Reprogramming Cellular Fate using Defined Factors

By

Ruth K. Foreman

B.A. Biochemistry, History and Sociology of Science
University of Pennsylvania, College of Arts and Sciences, 2001

SUBMITTED TO THE DEPARTMENT OF BIOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MARCH 2009

Signature of Author:_____
Department of Biology
March 13, 2009

Certified by: _____
Rudolf Jaenisch
Professor of Biology
Thesis Supervisor

Accepted by: _____
Stephen P. Bell
Professor of Biology
Chairman, Committee for Graduate Students

I would like to dedicate this work to my father, the late Thomas K. Foreman, whose curiosity and love of science captured my attention at a young age and continues to inspire me.

Acknowledgments

First, I would like to thank Rudolf, for giving me the chance to be a part of such an exciting time in the field of stem cell biology. I have learned a tremendous amount in my tenure in the Jaenisch lab and am grateful for the opportunity.

I would also like to thank my thesis committee members, Richard Hynes and Tyler Jacks for their helpful discussions and constructive feedback on my projects. Thanks also to Rick Young, who, while not an official member of my committee until the very end, was party to numerous discussions over the last few years. Thank you to Chad Cowan who generously agreed to serve as the outside member of my committee.

I would like to give special thanks to my collaborators, including Tobi Brambrink, Grant Welstead, Alex Marson, Brett Chevalier, Judith Staerk, Costas Lyssiotis, Marius Wernig and Alex Meissner, for helping me to accomplish this work.

Special thanks to Caroline Beard for her scientific wisdom and sage advice throughout my tenure in the lab, as well as for her constructive criticism of this thesis.

Thanks:

To Todd Meyerrose for critically reviewing this thesis.

To Divya Mathur for being a terrific lab mate and fantastic friend.

To Konrad Hochedlinger and Teresa Holm for getting me started in the lab.

To Jessie Dausman for all of the wonderful time spent in the mouse house.

To all the members of the Jaenisch lab, past and present. I would also like to thank everyone who helped me at MIT and the Whitehead, especially the Young lab members and Stacie Bumgarner and Jana Hersch.

I am grateful for my friendships with Tobi, Grant, Alex and Brett, which helped me through a particularly difficult time in the lab.

My family and friends have been supportive and loving throughout this process and I love them for it.

The most special thanks goes to my husband, Dom, who has had unending patience and support for me and my career path. You deserve an honorary degree.

Reprogramming Cellular Fate using Defined Factors

by

Ruth K. Foreman

Submitted to the Department of Biology
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy in Biology

Abstract

Embryonic stem (ES) cells have a vast therapeutic potential given their pluripotency, or the ability to differentiate into tissues from all three germ layers. One of the ultimate goals of regenerative medicine is to isolate pluripotent stem cells from patients. Nuclear reprogramming offers the possibility of creating patient-specific cell lines, thus abrogating the need for immunosuppressants following cell transplantation therapy. It was recently reported that the forced expression of four transcription factors, Oct4, Sox2, c-Myc and Klf4 can induce a pluripotent state in somatic cells, without the need for embryo destruction. The work presented here aims to characterize reprogramming using defined factors and provide insight into the mechanisms governing this process. It also seeks to identify transient cues to induce reprogramming in somatic cells, alleviating the need for virally transduced transcription factors that hinder its eventual clinical use.

Thesis Supervisor: Dr. Rudolf Jaenisch
Title: Professor of Biology

Table of Contents

Chapter 1

Introduction

Adult, multicellular organisms are heterogeneous collections of specialized cell types arising from a totipotent zygote. As that single cell zygote divides and proceeds through development it undergoes irreversible differentiation events that result in a progressive loss of potency. For example, embryonic stem (ES) cells are pluripotent cells derived from the inner cell mass (ICM) of a blastocyst stage embryo that are still able to self-renew indefinitely *in vitro* and differentiate into tissue types from all three primordial germ layers: mesoderm, ectoderm and endoderm. In contrast, cells isolated from the next stage of development, the gastrula, are multipotent and only able to produce tissues from one of each germ layer. This hierarchical organization continues, terminating with post-mitotic somatic cells that no longer have the ability to create other cells. While it was once believed that this loss of potency was due to the progressive removal of genetic material, cloning experiments proved this idea to be false, showing that the genetic information of an adult cell is equivalent to an embryonic cell. Rather than genetic changes, developmental loss of potency is due to heritable epigenetic events, which are chemical modifications of DNA regulating the pattern of genes that are expressed and silenced. This pattern of gene expression offers a unique signaling “program” to a given cell type, and explains the progressive loss of potency through development without a loss of genetic material.

The absence of these acquired epigenetic marks in ES cells, and the associated enhanced developmental potential, offers the possibility for their therapeutic use

as donor sources for tissue transplants in patients with degenerative diseases. Unfortunately, efforts to use ES cells in a therapeutic setting have many obstacles, including the ethics of using embryonic material and immune rejection of the transplanted material. As an alternative solution, methods have been developed by which epigenetic regulation in terminally differentiated somatic cells can be removed, thus “reprogramming” these cells to a more primitive state. The reprogramming of somatic cells in this way may provide a solution to the current obstacles hampering application of ES cells in a therapeutic setting, as they do not require embryo-derived material and could be made patient-specific, thus alleviating immune rejection issues. Furthermore, these reprogrammed cells provide a model for investigation of normal molecular regulation during early development, without the ethical problems of ES material, or the contamination and chromosomal abnormalities of currently available ES lines.

The work in this thesis focuses on the characterization and mechanics of somatic cell reprogramming through the forced expression of the transcription factors Oct4, Sox2, Klf4, and c-Myc. To put this work into proper context, I will discuss the biology of ES cells and what is currently known about the maintenance of pluripotency. I will also describe other methods to reprogram somatic cells, including nuclear transfer and cell fusion.

Embryonic stem cells

Embryonic stem (ES) cells are derived from the inner cell mass (ICM) of a blastocyst stage embryo and are able to self-renew indefinitely *in vitro* while maintaining pluripotency, or the ability to differentiate into all tissues of the embryo proper. ES cells were initially isolated from mouse blastocysts, then later generated from human blastocysts (Evans and Kaufman, 1981; Martin, 1981; Thomson et al., 1998). The maintenance of pluripotency is regulated by the interplay between transcription factors, signaling pathways and epigenetic regulators, some of which are described below.

