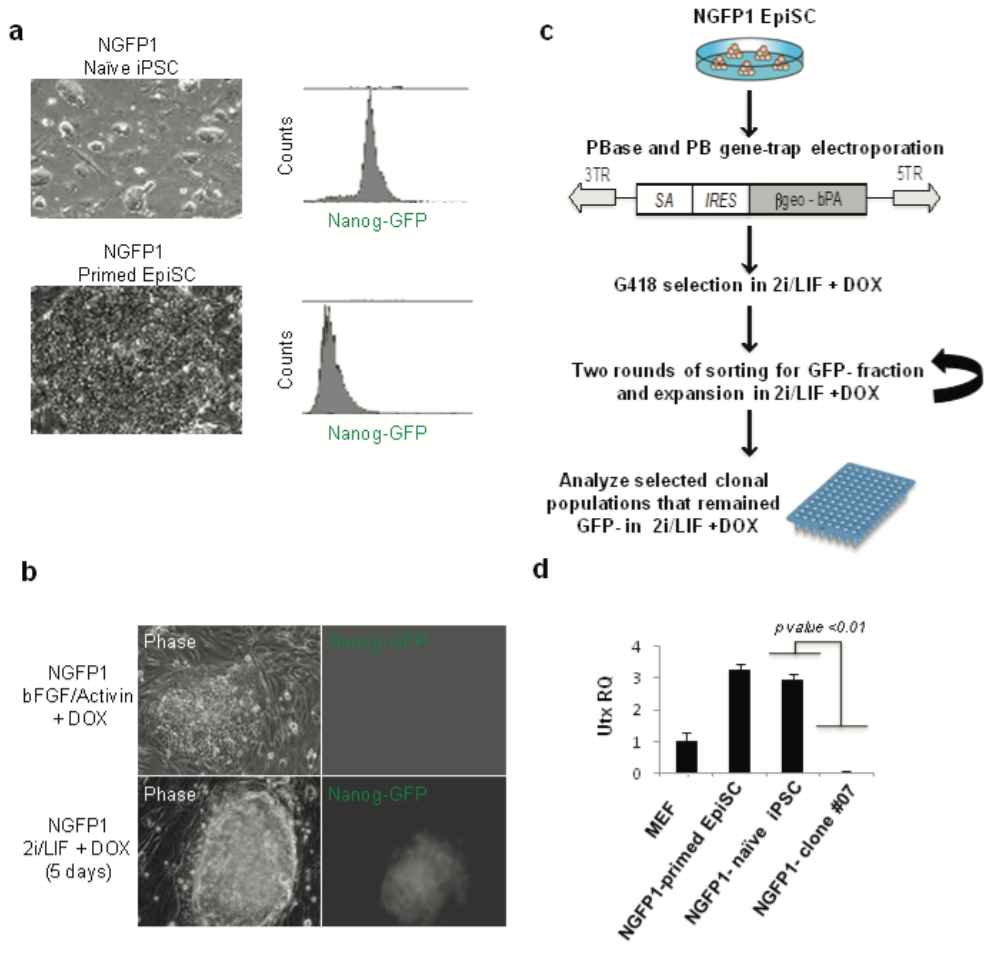

Figure 1. Experimental platform for epigenetic reversion of murine EpiSCs back to naïve ground state pluripotency. a, NGFP1 naïve-iPSCs (40XY) expanded in 2i/LIF defined naïve conditions had typical ES-like domed morphology and expressed the endogenous Nanog-GFP reporter. Genetically identical NGFP1-EpiSCs were derived from day E6.5 embryos and expanded in bFGF/Activin supplemented conditions. Cells had typical EpiSC-like flat morphology and lacked expression of Nanog-GFP reporter (consistent with previous reports.

Silva et al. Cell 2009). Note that naïve murine cells grown in N2B27 2i/LIF conditions lack the previously described oscillatory expression of Nanog in murine ESCs grown in bovine serum + LIF conditions. **b,** NGFP1-reverted into naïve Nanog+ cells upon exposure to 2i/LIF conditions in addition to DOX with >70% detectable efficiency already within 6 days (Fig. 1a). 2i-LIF in the growth conditions were required to facilitate Nanog-GFP reactivation. **c**. Schematic representation of piggyBac screening for inhibition of reverting primed EpiSCs towards Nanog-GFP+ naïve cells *in vitro*. **d**, Several of the clonal populations analyzed, likely of identical origin, carried an identical PB integration in the second intron of the *Utx* gene, which resulted in the abrogation of the full-length *Utx* transcript. RT-PCR analysis for *Utx* transcript expression in cell samples and mutant clone #7, presented as mean value +/- s.d.

Generating and characterizing *Utx* knock-out cell lines

To exclude the possibility that the *Utx* mutant related phenotype was influenced by other unidentified insertions, I knocked-out *Utx* in mouse ESCs and examined their pluripotent characteristics and differentiation potency (Fig. 2). V6.5 $Utx^{\Delta/Y}$ ESCs were pluripotent as evident by immunostaining for pluripotency markers, lack of changes in global gene expression of core pluripotency regulators, normal growth rate, lineage commitment marker upregulation upon differentiation into embryonic bodies and mature teratoma formation *in vivo* (Fig. 3). While Utx deficient embryos predominantly die around E10.5, and at E8.5-E9.5 following tetraploid complementation assay, viable high contribution adult chimeras were readily obtained from $Utx^{\Delta/Y}$ ESCs, and displayed wide tissue contribution and uncompromised long term survival (Fig. 4). The above results demonstrate that Utx is dispensable for pluripotent state maintenance, and that functional commitment into multiple somatic lineages can be achieved in the absence of this gene.

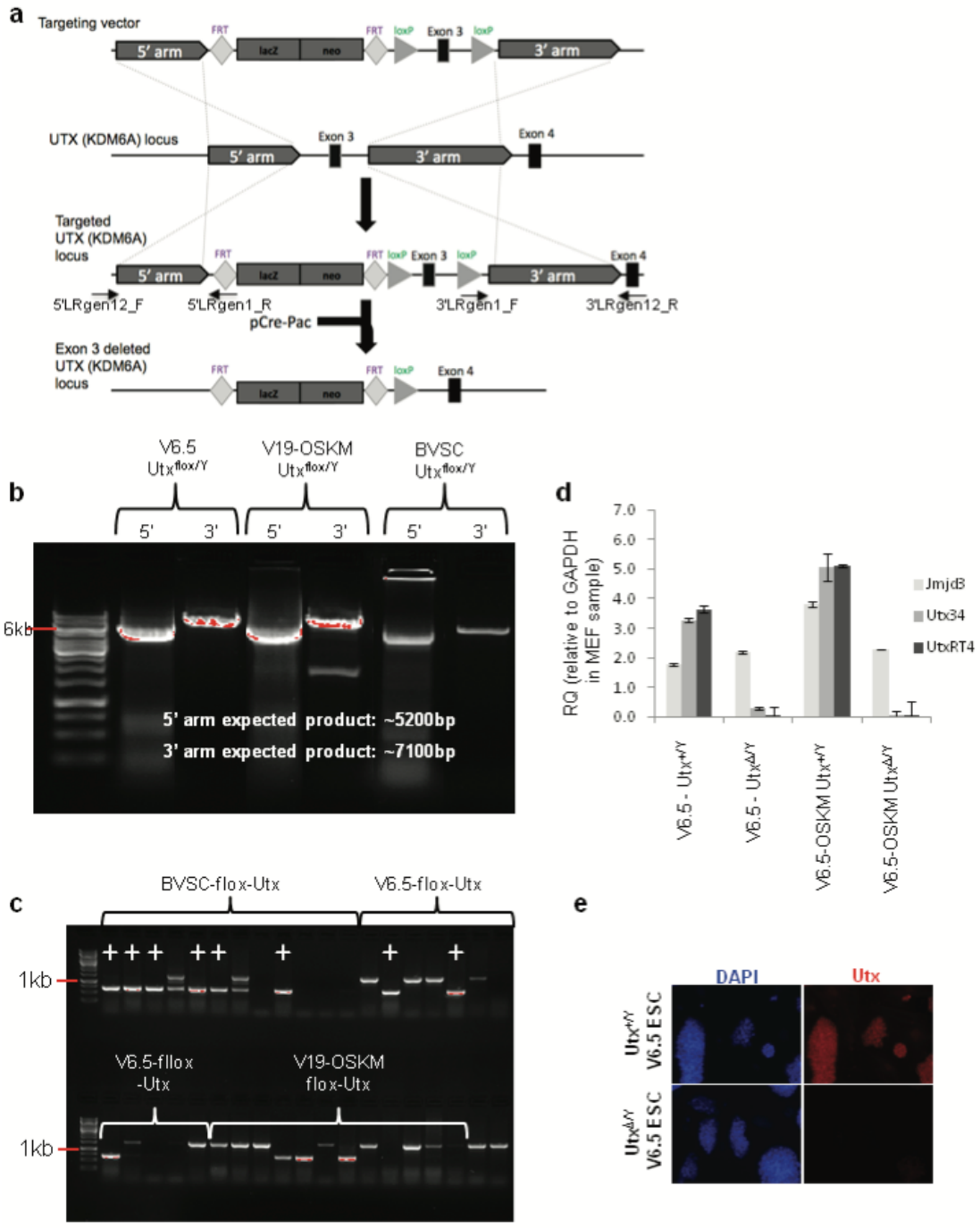

Figure 2. Gene targeting of murine Utx locus in pluripotent stem cells. a, Scheme depicting targeting construct and strategy to generate conditional deletion of *Utx* by targeting exon 3. **b,** PCR strategy and analysis confirming correct targeting of the endogenous Utx locus in 3 different pluripotent stem cell lines. **c**, Different $Utx^{\Delta/Y}$ ESCs or iPSCs were transiently transfected with pPac-Cre construct and sub cloned. PCR analysis on genomic DNA from selected clones indicated successful deletion of *Utx* exon 3 (as indicated by +). **d,** RT-PCR analysis demonstrating knockout of exon3 transcript by using 2 different sets of primers. Comparable transcript measurement of Jmjd3 was obtained, demonstrating the specificity of the Utx knockout model used herein. **e**, Representative immunostaining for Utx protein and DAPI marker demonstrating lack of protein expression following targeting in V6.5 ESCs.

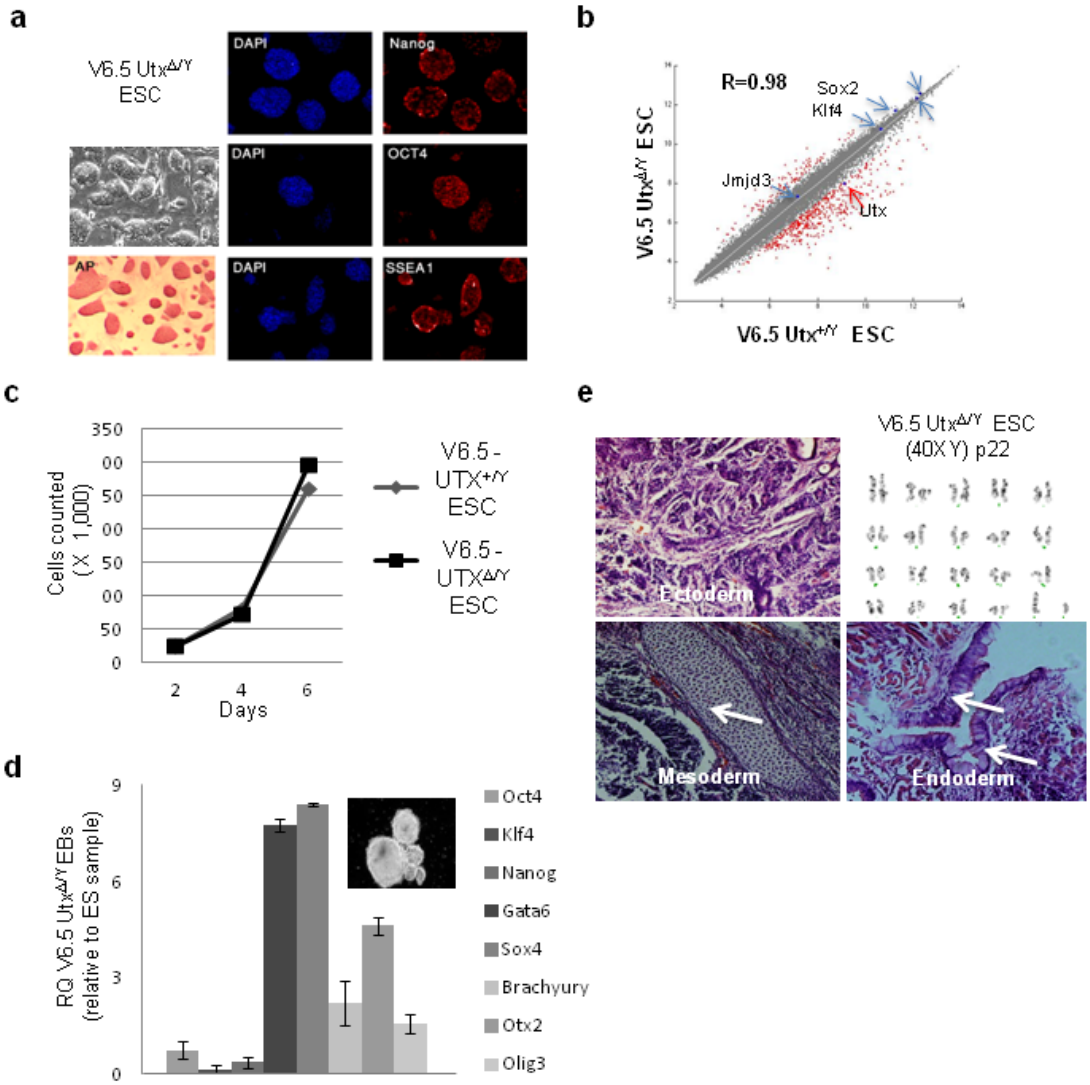

Figure 3. Utx is dispensable for maintenance of the pluripotency in mouse ESCs and already established iPSCs. a, Phase contrast and immunofluorescent staining images demonstrating homogenous expression of alkaline phosphates (AP), Nanog, Oct4 and SSEA1 on V6.5 $Utx^{\Delta/Y}$. **b,** Scatter plot of transcriptional levels measured in $Utx^{+/Y}$ (x-axis) and $Utx^{\Delta/Y}$ (y-axis) ESCs. Transcription levels are shown in log2 scale. Rank correlation between the samples is shown. **c,** Growth curves of V6.5 $Utx^{+/Y}$ and $Utx^{\Delta/Y}$ ESCs, demonstrating normal and comparable expansion capacity in vitro (1 out of 2 replicates is shown). **d,** V6.5 $Utx^{\Delta/Y}$ ESCs can differentiate *in vitro* into embryonic bodies and molecularly down-regulate Oct4, Klf4 and Nanog pluripotency markers. In parallel, they up-regulate multiple lineage commitment markers corresponding to the three lineages of development (mesoderm, ectoderm and endoderm). **e,** Representative histological analysis of V6.5 $Utx^{\Delta/Y}$ ESC derived teratomas. The cells displayed normal karyotype 40XY at the indicated passages (abbreviated as p).

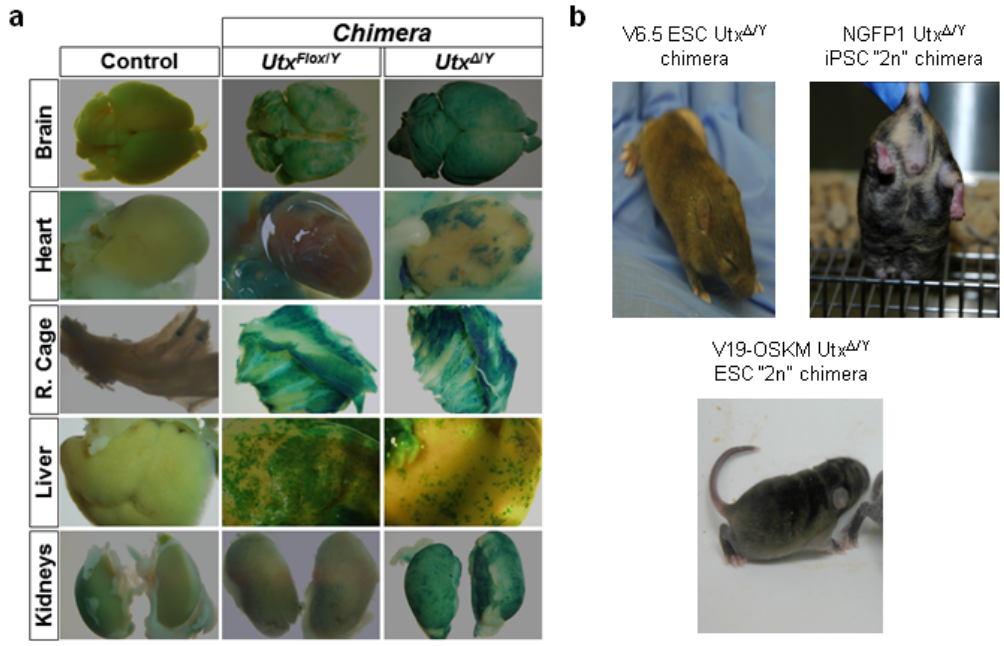

Figure 4. ***In vivo*** **developmental potential of Utx knockout pluripotent stem cells. a,** Whole mount X-Gal staining of indicated tissues extracted from 12 weeks old adult mice, demonstrating *in vivo* differentiation of $Utx^{\Delta/Y}$ ESCs. **b,** Adult chimeric animals obtained from V6.5 $Utx^{\Delta/Y}$, NGFP1 $Utx^{\Delta/Y}$ iPSC and V19-OSKM $Utx^{\Delta/Y}$ ESCs following microinjection into host blastocysts. For clarification, NGFP1 $Utx^{\Delta/Y}$ cells were generated by gene targeting in already established naïve NGFP1 iPSCs. Chimerism in adult animals is evident by the agouti coat color.

In order to test whether *Utx* is required for the re-establishment pluripotency, rather than its maintenance, we knocked out *Utx* in two different "secondary reprogramming" pluripotent stem cell lines (NGFP1 induced pluripotent stem cells (iPSCs) and V19 ESCs) carrying M2rTta and DOX inducible OSKM transgenes [55]. Genetically matched $Utx^{\Delta/Y}$ and $Utx^{flox/Y}$ cells were injected into host blastocysts, and differentiated mouse embryonic somatic fibroblasts (MEFs) were purified from chimeric animals (Fig 5a). Induction of pluripotency in these cells by DOX demonstrated a dramatic block in retrieving iPSCs from $Utx^{\Delta/Y}$ cells obtained from two independent reprogrammable transgenic systems with distinct OSKM stoichiometry. This phenotype was not restricted to embryonic fibroblasts as the reprogramming of $Utx^{\Delta/Y}$ B lymphocytes yielded a similar outcome. The restricted reprogramming capability of $Utx^{\Delta/Y}$ cells was also recapitulated in primed bFGF/Activin-containing growth conditions, in which somatic cells are reprogrammed into EpiSC-like state [56] (Fig. 5b). OSKM induced $Utx^{\Delta/Y}$ MEFs proliferated vigorously *in vitro* upon extended culturing with DOX and maintained fibroblast-like morphology (Fig. 4c). $Utx^{\Delta/Y}$ somatic cells had similar growth rates and background apoptosis levels compared to control $Utx^{+/Y}$ cells upon OSKM induction, thus excluding cell proliferation and apoptosis as potential factors underlying the observed deficiency in iPSC derivation. Long term follow-up for reprogramming of monoclonal populations of donor somatic $Utx^{\Delta/Y}$ B cells, indicated failure

to detect Nanog-GFP reactivation before 6 weeks of DOX induction. Notably, reprogrammed Nanog-GFP+ cells that could be derived at exceedingly low efficiency, lacked typical ESC morphology, retained a sporadic Nanog-GFP expression in serum-free 2i/LIF conditions, and were incapable of generating high-contribution chimeric animals, highly suggestive of aberrant reprogramming (Fig 4d). Taken together, these results demonstrate that absence of Utx demethylase constitutes a major roadblock for reprogramming, and suggests that Utx plays a critical role in re-establishing pluripotency *in vitro*.

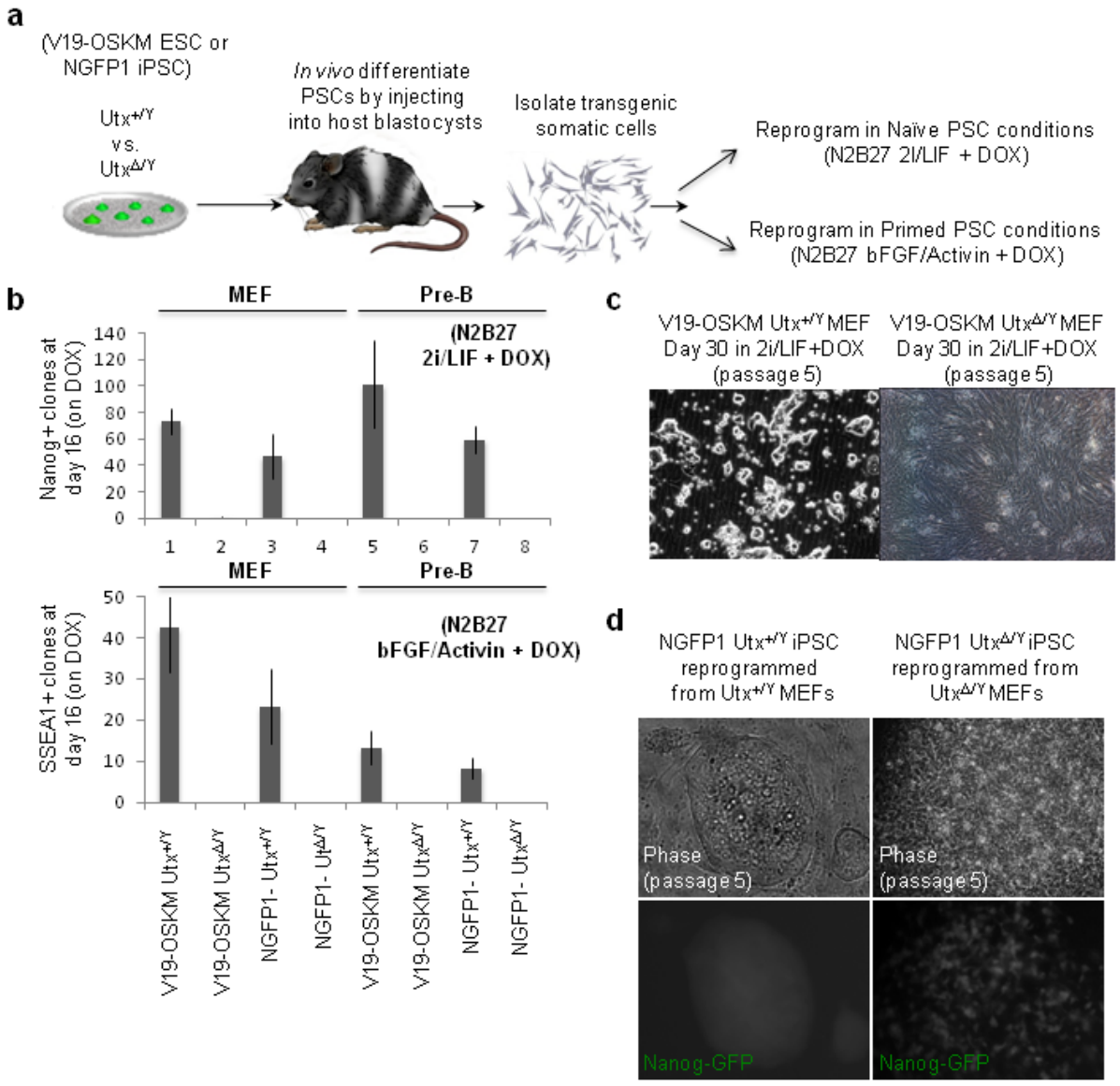

Figure 5. Comparative analysis of Utx$^{\Delta/Y}$ somatic cell properties during reprogramming. a, Scheme for secondary reprogramming conducted in our study. **b,** Reprogramming of 5000 secondary OSKM-transgenic MEF or Pre-B donor cells after 16 days on DOX in naïve growth conditions quantifying Nanog+ iPSC colonies (top) and in primed growth conditions quantifying SSEA1+ colonies (bottom). One representative experiment out of 3 performed is shown. **c,** Representative microscope images of somatic MEFs undergoing reprogramming. Note the persistence of fibroblast cells and absence of ES-like cells in Utx$^{\Delta/Y}$ cells despite

extended passaging. **d,** Microscope images indicating obtained Nanog-GFP+ cells from $Utx^{+/Y}$ and $Utx^{\Delta/Y}$ NGFP1 donor fibroblasts. Note the round ES like morphology and homogenous Nanog-GFP expression in 2i/LIF conditions (and absence of FBS) for $Utx^{+/Y}$ reprogrammed cells, while $Utx^{\Delta/Y}$ cells had flatted morphology and patchy Nanog-GFP expression despite the lack of serum and their expansion in defined N2B27 2i/LIF conditions.

Defining the epigenetic remodeling dynamics

To define the molecular basis of epigenetic remodeling dynamics during reprogramming facilitated by Utx, we conducted a global gene expression analysis of somatic cells during reprogramming and compared them to patterns observed in ESCs. The effect of *Utx* deletion on the induction, rather than on the maintenance, of the pluripotent state was evident at the transcriptional level. The number of genes affected by Utx knockout was 6.6 times higher in MEF samples following 8 days of DOX induction (3493 genes > 2-fold), compared to ESCs (527 genes > 2-fold) (Fig. 6a). Hierarchical clustering and Principle Component Analysis clustered $Utx^{\Delta/Y}$ DOX induced MEFs together with untreated $Utx^{\Delta/Y}$ and $Utx^{+/Y}$ MEFs (Fig. 6b,c), indicating that $Utx^{\Delta/Y}$ MEFs fail to profoundly initiate molecular changes typically detected following 8 days of OSKM induction, and remain highly similar to the original donor somatic cells (Pearson correlation = 0.97). Importantly, the genes which are activated (>2-fold) during reprogramming in $Utx^{+/Y}$ MEFs are enriched (FDR <1%) for targets of Klf4, Oct4 and Sox2 factors, as was measured in ESCs or during iPSC formation following OSKM induction in MEFs [57]. On the contrary, known targets of the latter reprogramming factors were preferentially enriched among the down regulated gene subgroup in $Utx^{\Delta/Y}$ DOX induced MEFs (>2-fold) (Fig. 6d). In both cases, the affected genes were highly enriched for H3K27me3 chromatin marks that were aberrantly accumulated in $Utx^{\Delta/Y}$ MEFs following DOX induction.

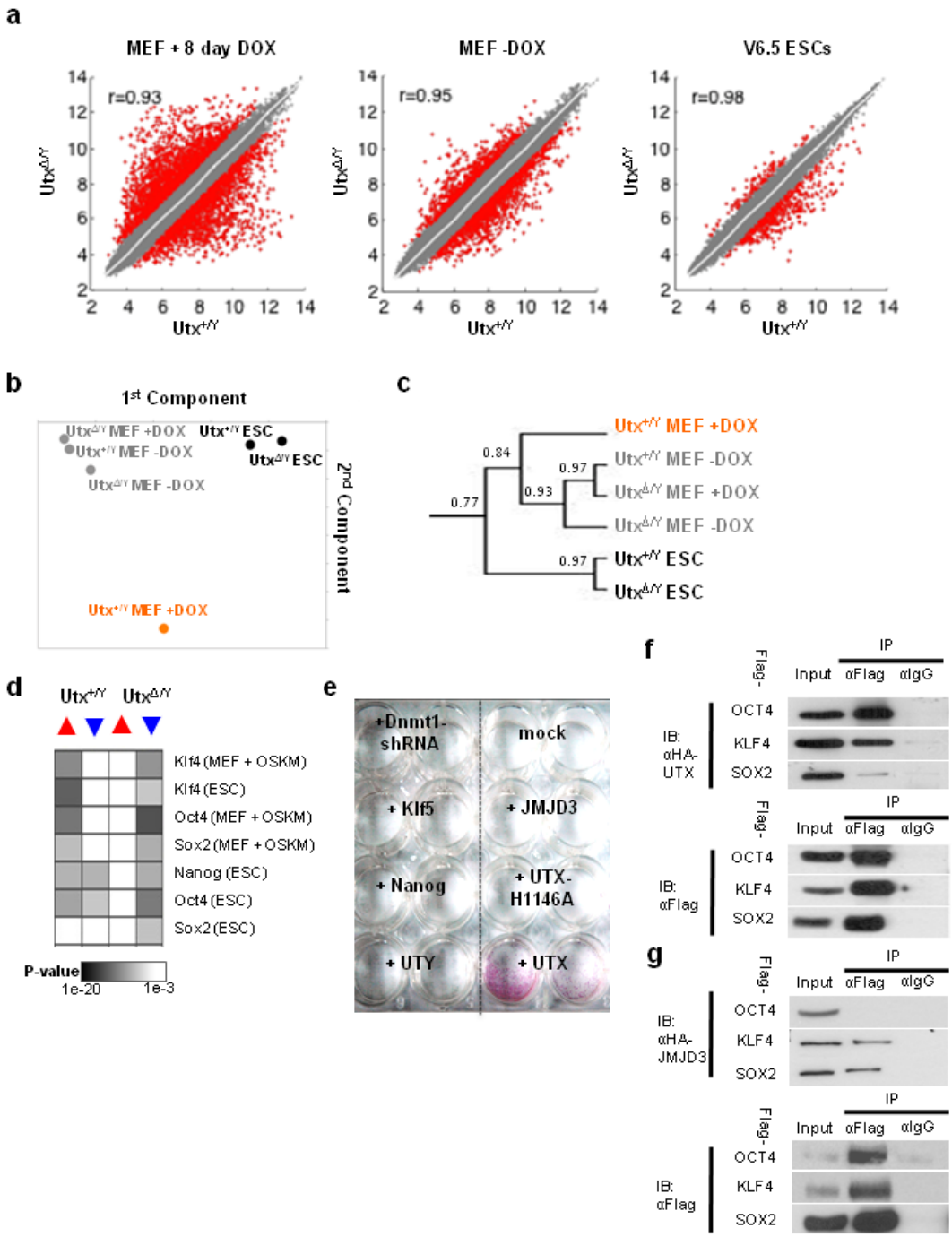

Figure 6. Utx-KO cells transcriptional profiles and genomic dissection. a, Scatter plot of transcriptional levels measured for 35,558 transcripts in $Utx^{+/Y}$ (x-axis), compared to $Utx^{\Delta/Y}$ (y-axis). Left – comparison in MEF+DOX, middle – comparison in MEF-DOX, right – comparison in V6.5 ESCs. Red – transcripts above 2-fold change. Transcription levels are shown in $\log_2$ scale. Rank correlation between the samples is shown at the top left corner of each plot. The number of genes that are differentially expressed (>2-fold) between $Utx^{+/Y}$ and $Utx^{\Delta/Y}$ is 1.8 times higher in MEF+DOX (3493) compared to MEF-DOX (1918), and 6.6 higher compared to ESCs (527). This indicates that the knockout affects the transcription during reprogramming, and in much less extent in MEF or in stable ESCs. **b,** Principle Component Analysis of transcriptional profiles of the indicated samples. Colors indicate distinct clusters: ESCs (black); MEF and MEF-like (gray); and DOX induced partially

reprogrammed pre-iPSCs (orange). **c,** Hierarchical clustering analysis of transcriptional profiles of the 6 indicated cell samples. Colors indicate distinct clusters: ESCs (black) MEF and MEF-like (gray) and DOX induced partially reprogrammed pre-iPSCs (orange). **d,** Selected functional categories enriched in induced (red) and repressed (blue) genes (> 2-fold), following 8 days of DOX in $Utx^{+/Y}$ and $Utx^{\Delta/Y}$ MEFs. Gray intensity correlates with *P*-value. **e,** $Utx^{\Delta/Y}$ somatic cells on DOX for 7 days were transduced with different rescue over-expression vectors and analyzed 7 days later for iPSC formation via staining for alkaline phosphatase staining. **f-g,** Co-immunoprecipitation (IP) of Flag-tagged OCT4, KLF4, or SOX2 with HA-tagged UTX or JMJD3, followed by an immunoblot analysis (IB).

Comparing Utx to other known H3K27 known demethylases

We next evaluated potential functional overlap in our system between Utx and other known H3K27 demethylases. Specific knockdown of Jmjd3 did not significantly reduce iPSC formation in mice and in humans. Further, Utx deficient somatic cells reprogramming could not be rescued by over-expression of JMJD3 or UTY, suggesting a distinct requirement for Utx in this process. Over-expression of Nanog and Klf5, that are required for pluripotency re-establishment, rather than maintenance [58, 59], was not sufficient to compensate for the absence of Utx during iPSC formation and suggesting that Utx does not exert its function through regulating only these genes. Utx has been shown to possess chromatin-remodeling histone demethylation independent functions [60], however $Utx^{\Delta/Y}$ iPSC formation rescue was observed by over-expressing wild-type UTX, but not UTX-H1146A allele that carries a mutation in the JmjC catalytic domain abrogating its H3K27 demethylase activity [32] (Fig. 6e).

Genome wide ChIP-seq analyses of active and repressive chromatin marks

We next mapped H3K27me3 and H3K4me3 chromatin modifications by genome-wide ChIP-seq analysis in MEFs before and after DOX induction, in genetically matched $Utx^{+/Y}$ and $Utx^{\Delta/Y}$ cell samples and in established ESCs. The perturbed active (H3K4me3) and repressive (H3K27me3) chromatin marks correlated with global active and repressed transcriptional levels, respectively [61]. Remarkably, the effect of *Utx* deletion during iPSC reprogramming was also apparent at the global chromatin level. While 73% of 14619 mapped genes had the same modifications in $Utx^{+/Y}$ and $Utx^{\Delta/Y}$ (Fig. 7a, rightmost column), the remaining 27% had aberrant H3K27me3 and/or H3K4me3 modifications in $Utx^{\Delta/Y}$ cells. A total of 957 genes retained or aberrantly gained H3K27me3 repressive mark (Fig. 7a, dark green and light green bars, respectively) during reprogramming in $Utx^{\Delta/Y}$ and not in $Utx^{+/Y}$ MEFs. More than 500 genes became potentially active by losing H3K27me3 and gaining or maintaining H3K4me3 active chromatin mark in $Utx^{+/Y}$, but not in $Utx^{\Delta/Y}$ DOX induced MEFs (Fig. 7a, purple bars). Notably, *Utx* deletion resulted in global hypermethylation of

H3K27me3 in DOX induced MEFs but not in ESCs. This trend is consistent with our functional data indicating that Utx predominantly regulates pluripotency re-establishment rather than maintenance. H3K4me3 was under-methylated in DOX induced MEFs but not in ESCs, possibly resulting from indirect effect of perturbation in H3K27me3 repressed genes and/or from the fact that Utx partners with Mll2/3 H3K4me3 transferase complex [32, 33], and absence of Utx might alter its activity. Consistent with the gene-expression enrichment analysis in Fig. 6d, OSKM and Nanog target genes in ESCs and MEFs undergoing reprogramming [57] were found to be specifically hypermethylated for H3K27me3 and under-methylated for H3K4me3 in $Utx^{\Delta/Y}$ DOX induced MEFs in comparison to control genetically matched cells (Fig 7b).

Utx genome wide binding sites

ChIP-seq analysis for Utx binding in DOX induced MEFs, identified a global increase in Utx binding, compared to non-induced MEFs. Utx specifically bound 1845 target genes that are enriched for H3K4me3 active chromatin marked genes as well as for targets of Nanog and Oct4 (Fig 7c). These results were detected despite the fact that DOX induced MEFs are highly heterogeneous cell population with only 5% of the cells undergoing reprogramming. Notably, ChIP-seq analysis for Utx binding in already established ESCs identified 2416 bound genes that are enriched for H3K4me3 active chromatin marked genes as well as for targets of Oct4, Sox2 and Klf4 (Fig 7c). Motif search analysis in Utx binding locations identified 222 motifs, where the motifs of Klf4, Sox2, Oct4 and Nanog are the four most highly enriched (Fig 7d).

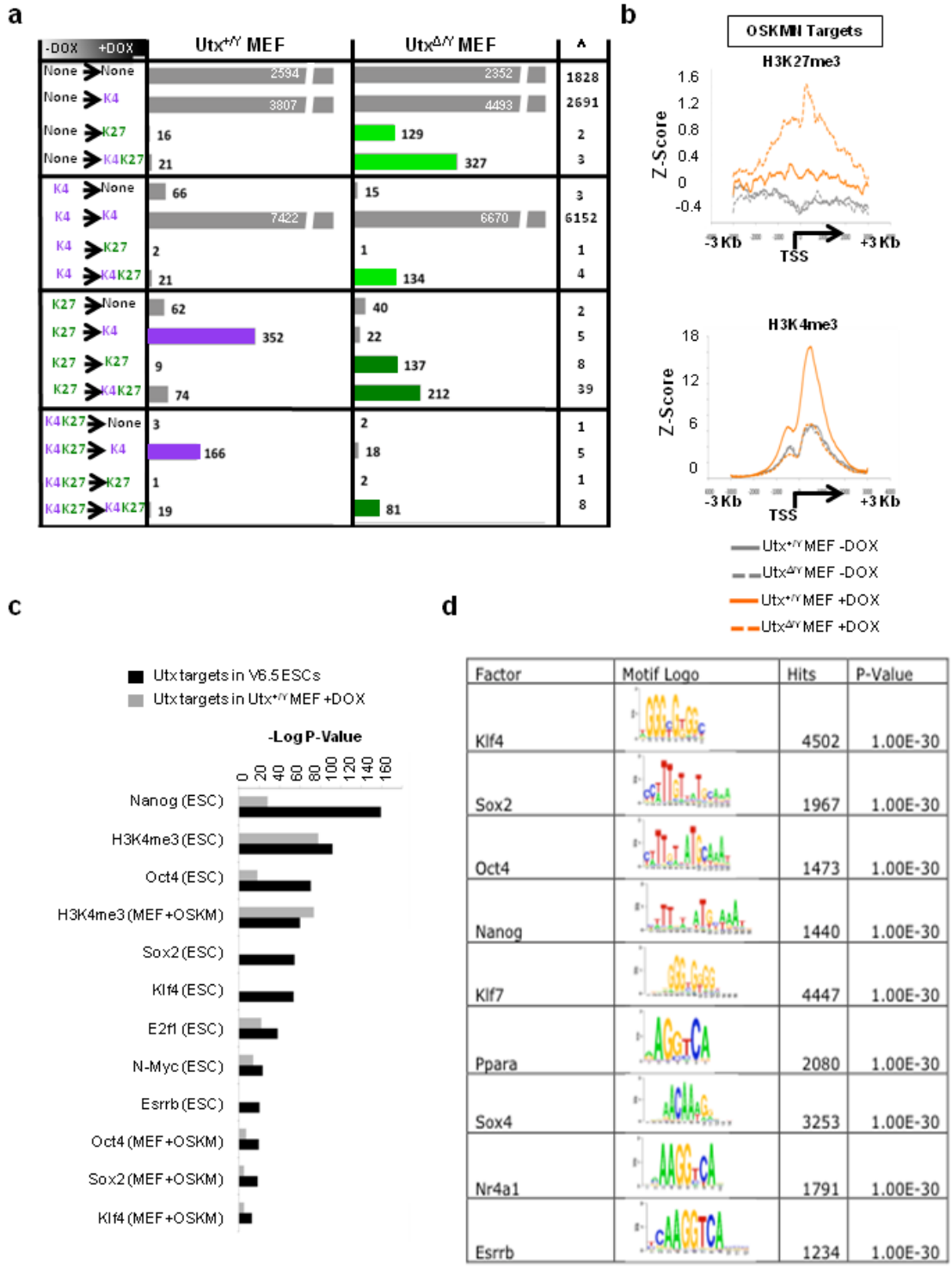

Figure 7. Chromatin and Utx genomic binding profiles in pluripotency and reprogramming. a, Number of genes with H3K4me3 or H3K27me3 before and after DOX induction in $Utx^{+/Y}$ and $Utx^{\Delta/Y}$ MEFs. Light Green: gene number that gained H3K27me3 in $Utx^{\Delta/Y}$ only, Dark green: gene number that retained with H3K27me3 in $Utx^{\Delta/Y}$ only. Purple: gene number that gained/maintained H3K4me3 in $Utx^{+/Y}$ only. Rightmost column: overlapping gene number in each category. **b,** Binding profile around the TSS of H3K4me3 and H3K27me3. Profiles were calculated over 805 genes that are bound by at least 3 of the following factors: Oct4, Sox2, Klf4, c-Myc and Nanog. Z-score of the mean profile is plotted (y-axis) against a sliding window around the TSS (x-axis). **c,** Top categories enriched among genes bound by Utx in ESCs (black) or in MEFs+DOX (gray), following genome-wide ChIP-seq analysis conducted either on ESCs or MEFs following OSKM transduction (as indicated in brackets). –Log *P*-value is presented for each category. **d,** Top motifs abundant among genes that are bound by Utx in ESCs. Shown are the sequence logos of the motifs, along with the associated factors, the number of hits and the p-values. The motifs were inferred using SeqPos software in Cistrome package.

Identifying Utx biochemical partners

The above findings prompted us to examine whether UTX physically associates with the OSK reprogramming proteins. Using HA-tagged UTX, we saw that Flag-Tagged OCT4, KLF4, and to lesser extent, SOX2, were co-immunoprecipitated with UTX (Fig. 6f), whereas co-immunoprecipitation was not observed with a Flag-tagged NANOG protein. These data indicate that UTX can associate with OKS by a protein–protein interaction and thereby might promote the adequate reactivation of their target genes during reprogramming. Overall, the above analysis provides evidence for aberrant transcriptional and epigenetic reprogramming dynamics upon OSKM induction in Utx knockout somatic cells, and suggests that genes regulated by the pluripotency reprogramming machinery cannot be adequately reactivated in the absence of Utx during this process. JMJD3 was not able to interact with OCT4, partially explaining the distinct influence for different H3K27 demethylases on iPSC formation (Fig. 6g).

Derepression of key pluripotency genes

We then took unbiased approach to reveal key genes that are deregulated early during the reprogramming process by the loss of Utx. We identified 1,430 genes that were transcriptionally upregulated (>2-fold) in $Utx^{+/Y}$ and not in $Utx^{\Delta/Y}$ MEFs undergoing reprogramming. Of these, 98 genes had aberrantly upregulated H3K27me3 marks in $Utx^{\Delta/Y}$ doxycycline-induced MEFs (Fig. 8a), suggesting that their deregulation is likely to be caused by H3K27me3 retention. The signature of these deregulated genes is enriched (FDR<1%) for stem- cell maintenance factors (for example, Fgf4, Sall4, Sall1 and Utf1), for validated direct targets of Klf4 and Oct4 (for example, Sall1, Crb3 and Nef1), and for H3K27me3-marked genes. Of these, we focused on three transcription factors, Sal1l, Sall4 and Utf1, previously shown to undergo reactivation after OSKM induction in wild-type cells [61]. These genes were directly bound by Utx in ESCs and in doxycycline-induced MEFs, and aberrantly retained H3K27me3 in $Utx^{\Delta/Y}$ cells, as was confirmed by ChIP-seq and ChIP–PCR analyses (Fig. 8b). This suggested that Utx can act during early reprogramming by directly regulating the reactivation of a variety of pluripotency promoting factors including Sall1, Sall4 and Utf1. Indeed, inhibiting the endogenous expression of Sall1, Sall4 and Utf1 in $Utx^{+/Y}$ MEFs through specific knockdown during reprogramming, recapitulates the block in iPSC formation observed in $Utx^{\Delta/Y}$ cells (Fig. 8c). Overexpression of Sall1, Sall4 and Utf1 in $Utx^{\Delta/Y}$ MEFs promoted the activation of early reprogramming marker SSEA1, but did not result in reactivation of Nanog–GFP, suggesting that other Utx-regulated pluripotency factors need to be reactivated to complete the reprogramming progression. This is consistent with the role of

Utx as global regulator that targets multiple genes (Fig. 7c). Remarkably, consistent with a profound pluripotency promoting potential for Sall1, Sall4 and Utf1, their combined overexpression with Nanog and Myc in wild-type $Utx^{+/Y}$ MEFs replaces exogenous OSK expression, and generated chimaera- formation-competent iPSC clones (Fig. 8d).

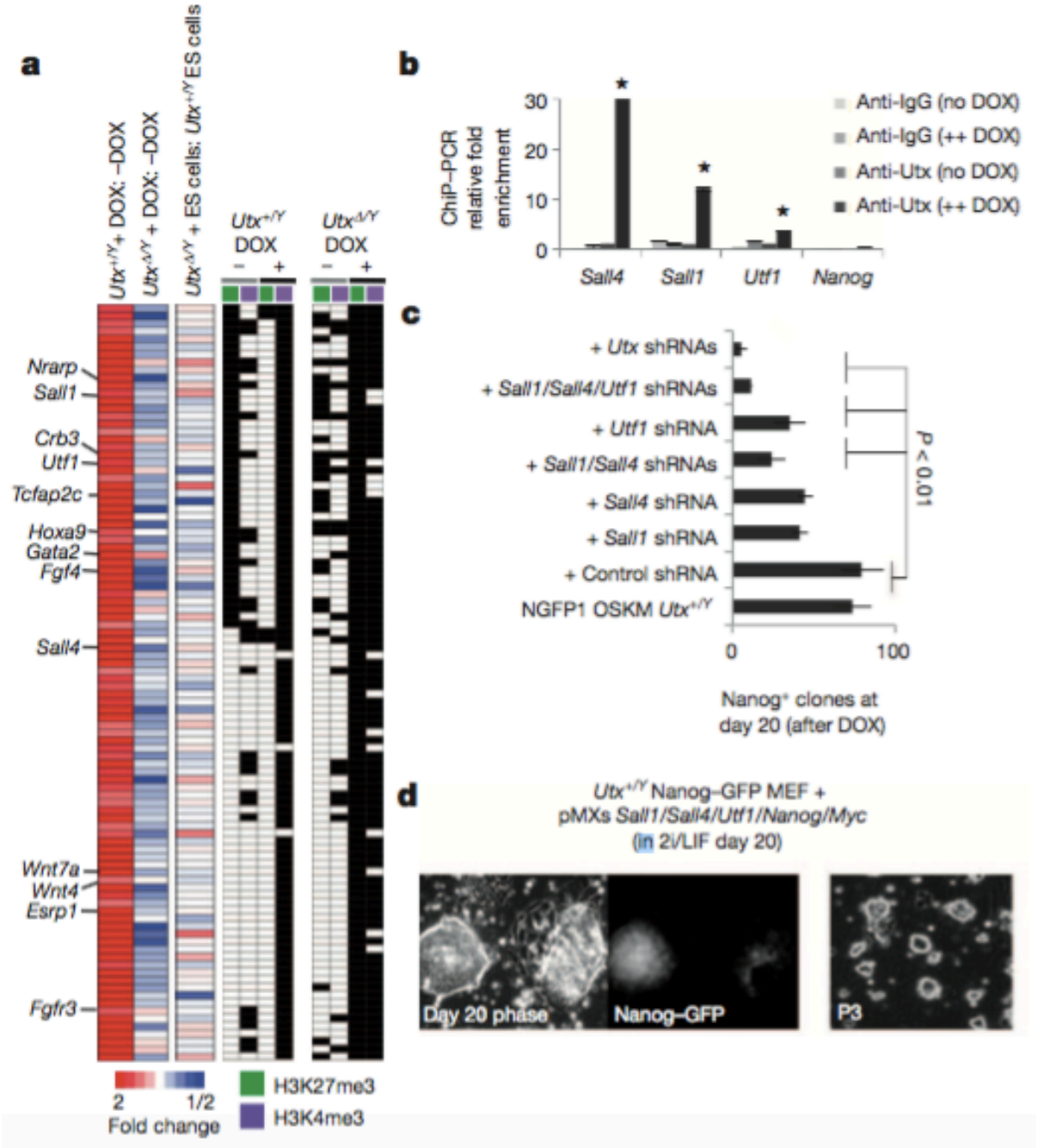

Figure 8. Downstream effectors of Utx that promote iPSC reprogramming. a, Unbiased identification of 98 deregulated genes that were induced (.2-fold) in $Utx^{+/Y}$ but not in $Utx^{\Delta/Y}$. Left, expression change before and after 8 days of doxycycline (DOX) induction in $Utx^{+/Y}$ and $Utx^{\Delta/Y}$ MEFs, and in ESCs (comparing $Utx^{\Delta/Y}$ to $Utx^{+/Y}$). Right, chromatin marks before and after doxycycline induction. Purple indicates H3K4me3 modifications; green indicates H3K27me3 modifications; black boxes, methylated; white boxes, unmethylated. **b,** Anti-Utx and control anti-IgG ChIP–PCR relative fold enrichment of Sall4, Sall1, Utf1 and Nanog loci in $Utx^{+/Y}$ and $Utx^{\Delta/Y}$ MEFs before and after doxycycline treatment (normalized to $Utx^{+/Y}$ without doxycycline) (n=3). **c,** Reprogramming following individual and combined knockdown with Sall4, Sall1 and Utf1 were

introduced in $Utx^{+/Y}$ Nanog–GFP secondary MEFs (n=2). *P value <0.01 in comparison to wild-type samples. Eror bars indicate s.d. **d,** Overexpression of Sall4, Sall1, Utf1, Nanog and Myc reproducibly gives rise to Nanog–GFP exogenous OSK. P3, passage 3.

We next investigated whether Utx has a role in epigenetic reprogramming during early development. Whereas ESCs could be readily derived from $Utx^{\Delta/Y}$ blastocysts, we noticed that chimeric animals generated from three different $Utx^{\Delta/Y}$ iPSC and ESC clones reproducibly failed to demonstrate germline transmission after extensive breeding (Fig. 9a). Importantly, multiple adult chimaeras generated from three different corrected transgenic lines with catalytically active UTX, but not the mutant UTX^{H1146A} allele, robustly achieved germline transmission (Fig. 9a).

Primordial germ cells (PGCs) increase global levels of H3K27me3 by E8.5–E9.5, and then undergo a second wave of epigenetic reprogramming at E10.5–E11 that entails global loss of the H3K27me3 mark [62]. Thus we tested whether Utx deficiency results in perturbed PGCs development and reprogramming *in vivo* during this critical developmental window. To reproducibly study PGCs development beyond E9.5–E10.5, we analyzed chimeric animals after microinjection of either V6.5 $Utx^{\Delta/Y}$ ESCs labelled with LacZ, or the BVSC ESC line [63] knocked out for Utx ($Utx^{\Delta/Y}$; BVSC), which carries Blimp1–mVenus (Blimp1 also known as Prdm1) and Stella–CFP (Stella also known as Dppa3) early PGCs reporters. $Utx^{\Delta/Y}$ PGCs were found to express the PGC markers Oct4, Nanog, Sall4 and SSEA1 normally at E10.75 (Fig. 9b). However, only Utx-null PGCs did not undergo H3K27me3 demethylation during this developmental window (Fig. 9c). The remaining chromatin marks tested were not different at this stage, suggesting the aberrant accumulation of H3K27me3 at this critical developmental stage may be an early identifiable trigger for subsequent PGCs developmental defects. Indeed, despite the adequate localization of Blimp1- and Stella-positive $Utx^{\Delta/Y}$ PGCs in the gonads at E12.5, they fail to retain the expression of several PGC pluripotency markers (Nanog, Sall4, Oct4 and SSEA1) or show reactivation of the late PGC marker Mvh (Fig. 9b). In addition, $Utx^{\Delta/Y}$ PGCs show aberrant chromatin dynamics at E12.5, as evident from the low chromatin levels of H3K4me2/3 and H3K27me3, and reduced nuclear size in comparison to adjacent wild-type PGCs (Fig. 9c). Finally, we find that $Utx^{\Delta/Y}$ PGCs isolated from E8.5 chimeras are deficient in their ability to revert to ground state pluripotency and form ES-like embryonic germ cells.