Genetic control of pluripotency

Oct4

Oct4, or Pou5f1, is a member of the POU family of transcription factors, which binds to the octamer motif ATT(A/T)GCAT. Oct4 is important in early embryonic development and is expressed in blastomeres, ICM, epiblast, and germ cells, as well as in pluriptotent stem cells, including ES cells, embryonic germ (EG) cells, and embryonic carcinoma (EC) cells (Okamoto et al., 1990; Rosner et al., 1990; Scholer et al., 1990). Embryos null for Oct4 die around implantation and ES cells cannot be derived from the blastocysts (Nichols et al., 1998). Oct4 has been shown to act cooperatively with other genes in the nucleus, including Sox2 (Boiani and Scholer, 2005). Overexpression of Oct4 (by 50%) causes differentiation into primitive endoderm and mesoderm in mouse ES cells (Niwa et al., 2000). In vivo activation of Oct4 in epithelial tissues causes dysplasia and an

expansion of progenitor cell populations (Hochedlinger et al., 2005), showing that silencing of Oct4 in adult cell populations is crucial to their differentiation capacities as well as showing that certain populations may be able to respond to embryonic signaling.

Sox2

Sox2, a member of the **S**RY-related HMG-b**ox** family of transcription factors, binds to the DNA sequence (A/T)(A/T)CAA(T/A)G, thereby regulating its target genes. Sox2 is expressed in the ICM, epiblast and germ cells of the developing embryo, as well as the extraembryonic ectoderm and neural stem cells and neural progenitor cells. Embryos null for Sox2 die at implantation due to failure of the epiblast (Avilion et al., 2003) and knock-out ES cells cannot be established. There is evidence that SOX proteins cooperate with other proteins in order to regulate gene expression, and Sox2 is known to heterodimerize with Oct4 (Kamachi et al., 2000).

Nanog

Nanog is a homeodomain transcription factor that plays a crucial role in early embryonic development. Nanog is expressed in the developing morula, ICM and epiblast as well as ES cells, EC cells and germ cells (Chambers et al., 2003; Mitsui et al., 2003). Embryos null for Nanog die shortly after implantation at E5.5 due to a lack of epiblast tissue (Mitsui et al., 2003). Knockout ES cells can be established but spontaneously differentiate into primitive endoderm.

Overexpression of Nanog in ES cells induces independence from the cytokine leukemia inhibitory factor (LIF) (see below). It has also been reported that hybrid cells formed by the fusion of neural stem cells and ES cells overexpressing Nanog reprogram with a higher efficiency, providing further evidence for Nanog's role in the maintenance of pluripotency (Silva et al., 2006).

Core transcriptional circuitry in ES cells

Given the importance of Oct4, Sox2 and Nanog in the maintenance of pluripotency, numerous studies have analyzed their downstream targets to gain a more detailed understanding of their molecular mechanisms of action. Genome-wide location analysis for OCT4, NANOG, and SOX2 in both mouse and human ES cells has demonstrated that these three genes form a core transcriptional regulatory network that governs both pluripotency and the control of differentiation in ES cells (Boyer et al., 2005; Loh et al., 2006). By binding to their own promoters and the promoters of the other two factors, Oct4, Sox2 and Nanog form an interconnected, autoregulatory loop (**Figure 1**). Coordinate reinforcement of their own expression allows for maintenance of the pluripotent cell "program" and provides evidence for the mechanism of self-renewal in ES cells. Studies using an inducible Sox2-null ES line suggest that this is indeed the case (Masui et al., 2007). ES cells that have lost Sox2 expression can be rescued in two ways: 1) by the overexpression of Sox2, or 2) by the overexpression of Oct4. This suggests that the function of Sox2 is, in part, to regulate Oct4 expression.

In addition to binding to each others' promoters Oct4, Sox2 and Nanog share numerous other target genes, co-occupying both active and inactive genes, suggesting that transcriptional regulation is a concerted process requiring the interaction of multiple proteins (**Figure 1**). Consistent with this hypothesis, sequential immunoprecipitation experiments identified a protein-interaction network for Nanog that consists of known regulators of pluripotency, including Oct4 (Wang et al., 2006).