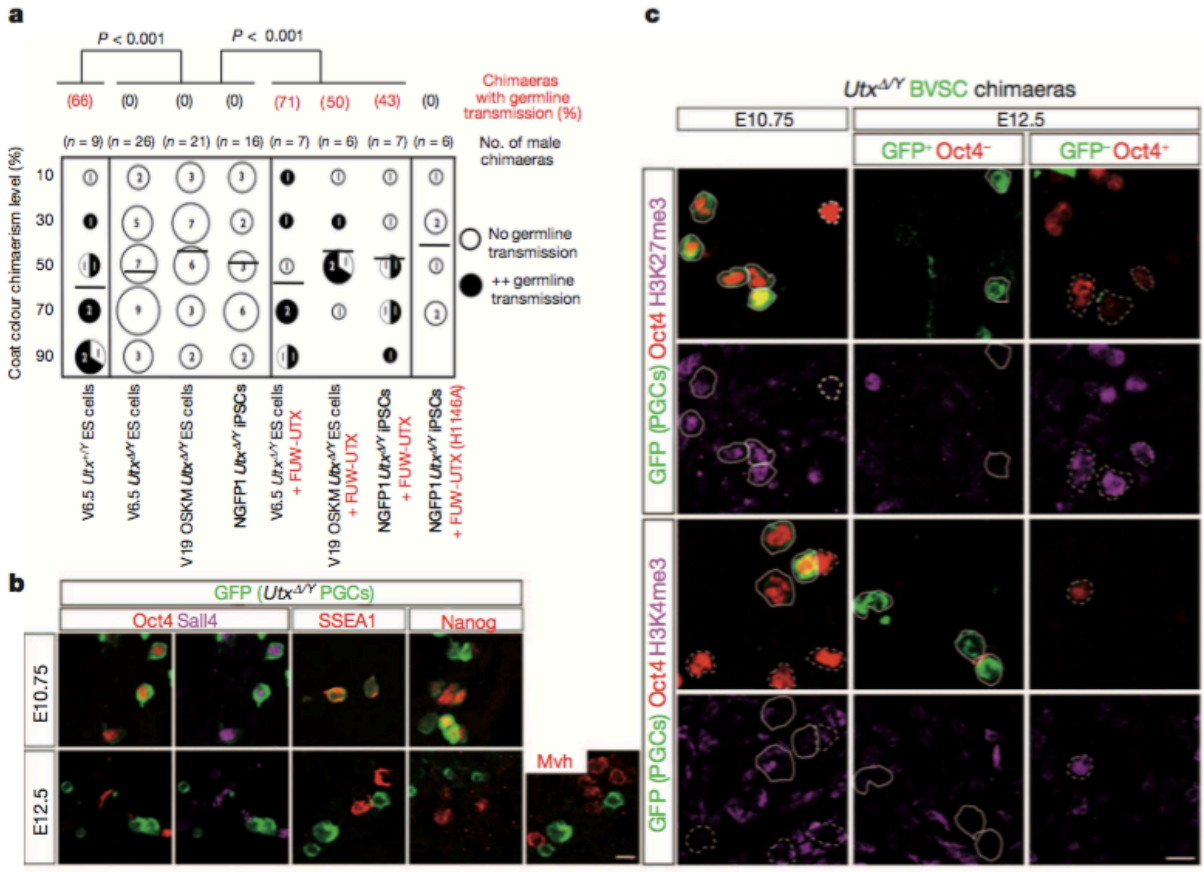

Figure 9. Utx regulates epigenetic reprogramming and development of PGCs. a, Quantification of coat color chimerism in male mice derived from the indicated cell lines. Each pie chart indicates a group with the indicated chimerism levels. Black areas indicate successful germline transmission; white areas indicate germline transmission failure. P values are indicated. **b, c,** Confocal images represent gonad sections from E10.75 or E12.5 chimeric embryos injected with $Utx^{\Delta/Y}$ BVSC ESCs and immunostained for the indicated proteins. **b,** Immunostaining for GFP (labelling BVSC $Utx^{\Delta/Y}$ injected cells) and the indicated PGC markers revealed abnormal *in vivo* differentiation of Utx-deficient PGCs at E12.5 but not at E10.75. **c,** Immunofluorescence staining for epigenetic markers on the genital ridge of E10.75 embryos demonstrates that high levels of H3K27me3 are retained in Utx-deficient PGCs (unbroken circles) compared with adjacent wild-type PGCs (dashed circles). At E12.5 Utx- deficient PGCs have diminished levels of all tested epigenetic marks. Scale bars, 10 µm.

Part II: Novel human naïve state growth conditions enables human-mouse interspecies chimerism

Most of the work describes in this chapter was published in:
"Derivation of novel human ground state naive pluripotent stem cells". Nature, 2013 [64].

Defining human naïve PSC conditions

It is feasible that different exogenous factor combinations, in addition to 2i/LIF, may allow *in vitro* stabilization of transgene-independent and indefinitely stable human naive pluripotent cells [40]. In order to define such conditions we used 'secondary' human C1.2 iPSC line containing doxycycline (dox)-inducible lentiviral transgenes encoding OCT4, SOX2, KLF4 and c-MYC (OSKM) reprogramming factors, that was targeted with an OCT4–GFP–2A–PURO knock-in reporter construct (Fig. 10a) [42]. As previously established, exogenous expression of reprogramming factors encoding transgenes by dox supplementation allows maintenance of cells that are morphologically similar to mouse ESCs, while retaining approximately 60% OCT4–GFP$^+$ cell fraction in 2i/LIF conditions [42]. This cell line was used to screen for components that, upon dox withdrawal, could stabilize C1.2 human iPSCs in 2i/LIF indefinitely with homogenous OCT4–GFP$^+$ expression (in nearly 100% of the cells) (Fig. 10a).

Whereas C1.2 cells rapidly differentiated in 2i/LIF only conditions, the combined action of 16 factor conditions (16F, divided into pool 1 and pool 2 subgroups) attenuated the differentiation propensity and allowed retaining of 32% OCT4–GFP$^+$ cells as measured at day 14 after dox withdrawal (Fig. 10b). This indicated that the 16F combination contains factors that cooperatively promote human pluripotency maintenance in 2i/LIF conditions. When pool 2 components were removed, OCT4–GFP$^+$ cell fraction increased relative to 16F combination, indicating that pool 2 contained factors that were negatively influencing GFP cell maintenance (Fig. 10b). We propose that FGFR and TGFR pathway inhibition was detrimental to growing human pluripotent cells in 2i/LIF conditions, as FGF2 and TGFβ cytokines have an evolutionary divergent function in promoting pluripotency in humans by inducing naive pluripotency KLF4 and NANOG transcription factor expression in human ESCs [65], but not in murine EpiSCs, where they promote murine pluripotency priming. Indeed, removing both TGFR and FGFR inhibitors recapitulated the phenotype obtained when removing dox and pool 2 components (Fig. 10b). Moreover, pool 1 components supplemented with exogenous FGF2 and TGFβ cytokines resulted in homogenous OCT4–GFP$^+$ detection in >95% of cells independent of dox in 2i/LIF containing conditions (Fig.

10b). We then tested which of the pool 1 components were essential, and found that 2i/LIF, p38i, JNKi together with FGF2 and TGFβ cytokine supplementation were essential to maintain exogenous transgene-independent GFP C1.2 clones (Fig. 10c). Secondary optimization identified Rho-associated coiled-coil kinases (ROCK) [66] and protein kinase C (PKC) [67] inhibitors (included in pool 2, Fig. 10a) as optional beneficial boosters of naive cell viability and growth, and resulted in optimized chemically defined conditions termed NHSM (naive human stem cell medium) (Fig. 10d). NHSM conditions enabled expansion of karyotypically normal OCT4–GFP^{+} C1.2 human iPSCs for over 50 passages independent of exogenous transgene activation (Fig. 10e). Remarkably, hPSCs grown in NHSM have domed-shaped colonies resembling murine naive cells.

Derivation of human ESCs with NHSM

We examined whether NHSM conditions allow derivation of new human ESC lines from human blastocysts. Human-blastocyst-derived inner cell masses were plated in NHSM conditions, and successfully generated domed cell outgrowths after 6–8 days (Fig. 10f). We were able to establish four newly derived naïve human ESC lines termed LIS1, LIS2, WIS1 and WIS2 (Fig. 10f). Several conventional (hereafter will be named 'primed') already established human ESC lines (H1, H9, BGO1, WIBR1, WIBR2 and WIBR3) and human iPSC lines (C1 and C2) were plated on gelatin/vitronectin coated dishes in NHSM medium. After 4–8 days, dome-shaped colonies with packed round cell morphology could be readily isolated and further expanded (Fig. 10g). Human fibroblast cells were reprogrammed to human iPSCs in NHSM following reprogramming factor transduction. All human ESC and iPSC lines expanded in NHSM conditions were positive for pluripotent markers and formed mature teratomas *in vivo* (Fig. 11). Human naive pluripotent lines maintained normal karyotype after extended passaging following trypsinization and expansion on irradiated MEF feeder cells or on vitronectin/gelatin-coated plates (Fig. 11a). The average doubling time was significantly reduced from 26 h for primed hPSCs down to 14 h for naive hPSCs. Naive hPSCs displayed up to 35% single-cell cloning efficiency after trypsinization and sorting (without the use of ROCK inhibitors), whereas primed hPSCs largely did not survive single-cell cloning. In the presence of ROCK inhibitor, naive hPSCs had single-cell cloning efficiency of up to 88%, whereas that of primed cells increased only up to 22%.

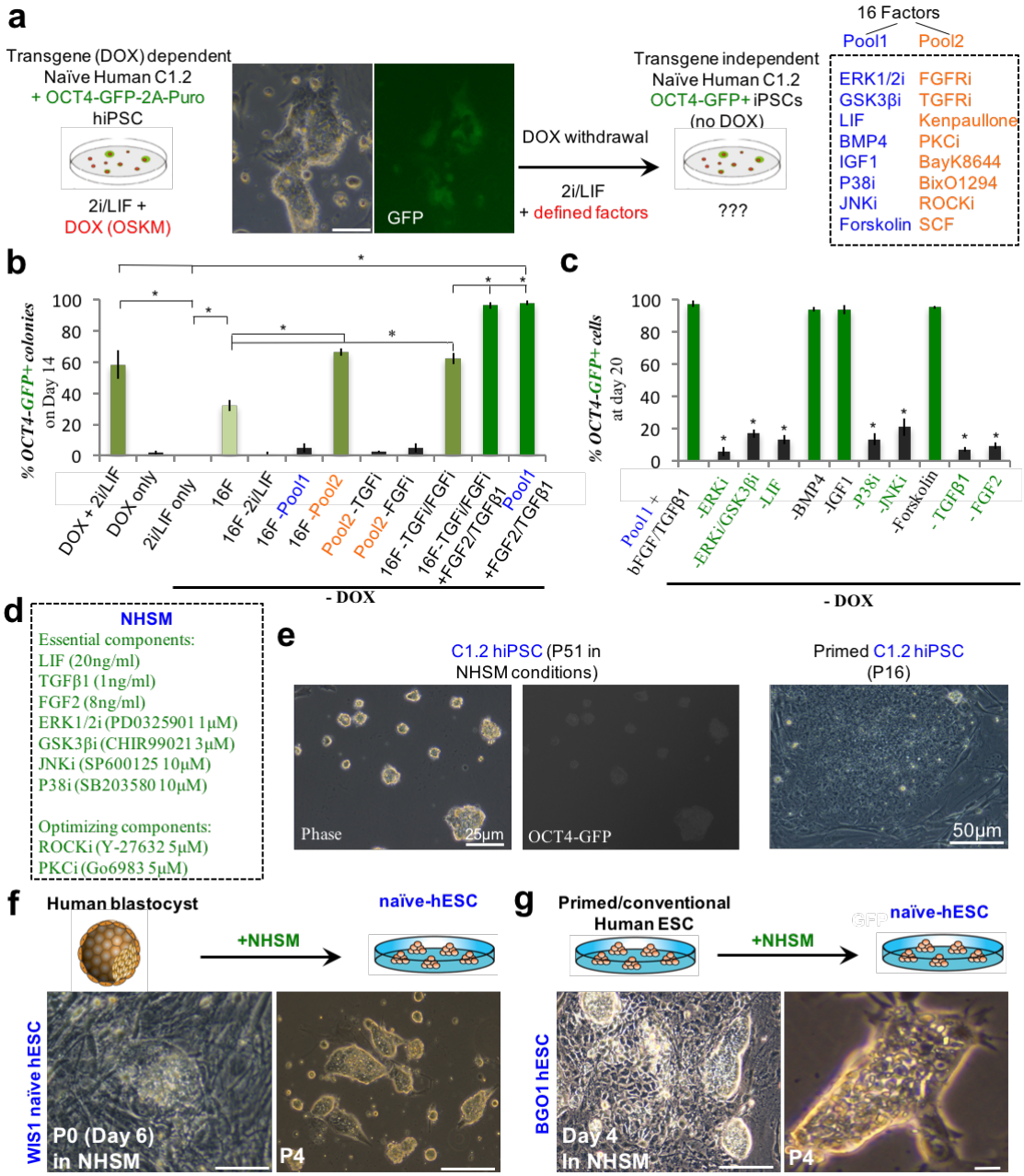

Figure 10. Capturing human naive pluripotency. a, Schematic drawing of the strategy used for calibrating conditions to isolate naive transgene-independent human iPS cells (hiPSCs) in the presence of 2i/LIF. Inhibitor is abbreviated as 'i' in the dotted box. Scale bar, 200 μm. **b,** The percentage of OCT4–GFP$^+$ colonies obtained in different combinations of indicated factors as measured at day 14 ('2' indicates without, '1' indicates with). Error bars indicate s.d. (n=3). *t-test P value, 0.01. c, Cells were expanded in the presence of FGF2/TGF-b1 and pool 1 factors, and the essential factors required for maintaining OCT4–GFP$^+$ cells in the absence of dox were determined by screening for loss of GFP upon withdrawal of these components (green font). *t-test P value<0.01. Error bars indicate s.d. (n=4). d, Components of optimized NHSM conditions. e, Representative large-field view of OCT4–GFP$^+$ human iPSC colonies grown in NHSM (left) and, for comparison, primed/ conventional iPSCs (right). Scale bars, left, 200 mm; right, 50 mm. **f,** Human-blastocyst- derived inner cell masses were plated on feeder cells in NHSM conditions. At day 6–8, the original outgrowth was trypsinized, and naive pluripotent cell lines were established in NHSM conditions (representative images of established line at the indicated passage (P) are shown). hESC, human embryonic stem cell. Scale bars, left, 50 μm; right, 200 μm. **g,** Epigenetic reversion of already established primed/conventional hESCs to naive ground state. Representative images of BGO1 hESC line are shown. Scale bars, left, 200 μm; right, 30 μm.

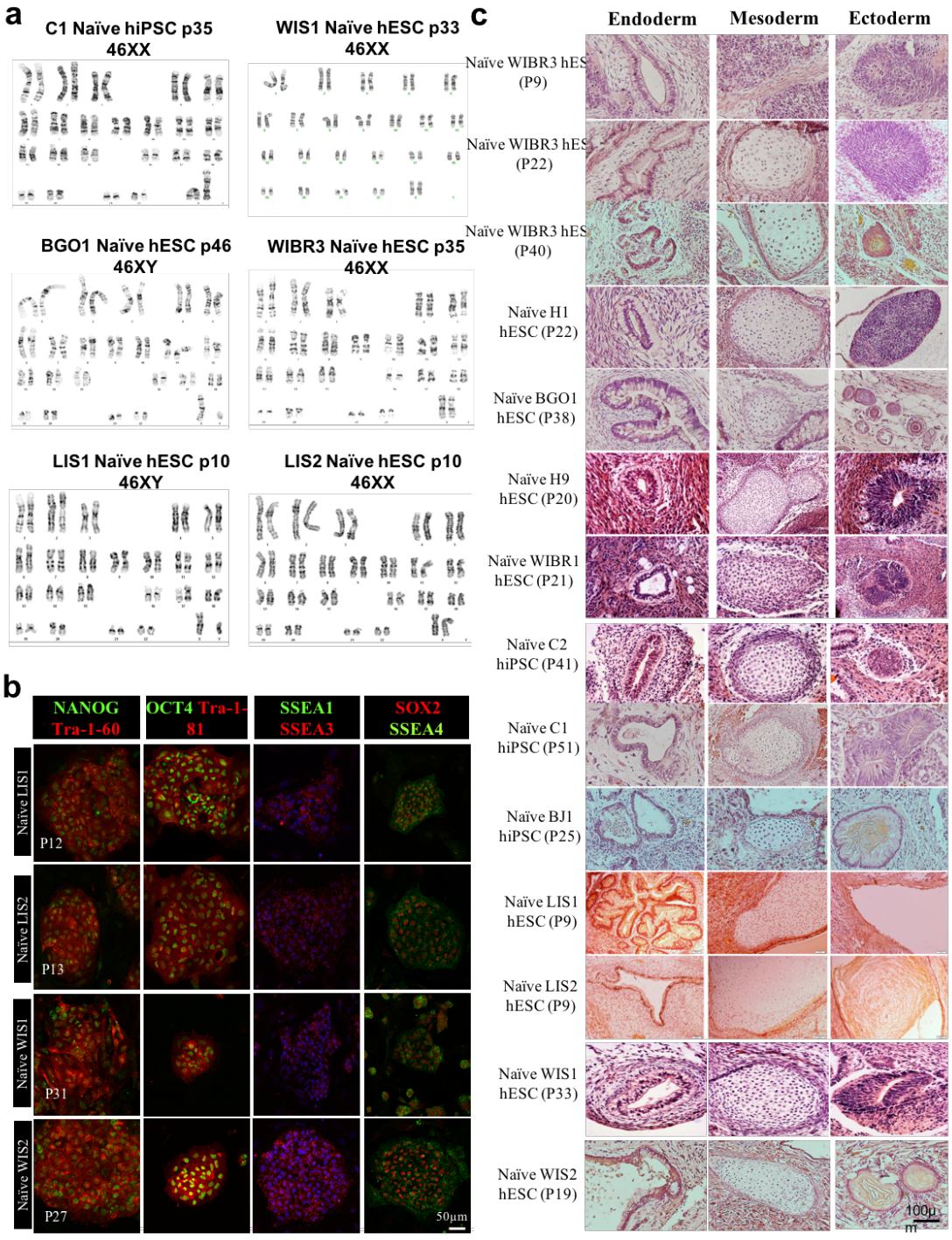

Figure 11. Validation of human naive pluripotency. a, Karyotype analysis results indicating normal karyotypes of different naive hESCs and hiPSCs. The passage number at which cells were collected for karyotyping is indicated. **b,** The newly derived naive LIS1, LIS2, WIS1, and WIS2 human ESC lines established from human inner cell mass in NHSM conditions were analyzed, and showed a strong uniform staining for all indicated human pluripotency markers (OCT4, NANOG, SSEA3, SSEA4, TRA1-60, TRA1-81 and SOX2). Notably, SSEA1, which is specific for mouse (both naive and primed stem cells) and not human pluripotent cells, is not expressed on naive hESCs. Similar results were validated for all human naive PSCs derived in this study (data not shown). **c,** *In vivo* teratoma differentiation of naive hESCs and iPSCs. Naive pluripotent stem cell lines were expanded in NHSM conditions for the indicated number of passages. All tested lines gave rise to well differentiated mature teratomas with cells from the three germ lineages: endoderm, ectoderm and mesoderm. P indicates passage number in NHSM conditions at which cells were collected and injected.

Epigenetic features of human PSCs grown in NHSM

We next aimed to characterize epigenetic features in naive human iPSCs and human ESCs established in NHSM conditions. As described in previous reports, primed human ESCs show preferential activation of the *OCT4* proximal enhancer (PE) element as typically seen in murine EpiSCs [42, 51]. Utilization of the *OCT4* distal enhancer (DE) is detected in naive human ESCs and human iPSCs. To further substantiate these findings, WIBR3 human ESCs were stably transfected with full-length OCT4–GFP– 2A–PURO, ΔPE–OCT4–GFP–2A–PURO or ΔDE–OCT4–GFP–2A– PURO engineered BAC reporter constructs. The wild-type OCT4–GFP reporter was specifically active in both naive and primed conditions (Fig. 12a). The DPE–OCT4–GFP reporter was predominantly active in naive growth conditions whereas the DDE–OCT4–GFP reporter was more active in primed pluripotent cells (Fig. 12a). We then analyzed the frequency and properties of X inactivation state[12] in several naive human ESC/iPSC lines. Naive hPSCs captured in NHSM maintain a pre-X inactivation state as evident by nearly complete lack of H3K27me3 nuclear foci and down regulation of XIST transcription. The majority of primed hPSCs demonstrated X inactivation as evident by the presence of H3K27me3 nuclear foci and methylation of one of the XIST gene alleles (Fig. 12b-f) [68]. Genome-wide mapping of H3K9me3 by chromatin immunoprecipitation, followed by deep sequencing (ChIP-seq) in female and male naive and primed human pluripotent cells, indicated a significant increase ($P < 3.8x10^{-63}$) in this mark on the X chromosome only in female primed pluripotent cells (Fig. 12e). XIST promoter alleles are demethylated in male and female naive ESCs (Fig. 12f). Upon differentiation of female naive human ESCs/iPSCs, inactivation of one of the X chromosomes alleles becomes evident as the cells demonstrate H3K27me3 clouds, upregulate XIST transcription simultaneously with methylation of one of the promoters of XIST alleles (Fig. 2b-f).

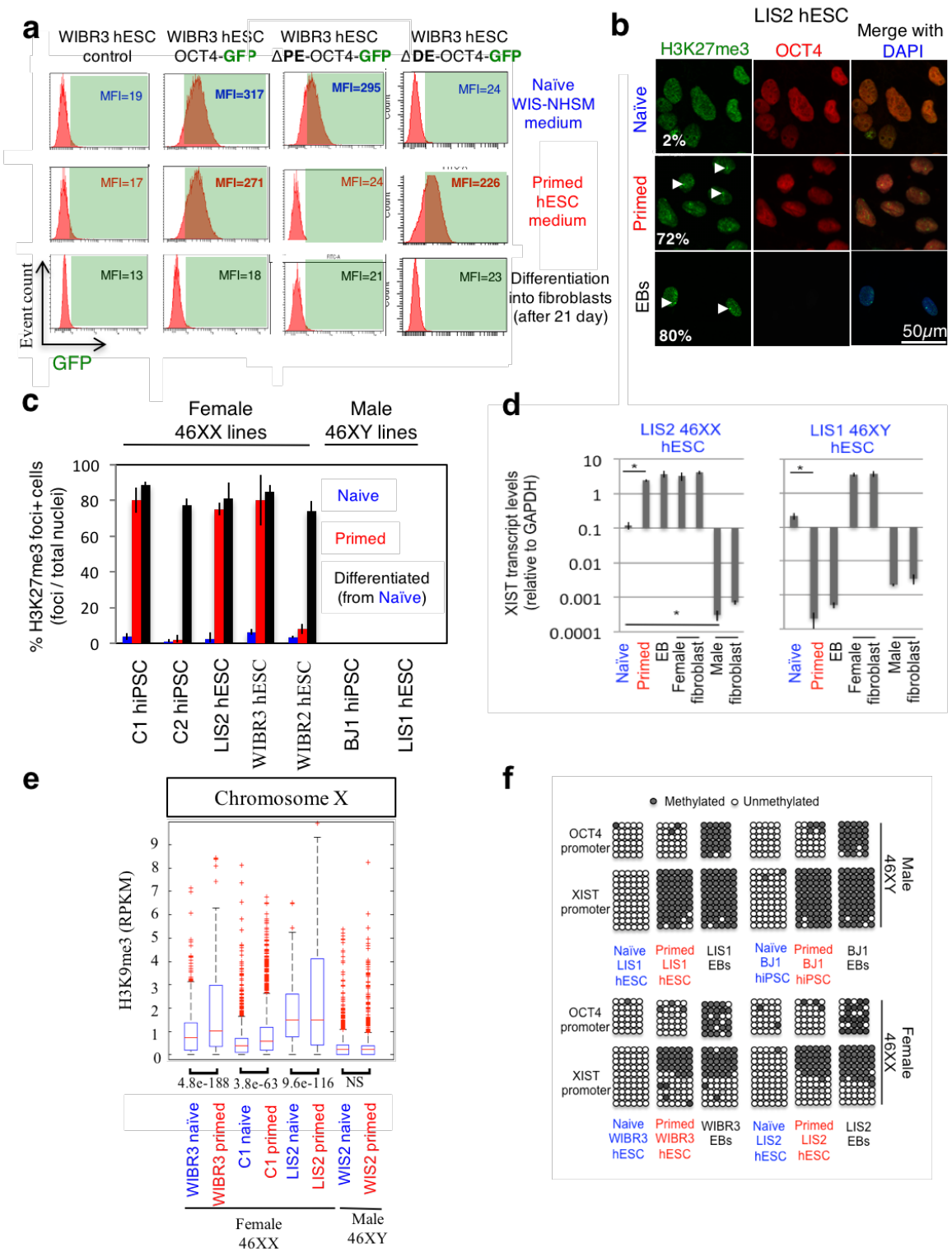

Figure 12. Epigenetic properties of human naive state pluripotency. a, WIBR3 human ESCs were stably transfected with the different engineered BAC reporter constructs indicated. GFP levels were evaluated in genetically matched cells by flow cytometry. MFI, mean fluorescence intensity. Average MFI values of 3 biological replicates per sample are indicated. **b,** Representative confocal images obtained after double immunostaining for OCT4 and H3K27me3 on naive, primed and differentiated samples (obtained from naive cells). Average percentages of H3K27me3 foci-positive nuclei out of 150–200 cells per sample (arrowheads) are indicated. EBs, embryonic body differentiated cells. **c,** Fraction of cells with nuclear H3K27me3 foci. Confocal images, obtained after double immunostaining for OCT4 and H3K27me3 on naive, primed and differentiated samples, were analyzed and quantified for the fraction of cells with nuclear H3K27me3 foci that mark the

inactive X allele. Average percentages of 150–200 individual cells counted per sample from independent frames are shown. Error bars indicate s.d. (n=10). *t-test P value < 0.01 when comparing naive pluripotent cells and their derived differentiated cells. All 5 naive female pluripotent cells are predominantly devoid of H3K27me3 foci. Male lines do not exhibit H3K27me3 foci/clouds in any of the states as expected. Differentiated fibroblasts from naive cell lines uniformly acquired H3K27me3 foci. In most (3 out of 5) primed human ES/iPS cell lines tested, H3K27me3 foci (one per nucleus) became clearly detectable in the majority of cells consistent with XaXi configuration. Two subcloned primed cell lines (C2 human iPS cell and WIBR2 human ES cell) showed low H3K27me3 foci abundance, consistent with XaXa or XaXe configurations known to also be acquired in human primed cell lines [68]. **d,** qRT– PCR analysis indicates no/low expression levels of XIST in naive hPSCs, in comparison to female differentiated fibroblast cells that upregulate XIST expression. Error bars indicate s.d. (n=3). **e,** Distributions of H3K9me3 RPKM (reads-per-kilobase-per-million reads) levels in chromosome X genes, measured in four different human cell lines. Boxes, 25th and 75th percentiles; horizontal lines, median; crosses, outliers. RPKM were measured for each gene on the region starting 1 kb upstream of the TSS and ending at the TES. Lines WIBR3, C1 and LIS2 are female and WIS2 is a male cell line. Distributions of H3K9me3 in female primed cells are significantly higher compared to their naive counterparts, whereas in male cells they are the same. P values were calculated with one-tail paired-sample t-test. NS, not significant. **f,** Bisulphite sequencing analysis of six CpG sites in single clones of an XIST promoter amplicon is shown. Note that in female human naive cells both alleles are demethylated, whereas in primed cells one of the alleles becomes methylated. Methylation of OCT4 locus is shown for comparison.