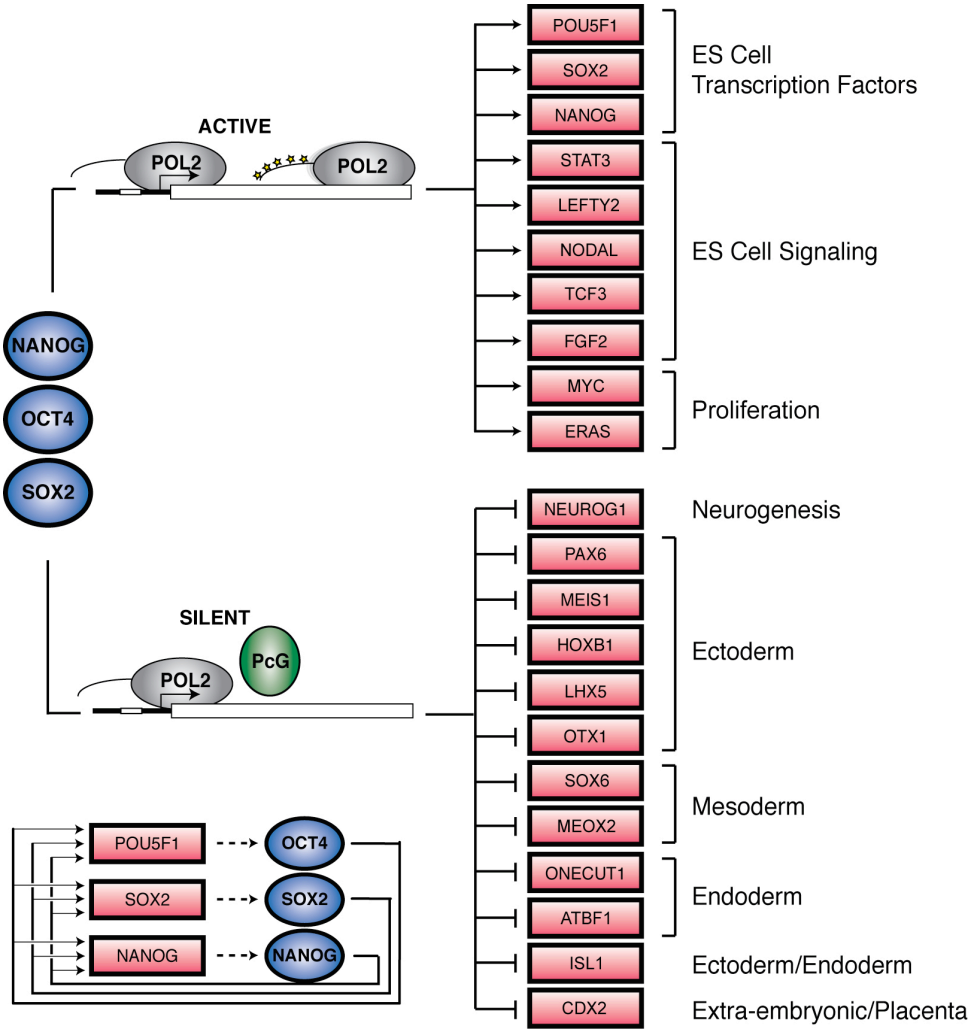


Figure 1. Core transcriptional circuitry of ES cells. The Oct4, Sox2, and Nanog transcription factors (blue circles) (i) participate in an interconnected autoregulatory loop (bottom left), (ii) generally co-occupy the same sites in the promoters of their target genes, and (iii) collectively target two sets of genes, one that is actively expressed in ES cells and another that is silent in ES cells, but remains poised for subsequent expression during differentiation (Boyer et al., 2005). Polycomb Group proteins (PcG, green circle) co-occupy a set of silent developmental regulators in ES cells that are expressed in differentiated cells in a lineage-specific pattern. These genes show evidence of transcriptional initiation by PNA Polymerase II (Pol2, gray oval), but are not productively transcribed in ES cells.

A surprising number of genes bound by Oct4, Sox2 and Nanog are transcriptionally silent in pluripotent ES cells. Many of these inactive genes are important transcription factors that are associated with lineage specification. This suggests that Oct4, Sox2, and Nanog maintain pluripotency not only by activating genes responsible for self-renewal, but also by directly repressing genes associated with differentiation.

The genome-wide location analysis performed in the studies described above utilized chromatin immunoprecipitation coupled to a microarray platform (ChIP-chip for human cells) (Boyer et al., 2005) or sequencing analysis (ChIP-PET for mouse cells) (Loh et al., 2006). While there were many overlapping gene targets between these two analyses, including the promoters of Oct4, Sox2 and Nanog, there were also significant differences between the studies. For example, some genes were identified as target genes in only human ES cells, such as *Hand1*, while others were recognized only in mouse ES cells, like *Essrrb*. It is known that human and mouse ES cells respond differently to some signaling molecules (see below). Therefore, these discrepancies may reflect those differences in the pathways responsible for pluripotency between mouse and human ES cells (Brons et al., 2007; Rossant, 2008; Tesar et al., 2007). Another possibility, however, is that the different genomic platforms used in the studies created bias in the identified target genes. Chromatin immunoprecipitation followed by direct sequencing of bound DNA (ChIP-seq) could remove the bias associated with the genomic platforms, but not address technical issues surrounding the

immunoprecipitation, such as antibody specificity. Genetic manipulation of Oct4, Sox2 and Nanog combined with expression analysis will begin to unravel the key targets responsible for self-renewal (Loh et al., 2006).

Epigenetic control of pluripotency

ES cells display epigenetic hallmarks of plasticity, namely an “open” chromatin (euchromatin) structure typified by acetylated histone modifications, while cells undergoing differentiation exhibit characteristics of “closed” chromatin (heterochromatin) such as decreased histone acetylation. In addition, specific histone modifications, including methylation and acetylation, are associated with euchromatin or heterochromatin. The dynamics of chromatin structure and the heritability of chromatin marks play key roles in development and may therefore have implications on reprogramming.

The developmental regulators bound by Oct4, Sox2 and Nanog are co-occupied by Polycomb group (PcG) proteins, chromatin modifiers that are associated with epigenetic silencing (**Figure 1**) (Bernstein et al., 2006; Boyer et al., 2006; Lee et al., 2006). PcG proteins form two repressive complexes: PRC1 and PRC2. PRC2 catalyzes H3K27me3, a repressive mark that may allow PRC1 binding. Once bound, PRC1 facilitates chromatin condensation and gene silencing. H3K27me3 was found associated with the developmental regulators, and loss of PRC2 components showed de-repression of the target genes, suggesting a functional

link between PRC2 proteins and silencing of developmental regulators in ES cells.

In addition to the H3K27me3 repressive mark, many developmental regulators in ES cells also contain the active H3K4me3 mark, giving their promoters “bivalent” domains (Bernstein et al., 2006). In addition to bivalent domains, the developmental regulators bound by PcG undergo transcriptional initiation, but not elongation (Guenther et al., 2007). This suggests that developmental regulators are uniquely poised for expression in ES cells, supporting the ability of ES to easily differentiate if given the proper cue(s). This also highlights the important balance that must be maintained in ES cell cultures: to continually self-renew, yet remain primed for differentiation.

Signaling pathways in pluripotency

The property of ES cell self-renewal is not entirely intrinsic, but also depends upon extracellular signals. MEF feeder layers, as well as conditioned media, confer the ability of ES cells to self-renew, suggesting they secrete factors involved in the maintenance of pluripotency. Indeed, the cytokine LIF is necessary for the maintenance of the undifferentiated state in mouse ES cells and functions by activating the transcription factor Stat3 (Niwa et al., 1998; Smith et al., 1988; Williams et al., 1988). Ying et al. showed that serum is dispensible in mouse ES media if it is supplemented with BMP4 protein, showing a clear role of BMP signaling in the self-renewing properties of ES cells (Ying et al., 2003).

Inhibitor of differentiation (ID) proteins downstream of BMP4 suppress lineage commitment in ES cells by blocking of FGF/ERK signaling (Silva and Smith, 2008), which is important for the initial stages of differentiation (Kunath et al., 2007). Simultaneous inhibition of ERK and glycogen synthase kinase (GSK)-3 signaling in mouse ES cells renders them independent of LIF and BMP signaling, highlighting the parallel signaling pathways involved in ES self-renewal (Ying et al., 2008).

The Wnt signaling pathway promotes self-renewal in mouse and human ES cells (Cai et al., 2007; Ogawa et al., 2006; Sato et al., 2004; Singla et al., 2006). Canonical Wnt signaling leads to GSK-3 phosphorylation and sequestration, allowing for β-catenin accumulation in the nucleus, where it associates with Tcf/Lef family transcription factors and activation and/or repression of gene expression (**Figure 2**). Tcf3 is the most abundant member of this transcription factor family in ES cells and was recently shown to bind the same genomic targets as Oct4, Sox2, and Nanog, indicating that Wnt signaling feeds into the core transcriptional circuitry of pluripotency (Cole et al., 2008).