We next mapped H3K4me3 and H3K27me3 chromatin marks by using ChIP-seq in mouse and human naive and primed pluripotent cells. Whereas distribution of both epigenetic markers over all genes promoters and bodies in mouse ESCs showed a significant ($P < 2x10^{-37}$) decrease in naive conditions (Fig. 13d), an even more notable decrease of both marks in mouse ESCs was observed over developmental genes ($P < 1.6x10^{-71}$; Fig. 13a, right) [45]. Similarly, there was a marked ($P < 8.6x\ 10^{-61}$) reduction of H3K27me3 in promoter and gene-body region over developmental genes (n=5,922) in human naive cells compared to primed (Fig. 13a, b). The reduction of H3K27me3 mark to nearly background levels over developmental genes in human naive cells was also reflected in the number of genes with bivalent marks near their TSS, which is more than 13-fold higher in primed (3,013 genes) compared to naive cells (226 genes). We next measured H3K4me1 and H3K27ac marks using ChIP-seq, and globally mapped enhancers of class I and class II in naive and primed pluripotency states. Finally, reminiscent of recent reports on murine stem cells [69] human naive ESCs/iPSCs expanded in NHSM conditions demonstrated a marked downregulation in DNMT3A, DNMT3B and DNMT3L de novo DNA methyltransferase enzymes, but not in DNMT1 and TET enzymes (Fig. 13c). Sampling the DNA methylation states of human naive and primed pluripotent cells by reduced representation bisulphite sequencing (RRBS) showed a significant reduction (paired sample t-test P value < 0) in CpG methylation in human naive WIBR3 ESCs after expansion in NHSM conditions for 17 days (Fig. 13e).

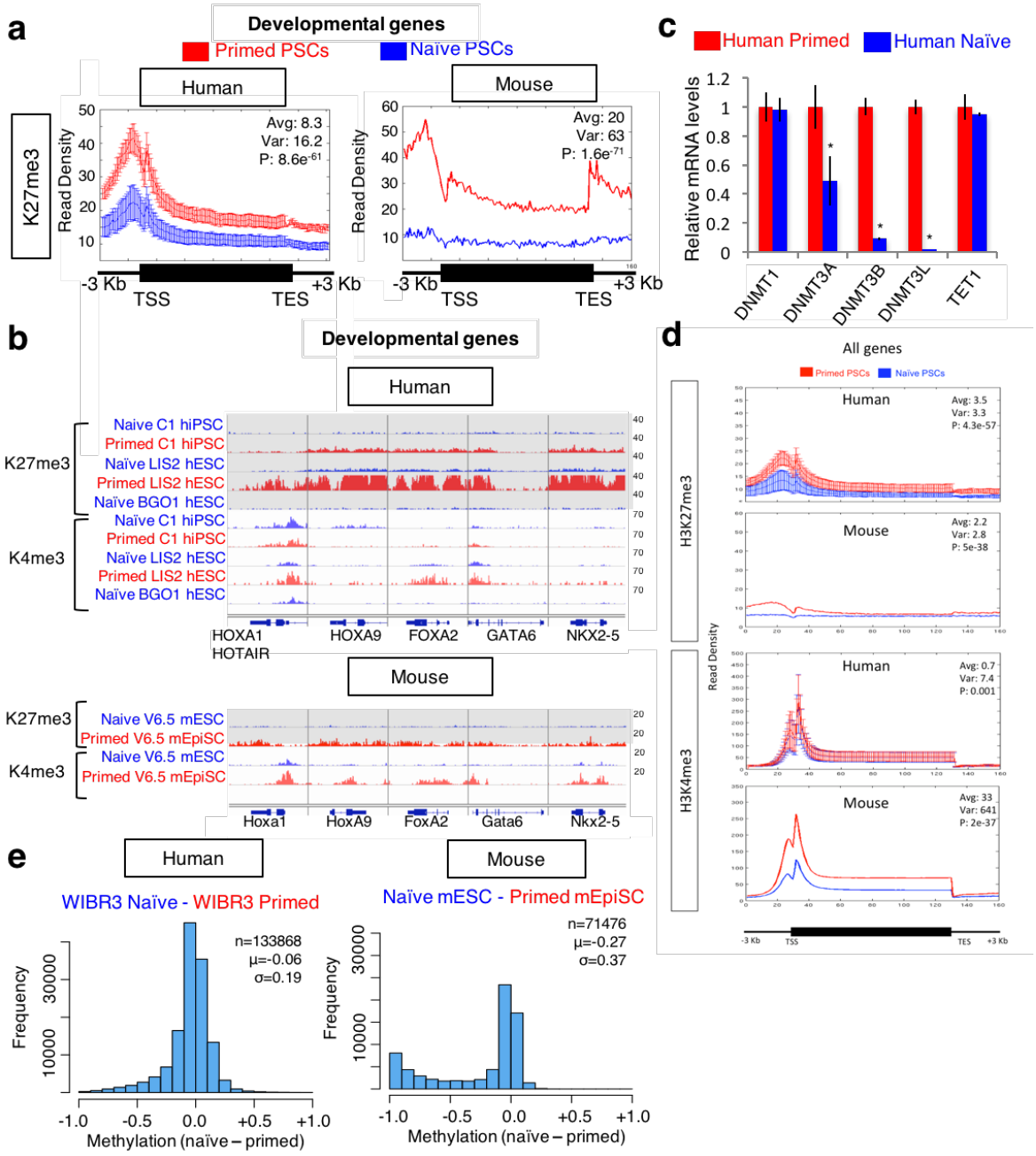

Figure 13. Epigenetic landscape of human naive pluripotency. a, Profiles of H3K27me3 chromatin mark over developmental genes in human and mouse, represented as normalized read-density. Blue, naive; red, primed. Human profiles indicate average and s.d. (error bars) calculated over 5 different cell lines (C1, LIS2, WIBR3, WIBR3-MBD3^{mut} and BGO1). Average difference between plots is indicated alongside variance and P-values (calculated with paired- sample t-test). **b,** Chromatin landscape of selected 5 developmental regulatory genes. H3K27me3 and H3K4me3 marks are shown for naive (blue) and primed (red) cell lines, in both human and mouse, showing high consistency between both species. **c,** Representative relative transcript levels in human primed and naive WIBR3 human ESCs after 48 h in NHSM conditions. *t-test P value < 0.01. Error bars indicate s.d. (n=3). **d,** Global H3K27me3 and H3K4me3 deposition in naive and primed pluripotent cells. Profiles of H3K27me3 and H3K4me3 chromatin modifications of all RefSeq genes in human (n=43,463) and mouse (n=30,480), naive (blue) and primed (red), represented as normalized read-density. Human profiles indicate average and s.d. (error-bars) calculated over 5 different cell lines (C1, LIS2, WIBR3, WIBR3-MBD3^{mut} and BGO1). Average difference between plots is indicated alongside variance and P values (calculated with paired-sample t-test). **e,** Histograms of the change in methylation between primed and naive

samples in human and mouse. The histograms depict the distribution of the per-CpG difference in methylation (naive methylation – primed methylation), calculated for all differentially methylated CpGs residing in CpG-rich regions (>4% CpG content) and having a coverage of ≥ x10 in both analyzed samples. n, the number of covered, differentially methylated CpGs. μ, s, the mean and standard deviation of the distribution, respectively.

Gene expression pattern in human naïve PSCs grown in NHSM

Next we compared global gene expression patterns between naive and primed human ESCs and human iPSCs. Unbiased clustering of genome-wide expression profiles demonstrated that naive hESCs and hiPSCs possess a distinct gene expression pattern and clustered separately from primed hESCs and hiPSCs (Fig. 14a). Multiple transcripts associated with naive pluripotency were significantly upregulated in naive cells, including members of the NANOG and DUSP gene families (Fig. 14b) [70]. Importantly, naive pluripotent cells had profound down-regulation of transcripts associated with lineage-commitment genes, including ZIC1, SOX6 and SOX11 that were expressed at low, but appreciable, levels in primed human ESCs (Fig. 14b) [45]. Functional annotation analysis of differentially expressed genes with Gene Ontology (GO) revealed that genes downregulated in mouse and human naive pluripotency were significantly enriched for GO terms linked to developmental processes (Fig. 14c). Furthermore, hierarchical clustering showed that human inner cell mass samples are transcriptionally more similar (not identical) to naive human ESCs than to primed cells (Fig. 14d).

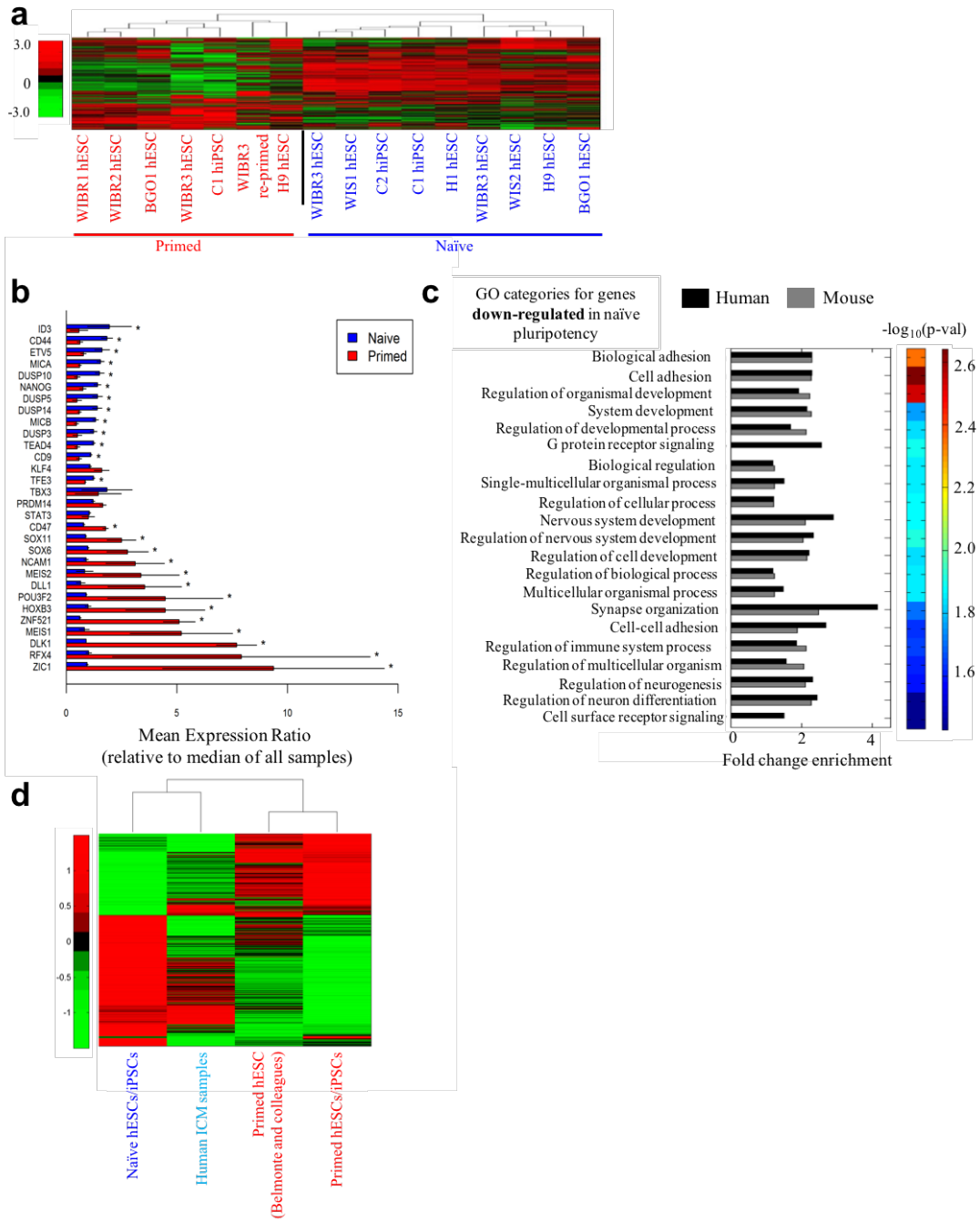

Figure 14. Distinct transcriptome for human naive ESCs/hiPSCs. a, Genome-wide (n=12,071) transcriptional profile of the indicated primed and naive human ESC/iPSC lines measured by Affymetrix microarrays and hierarchically clustered using Pearson correlation. Heat map shows row-normalized expression levels (log2 scale) with red and green colors representing up and downregulated genes, respectively. **b,** Transcriptional profile of selected pluripotency and lineage-specific marker genes, their mean expression ratio in primed human ESCs and naive human ESCs relative to the median of all samples is presented. Values are shown in natural scale. Error bars represent s.e.m of each gene. Asterisks denote statistically significant differentially expressed genes in which the false discovery rate was < 0.05 between the naive and primed groups of samples. **c,** GO (gene ontology) categories significantly enriched for genes downregulated in naive compared to primed human cell lines, with their (–log10) FDR-corrected P-value. The respective mouse fold-enrichment values for the categories that are also significant in mouse are indicated. **d,** Transcriptional comparison of *in vitro*- and *in vivo*-isolated human pluripotent cells. Hierarchical clustering of the mean expression profile of differentially expressed genes between naive and primed samples (FDR < 0.05), in the different groups (naive (this study), primed (this study), Belmonte's primed ESCs, and human inner cell masses (Belmonte and

colleagues [71])), using Spearman correlation. Note that while primed ESC samples previously derived by Belmonte and colleagues cluster with the primed samples derived herein, human inner cell mass samples cluster with the naive samples expanded in NHSM conditions.

Primed human ESCs demonstrate intermediate expression levels of MHC class I surface antigen in comparison to somatic cells, whereas naive hESCs/hiPSCs express only trace levels of MHC class I (Fig. 15a) [42]. Moreover, although E-CADHERIN is expressed in primed human ESCs, the surface expression pattern becomes more prominent in naive human ESC colonies (Fig. 15b). Next we conducted an unbiased cross-species hierarchical clustering of the globally measured transcriptome to evaluate whether the human primed and naive pluripotent cells, described herein, globally correspond to those established in mice. By applying an algorithmic cross-species gene expression analysis on all 9,803 mouse–human orthologous genes found in our gene expression data sets [42, 49], we found that whereas primed hESCs and hiPSCs clustered with murine EpiSCs, all naive hESCs/hiPSCs clustered with naive mouse ESCs/iPSCs independent of the genetic background or naive growth conditions used (Fig. 16a). Finally, nuclear localization of the transcription factor TFE3 was recently shown to be enhanced in naive mouse ESCs, and compromised upon pluripotency priming [72]. Remarkably, a similar nuclear enrichment pattern for TFE3 was evident in naive hESCs, and relative enrichment was compromised in primed hESCs (Fig. 16b,c).

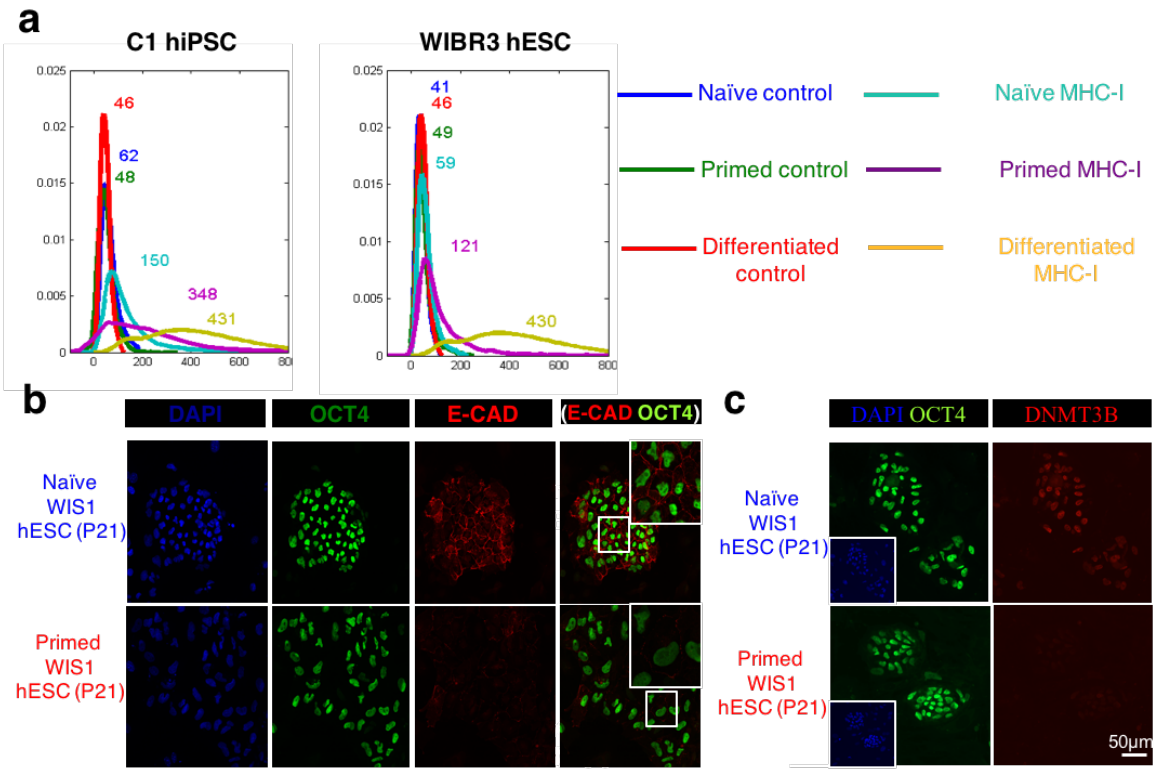

Figure 15. Transcriptional and molecular features of human ground state naive pluripotency. a, Histogram of surface expression of MHC class I using FACS analysis on the indicated naive, primed and

differentiated cell lines. These results show that naive human ESCs and human iPSCs express rudimentary or low levels of MHC class I on their cell surface, whereas primed cells upregulate MHC class I expression to intermediate levels (in comparison to differentiated 293 cells that express high levels of MHC class I). These results are identical to patterns observed in murine naive and primed cells [42]. **b,** Representative confocal immunostaining images for the expression of E-CADHERIN and OCT4 on genetically matched naive and primed human ESCs. The samples were processed simultaneously and analyzed under identical conditions. Insets represent enlargements of boxed areas. Whereas E-CADHERIN is expressed in primed human ESCs, its expression becomes homogenously distributed and more enhanced in naive human ESCs expanded in NHSM conditions. **c,** Representative immunostaining for OCT4 and DNMT3B in human primed and naive pluripotent cells, showing DNMT3B down regulation in the human naive ground state of pluripotency. This effect was highly dependent on efficient p38 inhibition.

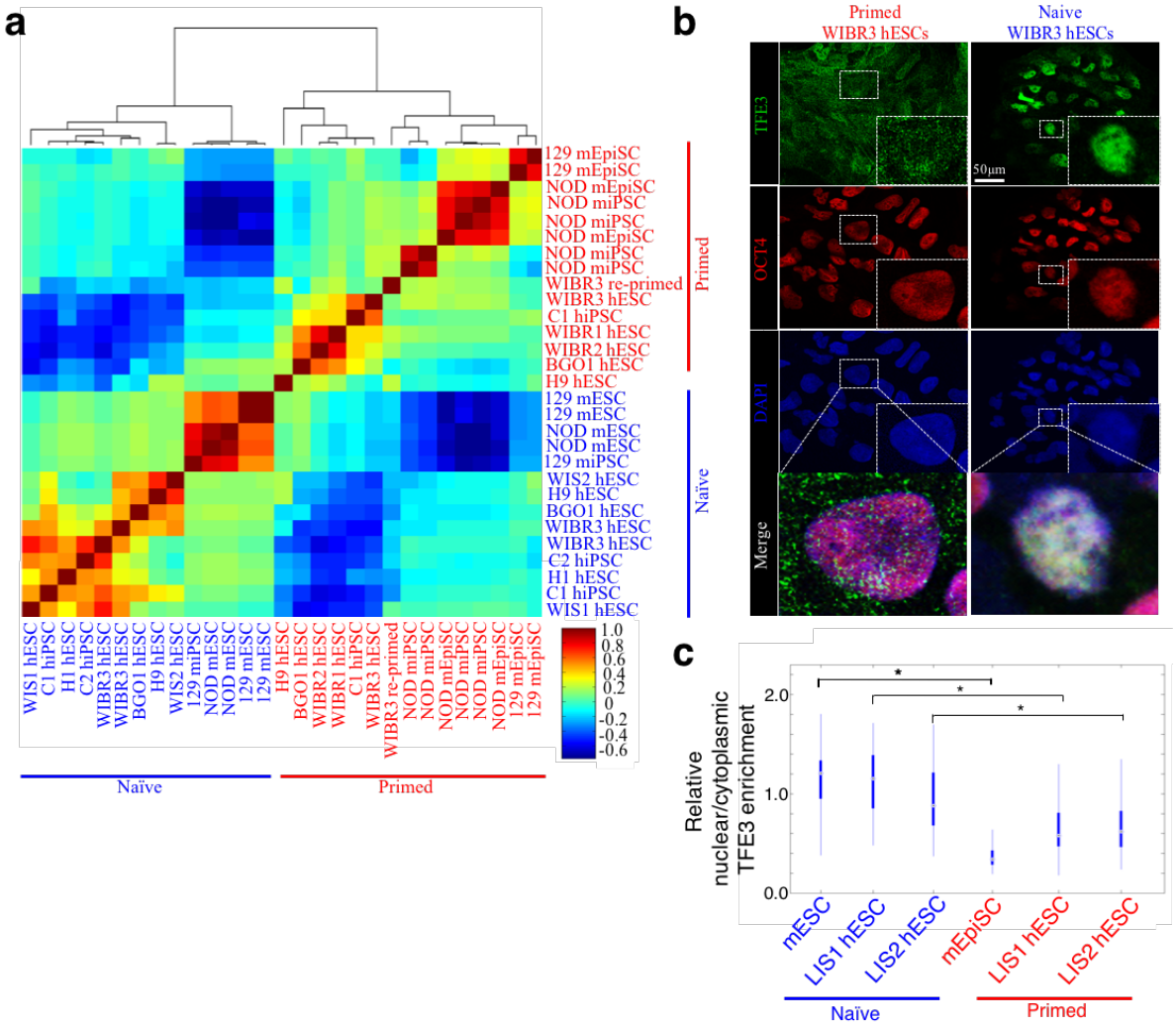

Figure 16. Naive human stem cells share defining molecular features with mouse ESCs. a, Unbiased global transcriptional cross-species hierarchical clustering of naive and primed pluripotent cells from mice and humans. Correlation matrix of gene expression was clustered using Spearman correlation and average linkage. Color bar indicates correlation strength. Each row/column represents an independent cell line or clone. mEpiSCs, murine EpiSCs; mESC, mouse ES cell; miPSC, mouse iPS cell. NOD indicates non-obese diabetic ICR mice. **b,** Naive and primed human ESCs were double immunostained for TFE3 and OCT4. Representative confocal images are shown for WIBR3 human ESCs. Insets are enlargements of the dashed boxes. Scale bars, 50 mm. **c,** Box and whisker plots of nuclear/ cytoplasmic TFE3 ratios in naive and primed mouse and human ESCs are shown. Quantitative unbiased imaging analysis for preferential nuclear localization was conducted on randomly selected 200 cells from independent image frames per sample. Naive human ESCs showed distributions similar to those in naive mouse ESCs, and the nuclear enrichment was lost in primed human and mouse ESCs. *t-test P values < $1x10^{-100}$.

Human PSCs grown in NHSM are able to contribute to human mouse interspecies chimeras

One of the main naïve pluripotency traits is the ability to generate a chimeric animal once cells are injected into host blastocysts. We moved to test this on our human naïve iPSCs that were constitutively labeled with GFP (by ZFN-mediated targeting of AAVS1 locus with CAGGS-GFP knock-in reporter) and grown on NHSM. Human cells were validated before injection with immunocytochemistry positive staining of GFP, SSEA4 (specifically stains hPSCs), and OCT4. Importantly, Cells showed positive stain for SSEA4, but there was no stain for SSEA1, which specifically stains mouse pluripotent cells and not human PSCs (Fig 17b,c). Immunostaining with anti-human nuclei antibody was done to acquire an additional tool for identification of human cells amongst mouse cells. Staining was done on human cells plated on mouse embryonic fibroblasts (Fig. 17a). Human iPSCs intended for injection showed evident GFP prior injection (Fig 17b). GFP signal was also detected immediately after injection to mouse morulas and at 1-day post injection in *ex-vivo* culture, before blastocysts transfer to host mothers (Fig. 17b; lower panel). To increase injected cells survival, 10μM ROCKi was added to M2 injection medium for the whole injecting procedure. Moreover, same concentration of ROCKi should be used for KSOM *ex-vivo* culture medium for three hours incubation post injection, then the morulas were transferred to KSOM without ROCKi for over-night incubation until they reached the blastocyst stage. Immunostaining 24 hours post injection of mouse morulas and blastocysts with anti-OCT4 and anti-CDX2 (trophoblast marker) antibodies revealed co-localization with OCT4 but not with CDX2, indicating that the injected cells specifically integrated in the ICM (Fig. 17c).

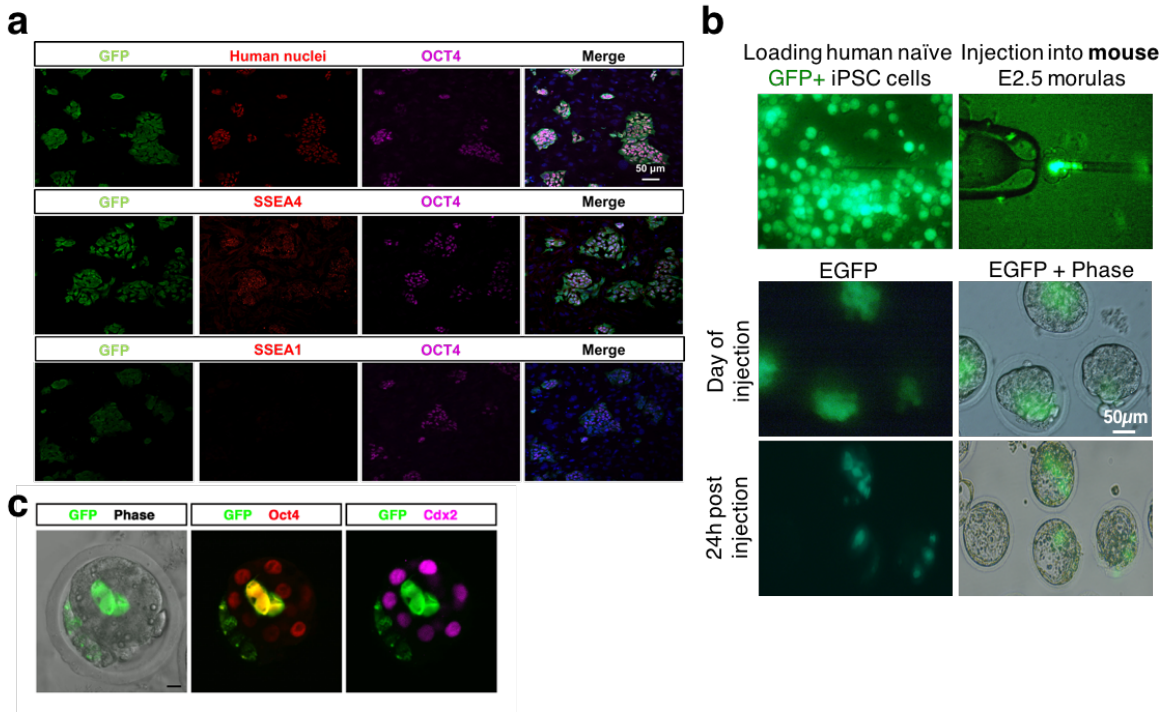

Figure 17. Mouse morula microinjection of human naive iPSCs. a, C1 human naive iPSCs targeted with constitutively CAGGS promoter driven EGFP into the human AAVS1 locus via ZFN utilization. Injected cells were validated through immunostaining with Anti-GFP (green) and OCT4 (magenta) together with Human nuclei (red) (upper panel), SSEA4 (red) (middle panel) or SSEA1 (red) (bottom panel) antibodies. Representative confocal images show positive stain for GFP, OCT4, human nuclei and SSEA4, but did not stain for SSEA1. Scale bar 50µm. **b,** Human iPSCs were microinjected into E2.5 mouse early morulas (upper panel), and micro-manipulated embryos were allowed to recover and develop to blastocyst stage *ex-vivo* for an additional 24 h (lower panel). Images show specific GFP^+ human cell survival and integration in mouse pre-implantation embryos. Scale bar 50µm. **c,** Representative confocal analysis following immunostaining for GFP (green), OCT4 (red) and CDX2 (magenta) was done 24 h following mouse morula microinjections with GFP-labelled naive human iPSCs. Note surviving GFP cells that specifically integrate and stain positive for OCT4, but not CDX2 (no co-localization between GFP and CDX2 was observed).

Identification of injected GFP-human iPSCs was fairly challenging; as the embryo develops and gains more tissues and layers, conventional microscopy methods fail to detect low levels or sparse signals of fluorescence. To address this issue, we applied the *in-toto* live imaging technique that was described in Massarwa et al. 2013 [32]. Using this method, we were able to detect GFP positive cells in different regions and at different developmental stages of the mouse embryo. At embryonic day E10.5, GFP positive cells were detected at different regions, mainly in the head region including the optic vesicle, where the adult eye will develop, and in the otic pit, where the adult internal ear will develop (Fig. 18). The contribution of injected human iPSCs was aberrant in several different embryos, yet they were present in only up to 3-20% of each litter approximately. To assess their contribution in later stages, we preformed the same experiment on injected embryos at days E11.5-E13.5. We were able to identify GFP positive cells in different embryonic tissues, including forelimb

bud, and mid body somites in E11.5 embryos, in the trunk of E12.5 embryo, and in the proximal forelimb of E13.5 embryo (Fig. 19). Analysis of embryos past E13.5 was not efficient using this technique due to the size and thickness of the whole embryo. Thus, a different approach should be taken in order to be able to detect GFP positive cells in older embryos. Amazingly, GFP human injected iPSCs survive during development of the mouse embryo, and can be detected *in-toto* until embryo day E13.5.

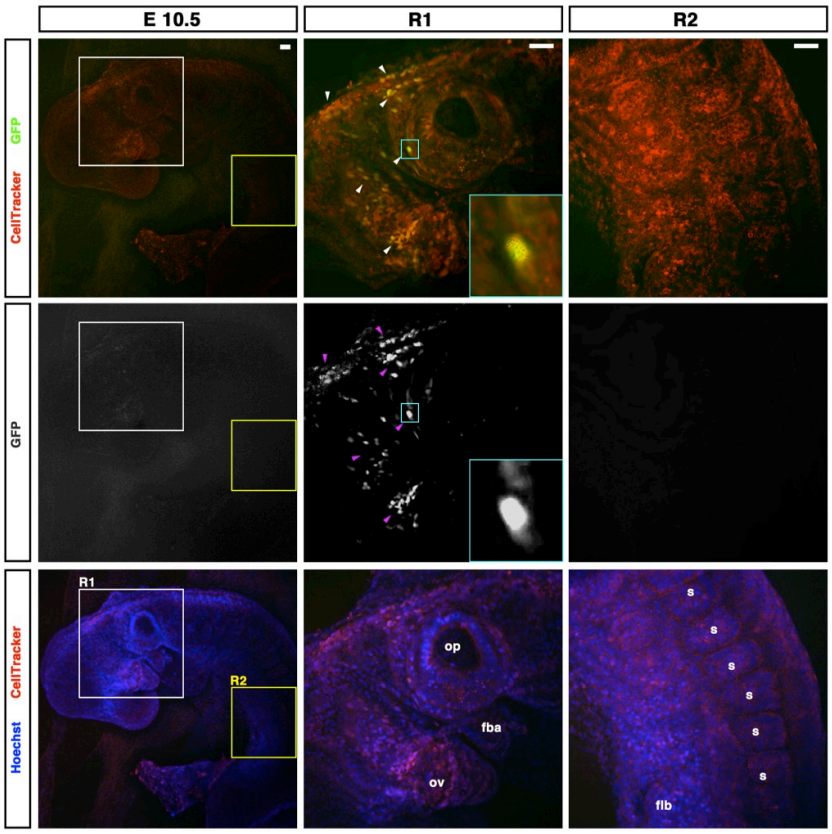

Figure 18. Robust generation of interspecies chimeric humanized mice following naive human iPSCs microinjection into mouse morulas. Representative images showing widespread integration of GFP-labeled human naive iPSCs into different locations in the anterior part of an E10.5 mouse embryo. Hoechst (blue) and CellTracker (red) were used for counterstaining. The first column shows the whole embryo (z-stack interval 30 µm, 18 focal planes in total). The second column shows magnified images focusing on the head region (white square R1) where the human iPSCs (GFP-positive cells) are pointed out (arrowheads, z-stack interval 20 µm, 11 focal planes in total). The third column shows the posterior part of the embryo (yellow square R2), where no GFP-positive cells were detected (z-stack interval 20 µm, 9 focal planes in total). Blue squares in the first two images in the second column represent the area shown in the insets at the corner of each image. ov, optic vesicle; op, otic pit; fba, first branchial arch; flb, forelimb bud; s, somite. Scale bars, 50 µm.

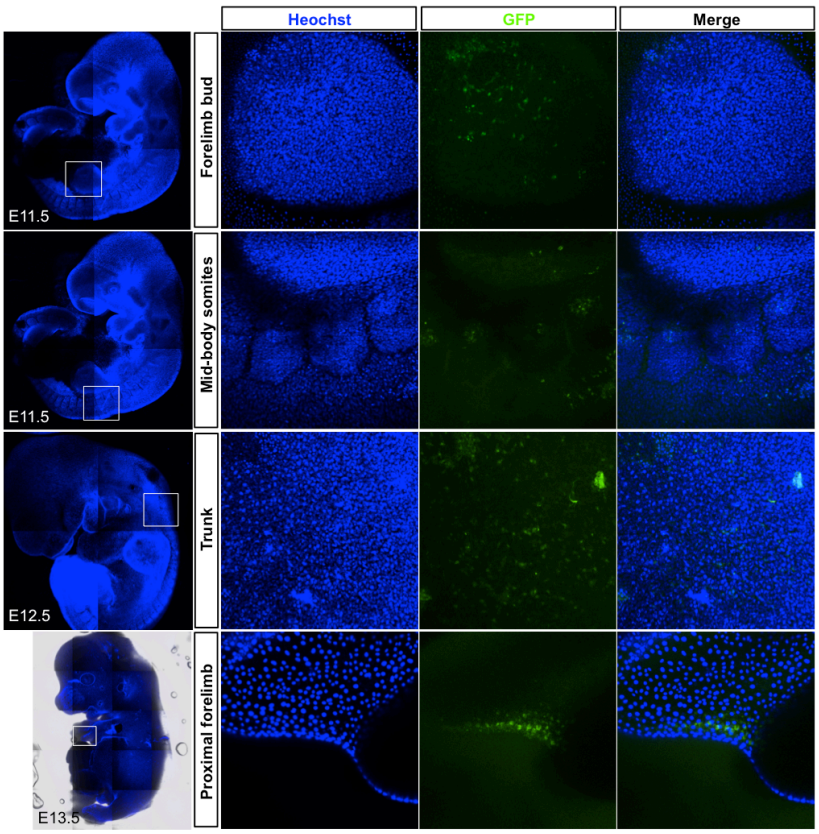

Figure 19. Interspecies chimeric humanized mice at different developmental stages during mouse embryogenesis. Different embryos at a variety of developmental stages show significant contribution of human naïve GFP positive iPSCs into different regions. Overviews of embryos at different developmental stages are in the left panel. The right panels represent magnified image of embryonic areas included in the white boxes on the left panels (z-stack interval 20 μm, 9 focal planes in total).

Contribution of injected GFP-human cells to the mouse pulmonary system

In order to test whether the injected human iPSCs not only survive and proliferate in the mouse embryo, but also differentiate and express specific markers in various tissues, we tested the mouse pulmonary system for survival and integration of the injected GFP-human cells. GFP-positive cells were detected in a scatter distribution all over the mouse embryonic lung at embryonic day E17.5. To verify that in fact human cells are involved, we stained the tissue with anti-human nuclei antibody, and indeed, all of the cells that showed positive GFP stain were also positive for human nuclei stain. Moreover, a few cells showed only human nuclei positive stain indicating that some human cells lose their GFP signal during development (Fig. 20, wild type lungs showed neither human nuclei nor GFP stain). To test if these cells are differentiated into different pulmonary subpopulations, we stained several sections with lung specific markers including bronchiolar Clara cell secretory protein (CC10), which is a marker for bronchiolar exocrine cells in the proximal airways of the lung [73]. We found that the injected human GFP positive cells were also stained for CC10 (Fig. 21). To test if the injected cells are able to contribute and function in the distal airways of the lung,

we carried out immunostaining experiments using antibodies against pulmonary surfactants associated protein C (Pro-SpC) [73] and Aquaporin 5 (Aqp5), which is expressed in alveolar epithelial cells [74]. In both cases, it was clear that some of the GFP positive cells co-express both Pro-SpC and Aqp5 as well (Fig. 22 and 23, respectively).

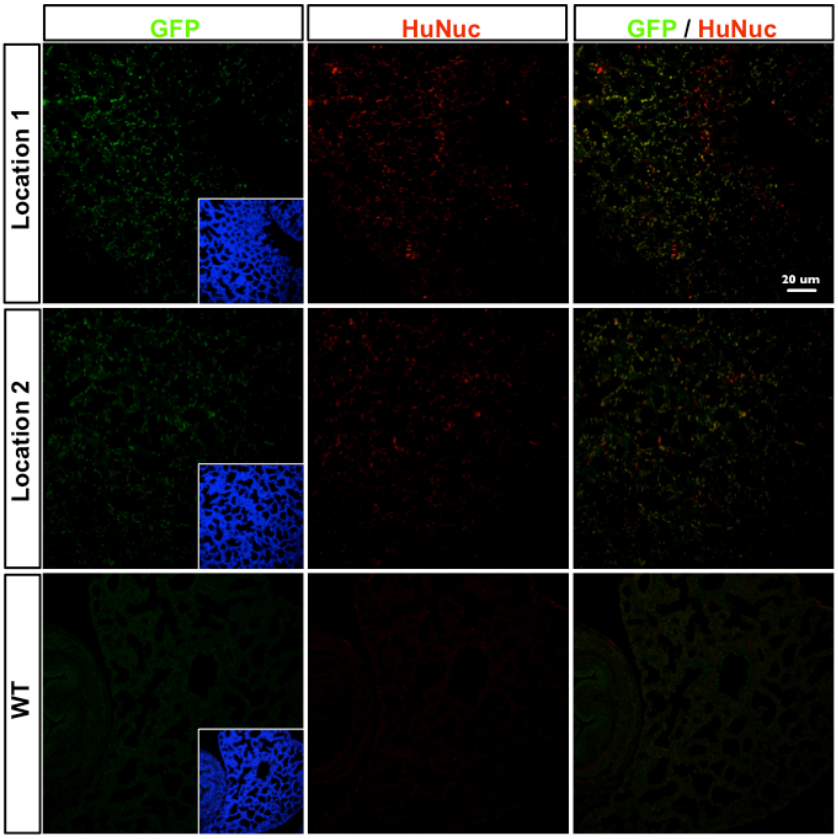

Figure 20. Injected hiPSCs contribute to pulmonary tissues in the chimeric mouse embryo and are stained for human nuclei. Lung paraffin sections of an E17.5 injected and WT embryos were immunostained for GFP and Human nuclei. Co-expression of GFP (green) and Human nuclei (red) is shown in two different locations (top and middle panels) of the chimeric mouse lung but not in WT mouse lung (lower panel). DAPI (blue) counterstaining is shown in the insets. Scale bar 20μm.

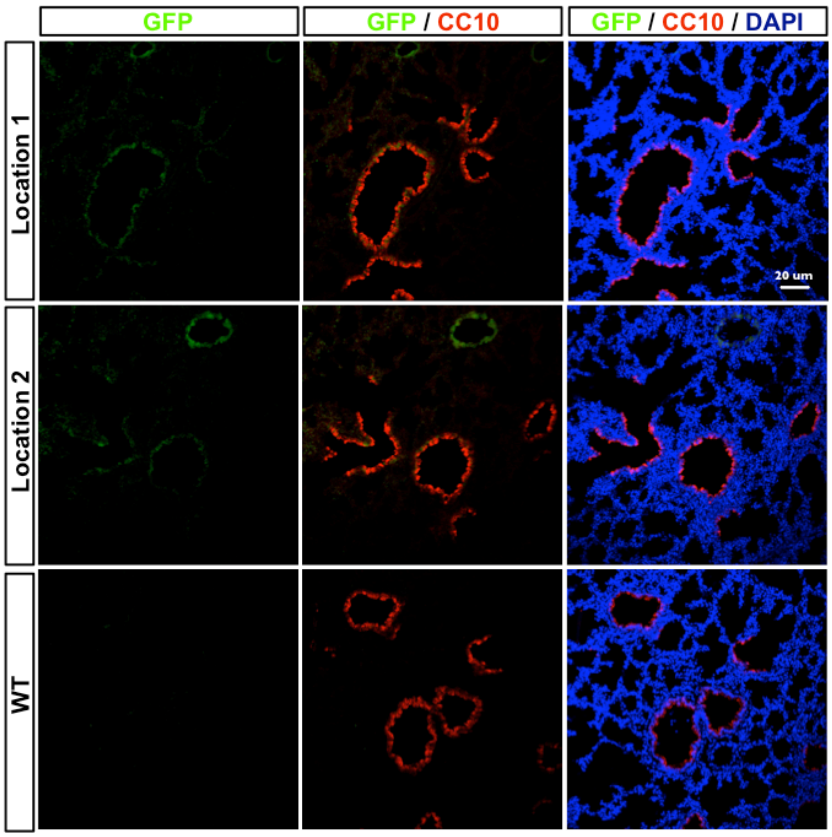

Figure 21. Injected hiPSCs contribute to pulmonary tissues and express proximal airway epithelium marker in the chimeric mouse embryo. Lung paraffin sections of an E17.5 injected and WT embryos were immunostained for GFP and Clara Cell 10 (CC10), a marker for proximal airways. Co-expression of GFP (green) and CC10 (red) is shown in two different locations (top and middle panels) of the chimeric mouse lung. WT lung shows evident CC10 staining but no GFP stain was detected (lower panel) as expected. DAPI (blue) was used for counterstaining. Scale bar 20μm.

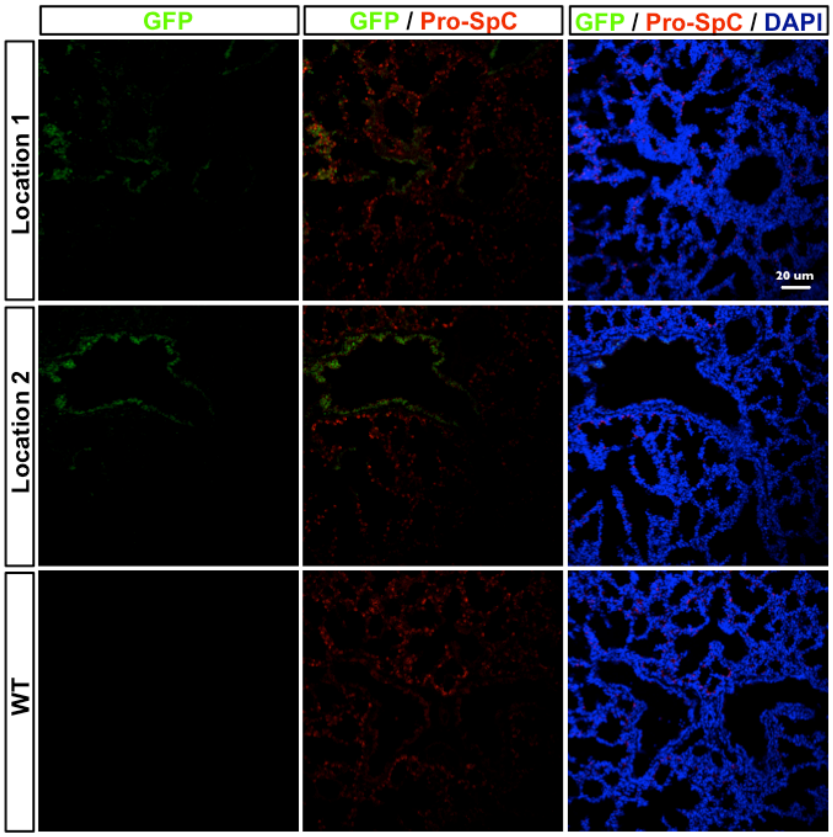

Figure 21. Injected hiPSCs contribute to pulmonary tissues and express distal airway epithelium marker Pro-SpC in the chimeric mouse embryo. Lung paraffin sections of an E17.5 injected and WT embryos were immunostained for GFP and pro-surfactant C (Pro-SpC). Co-expression of GFP (green) and Pro-SpC (red) is shown in two different locations (top and middle panels) of the chimeric mouse lung. WT lung shows evident Pro-SpC staining but no GFP stain was detected (lower panel) as expected. DAPI (blue) was used for counterstaining. Scale bar 20µm.

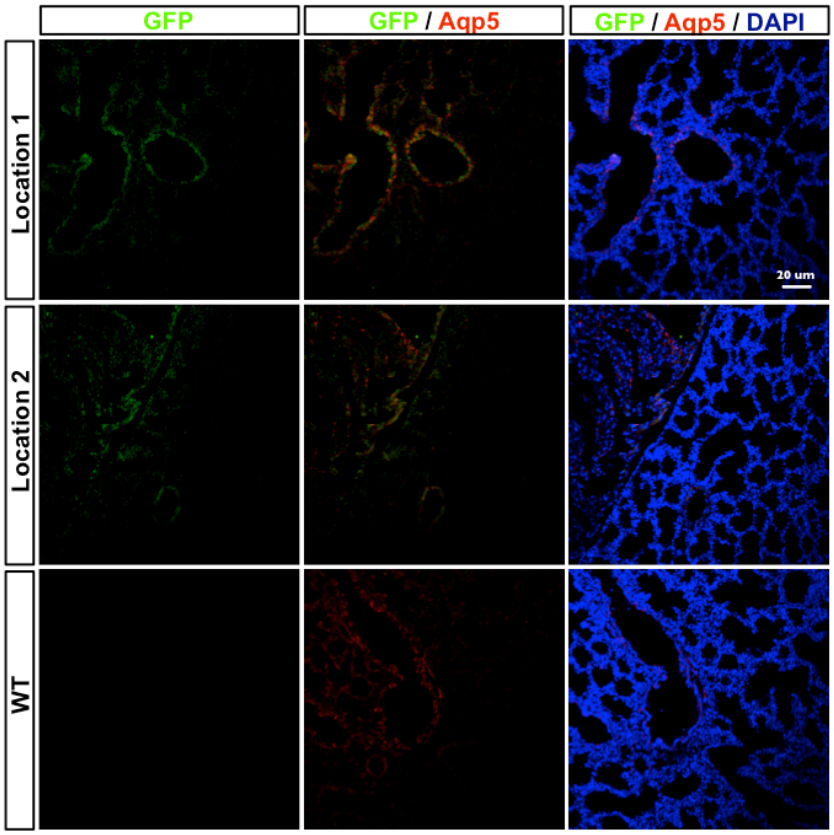

Figure 22. Injected hiPSCs contribute to pulmonary tissues and express alveolar epithelial cells marker, Aquaporin 5, in the chimeric mouse embryo. Lung paraffin sections of an E17.5 injected and WT embryos were immunostained for GFP and Aquaporin 5 (Aqp5). Cells in the distal area of the chimeric mouse lung show co-expression of GFP (green) and Aqp5 (red) in both locations (top and middle panels). WT lung shows Aqp5 staining, while no GFP stain was detected (lower panel) as expected. DAPI (blue) was used for counterstaining. Scale bar 20µm.

Human mouse interspecies chimerism using human P53 knockout 'cheater' cells

The expected low chimerism efficiency led us to examine ways to improve the injected human iPSCs contribution ratios in our system. Few relevant cell competition studies showed that P53 knockout or knockdown cells have a clear advantage in proliferation and overall survival over wild type cells [29, 31]. These studies have prompted us to establish human PSC lines that contain P53 knockout in addition to the EGFP constitutive reporter. For this purpose, we used CRISPR mediated targeting to the fourth exon of human P53. Several clones showed complete ablation of P53. The two alleles of C2 clone showed the same deletion as was evident by western blot, normal hESC morphology and normal karyotype and sustained EGFP expression (Fig. 23). This cell line was injected into mouse morulas in the same manner as described above. To test the competition

effect, the injected embryos were dissected at E9.5-E10.5 and examined under stereomicroscope. Several embryos showed significant human cells integration to the entire embryo (Fig. 24a). The embryos were then trypsinized to a single cell suspension and analyzed for GFP signal using FACS. The chimerism level in each chimeric embryo was dramatically elevated (Fig. 24b).