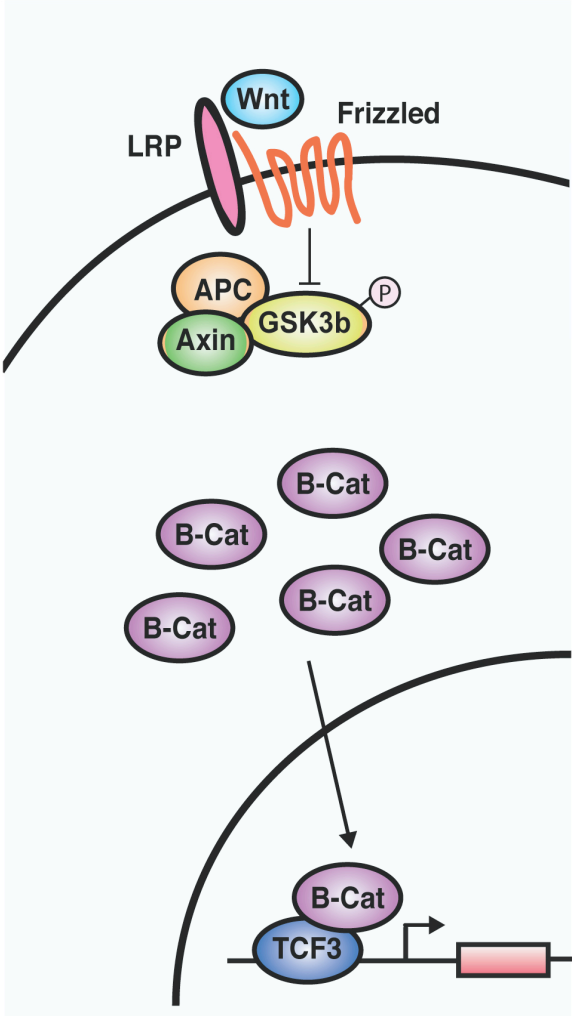


Figure 2. Canonical Wnt signaling pathway. Wnt ligand promotes the association of Frizzled and LRP, leading to phosphorylation and inactivation of GSK-3β (glycogen synthase linase 3β). β-catenin (B-cat) accumulates and migrates to the nucleus where it associates with Tcf/Lef proteins, including Tcf3, shown here, to target genes to regulate transcription.

Interestingly, human ES cells have different signaling requirements than mouse. For instance, unlike mouse ES cells, human ES cells do not require LIF/Stat3 signaling to maintain pluripotency and BMP signaling in human ES cells drives differentiation rather than self-renewal (Yu and Thomson, 2008). Furthermore, human ES cell pluripotency is maintained through FGF and TGFβ/Activin/Nodal signaling (Yu and Thomson, 2008). These differences between human and mouse ES cells could simply be due to differences in species-specific signaling or could be due to a difference in the developmental stage of ES cell isolation. It has been postulated that human ES cells more closely represent a primitive streak stage of development, rather than a blastocyst stage (Brons et al., 2007; Rossant, 2008; Tesar et al., 2007).

Laboratory use of ES cells

The isolation and use of ES cells as a laboratory tool has aided in the study of developmental genetics and has led to numerous mouse models of human disease. Animals in which a specific gene is deleted, or "knock-outs," can be engineered with the use of ES cells. A DNA construct is first generated with one or more of a gene's exons deleted, then is targeted to the corresponding genomic locus in ES cells by homologous recombination. The engineered ES cells are subsequently injected into blastocysts and because of their pluripotency, can contribute to all somatic lineages and the germline in the resulting chimeras. The chimeras are bred to wild-type mice to generate stable transgenic lines that contain the deleted gene. One example of a knock-out

mouse that had important implications for understanding human disease processes is the knock-out of the p53 tumor suppressor (Harvey et al., 1993; Jacks et al., 1994). Although adult homozygous animals were obtained, they were highly susceptible to tumor formation, providing key insight into cancer biology. Targeted ES cells can also be engineered to carry a visible marker such as GFP or lacZ downstream of a promoter to allow for lineage tracing within cell populations. Targeted ES cells can be engineered to overexpress proteins, resulting in transgenic animals with a more predictable expression pattern than traditional transgenics (Beard et al., 2006).

The *in vitro* differentiation of ES cells into mesodermal, ectodermal and endodermal cell types has been shown to recapitulate normal embryonic development, making ES cells a unique *in vitro* system for studying differentiation into specific tissue lineages. For instance, it has become possible to model the process of gastrulation with ES cultures, providing a convenient *in vitro* system to study the effects of genetic and epigenetic manipulation [reviewed in (Murry and Keller, 2008).

Clinical use of ES cells

Beyond being used as a laboratory technology, ES cells offer the possibility for therapeutic use as donor sources for tissue transplants in patients with degenerative diseases. From ES cells one could potentially derive dopaminergic neural cells to transplant into patients suffering from Parkinson's disease,

cardiomyocytes to treat heart failure patients, motor neurons or oligodendrocytes to repair injured spinal cords, or any multitude of other cell types for a variety of diseases (Murry and Keller, 2008). Despite their therapeutic potential, there are restrictions to the current clinical use of ES cells. The first is a technical limitation. We must be able to differentiate ES cells into the desired cell type at high efficiency (or be able to purify out specific populations), transplant those cells and have them incorporate physiologically into an organ. Care must be taken not to introduce undifferentiated stem cells that are known to be tumorigenic. There is continued progress regarding these issues and they may ultimately be overcome with further understanding of disease processes and *in vitro* culturing systems.

Beyond these technical issues, other critical impediments to stem cell therapy remain, including immune rejection of allogeneic transplanted cells and the ethical use of human embryos to derive ES cells. “Reprogramming” a differentiated cell to an ES cell would circumvent the latter two issues, providing a patient-specific donor source for cell transplantation. There have been different methods used to reprogram cells to a more embryonic state, including nuclear transfer, cellular fusion, and the more recently identified transduction of defined factors (**Figure 3**).