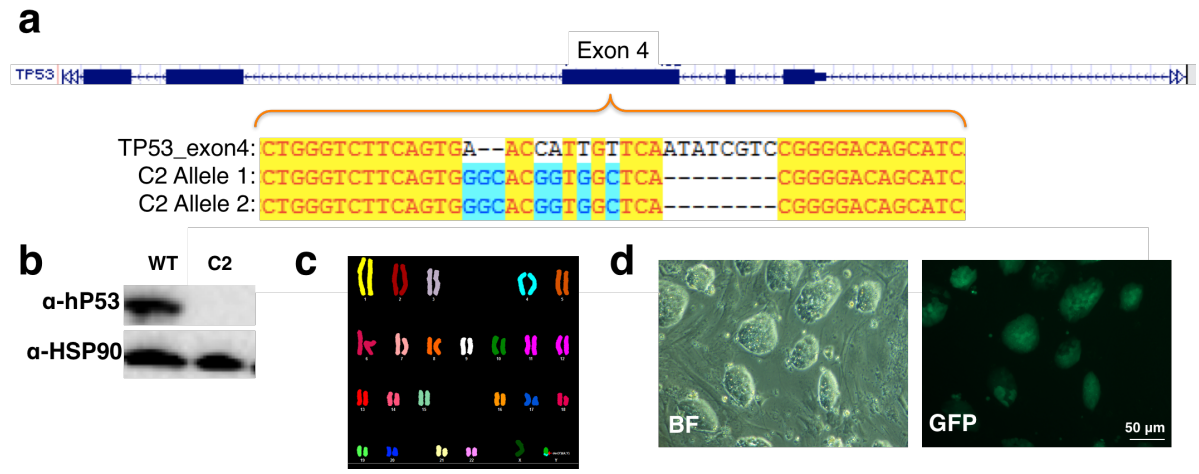

Figure 23. LIS 38 tg-lck GFP P53 CRISPR KO cell line generation. a, C2 clone vs. WT LIS 2 P53 CRISPR targeted region sequence alignment. **b,** P53 Immunoblot analysis showing the lack of P53 protein in comparison to WT protein. **c,** LIS 2 tg-pTrip lck EGFP P53 CRISPR karyotype analysis. **d,** LIS 2 tg-lck GFP P53 KO C2 clone representative bright field (left) and GFP (right) images. Scale bar 50µm.

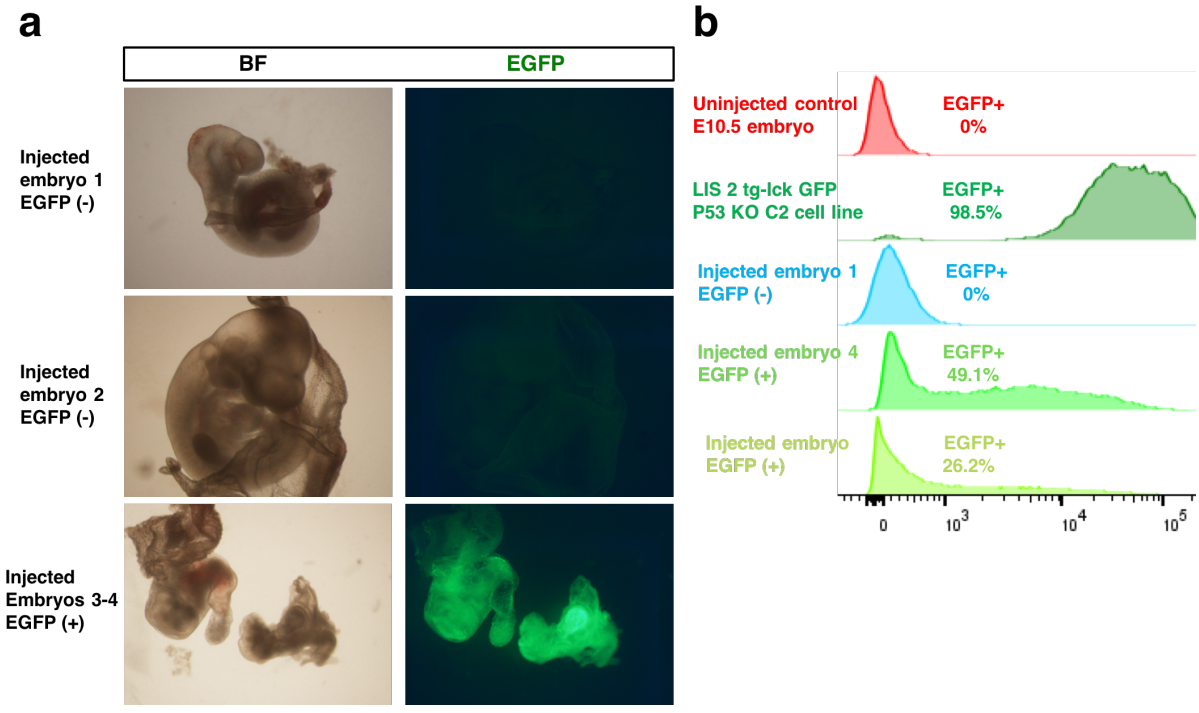

Figure 24. LIS 2 tg-lck GFP P53 CRISPR KO cells integrated more efficiently into the mouse embryo. a, Representative bright field (BF; left column) and EGFP (right column) images of negative EGFP ((-); embryo 1-2; top two rows) and positive EGFP ((+); embryo 3-4; bottom row). **b,** FACS histogram plot describing the different embryos at E10.5: uninjected (red), LIS 2 tg-lck EGFP P53 CRISPR KO cell line (Dark green), non chimeric (cyan), chimeric embryo 1 (mild green) and chimeric embryo 2 (light green).

Next, we set out to examine C2 clone integration ability in various parts of the injected embryo. Injected human PSCs carrying P53 knockout showed significantly higher integration to several embryos. Although we were only able to obtain six normally developed embryos at embryonic day E10.5 out of 220 injected morulas that were transferred back to host mothers, five of them showed a significant amount of GFP positive cells in the chimeric animal from head to tail in comparison to WT embryos, as was evident by *in-toto* live imaging (Fig. 25) [75].

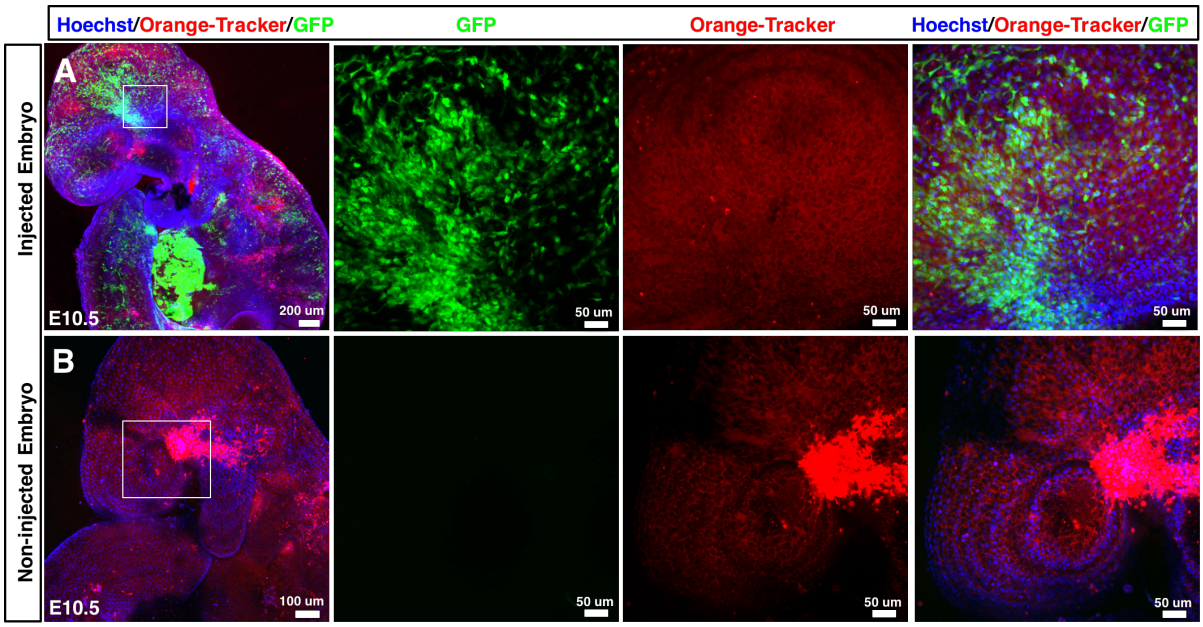

Figure 25. Contribution of LIS 2 tg-lck GFP P53 CRISPR KO injected cells to chimeric animals is significantly higher. Representative images of E10.5 embryos (out of 5 positive embryos) showing high contribution of LIS 2 tg-lck GFP P53 CRISPR KO injected cells (upper panel) in comparison to wild type; embryos were counterstained with Hoechst (blue) and Orange-tracker (red). Scale bar 50µm.

From another injection session of 220 morulas, embryos were dissected at embryonic day E15.5. This batch yielded only one chimeric embryo. However, once this embryo was analyzed under a stereomicroscope, the GFP signal was already evident (Fig. 26). To get a better insight on the injected cells contribution, this embryo was processed for paraffin sections and stained for GFP and NUMA, an antibody used to specifically detect human nuclei. GFP signal was evident in various tissues such as the extraocular muscles, spinal cord and more (Fig. 27 & Fig. 28). In addition, co-expression of GFP and NUMA was present in the GFP positive areas, strengthening the fact that the injected cells maintained their human identity. Taken together, these results suggest that injected P53 KO cells indeed exhibit higher survival and proliferation in the mouse embryo compared to regular human PSCs.

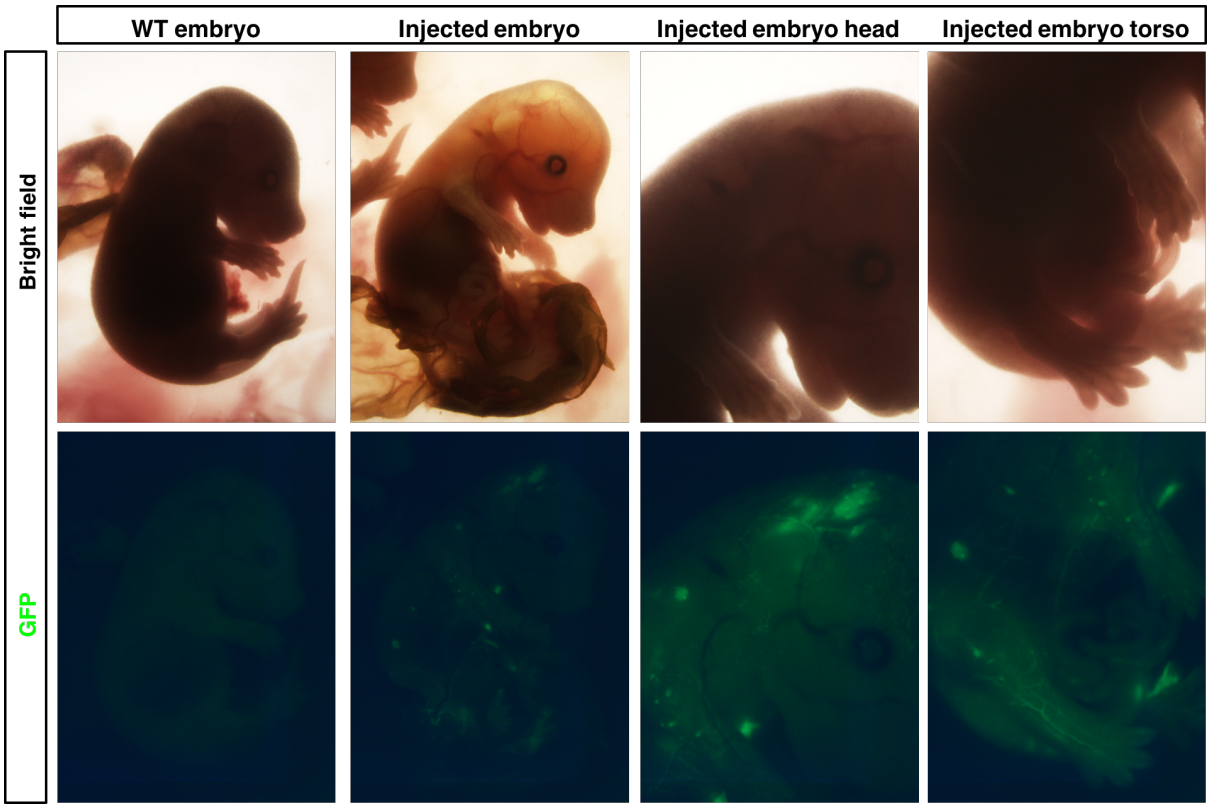

Figure 26. P53 KO injected human cells are evident at embryo day E15.5 using stereoscope. Injected human GFP positive cells were visible in various parts of the injected embryo at embryonic day E15.5 (second column from left). WT embryo did not show GFP signal (first left column). Head region image showing human GFP positive cells contribution in the brain, eye, neck and forelimbs (third column from the left). Torso region image showing human GFP positive cells contribution to the vascular system, tail and hind limbs (forth column from the left).

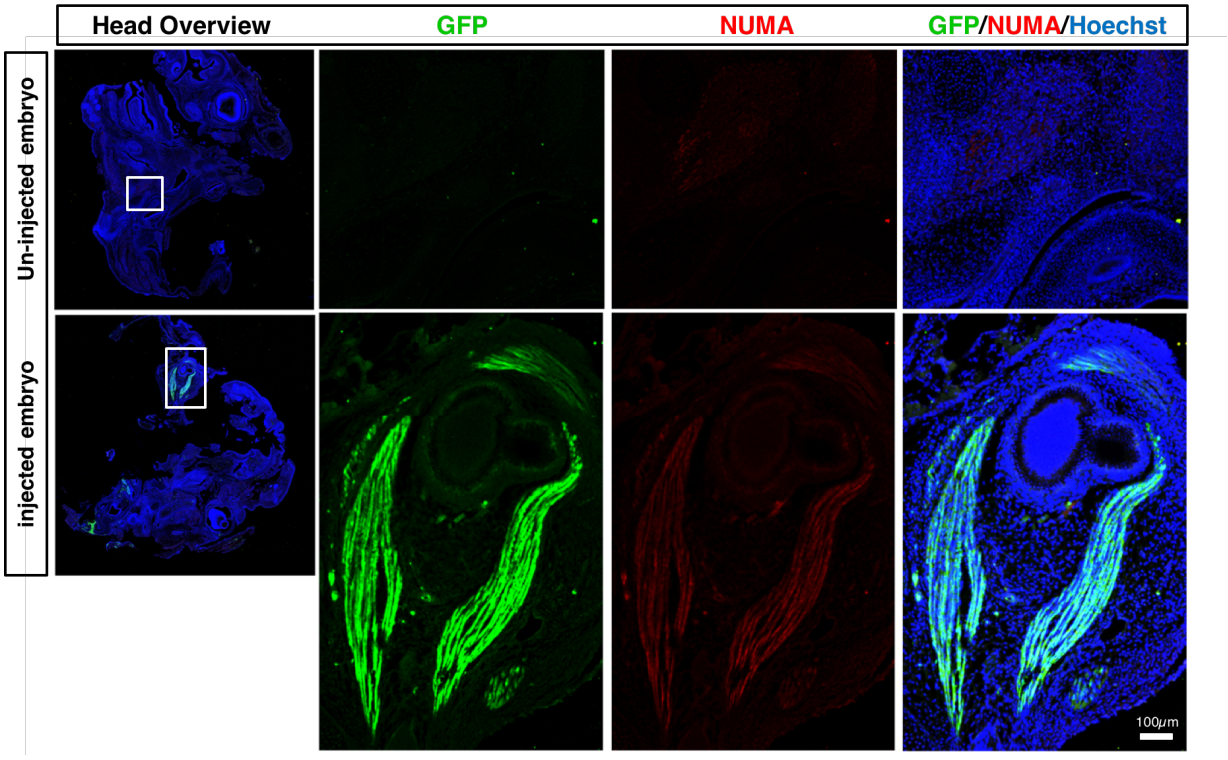

Figure 27. P53 KO injected human cells integrates into the mouse head tissues. Paraffin sections of an E15.5 WT (top panel) and injected (lower panel) embryos were stained for GFP (green) and NUMA (red). Most

left columns are overview images of the head sections, white squares represent the magnified regions shown in the images on the right panels. GFP positive cells integrated into the extraocular muscle are also positive for NUMA. Scale bar 100µm.

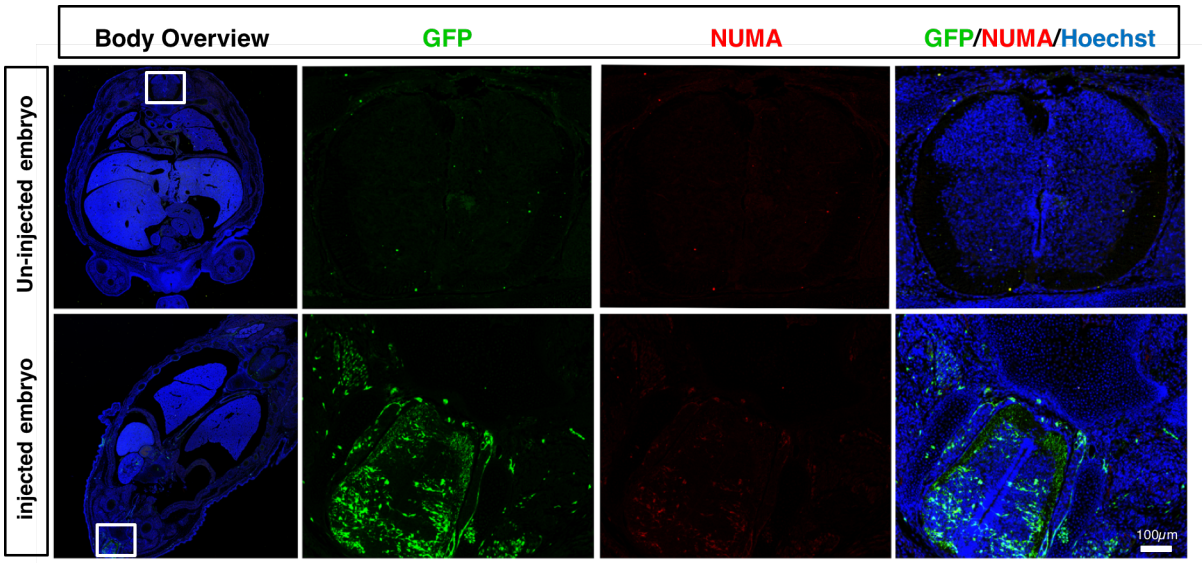

Figure 28. P53 KO injected human cells integrated into different regions of the embryo body. Paraffin sections of an E15.5 WT (top panel) and injected (lower panel) embryos were stained for GFP (green) and NUMA (red). Most left columns are overview images of the body sections, white squares represent the magnified regions shown in the images on the right panels. GFP positive cells integrated into the spinal cord are also positive for NUMA. Scale bar 100µm.

Discussion

Even though pluripotency is widely studied in both human and mice species during the past two decades, many signaling patterns and regulation pathways are still poorly understood. The work during my PhD started paving the way for better understanding of the pluripotency regulation circuitry.

In the first part of my PhD, we wished to discover potential inhibitors of naïve pluripotency reversion using a piggyBac gene trap screen on Nanog-GFP cell line. A few of the analyzed clonal populations failed to regain pluripotency due to a disruption in *Utx* expression, an H3K27me3 demethylase. Specific targeting to the *Utx* locus revealed that absence of *Utx* allowed the already established PSCs to maintain their pluripotent characteristics, and did not affect their differentiation potential. However, when *Utx* deficient "secondary reprogramming" pluripotent cell lines were injected to host blastocyst, the differentiated MEFs derived from chimeric animals failed to reprogram efficiently upon induction of OKSM. Taken together, these results demonstrate that absence of Utx demethylase constitutes a major roadblock for reprogramming, and suggest that Utx plays a critical role in re-establishment of pluripotency *in-vitro*. Furthermore, gene expression analyses presented a similar pattern of $Utx^{\Delta/Y}$ ESCs in comparison to WT ESCs, while aberrant differences were evident between $Utx^{\Delta/Y}$ MEFs and WT MEFs after 8 days of OKSM induction. Both hierarchical clustering and PCA clustered $Utx^{\Delta/Y}$ OKSM induced MEFs with untreated $Utx^{\Delta/Y}$ and $Utx^{+/Y}$ MEFs, which indicates that $Utx^{\Delta/Y}$ MEFs were not able to commence the reprogramming process.

The chromatin state in $Utx^{\Delta/Y}$ ESCs or MEFs was not different from their WT counterparts. However, significant accumulation of H3K27me3 repressive marks around the transcription start sites (TSSs) of OSKM and Nanog target genes were evident by ChIP-seq analysis in $Utx^{\Delta/Y}$ OSKM induced MEFs. These results support our previous data showing that Utx specific role as H3K27me3 demethylase could not be compensated by other demethylases such as Jmjd3 or Uty, or by Utx carrying a point mutation in its catalytic site. ChIP-seq analysis for H3K4me3 showed a dramatic decrease in this active mark that was apparent in the TSSs of OSKM and Nanog target genes. ChIP-seq analysis for Utx binding on OSKM induced MEFs resulted in enrichment of Oct4 and Nanog targets and more than 1800 sites of H3K4me3 active chromatin mark. Direct targets of Oct4 and Klf4, for example, Sal1l, Sall4 and Utf1, were hypermethylated for H3K27me3 in $Utx^{\Delta/Y}$ OSKM induced MEFs and specifically bound by Utx in $Utx^{+/Y}$ OSKM induced MEFs. Motif search analysis for Utx binding identified the OSK and Nanog consensus motifs as the four most highly enriched. In

addition, Co-immunoprecipitation assay showed that Utx could interact with OSK. Overall, these data indicate that Utx can associate with OKS by a protein–protein interaction and promotes the reactivation of potent pluripotency promoting modules that cooperatively facilitate iPSC formation.

Chimeric animals with both Utx deficient iPSCs and ESCs failed to exhibit germline transmission after breeding. This phenotype was rescued with transgenic insertions of WT UTX but not with catalytically inactive UTX^{H1146A}, indicating that the defect in germ-lineage development can be unambiguously assigned to the absence of catalytically active Utx.

Hajkova et al. have shown that murine PGCs globally increase H3K27me3 levels by E8.5-E9.5 embryonic stages, these cells later undergo *in-vivo* reprogramming by global loss of H3K27me3 repressive mark. PGCs of chimeric animals generated from Utx null ESCs blastocyst microinjection were stained for early PGC markers expression. Supporting Hajkova et al. data, PGCs lacking Utx accumulated H3K27me3 mark and failed to express Mvh, a late PGC marker. Finally, we find that $Utx^{\Delta/Y}$ PGCs isolated from E8.5 chimeric embryos are deficient in their ability to revert to ground state pluripotency and form ES-like embryonic germ cells. Taken together, these results provide direct evidence for cell-autonomous aberrant germ cell development *in vivo* and perturbed epigenetic reprogramming in the absence of Utx.

My study shed new light on the chromatin regulation mechanism in the re-establishment of pluripotency either from somatic cell reprogramming or during *in vivo* reprogramming of germ cells differentiation. Even though studies have shown that there are two functional H3K27me3 demethylases, Jmjd3 and Utx, only the latter is crucial for demethylation of pluripotency related genes during these processes.

In the second part of my work, we set out to define novel growth conditions for human naïve PSCs. Previous attempts have showed that 2i/LIF naïve mouse conditions are sufficient to maintain naïve characteristics in human PSCs, this is due to the fact that human naïve PSCs involve a more complex circuitry of signaling network. To pursue this goal, we carried out a screen for additional cytokines and inhibitors on top of the conventional 2i/LIF conditions. The screen was narrowed down to two main pools, one of them resulted in a more stable and coherent compound that was later termed NHSM, and several human ESC and IPSC lines were thoroughly examined and characterized in this medium.

Similar to conventional human primed state growth conditions, cells grown in NHSM maintained pluripotent traits as evident by teratoma assay, expression of key pluripotent

factors and derivation of hESCs from human blastocysts. In addition to these attributes, hPSCs maintained in NHSM exhibited a distinct gene expression profile that clustered closer to human ICM and mouse naïve cells grown in 2i/LIF rather than hPSCs grown in conventional medium. DNA methylation pattern of hPSCs grown in NHSM was significantly lower, especially in promoters of pluripotency related genes. H3K27me3 inhibitory mark and H3K4me3 active mark were both at low levels around the TSSs and in gene bodies. Re-activation of X chromosome in female lines was also indicated by low DNA methylation on the XIST promoter and low levels of H3K9me3 repressive mark. Collectively, these results indicate that the genome wide chromatin landscape of human PSCs grown in NHSM possess characteristics affiliated with naïve pluripotency.

Ultimately, we were able to generate interspecies chimeras with hiPSCs grown in NHSM. These cells were able to integrate in several locations in the mouse embryo after morula injection. The integration of the human injected cells to the mouse ICM strongly suggests that both populations were at similar stage. Furthermore, immunostaining for pulmonary specific markers showed that the injected cells not only proliferated and survived in the mouse embryo, but also differentiate and express tissue specific markers such as CC10, Pro-SpC and Aqp5. Human nucleus specific antibody strengthens the fact that the injected cells preserved their human identity, and excluded the possibility of contaminations in the injected population. These results suggest that the human injected iPSCs are able to proliferate, differentiate and survive in the mouse lung, while maintaining their human identity. Injection of P53 KO cheater cells into the mouse morula resulted in significantly higher contribution of human injected cells, probably due to higher proliferation and survival rate.

Following our naïve medium conditions, several groups have published their versions of naïve medium combinations. This phenomenon could imply on the importance of this study, the holy grail of human naïve pluripotency is of high importance as it could impact the world of pluripotency and regenerative medicine. Although many groups are on continuous pursue for these optimal human naïve growth conditions, the ability to generate interspecies chimera remained elusive. In my work I have shown that human cells grown in NHSM can integrate, proliferate and differentiate in the mouse embryo, an ability that may hold a great promise in regenerative medicine, where cells could differentiate in animal models instead of conventional *in vitro* environment.

Materials and methods

Murine stem cell lines and cell culture

Reprogramming and maintenance of murine pluripotent cells were conducted in serum-free chemically defined N2B27-based media: 240 ml DMEM/F12 (Biological Industries – custom made), 240 ml Neurobasal (Invitrogen; 21103), 5 ml N2 supplement (Invitrogen; 17502048), 3.33-5 ml B27 supplement (Invitrogen; 17504044), 1 mM glutamine (Invitrogen), 1% nonessential amino acids (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma), penicillin-streptomycin (Invitrogen), 5 mg/ml BSA (Sigma). Naïve conditions for murine PSCs included 5 µg recombinant human LIF (Millipore; LIF1005) and small-molecule inhibitors CHIR99021 (CH, 3 µM- Axon Medchem) and PD0325901 (PD, 1 µM - TOCRIS). Primed N2B27 media for murine cells (EpiSCs) contained 8 ng/ml recombinant human bFGF (Peprotech Asia) and 20 ng/ml recombinant human Activin (Peprotech Asia). Stem cell lines and mice deficient for Utx were generated by targeted disruption of the endogenous Utx locus via homologous recombination. The targeting strategy and construct (European Conditional Mouse Mutagenesis Program) introduced loxP sites spanning the third exon that would result in an out-of-frame and truncated product upon deletion, and introduced a LacZ reporter cassette driven by the endogenous Utx promoter. 50 µg DNA of the targeting construct was linearized and electroporated into 4 different male pluripotent cell lines: (i) V6.5 ESCs (ii) NGFP1 iPSC15, (iii) V19-OSKM ESC lines that were then subjected to selection with G418 (300microg/ml) and Gancyclovir. After 10 d of selection, double resistant clones were analyzed for correct targeting of 5' and 3' arms (Supplementary Table 5). $Utx^{flox/Y}$ ESCs showed slightly reduced levels of Utx mRNA and protein indicating that Utx^{flox} is a hypomorphic allele, but not a full null allele. Generation of Utx^{Δ} null allele was achieved following transfection with pPac-Cre construct, PSC picking and sub-cloning, followed by genomic DNA genotyping for detection of allele deletion by PCR (Supplementary Table 5). $Utx^{+/flox}$ females were crossed with PGK1-Cre^{m} transgenic male to generate $Utx^{+/\Delta}$; PGK-Cre+ adult females or $Utx^{\Delta/y}$ embryos. $Utx^{+/\Delta}$;PGK-Cre+ females were crossed with $Utx^{flox/y}$ male to obtain all expected genotypes.

Human naïve pluripotent stem cell lines and cell culture

The following conditions, termed NHSM (naive human stem cell medium) conditions were used to isolate, generate, derive and stabilize naive human pluripotent stem cells (iPSCs and ESCs). NHSM includes:

475ml knock-out DMEM (Invitrogen), 5g AlbuMAX I (Invitrogen), 50ml KSR (Thermo),1 mM glutamine (Invitrogen), 1% non-essential amino acids (Invitrogen), 0.1mM b-mercaptoethanol (Sigma), Penicillin-Streptomycin (Invitrogen).

Cytokines: 10µg of recombinant human LIF (Peprotech), 8ng/ml recombinant FGF2 (Peprotech) and 1ng/ml recombinant TGFβ1 (Peprotech).

Small molecule inhibitors: PD0325901 (1µM, ERK1/2i, Axon Medchem); CHIR99021 (3µM, GSK3βi, Axon Medchem); SP600125 (10µM, JNKi, TOCRIS) and SB203580 (10µM, p38i, TOCRIS). After further optimization we alternatively used SB202190 (5µM, p38i, Axon Medchem) or BIRB796 (2µM,p38i Axon Medchem) for enhanced p38 inhibition.

Rock inhibitor Y-27632 (5µM, Axon Medchem) was added to human cells 4-24 hours before and after splitting. PKC (protein kinase C) inhibitor Go6983 (5mM, PKCi, TOCRIS) can be permanently used in NHSM conditions and result in positive metabolomics effect and reduction in background apoptosis levels in the cell cultures.

Naive hESCs/hiPSCs were grown on gelatin (0.2%) + Vitronectin (1µg/ml) -coated plates or on mouse ICR-DR4 embryonic fibroblast (MEF). Media was replaced every 24h. Cells were passaged by single-cell trypsinization (0.05% trypsin EDTA) every 3–4 days.

The following already established human ESC and iPSC lines were used: human induced pluripotent stem cells C1, C2, FX17 human iPSC lines and the human embryonic stem cell (human ESC) lines BGO1 (National Institutes of Health ID code BG01; BresaGen), H1, H9, WIBR3, WIBR3-MBD3mut, WIBR3 OCT4-GFP, WIBR3 ΔPE-OCT4-GFP (gift from Prof. Rudolph Jaenisch lab), H9 NANOG-GFP (WICELL), HUES64 DNMT1-Tet off (Meissner lab). The following hESCs lines were derived from human embryos using NHSM medium: WIS1, WIS2, LIS1, LIS2.

Culture of conventional/primed human ESCs and iPSCs

Human ES cells were maintained on irradiated mouse embryonic fibroblast (MEF) feeder layers or gelatin & vitronectin-coated plates, in human ES cell medium:

425ml knock-out DMEM or DMEM-F12 (Invitrogen), 20% KSR (Invitrogen), 1mM glutamine (Invitrogen), 1% nonessential amino acids (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma). Cytokines: FGF2 (Peprotech, 8ng/ml). Recombinant human TGF-β1 (Peprotech, 1ng/ml) was added in feeder-free conditions. Cultures were passaged every 5–

7 days manually using Collagenase and scraping or by trypsinization (ROCK inhibitor at 10mM concentration was added 4-24h pre and 24h trypsinization). For transfection of hiPSC and hESC lines, cells were collected with 0.25% trypsin-EDTA solution (Invitrogen), and resuspended in PBS with 100mg DNA constructs (Gene Pulser Xcell System; Bio-Rad; 250 V, 500mF, 0.4-cm cuvettes). Antibiotics selection was applied 48h after transfection.

ESC derivation from human blastocysts

The use of human pre-implantation embryos for ESC derivation was performed in compliance with protocols approved by a Weizmann Institute and LIS hospital ESCRO committees, Lis hospital Institutional review committee and Israeli National Ethics Committee (7/04-043) and following the acceptance of a written informed consent. The couples participation in the study was voluntary after signing informed consent forms and there was no monetary compensation for their embryo donation. Inner cell masses were isolated mechanically by laser-assisted micromanipulation from spare *in vitro* fertilized embryos, at day 6–7 following fertilization. The intact inner cell mass clumps were placed on a feeder cell layer of irradiation treated DR4 mouse embryonic fibroblasts and cultured in NHSM media. Initial outgrowths of proliferating ESCs were evident by day 6, and were trypsinized into single cells, 6–10 days following inner cell mass plating. The newly established cell lines were further propagated by trypsin and then either frozen or used for further analysis.

Epigenetic reversion of murine primed epiblast cells

Naive NGFP1-iPSC [55] cells maintained in 2i/LIF conditions, were injected into BDF2 blastocysts and epiblast. Chimeric embryos were dissected at day E6.5 and explanted on fibronectin/vitronectin coated plates in N2B27 bFGF/Activin conditions supplemented with 1 µg/ml puromycin, allowing the isolation of NGFP1-EpiSCs. For epigenetic reversion of murine EpiSCs to naïve pluripotency, cells were passaged into N2B272i/LIF conditions on vitronectin (1 µg/ml) and gelatin (0.2%) coated plates with or without DOX as indicated. When epigenetic reversion assay involved single cell plating, EpiSC growth medium was supplemented with ROCK inhibitor for 24 hours before trypsinization.

Reprogramming of murine somatic cells and cell transfection

Virus-containing supernatants of the different reprogramming viruses (FUW-tetO-lox-hKLF4, FUW-tetO-lox-hOCT4 and FUW-tetO-lox-hSOX2; FUW-tetO-mKlf4, FUW-tetO-

mOct4, FUW-tetO-mSox2 and FUW-tetO-c-Myc) was supplemented with the FUW-lox-M2rtTA virus (when necessary) and an equal volume of fresh culture medium for infection. Mouse fibroblast and Pre-B cell reprogramming was conducted by plating 5000 cells per well in 6 well gelatin coated plates in N2B27 2i/LIF or bFGF/Activin conditions supplemented with Doxycylcine. shRNA and over-expression experiments were conducted on $Utx^{\Delta/Y}$ or $Utx^{+/Y}$ MEF cells as indicated, that were subjected to lentiviral shRNA infection or Amaxa transfection (Lonza) with mammalian expression vectors. The following lentiviral vectors (pLKO.1-puro) purchased from Sigma Aldrich were used: murine Utx shRNA #1-TRCN0000096243; human UTX shRNA #1 -TRCN0000107760; human UTX shRNA #2 TRCN0000107761; mouse Jmjd3 shRNA #1 TRCN0000095266; mouse Jmjd3 shRNA #2 - TRCN0000095268; human JMJD3 shRNA #1- TRCN0000236677, human JMJD3 shRNA #2 - TRCN0000236678.

Teratoma assay

For teratoma formation and analysis, naive human ESCs and iPSCs were collected by trypsinization before injection. Cells were injected sub-cutaneously into NSG mice (Jackson laboratories). Tumors generally developed within 4–6 weeks and animals were euthanized before tumor size exceeded 1.5 cm in diameter. All animal studies were conducted according to the guideline and following approval by the Weizmann Institute IACUC (approval 00960212-3). Tumor mass extracted and fixed in 4% paraformaldehyde overnight. Slides were prepared from the paraffin embedded fixed tissue, which were next Hematoxylin & Eosin stained and inspected for representation of all three germ layers.

DNA plasmids

The following mammalian constitutive over-expression vectors were used in somatic cells undergoing reprogramming following Amaxa Nucleoeffector transfection (Lonza): pCS2-UTY-F, PCS2-UTX-F, pCS2-JMJD3-F, pCMV-HA-JMJD3 (Addgene). Constitutively expressed lenti-viruses FUW-Utx, FUW-UTX and FUW-UTXH1146A were generated by insert cloning into EcoRI sites of FUW vector to generate constitutive expression following viral transduction and stable integration in somatic or PSC lines. pMXs moloney viruses encoding Nanog, Sall4, and Utf1 were obtained through Addgene.

Immunoprecipitation and immunoblotting analyses

HEK293T cells were transfected with each cDNA clones in an expression vector using jetPEI (Polyplus transfection) and were lysed 48 hours later in lysis buffer (5 mM Tris-HCl pH 7.4, 15 mM NaCl, 0.1% Triton and 1.5 mM EDTA). The following plasmids were used for

transfections in different combinations: pCMV-HA-UTX, pCMV-HA-JMJD3 pMSCV-Flag-OCT4, pMSCV-Flag-SOX2, pMSCV-Flag-KLF4 (obtained through addgene). 30 µl of anti-FlagM2 beads (Sigma) were incubated for 4hrs in the anti-flag fractions, where 5 µg of IgG were added to the anti-IgG fractions for 4hrs and then 30 µl of G-sepharose beads were added for another 4hrs. Subsequently, both fractions (the anti-flag and anti-IgG) were centrifuged and the beads were washed six times with lysis buffer. The binding proteins were eluted with 1 mg/ml of flag peptide buffer for the anti-flagM2 beads (except for OCT4 were the beads where boiled) or by boiling with sample buffer and analyzed by SDS–polyacrylamide gel electrophoresis and immunoblotting. The immunoblot analyses were performed using the following antibodies: anti-Flag (clone M2, F3165, Sigma), Anti-Flag (M2 affinity gel, A2220, Sigma), anti-HA (clone 16B12 MMS-101R, Covance), Anti-P53 (PAb421 and PAb248, Millipore).

Immunofluorescence staining.

For immunofluorescence of gonads, embryos from blastocyst injection of *Utx*$^{flox/Y}$ BVSC, Utx$^{\Delta/Y}$ BVSC and Utx$^{\Delta/Y}$ V6.5 ESCs were dissected at E12.5 or at E10.75. Embryos were further dissected at the forelimb level and the anterior half was stained for X-Gal to determine the chimeric embryos. The posterior half or isolated gonads (dissected from the posterior portion with attached mesonephros) of E10.75 and E12.5 embryos, respectively, were placed in mesh inserts (Corning, 3477), fixed in 4% PFA in 0.1M PB (pH 7.2) at 4ºC for 2–3h, washed three times at 4uC in PB for 10min, cryoprotected by sequential incubation in 10, 20 and 30% sucrose (each for at least 6 h) in PB at 4ºC, embedded in Tissue-Tek OCT compound (Sakura), frozen on dry ice, and stored at 270ºC. Sections (10–12 mm) were cut, mounted on Superfrost plus slides (Thermo Scientific), and stored up to 6 months at 270ºC. For immunofluorescence, sections were equilibrated to room temperature (25ºC), rinsed in PBS, and blocked and permeabilized in Blocking solution (5% normal donkey serum, 0.5% BSA, 0.5% Glycine, 0.1% Triton X-100 in PBS) for 1 h at room temperature (for epigenetic markers staining, sections were rinsed in 0.5% Triton/PBS before blocking step to improve permeabilization). Slides were then incubated in the appropriate primary antibody diluted in blocking solution at 4ºC overnight. Sections were then washed three times (5 min each) in PBS plus 0.1%

Mouse embryo micromanipulation and imaging.

Pluripotent mouse ESCs were injected into BDF2 diploid blastocysts, collected from hormone-primed BDF1 6-week-old females. Microinjection into E3.5 blastocysts placed in

M16 medium under mineral oil was done by a flat-tip microinjection pipette. A con-trolled number of 10–12 cells were injected into the blastocyst cavity. After injection, blastocysts were returned to KSOM media (Invitrogen) and placed at 37 ºC until transferred to recipient females. Ten to fifteen injected blastocysts were transferred to each uterine horn of 2.5 days post coitum pseudo-pregnant females. 4n tetraploid complementation assay was performed by fusing BDF2 embryos at two-cell stage, and subsequently allowing the embryos to develop until the blastocyst stage at day 3.5, and were then used for mouse ESCs microinjection. Embryos were allowed to develop into full term. For human naive iPSCs injection, C1 naive iPSC line was constitutively labelled by EGFP (by ZFN-mediated targeting of AAVS1 locus with CAGGS-EGFP knock-in reporter). Cells were trypsinized and microinjected into E2.5 BDF2 diploid morulas (10–15 cells per embryo), and allowed to develop ex vivo until E3.5 (in KSOM supplemented medium). Up to 20 injected blastocysts were transferred to each uterine horn of 2.5 days post coitum pseudo-pregnant females. Embryos were dissected at the indicated time points and imaged for GFP^{+} cell localization. The latter experiments were approved by Weizmann Institute IACUC (00330111-2) and by the Weizmann Institute Bioethics and ESCRO committee. *In toto* confocal live imaging of chimeric mouse embryos was conducted as previously described [75]. Briefly, embryos were removed from the uterus into Tyrode's solution, the decidua removed and the embryo carefully isolated with yolk sac intact. Embryos were moved into a droplet of culture media on a paper filter attached to a coverslip with vacuum grease. Starting farthest away from the embryo, the yolk sac was gently peeled open and lightly pressed onto the paper filter, which has adherence qualities, to anchor the embryo to the filter. To expose the tissue of interest, the embryo can be stretched and lightly pressed onto the filter. After mounting the embryos, a glass-bottomed dish was placed over the embryos using vacuum grease drops for spacing. The dish was inverted, filled with culture media and placed in the imaging chamber [75]. No blinding or randomization was conducted when testing outcome of microinjection experiments. Mounted embryos were placed in a heat- and humidity-controlled chamber (37 ºC; 5% O_2, 5% CO_2 and 90% N_2 gas mixture) on a motorized stage of an inverted Zeiss LSM700 confocal microscope. Images were acquired with a 20X/ 0.8 M27 lens with 0.5–0.6 digital zoom-out or 10X/0.3 Ph1 M27 lens with 0.5- 0.6 digital zoom-out. For GFP detection, a 448-nm laser (30–50% power) was used. CellTracker (red, Molecular Probes, used at 20 mM) was excited by a 555-nm laser (30% power). Hoechst (Sigma, used at 10 mg/ml [76]) was excited by a 405-nm laser (40% power). Images were acquired at 1,024 x 1,024 pixel resolution. The thickness of z-slices appears in the relevant figure legends. Post imaging

processing was done using Zeiss Zendesk software (Carl Zeiss) and Adobe Photoshop CS4. Non- injected mouse embryos were routinely used as negative controls for signal detection.

Reverse transcription and quantitative PCR analysis

Total RNA was isolated using the RNeasy Kit (Qiagen) or TRIZOL (Invitrogen). Three micrograms of total RNA were treated with DNase I to remove potential contamination of genomic DNA using a DNA Free RNA kit (Zymo Research). One microgram of DNase-I-treated RNA was reverse transcribed using a First Strand Synthesis kit (Invitrogen) and ultimately re-suspended in 100 µl of water. Quantitative PCR analysis was performed in triplicate using 1/50 of the reverse transcription reaction in an Viia7 platform (Applied Biosystems). Bars represent RQ (relative quantity). Error bars indicate standard deviation of triplicate measurements for each sample and represent one repeat of 3 biological repeats for each analysis.

qPCR primers:

Part I:

	Forward	Reverse
Utx_3'LRgen1	CCTTCTTGACGAGTTCTTCTGAGCG	AAGTGGAGGAAACCTAGGATGCACC
Utx_5'LRgen1	CTGTCTTGCCATCCTGAGTGTTGG	TGTTTGGGGCAAGTGTGGAGG
Utx3arm		GAACAGTTTTCCTTTACTGTGG
Actb	GGCTGTATTCCCCTCCATCGTG	TGGGGTACTTCAGGGTCAGGATACC
GapDH	CATTGTGGAAGGGCTCATGACCA	GCAGGGATGATGTTCTGGGCAG
Utx34	TGAAGGGAAAGTGGAGTCTGAT	CCTCTGGTATGCAGATAATGCTT
UtxRT4	GTTCGTGAGGTTTCATGAAGATGGC	AGATTAGAGATTCGTAGCAG
Jmjd3	GTATTCCCCTAGTCCTGCATCA	GAACTGAGATGACGAGGAAACC
Klf4	GCACACCTGCGAACTCACAC	CCGTCCCAGTCACAGTGGTAA
Nanog	CCTCCAGCAGATGCAAGAACTC	CTTCAACCACTGGTTTTTCTGCC
Oct4	GCTCACCCTGGGCGTTCTC	GGCCGCAGCTTACACATGTTC
c-Myc	CCACCAGCAGCGACTCTGA	TGCCTCTTCTCCACAGACACC
Gata6	GACGGCACCGGTCATTACC	ACAGTTGGCACAGGACAGTCC
Brachyury	GCTCTAAGGAACCACCGGTCATC	ATGGGACTGCAGCATGGACAG
Otx2	AAGACCCGGTACCCAGACATC	TTGGCGGCACTTAGCTCTTC
Olig3	CTGAATGATGATGGCACGTC	CTGTTTCCTGCACTGCACAT
Sox4	ACAGCGACAAGATTCCGTTCATC	CGTTGCCCGACTTCACCTTC
Sall1	TGAGCACGTGTGAAGCATCTC	CAATCCAGTGGCCAGAAACC

Sall4	TTGCAAGGAAGTTAGCCAATTAAA	CCTGATTCGGGAATGCCTAGA
Utf1	CAAATCCCAGCAACCACATG	GCTGACTTCAAACACACCAGAAGA
Nanog_RT5	CACAAGGACTGATCGGCAAA	CCCGAGACTGGCCTCACA

Part II:

	Forward	Reverse
HPRT	TGACACTGGCAAAACAATGCA	GGTCCTTTTCACCAGCAAGCT
XIST	AGGGAGCAGTTTGCCCTACT	CACATGCAGCGTGGTATCTT
DNMT3A	AGTACGACGACGACGGCTA	CACACTCCA CGCAAAAGCAC
DNMT3B	AGGGAAGACTCGATCCTCGTC	GTGTGTAGCTTAGCAGACTGG
DNMT3L	TGAACAAGGAAGACCTGGACG	CAGTGCCTGCTCCTTATGGCT

Gene expression data acquisition, processing and analysis

Total RNA was isolated from indicated cell lines for human cell lines also indicated primed and naive (in NHSM with 5mM SB203580). The concentration of RNA was quantified and subjected to quality control on Agilent Bioanalyzer. 250 ng of RNA was processed from each sample. cDNA was fragmented, labeled, and hybridized to Affymetrix Mouse Gene 1.0 ST GeneChip (Affymetrix, Santa Clara, CA), which contain 35,557 probes or Affymetrix Human Gene 1.0 ST GeneChip (Affymetrix). Transcripts levels were processed from image files using RMA method24, which corrects for non-biological sample variation using quantile normalization, implemented by the Affymetrix "Expression Console" software. In human samples, transcripts levels were processed from Affymetrix CEL files and CDF file (version V1.r3, which maps probes into 33,252 probe sets) using Robust Multi-Array Average method. All analyses in part II were done using Matlab (Version R2012b) and its Bioinformatics toolbox. Data in part I was further filtered to include probes that have at least one call higher than 64 (526), resulting in 21,811 probes that are mapped to 15,815 RefSeq transcripts (http://www.ncbi.nlm.nih.gov/RefSeq/), in part II data was further filtered to include probes that have at least one call higher than 64 (526), and mapped to unique Entrez IDs using Affymetrix annotations (NetAffx Annotation Files version 33.1) and NCBI sites. Probes targeting the same gene are represented by their median, resulting in 13,894 genes. Microarray data from part I are available at the National Center for Biotechnology Information Gene Expression Omnibus database under the series accession no. GSE37822. For gene expression statistical analysis, we used "Cluster" software

(http://rana.lbl.gov/EisenSoftware.htm) to run hierarchical clustering on the samples, using complete linkage centered correlation as a distance metric. We used Matlab version R2011a to run Principle Component Analysis that detects the principle components with the largest variation in the data. Differentially expressed gene signatures include transcripts that are above 2-fold change in $Utx^{\Delta/Y}$ compared to $Utx^{+/Y}$, or in MEF +DOX compared to MEF -DOX. For the functional analysis the signature genes were mapped from RefSeq transcripts to Entrez Gene IDs.

Microarray data are available at the series accession number GSE46872. The samples were hierarchically clustered using average linkage and either Spearman or Pearson correlation as a distance matrix, with similar results. To compare our samples to hICM cells, human pre-implantation data from [71] were included. This data, from Affymetrix HuGene 1.0 st microarrays described before, was passed through the same processing, resulting in 16,953 genes before and 12,062 genes after intersection with the previous gene list. Batch effect was corrected by normalization of the new data to the mean values of its ES C samples, making it consistent with the normalization of our data.

Cross-species gene expression analysis was conducted on human arrays described above and previously described mouse ES cell and EpiSC gene expression data sets on an Agilent 4344 k array platform (GSE15603) containing 45,018 probes. Mouse data were processed as described above, resulting in 17,885 unique genes [77]. Human–mouse orthology was downloaded from MGI (http://www. informatics.jax.org, on April 2013) containing 17,772 pairs of orthologous genes. Of these, 9,803 were mapped to our expression data. The expression values from mouse and human were transformed separately into relative abundance values: For each gene, the relative abundance value is the expression value divided by the mean of expression values within the same gene across samples in the same species. The resulting expression matrix was subjected to hierarchical clustering (Spearman correlation, average linkage).

Chromatin Immuno-precipitation and sequencing library preparation

Approximately 60*10^6 cells were cross-linked in formaldehyde (1% final concentration, 10 min at room temperature (RT)), and then quenched with glycine (5 min at RT). Fixed cells were lysed in 50mM HEPES KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% Glycerol, 0.5% NP-40 alternative, 0.25% Triton supplemented with protease inhibitor at 4°C (Roche, 04693159001), centrifuged at 950 x g for 10 min and re-suspended in 0.2% SDS, 10 mM EDTA, 140 mM NaCl and 10 mM Tris-HCL. Cells were then fragmented with a Branson Sonifier (model S-450D) at -4°C to a size range between 200 and 800 bp, and precipitated by

centrifugation. 10 µg of each antibody was pre-bound by incubating with Protein-G Dynabeads (Invitrogen100-07D) in blocking buffer (PBS supplemented with 0.5% TWEEN and 0.5% BSA) for 2 hours at room temperature. Washed beads were added to the chromatin lysate, and then incubated overnight. Samples were washed 5 times with RIPA buffer, twice with RIPA buffer supplemented with 500 mM NaCl, twice with LiCl buffer (10 mM TE, 250 mM LiCl, 0.5% NP-40, 0.5% DOC), once with TE (10 Mm Tris-HCl pH 8.0, 1 mM EDTA), and then eluted in 0.5% SDS, 300 mM NaCl, 5 mM EDTA, 10 mM Tris Hcl pH 8.0 at 65°C. Eluate was incubated in 65°C for 4 hours, and then treated sequentially with RNaseA (Roche, 11119915001) for 30 min and Proteinase K for two hours. DNA was purified with The Agencourt AMPure XP system (Beckman Coulter). Libraries of cross-reversed ChIP DNA samples were prepared according to a modified version of the Illumina Genomic DNA protocol, as described previously [78]. 5pmole of DNA library was then applied to each lane of the flow cell and sequenced on Illumina Hiseq2000 sequencer according to standard Illumina protocols, at the INCPM unit, Weizmann institute. The following antibodies were used for chromatin-IP experiments: Control IgG (ChIP grade, ab46540, abcam), Anti-Histone H3 tri methyl K4 (ChIP grade, ab8580, abcam), anti-Histone H3 (tri methyl K27 (ChIP grade, 07-449, Millipore), anti-KDM6A UTX (ab84190 abcam) and affinity purified anti-Utx serum as previously described32. In Part I, genomic localization and gene expression experiments, DOX treatment was conducted for 8 days.