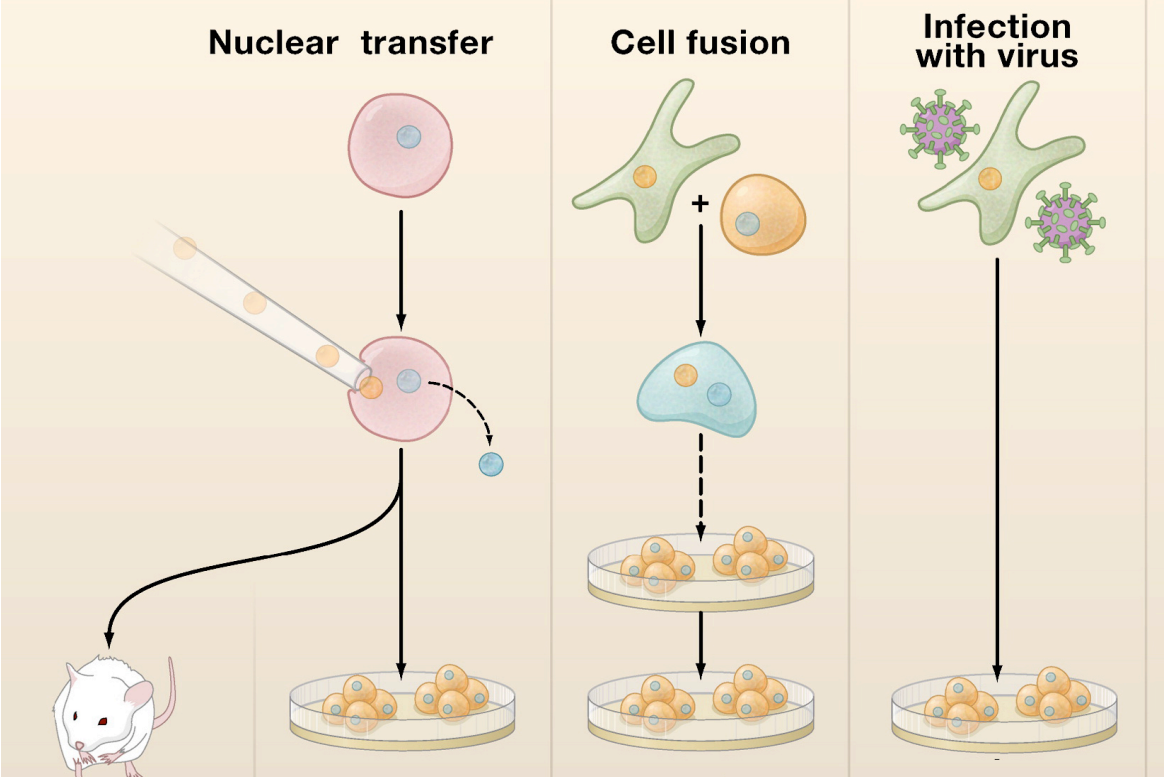


Figure 3. Strategies to induce reprogramming of somatic cells. Three methods for inducing reprogramming of somatic cells include nuclear transfer, fusion, and forced expression of defined factors. *Nuclear transfer* involves the injection of a somatic nucleus into an enucleated oocyte. If transferred to a surrogate mother, it can give rise to an animal clone, but if explanted into culture, it can give rise to genetically matched ES cells. *Cell fusion* of differentiated cells with pluripotent ES cells gives rise to hybrids that have characteristics of pluripotent cells. *Viral transduction* of four transcription factors, Oct4, Sox2, Klf4, and c-Myc can induce a pluripotent state.

Strategies for Reprogramming somatic cells

Nuclear transfer

Nuclear transfer involves the introduction of a nucleus from a donor cell into an enucleated oocyte to generate a cloned animal. Briggs and King successfully employed this method in the 1950s when they generated cloned tadpoles by transferring nuclei of blastula stage embryos into enucleated *Rana pipiens* eggs (King and Briggs, 1956). Their study further showed that while most nuclei from early stage cells could form tadpoles when transferred into oocytes, it was substantially more difficult to achieve with nuclei from later stages of development, including gastrula and neurula. Gurdon and colleagues later generated fertile, adult frogs when they transferred tadpole intestinal nuclei into enucleated *Xenopus laevis* oocytes [reviewed in (Gurdon and Byrne, 2003)]. Interestingly, when the donor nuclei were from adult somatic cells they could only generate tadpoles. Together these studies proved that cells at different developmental stages are genetically equivalent. Although the genetic material in a cell remains constant, epigenetic characteristics of DNA and chromatin influence gene expression patterns and a cell's developmental potency.

Nuclear transfer in mammals proved to be technically challenging, but was successfully accomplished when Wilmut and colleagues cloned Dolly the sheep from a differentiated mammary epithelial cell (Wilmut et al., 1997). Somatic cell nuclear transfer (SCNT) was subsequently used to clone other mammalian species, including mice (Wakayama et al., 1998), cows (Kato et al., 1998) and

pigs (Onishi et al., 2000; Polejaeva et al., 2000). Consistent with the findings in frogs, reproductive cloning in mice was easier with stem cells than with more differentiated cells, though terminally differentiated murine lymphocytes and post-mitotic olfactory neurons were successfully used as donor nuclei in nuclear transfer experiments (Eggan et al., 2004; Hochedlinger and Jaenisch, 2002). In order to obtain animals by nuclear transfer (NT) from terminally differentiated cells, however, a “two-step” cloning process was employed where ES cells were derived from NT embryos, followed by generation of mice from the NT ES cells. Disease-specific NT ES cells were also used to treat an immune-deficient mouse model through homologous recombination, providing proof-of-principle that reprogrammed stem cells could be used therapeutically (Rideout et al., 2002).

The generation of animals via reproductive cloning is highly inefficient, with most animals dying during gestation exhibiting a variety of maladies, including large offspring syndrome (Chavatte-Palmer et al., 2000; Young et al., 1998) and abnormal placentas (Hill et al., 2000; Wakayama and Yanagimachi, 1999). Faulty epigenetic reprogramming is hypothesized as a main cause of the inefficiency in nuclear transfer (Humpherys et al., 2002). Those animals that do survive suffer severe abnormalities, including premature death and obesity (Hochedlinger and Jaenisch, 2003; Yang et al., 2007). Interestingly, the derivation of ES cells from NT embryos is much more efficient and there are indications that NT ES cells are indistinguishable from embryo derived ES cells, which makes NT ES cells a good source of patient-specific cell lines (Brambrink et al., 2006; Wakayama et al.,

2006). Unfortunately, NT has not been successful to date with human cells. Even if the technical obstacles of human NT were resolved, the ethical issues of human oocyte usage and availability are further hindrances to its use as a therapeutic cell source.