Chromatin IP sequencing data analysis

In part I, we used bowtie software [79] version 0.12.5 to align reads to the mouse mm9 reference genome (University of California, Santa Cruz (UCSC), July 2007). We only considered reads that were uniquely aligned to the genome with up to a single mismatch, taking the single best match of each read. We identified enriched intervals of H3K4me3, H3K27me3 and Utx using MACS version 1.4.1 [28]. We used sequencing of whole cell extract as a control in order to define a background model. Duplicate reads aligned to the exact same location are excluded by MACS default configuration. Enriched intervals were mapped to genes if they overlapped a 3 Kb symmetric interval before and after their TSSs (taken from the UCSC mm9 Known Gene Table). Functional enrichment was calculated in the same way as for gene expression (described above). Reads land- scape was processed and visualized using IGV software version 2.0 [80]. ChIP-seq data on wild-type samples were highly reproducible in comparison to previous publications [61, 81] (data not shown). For motif detection, motifs that are enriched in Utx binding locations were detected using the SeqPos tool [82] in the Cistrome package (http://cistrome.org/ap/). Utx peaks in ESCs were

run against the Cistrome-curated motif database. For Utx coverage analysis, the mean density of Utx binding reads was calculated for each gene in the UCSC Known Genes Table (version mm9). The fold change between the mean density in MEFs with doxycycline (8 days) and MEFs without doxycycline (day 0) was then calculated for each gene. Binding profiles around the TSS were calculated using SitePro software in the Cistrome package (http://cistrome.org/ap/). The profiles over all mouse genes were calculated using the UCSC Known Genes Table (version mm9), including 55,419 regions that are mapped to 29,952 Entrez genes. The average number of reads in a 50-bp window was calculated 3 Kb before and after the TSS. These averages were then converted to a Z-score by normalizing each window with the mean and standard deviation of that window, as was calculated for each sample and over 10,000 random locations in the genome. The profiles over OSKMN targets were calculated in a similar manner, but over 805 genes that were previously found to be bound by at least 3 of the following factors: Oct4, Sox2, Klf4, Myc and Nanog [57]. The profiles around Utx binding peaks were calculated in a similar manner, but only over the peaks identified by MACS software. Chromatin heat- map was generated using the SeqMINER program directly from aligned read files, without further normalization.

In part II, the chromatin markers H3K27me3, H3K4me3, H3K4me1, H3K27Ac and H3K9me3 were measured in the different human pluripotent cell lines: C1, WIBR3, LIS2 (naive and primed), BGO1 and WIBR3-MBD3mut (naive). WIS1 naive and primed hESCs were profiled for the H3K9me3 mark only. In addition, H3K27me3, H3K4me3, H3K4me1 and H3K27Ac were measured in mouse V6.5 naive and 129 primed EpiSCs. H3K27me3 and H3K4me3 measurements in naive mouse ES were previously published [27]. Each sample was accompanied by control sequencing experiment of whole-cell extract input. We used bowtie software version 1.0.0 to align human reads to human reference genome hg19 (UCSC, February 2009) and mouse reads to mouse mm9 reference genome (UCSC, July 2007). To rule out sequencing depth bias, the aligned sequences were down-sampled such that all samples had the same number of aligned reads. Human samples of the marks H3K4me3, H3K27me3, H3K4me1 and H3K27Ac, as well as whole-cell extract, were down-sampled to include 3,750,000 aligned reads. H3K9me3 samples were down-sampled to 5,900,000 aligned reads. Mouse samples of the marks H3K4me3, H3K27me3, H3K4me1 and H3K27Ac, were down-sampled to 3,420,000 aligned reads. Chromatin profiles were calculated over all RefSeq genes (n=543,463), and over developmental genes in the following way: (1) Read densities were calculated between 3 kb upstream to TSS, and 3 kb downstream to TES. (2) Each gene was divided to 100 bins of identical size, and the sum of reads in each

bin was calculated. (3) Average profile was calculated over all genes, where the gene body, which is of changing size, was represented by 100 quantiles. Genes of size less than 1 kb were filtered out. (4) Profiles of human samples represent mean and s.d. (error bars) of primed and naive samples. Lastly, developmental genes were selected if they have a GO annotation that is related to development or differentiation. Using this criterion, we had 5,922 RefSeq human developmental genes, and 420 RefSeq mouse developmental genes. Concrete examples of genes were processed and visualized using IGV software version 2.0. To measure the distribution of H3K9me3 accumulation in chromosome X genes, RPKM (reads per kilobase per 5.9 million reads) was calculated for each gene (between 1kb upstream to TSS and TES). P-values between distributions were calculated with one-tail paired-sample t-test.

DNA methylation analysis

DNA was proteinase-K-treated and extracted, and 1 mg of DNA was subjected to conversion using the Qiagen EpiTect Bisulphite Kit. Promoter regions of OCT4 were amplified using previously described primers [42]: XIST forward primer (used on bisulphite treated DNA): TAAATTTTAAATTAATTAAATTAT; XIST reverse primer (used on bisulphite treated DNA): TGTTTTAGAAAGAATTTTAAGTGTAGAGA. PCR products were cloned using the pCR2.1-TOPO vector and sequenced using the M13 forward primer. Primers and methylation analysis approach for FMR1 promoter were previously described [83, 84].

Reduced representation bisulphite sequencing libraries for genome wide sequencing profile of DNA methylation.

RRBS libraries were generated as described previously [83]. Briefly, DNA was isolated from snap-frozen cell pellets using the Quick-gDNA mini prep kit (Zymo). Primed WIBR3 human ESC and naive WIBR3 human ESC (after 17 day reversion in NHSM) conditions were used for this analysis. For mouse pluripotent cell DNA methylation analysis, we used mouse 129 EpiSCs and naive ESCs in 2i/LIF for 90 days. Isolated DNA was then subjected to MspI digestion (NEB), followed by end repair using T4 PNK/T4 DNA polymerase mix (NEB), A-tailing using Klenow fragment (3'→5' exo-) (NEB), size selection for fragments shorter than 500 bp using SPRI beads (Beckman Coulter) and ligation into a plasmid using quick T4 DNA ligase (NEB). Plasmids were treated with sodium bisulphite using the EZ DNA Methylation-Gold kit (Zymo) and the product was PCR amplified using GoTaq Hot Start DNA polymerase (Promega). The PCR products were A-tailed using Klenow fragment, ligated to indexed Illumina adapters using quick T4 DNA ligase and PCR amplified using GoTaq DNA polymerase. The libraries were then size-selected to 200–500 bp by extended gel

electrophoresis using NuSieve 3:1 agarose (Lonza) and gel extrac- tion (Qiagen). Libraries were pooled and sequenced on an Illumina HiSeq 2500 system. The sequencing reads were aligned to either the Human Genome Build 37 (hg19) or the Mouse Genome Build 37 (mm9) using Bismark. Methylation levels were calculated for all CpGs that were covered by 10 or more distinct sequencing reads in both the primed and naive samples. For each such CpG, its methylation level in the primed sample was deducted from the methylation level in the naive sample, and only CpGs with a non-zero difference were included in the plotted histograms. Histograms of the per CpG difference in methylation were drawn for CpGs residing in CpG rich areas. CpGs were categorized as belonging to CpG rich area if the number of CpG dinucleotides found within a 500-bp window surrounding them exceeded 20 (that is, 4% of all nucleotides). The statistical significance of the reduction in DNA methylation was assessed using a paired-samples t-test over all CpGs with a coverage of 310 or more.

Generation of P53 knockout human ESC lines via CRISPR/Cas9

In order to knock out P53 gene, oligos encoding gRNAs targeting this gene were inserted into px335 vector [85]. Unique gRNA sequences were chosen with the help of Zhang Lab website http://www.genome-engineering.org/crispr/ 100 µg of resulting construct and 10 µg of mCherry expressing vector were electroporated into primed LIS2 hESC. 3 days later mCherry expressing cells were sorted by FACS and seeded at low density. 9 days after seeding, colonies were picked for each experiment and genomic DNA was extracted. DNA was analyzed by High Resolution Melt assay (HRM) using MeltDoctor reagent (Life Technologies) and the clones that showed reduced Tm for both alleles, compared to wild-type controls, were expanded for further validation analysis. In these selected clones targeted locus was amplified and sequenced. Primers for HRM of P53 locus were: Forward: CCTCTGACTGCTCTTTTCACCCATC; Reverse: TGGGAGCTTCATCTGGACCTGG. All relevant plasmids have been directly deposited by our group and are made available through Addgene.

References

1. Yu, J., et al., *Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells.* Science, 2007. **318**(5858): p. 1917-1920.
2. Yamanaka, S. and H.M. Blau, *Nuclear reprogramming to a pluripotent state by three approaches.* Nature, 2010. **465**(7299): p. 704-712.
3. Takahashi, K. and S. Yamanaka, *Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors.* Cell, 2006. **126**(4): p. 663-676.
4. Park, I.-H., et al., *Reprogramming of human somatic cells to pluripotency with defined factors.* Nature, 2008. **451**(7175): p. 141-146.
5. Orkin, Stuart H. and K. Hochedlinger, *Chromatin Connections to Pluripotency and Cellular Reprogramming.* Cell, 2011. **145**(6): p. 835-850.
6. Ang, Y.-S., et al., *Wdr5 Mediates Self-Renewal and Reprogramming via the Embryonic Stem Cell Core Transcriptional Network.* Cell, 2011. **145**(2): p. 183-197.
7. Barker, N., et al., *Identification of stem cells in small intestine and colon by marker gene Lgr5.* Nature, 2007. **449**(7165): p. 1003-1007.
8. Martin, G.R., *Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.* Proceedings of the National Academy of Sciences of the United States of America, 1981. **78**(12): p. 7634-7638.
9. Zhang, Y. and D. Reinberg, *Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.* Genes & Development, 2001. **15**(18): p. 2343-2360.
10. Smith, A.G., et al., *Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides.* Nature, 1988. **336**(6200): p. 688-690.
11. Thomson, J.A., et al., *Embryonic Stem Cell Lines Derived from Human Blastocysts.* Science, 1998. **282**(5391): p. 1145-1147.
12. Gurdon, J.B., *The Developmental Capacity of Nuclei taken from Intestinal Epithelium Cells of Feeding Tadpoles.* Journal of Embryology and Experimental Morphology, 1962. **10**(4): p. 622-640.
13. Niwa, H., J.-i. Miyazaki, and A.G. Smith, *Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells.* Nat Genet, 2000. **24**(4): p. 372-376.
14. Nichols, J., et al., *Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4.* Cell, 1998. **95**(3): p. 379-391.
15. Mitsui, K., et al., *The Homeoprotein Nanog Is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells.* Cell, 2003. **113**(5): p. 631-642.
16. Chambers, I., et al., *Functional Expression Cloning of Nanog, a Pluripotency Sustaining Factor in Embryonic Stem Cells.* Cell, 2003. **113**(5): p. 643-655.
17. Avilion, A.A., et al., *Multipotent cell lineages in early mouse development depend on SOX2 function.* Genes & Development, 2003. **17**(1): p. 126-140.
18. Goldberg, A.D., C.D. Allis, and E. Bernstein, *Epigenetics: A Landscape Takes Shape.* Cell, 2007. **128**(4): p. 635-638.
19. Shi, Y., *Histone lysine demethylases: emerging roles in development, physiology and disease.* Nat Rev Genet, 2007. **8**(11): p. 829-833.
20. Sanders, S.L., et al., *Methylation of Histone H4 Lysine 20 Controls Recruitment of Crb2 to Sites of DNA Damage.* Cell, 2004. **119**(5): p. 603-614.
21. Agger, K., et al., *The emerging functions of histone demethylases.* Current Opinion in Genetics & Development, 2008. **18**(2): p. 159-168.
22. Shi, Y., et al., *Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1.* Cell, 2004. **119**(7): p. 941-953.

23. Schuettengruber, B., et al., *Genome Regulation by Polycomb and Trithorax Proteins.* Cell, 2007. **128**(4): p. 735-745.
24. Lee, T.I., et al., *Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells.* Cell, 2006. **125**(2): p. 301-313.
25. Bracken, A.P., et al., *Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions.* Genes & Development, 2006. **20**(9): p. 1123-1136.
26. Boyer, L.A., et al., *Polycomb complexes repress developmental regulators in murine embryonic stem cells.* Nature, 2006. **441**(7091): p. 349-353.
27. Mansour, A.A., et al., *The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming.* Nature, 2012. **488**(7411): p. 409-413.
28. De Santa, F., et al., *The Histone H3 Lysine-27 Demethylase Jmjd3 Links Inflammation to Inhibition of Polycomb-Mediated Gene Silencing.* Cell, 2007. **130**(6): p. 1083-1094.
29. Sen, G.L., et al., *Control of differentiation in a self-renewing mammalian tissue by the histone demethylase JMJD3.* Genes & Development, 2008. **22**(14): p. 1865-1870.
30. Lee, M.G., et al., *Demethylation of H3K27 Regulates Polycomb Recruitment and H2A Ubiquitination.* Science, 2007. **318**(5849): p. 447-450.
31. Lan, F., et al., *A histone H3 lysine 27 demethylase regulates animal posterior development.* Nature, 2007. **449**(7163): p. 689-694.
32. Hong, S., et al., *Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases.* Proceedings of the National Academy of Sciences, 2007. **104**(47): p. 18439-18444.
33. Agger, K., et al., *UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development.* Nature, 2007. **449**(7163): p. 731-734.
34. Lee, S., Jae W. Lee, and S.-K. Lee, *UTX, a Histone H3-Lysine 27 Demethylase, Acts as a Critical Switch to Activate the Cardiac Developmental Program.* Developmental Cell, 2011.
35. Tesar, P.J., et al., *New cell lines from mouse epiblast share defining features with human embryonic stem cells.* Nature, 2007. **448**(7150): p. 196-199.
36. Brons, I.G.M., et al., *Derivation of pluripotent epiblast stem cells from mammalian embryos.* Nature, 2007. **448**(7150): p. 191-195.
37. Dutta, D., et al., *Self-Renewal Versus Lineage Commitment of Embryonic Stem Cells: Protein Kinase C Signaling Shifts the Balance.* STEM CELLS, 2011. **29**(4): p. 618-628.
38. Shimizu, T., et al., *Dual Inhibition of Src and GSK3 Maintains Mouse Embryonic Stem Cells, Whose Differentiation Is Mechanically Regulated by Src Signaling.* STEM CELLS, 2012. **30**(7): p. 1394-1404.
39. Acampora, D., L.G. Di Giovannantonio, and A. Simeone, *Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition.* Development, 2012. **140**(1): p. 43-55.
40. Hanna, J.H., K. Saha, and R. Jaenisch, *Pluripotency and Cellular Reprogramming: Facts, Hypotheses, Unresolved Issues.* Cell, 2010. **143**(4): p. 508-525.
41. Bernemann, C., et al., *Distinct Developmental Ground States of Epiblast Stem Cell Lines Determine Different Pluripotency Features.* STEM CELLS, 2011. **29**(10): p. 1496-1503.
42. Hanna, J., et al., *Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs.* Proceedings of the National Academy of Sciences, 2010.
43. Bao, S., et al., *Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells.* Nature, 2009. **461**(7268): p. 1292-1295.

44. Festuccia, N., et al., *The role of pluripotency gene regulatory network components in mediating transitions between pluripotent cell states.* Current Opinion in Genetics & Development, 2013. **23**(5): p. 504-511.
45. Marks, H., et al., *The Transcriptional and Epigenomic Foundations of Ground State Pluripotency.* Cell, 2012. **149**(3): p. 590-604.
46. Evans, M.J. and B. Ephrussi, *Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma.* J Natl Cancer Inst USA,, 1970. **44**(5): p. 1015-1036.
47. Creyghton, M.P., et al., *Histone H3K27ac separates active from poised enhancers and predicts developmental state.* Proceedings of the National Academy of Sciences, 2010.
48. Yeom, Y.I., et al., *Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells.* Development, 1996. **122**(3): p. 881-894.
49. Hanna, J., et al., *Metastable Pluripotent States in NOD-Mouse-Derived ESCs.* Cell Stem Cell, 2009. **4**(6): p. 513-524.
50. Buehr, M., et al., *Capture of Authentic Embryonic Stem Cells from Rat Blastocysts.* Cell, 2008. **135**(7): p. 1287-1298.
51. Chia, N.-Y., et al., *A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity.* Nature, 2010. **468**(7321): p. 316-320.
52. Nichols, J. and A. Smith, *Naive and Primed Pluripotent States.* Cell Stem Cell, 2009. **4**(6): p. 487-492.
53. Li, W. and S. Ding, *Human Pluripotent Stem Cells: Decoding the Naïve State.* Science Translational Medicine, 2011. **3**(76): p. 76ps10-76ps10.
54. Guo, G., et al., *A <italic>PiggyBac</italic>-Based Recessive Screening Method to Identify Pluripotency Regulators.* PLoS ONE, 2011. **6**(4): p. e18189.
55. Hanna, J., et al., *Direct cell reprogramming is a stochastic process amenable to acceleration.* Nature, 2009. **462**(7273): p. 595-601.
56. Han, D.W., et al., *Direct reprogramming of fibroblasts into epiblast stem cells.* Nat Cell Biol, 2011. **13**(1): p. 66-71.
57. Sridharan, R., et al., *Role of the Murine Reprogramming Factors in the Induction of Pluripotency.* Cell, 2009. **136**(2): p. 364-377.
58. Silva, J., et al., *Nanog Is the Gateway to the Pluripotent Ground State.* Cell, 2009. **138**(4): p. 722-737.
59. Ema, M., et al., *Krüppel-like factor 5 Is Essential for Blastocyst Development and the Normal Self-Renewal of Mouse ESCs.* Cell Stem Cell, 2008. **3**(5): p. 555-567.
60. Miller, S.A., S.E. Mohn, and A.S. Weinmann, *Jmjd3 and UTX Play a Demethylase-Independent Role in Chromatin Remodeling to Regulate T-Box Family Member-Dependent Gene Expression.* Molecular Cell, 2010. **40**(4): p. 594-605.
61. Koche, R.P., et al., *Reprogramming Factor Expression Initiates Widespread Targeted Chromatin Remodeling.* Cell Stem Cell, 2011. **8**(1): p. 96-105.
62. Hajkova, P., et al., *Chromatin dynamics during epigenetic reprogramming in the mouse germ line.* Nature, 2008. **452**(7189): p. 877-881.
63. Hayashi, K., et al., *Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by Pluripotent Stem Cells.* Cell, 2011. **146**(4): p. 519-532.
64. Gafni, O., et al., *Derivation of novel human ground state naive pluripotent stem cells.* Nature, 2013. **504**(7479): p. 282-286.
65. Greber, B., et al., *Conserved and Divergent Roles of FGF Signaling in Mouse Epiblast Stem Cells and Human Embryonic Stem Cells.* Cell Stem Cell, 2010. **6**(3): p. 215-226.
66. Watanabe, K., et al., *A ROCK inhibitor permits survival of dissociated human embryonic stem cells.* Nat Biotech, 2007. **25**(6): p. 681-686.

67. Rajendran, G., et al., *Inhibition of protein kinase C signaling maintains rat embryonic stem cell pluripotency.* Journal of Biological Chemistry, 2013.
68. Mekhoubad, S., et al., *Erosion of Dosage Compensation Impacts Human iPSC Disease Modeling.* Cell Stem Cell, 2012. **10**(5): p. 595-609.
69. Ficz, G., et al., *FGF Signaling Inhibition in ESCs Drives Rapid Genome-wide Demethylation to the Epigenetic Ground State of Pluripotency.* Cell Stem Cell. **13**(3): p. 351-359.
70. Chappell, J., et al., *MYC/MAX control ERK signaling and pluripotency by regulation of dual-specificity phosphatases 2 and 7.* Genes & Development, 2013. **27**(7): p. 725-733.
71. Vassena, R., et al., *Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development.* Development, 2011. **138**(17): p. 3699-3709.
72. Betschinger, J., et al., *Exit from Pluripotency Is Gated by Intracellular Redistribution of the bHLH Transcription Factor Tfe3.* Cell, 2013. **153**(2): p. 335-347.
73. Bellusci, S., et al., *Evidence from normal expression and targeted misexpression that bone morphogenetic protein (Bmp-4) plays a role in mouse embryonic lung morphogenesis.* Development, 1996. **122**(6): p. 1693-1702.
74. Borok, Z., et al., *Keratinocyte Growth Factor Modulates Alveolar Epithelial Cell Phenotype In Vitro: Expression of Aquaporin 5.* American Journal of Respiratory Cell and Molecular Biology, 1998. **18**(4): p. 554-561.
75. Massarwa, R.a. and L. Niswander, *In toto live imaging of mouse morphogenesis and new insights into neural tube closure.* Development, 2013. **140**(1): p. 226-236.
76. Tachibana, M., et al., *Generation of Chimeric Rhesus Monkeys.* Cell, 2011. **148**(1): p. 285-295.
77. James, D., et al., *Contribution of human embryonic stem cells to mouse blastocysts.* Developmental Biology, 2006. **295**(1): p. 90-102.
78. Barrett, A., et al., *PLU-1 nuclear protein, which is upregulated in breast cancer, shows restricted expression in normal human adult tissues: A new cancer/testis antigen?* International Journal of Cancer, 2002. **101**(6): p. 581-588.
79. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biology, 2009. **10**(3): p. R25.
80. Robinson, J.T., et al., *Integrative Genomics Viewer.* Nature biotechnology, 2011. **29**(1): p. 24-26.
81. Mikkelsen, T.S., et al., *Dissecting direct reprogramming through integrative genomic analysis.* Nature, 2008. **454**(7200): p. 49-55.
82. He, H.H., et al., *Nucleosome dynamics define transcriptional enhancers.* Nat Genet, 2010. **42**(4): p. 343-347.
83. Rais, Y., et al., *Deterministic direct reprogramming of somatic cells to pluripotency.* Nature, 2013. **502**(7469): p. 65-70.
84. Pietrobono, R., et al., *Quantitative analysis of DNA demethylation and transcriptional reactivation of the FMR1 gene in fragile X cells treated with 5‐azadeoxycytidine.* Nucleic Acids Research, 2002. **30**(14): p. 3278-3285.
85. Cong, L., et al., *Multiplex Genome Engineering Using CRISPR/Cas Systems.* Science, 2013. **339**(6121): p. 819.

List of publications

1) Weiner A., Lara-Astiaso D., Krupalnik V., **Gafni O.,** David E., Winter D. R., Hanna J. H., Amit I. **Co-ChIP enables genome-wide mapping of histone mark co-occurrence at single-molecule resolution.** ***Nature Biotechnology,*** 2016. Online publication.

2) Dhar S., Lee S.H., Chen K., Zhu G., Oh W., Allton K., **Gafni O.,** Kim Y.Z., Barton M., Hanna J.H., Wang Z., Li W., Lee M.G. **An essential role for UTX in resolution and activation of bivalent promoters.** ***Nucleic Acid Research,*** 2016. Online publication.

3) **Gafni O*.,** Weinberger L*., Mansour A.A*, Manor Y.S*., Chomsky E*., Ben-Yosef D., Kalma Y., Viukov S., Maza I., Zviran A., Rais Y., Shipony Z., Mukamel Z., Krupalnik V., Zerbib M., Geula S., Caspi I., Schneir D., Shwartz T., Gilad S., Zalcenstein D.A., Benjamin S., Amit I., Tanay A., Massarwa R., Novershtern N. & Hanna J.H. **Derivation of novel human ground state naive pluripotent stem cells.** ***Nature,*** 2013. 504: 282-286.

4) Rais Y., Zviran A., Geula S., **Gafni O.,** Chomsky E., Viukov S., Mansour A.A., Caspi I., Krupalnik V., Zerbib M., Maza I., Mor N., Baran D., Weinberger L., Jaitin D.A., Astiaso D.L., Gonen R.B.., Shipony Z., Mukamel Z., Hagai T., Gilad S., Zalcenstein D.A., Tanay A., Amit I., Novershtern N. & Hanna J.H. **Deterministic direct reprogramming of somatic cells to pluripotency.** ***Nature,*** 2013. 502: 65-70.

5) Mansour A.A.*, **Gafni O.*,** Weinberger L., Zviran A., Ayyash M., Rais Y., Krupalnik V., Zerbib M., Amann-Zalcenstien D., Maza I., Geula S., Viukov S., Holtzman L., Pribluda A., Canaani E., Horn-Saban S., Amit I., Novershtern N. and Hanna J.H. The H3K27 demethylase **Utx regulates somatic and germ cell epigenetic reprogramming.** ***Nature,*** 2012. 488: 409-413.

* Equal Contribution.

Student declaration

I hereby declare that the thesis presented summarizes my independent research work under the supervision of Dr. Jacob H. Hanna at the Department of Molecular Genetics, the Weizmann Institute of Science. Tissue culture work was done by me, including cell line generation, immunostaining, imaging, FACS analyses, ChIP-seq, ChIP-PCR and RNA purification. ChIP-seq and microarray analysis were done by Dr. Noa Novershtern. Mouse microinjection was done jointly with Mirie Zerbib. Imaging of mouse gonads and LacZ staining were done by Dr. Abed El-Fatah Mansour. In part two, I was intensively involved in the human mouse interspecies chimerism part. Moreover, I was involved in ChIP-seq and tissue culture work, including optimization of human naïve medium together with Dr. Leehee Weinberger. Analysis of gene expression data was done by Yair Manor. Analysis of ChIP-Seq data was done by Dr. Noa Novershtern. RRBS and DNA methylation analysis was done by Elad Chomsky. Planning, cloning and construction of reporter cell lines and mutations were done jointly with Sergey Viukov. Microinjections of human iPSCs to mouse embryos was done by me together with Mirie Zerbib. Immunofluorescence staining and imaging technics were done in close collaboration jointly with Dr. Rada Massarwa. Derivation of new hESCs lines was done jointly in LIS maternity hospital with Dr. Dalit Ben Yosef.

Acknowledgments

After five years of extensive and hard work it is time to thank all of the amazing people wjo helped, mentored and guided me through these past years. I would like to thank my mentor Dr. Jacob H. Hanna who believed in me all along. To Mirie zerbib, Dr. Rada Massarwa, Dr. Leehee Weinberger, Dr. Yoach Rais, Dr. Shay Geula, Vladik Krupalnik, Nofar Mor, Yair Manor, Dr. Itay Maza, Sergey Viyokov, Dr. Abed Al-Fatah Mansour and Dr. Noa Noverstern, thank you for the inspiring, teaching, listening and just creating a great atmosphere in the lab, you all made this journey much more convenient. To my one and only, my wife Moran, thank you for putting up with these past five years in which we got engaged, married, had two beautiful children, May and Dor. Thank you for understanding all the long days, nights and weekends, I could not do it without your support.