Reprogramming via cell fusion

Reprogramming of somatic cells by fusion was first demonstrated with mouse thymocytes and EC cells (Miller and Ruddle, 1976) and later shown with mouse ES cells (Tada et al., 2001). Fusion of human fibroblasts or hematopoietic cells with hES cells also resulted in reprogrammed hybrids (Cowan et al., 2005; Yu et al., 2006). While hybrids gain many hallmarks of reprogramming, including reactivation of pluripotency genes in the somatic, differentiated genome, it is unclear that the hybrid's somatic genome is fully reprogrammed (Kimura et al., 2004).

The molecular mechanisms governing reprogramming by fusion are poorly understood. Both the cytoblast (Strelchenko et al., 2006) and the karyoblast (Do and Scholer, 2004) of the ES cell have been implicated as the necessary cellular compartment that contains the reprogramming "factors." Of greater concern, however, is the use of tetraploid cells for human cell-based therapy. The ES cell genome will have to be removed to generate diploid cells before they can be used therapeutically. While selective elimination of some ES cell chromosomes was achieved using a chromosomal deletion cassette, the use of this technique

to eliminate an entire genome poses significant risks of generating chromosomal instability (Matsumura et al., 2007).

Transcription factor mediated reprogramming

A breakthrough in the field of reprogramming was achieved when Takahashi and Yamanaka reported the generation of induced pluripotent stem (iPS) cells from mouse embryonic and adult fibroblasts (Takahashi and Yamanaka, 2006) (**Figure 4**). Based on the hypothesis that ES cells contain reprogramming factors, 24 candidate factors were screened in combination for the ability to re-activate Fbx15, a downstream target of Oct4 and Sox2 (Tokuzawa et al., 2003). Retroviral-mediated transduction of four transcription factors (Oct4, Sox2, c-myc, and Klf4) in fibroblasts, followed by selection for Fbx15 expression, generated iPS cells that formed teratomas and contributed to chimeric embryos when injected into blastocysts (Takahashi and Yamanaka, 2006). Fbx15-selected iPS cells did not appear to be fully reprogrammed, however, as their gene expression and DNA methylation patterns differed significantly from ES cells, showing little endogenous expression of Oct4 and Nanog, among other genes. Furthermore, Fbx15-selected iPS could not contribute to adult chimeras.

This pivotal study raised numerous questions, including the role of selection in iPS cell generation, the role of viruses and the continued expression of the four transcription factors, the epigenetic state of the iPS cells, the role for transcription factor mediated reprogramming in human cells, and the mechanism of action of

iPS formation. Subsequent studies, including those in this thesis, have begun to address these questions. Although Fbx15 is a downstream target of Oct4 and Sox2, it is not an essential gene for development or pluripotency (Tokuzawa et al., 2003). When Oct4 or Nanog re-activation was used as a more stringent selection criterion to generate iPS cells through retroviral-mediated four-factor transduction, these iPS more closely resembled ES cells by numerous criteria, including: (i) global gene expression and chromatin configuration, (ii) contribution to adult, germline-competent chimeras as well as live, late-term embryos generated by tetraploid complementation, (iii) re-expression of endogenous *Oct4* and *Nanog*, and (iv) silencing of the Moloney-based viral transduced factors (Chapter 2) (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007).

Some key questions regarding the use of selection markers in iPS cell generation remain. Are the partially reprogrammed Fbx15-iPS cells intermediates because of the timing of selection or because Fbx15 does not select for pluripotent cells? It was subsequently determined that iPS cells can be isolated from culture without selection by morphological criteria alone (Blelloch et al., 2007; Meissner et al., 2007) suggesting that the Fbx15-iPS intermediates arose in the first reprogramming report due to early selection and could have become fully reprogrammed given enough time. However, the possibility does remain that Fbx15 selection does maintain a different cell state in culture. Additional studies are needed to sufficiently address this question.

Several groups have more recently achieved reprogramming of other murine cell types, including hepatocytes (Aoi et al., 2008) and progenitor B cells (Hanna et al., 2008). Interestingly, reprogramming of mature B cells to iPS cells could only be achieved with the addition of C-EBPα, a transcription factor known to facilitate reprogramming of B cells into macrophages (Xie et al., 2004). This suggests that, as in nuclear transfer, the efficiency of reprogramming of terminally differentiated cells is lower than somatic stem cells, although further analysis needs to be done to solidify this assertion.

Therapeutic use of iPS cells

iPS cells could provide patient-specific donor sources for therapeutic cell transplantation, by substituting for pluripotent ES cells. iPS cells were used to cure a mouse model of sickle cell anemia (Hanna et al., 2007) and iPS-derived neurons transplanted into mouse model of Parkinson's disease were able to alleviate some of the symptoms associated with the disease (Wernig et al., 2008b). These studies show great promise for iPS cells in patient-specific transplantation therapy. The initial use of drug-dependent selection in the generation of iPS cells hindered its applicability to human cells, as genetically modified donor cells would be undesirable. The isolation of iPS cells based solely on morphologic criteria proved these initial fears unfounded (Blelloch et al., 2007; Meissner et al., 2007). Indeed, reprogramming of human fibroblasts and keratinocytes has been reported (Park et al., 2008; Takahashi et al., 2007; Yu et

al., 2007), both with the original four-factor combination (Oct4, Sox2, Klf4, and c-Myc) and with Oct4, Sox2, Nanog, and Lin28 (**Figure 4**).

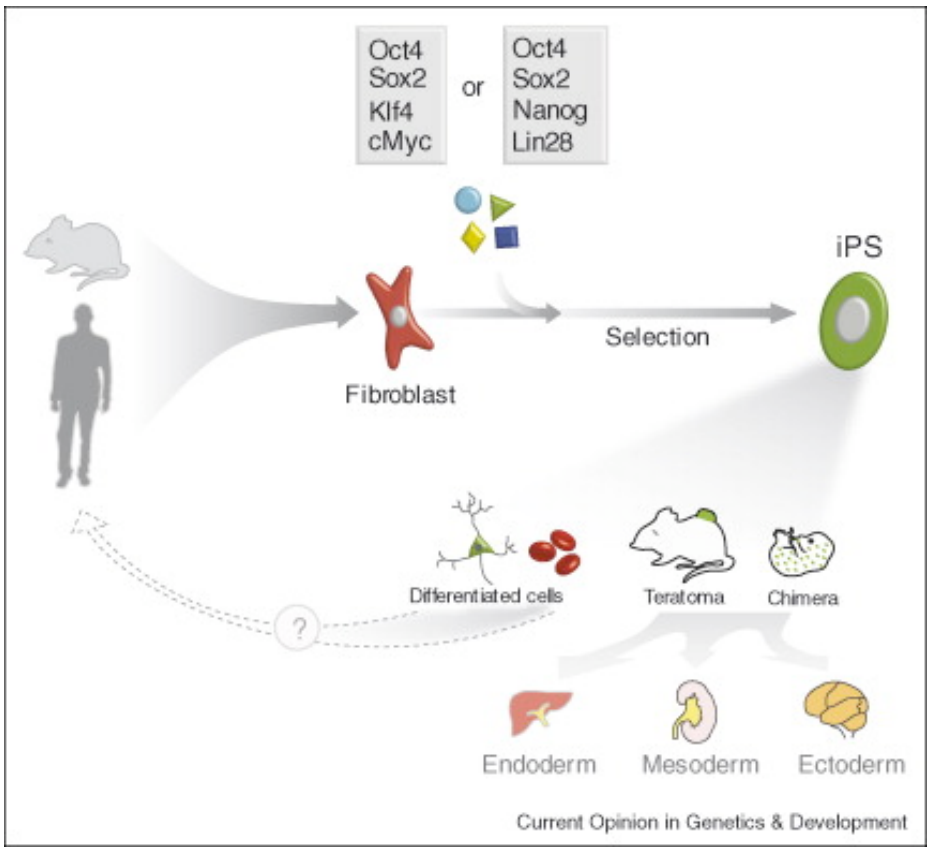


Figure 4. Reprogramming using defined factors. Fibroblasts derived from either murine or human sources are transduced with Oct4, Sox2, Klf4, and c-Myc or Oct4, Sox2, Nanog, and Lin28. induced pluripotent stem (iPS) cells are clonally derived following selection. These cells can give rise to derivatives of all three germ layers as determined by teratoma formation and chimera contribution. iPS cells have the potential to be differentiated into cell types that can be used as potential therapies.

Despite these breakthroughs in the field of reprogramming, there are key concerns with using iPS cells therapeutically. One is the efficiency of reprogramming, which is extremely low with reports ranging from roughly 0.05% to 0.5% in mouse cells (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007) and 0.02% to 0.1% in human cells (Park et al., 2008; Takahashi et al., 2007). The addition of known epigenetic modulators, including the DNA methylation inhibitor 5-aza-cytidine and the HDAC inhibitor valproic acid (VPA), were shown to increase the efficiency of reprogramming (Huangfu et al., 2008a; Mikkelsen et al., 2008). Interestingly, VPA can substitute for viral transduction of Klf4 and c-Myc in human cells, highlighting the importance of Oct4 and Sox2 in the reprogramming process, as well as the possible differences between reprogramming mouse and human cells (Huangfu et al., 2008b).

The main problem in using iPS cells therapeutically is the retroviral-mediated transduction of transcription factors to achieve reprogramming, leading to permanent transgene insertions in the iPS cell genome. It has been reported that reactivation of the viral c-Myc transgene causes tumorigenesis in chimeras generated from iPS cells (Okita et al., 2007). iPS cell formation was subsequently achieved without c-Myc (Nakagawa et al., 2008; Wernig et al., 2008a), but the efficiency of reprogramming was severely reduced. Furthermore, it remains unclear whether the other transgenes, especially Klf4 (Foster et al., 2005), and Oct4 (Hochedlinger et al., 2005), could cause tumors if reactivated.

Recent studies have explored strategies to derive iPS cells with minimal or no genetic manipulations. One study used a single virus carrying all four reprogramming factors to infect and subsequently reprogram fibroblasts (Carey et al., 2009). Another study was able to reprogram somatic cells by transiently transfecting a single vector containing all four reprogramming factors (Okita et al., 2008). A third group reported the use of non-integrating adenovirus to overexpress the reprogramming factors (Stadtfeld et al., 2008b). While these reports are a small step toward achieving reprogramming without genetic manipulations, the efficiencies of reprogramming in these studies are a severe hindrance to their continued use.

Small molecules and other transient cues may provide an alternative to viruses in the future. Indeed, several studies, including two included in this thesis (Chapters 4 and 5), have found a role for small molecules in reprogramming. As previously mentioned, the use of valproic acid and 5-aza-cytidine, has been shown to increase the efficiency of reprogramming (Huangfu et al., 2008a; Mikkelsen et al., 2008). BIX, a G9a histone methyltransferase inhibitor, was identified in a chemical screen to increase the reprogramming efficiency of Oct4 and Klf4 infected neural precursor cells (Shi et al., 2008b). Oct4 expression is rapidly downregulated upon differentiation and the *Oct4* locus gains the heterochromatin-associated marks H3K9me2 and H3K9me3, which is dependent on the histone methyltransferase G9a (Feldman et al., 2006). Inhibition of G9a could therefore promote *Oct4* re-expression. BIX was subsequently used to

replace Oct4 in Sox2, Klf4, and c-Myc infected neural precursor cells (Shi et al., 2008b) and later used in combination with another chemical, BayK, a L-type calcium channel agonist, to reprogram Oct4 and Klf4 infected fibroblasts (Shi et al., 2008a).

Many of these initial chemical compounds were identified through a candidate approach. The mechanism(s) by which these compounds act has been assumed and not explored in detail. Knowledge of the mechanisms governing reprogramming can inform the discovery of small molecules and vice versa. High-throughput chemical screens offer the potential to isolate unique chemical modulators that can provide key mechanistic insight into the process of reprogramming, as well as functionally replace virally transduced transcription factors.

Mechanisms of reprogramming to iPS cells

Initial studies exploring the mechanisms of reprogramming using defined factors show that it is a gradual, sequential process, where the somatic cellular program is converted to a pluripotent ES-like one (Chapter 3) (Brambrink et al., 2008; Stadtfeld et al., 2008a). Our current knowledge of the biology of ES cells and pluripotency may offer some comprehension of the mechanisms governing reprogramming. Reprogramming is likely achieved by the reactivation of the pluripotent autoregulatory loop, in part induced by Oct4 and Sox2 overexpression. Exogenous (viral) Oct4 and Sox2 can bind to the endogenous

Oct4, Sox2 and *Nanog* loci, and once the feed-forward loop is active, the need for the virally encoded transcription factors ceases (**Figure 5**). Oct4 and Sox2 were found to bind genes that encode known epigenetic modulators, such as Smarcad1, Myst3, Jmjd1a and Jmjd2c (Loh et al., 2006; Loh et al., 2007), suggesting a role for Oct4 and Sox2 in the epigenetic changes that occur during the reprogramming process.

It was initially postulated that Klf4 and c-Myc might function together to transform somatic cells and induce proliferation (Yamanaka, 2007). However, this may not be the case. Recently, Klf4 was identified to bind to many of the same target genes as Oct4, Sox2 and Nanog (Chen et al., 2008; Jiang et al., 2008; Kim et al., 2008), suggesting that its role in reprogramming could be to cooperatively bind with Sox2 and Oct4 to turn on the autoregulatory circuitry of pluripotency. The difference in overlap of c-Myc target genes with those of Oct4, Sox2, and Klf4 have led to the hypothesis that the function of c-Myc in reprogramming is separate from the other three factors (Chen et al., 2008; Kim et al., 2008). A recent study supports this claim, providing evidence for c-Myc's role in the initial phases of reprogramming, including the downregulation of fibroblast-specific genes (Sridharan et al., 2009). In addition, c-Myc might function to modify the chromatin state of fibroblasts (Yamanaka, 2007) or promote DNA replication (Dominguez-Sola et al., 2007), allowing for chromatin reorganization. Despite some indication as to the mechanisms of reprogramming, these studies have only begun to address how reprogramming of somatic cells is achieved.

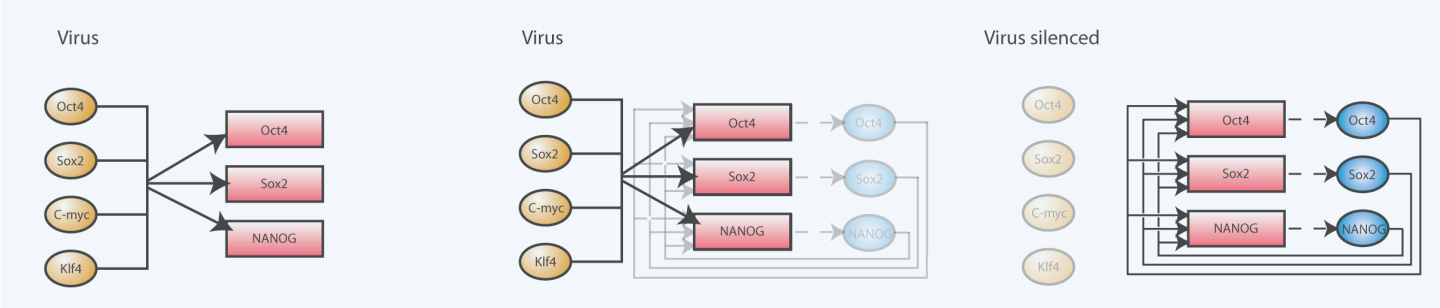


Figure 5. Model of the molecular circuitry during reprogramming. The transduced factors likely bind to the endogenous pluripotency genes, including Oct4, Sox2, and Nanog, gradually activating the autoregulatory loop of pluripotency. Once the endogenous circuitry is active the virally transduced factors are silenced by the de novo methyltransferases Dnmt3a and 3b.

The isolation of iPS cells holds great promise for the field of regenerative biology, but the current inability to achieve efficient, viral-free reprogramming presents a significant hurdle to its clinical use. We also have a superficial understanding of the mechanisms by which somatic cells are reprogrammed to iPS cells. A more detailed knowledge of the mechanisms underlying the process of reprogramming may allow the therapeutic use of iPS cells to become a reality.

The work in this thesis begins to address these shortcomings in our knowledge of reprogramming. Chapter 2 builds on Takahashi and Yamanaka's breakthrough study, answering some of the questions that arose from this pivotal work. Chapter 3 aims to provide some insight into the mechanism of reprogramming, including defining the minimal expression requirements and temporal kinetics of reprogramming. Chapter 4 analyzes the role of Wnt signaling in the reprogramming of somatic cells. Chapter 5 details the replacement of viral Klf4 with a small molecule identified from a high-throughput chemical screen.

References

Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T., and Yamanaka, S. (2008). Generation of pluripotent stem cells from adult mouse liver and stomach cells. Science *321*, 699-702.

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev *17*, 126-140.

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

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Blelloch, R., Venere, M., Yen, J., and Ramalho-Santos, M. (2007). Generation of induced pluripotent stem cells in the absence of drug selection. Cell Stem Cell *1*, 245-247.

Boiani, M., and Scholer, H. R. (2005). Regulatory networks in embryo-derived pluripotent stem cells. Nat Rev Mol Cell Biol *6*, 872-884.

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G.*, et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956.

Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Brambrink, T., Foreman, R., Welstead, G. G., Lengner, C. J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell *2*, 151-159.

Brambrink, T., Hochedlinger, K., Bell, G., and Jaenisch, R. (2006). ES cells derived from cloned and fertilized blastocysts are transcriptionally and functionally indistinguishable. Proc Natl Acad Sci U S A *103*, 933-938.

Brons, I. G., Smithers, L. E., Trotter, M. W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S. M., Howlett, S. K., Clarkson, A., Ahrlund-Richter, L., Pedersen,

R. A., and Vallier, L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature *448*, 191-195.

Cai, L., Ye, Z., Zhou, B. Y., Mali, P., Zhou, C., and Cheng, L. (2007). Promoting human embryonic stem cell renewal or differentiation by modulating Wnt signal and culture conditions. Cell Res *17*, 62-72.

Carey, B. W., Markoulaki, S., Hanna, J., Saha, K., Gao, Q., Mitalipova, M., and Jaenisch, R. (2009). Reprogramming of murine and human somatic cells using a single polycistronic vector. Proc Natl Acad Sci U S A *106*, 157-162.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643-655.

Chavatte-Palmer, P., Heyman, Y., and Renard, J. P. (2000). [Cloning and associated physiopathology of gestation]. Gynecol Obstet Fertil *28*, 633-642.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V. B., Wong, E., Orlov, Y. L., Zhang, W., Jiang, J.*, et al.* (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106-1117.

Cole, M. F., Johnstone, S. E., Newman, J. J., Kagey, M. H., and Young, R. A. (2008). Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev *22*, 746-755.

Cowan, C. A., Atienza, J., Melton, D. A., and Eggan, K. (2005). Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. Science *309*, 1369-1373.

Do, J. T., and Scholer, H. R. (2004). Nuclei of embryonic stem cells reprogram somatic cells. Stem Cells *22*, 941-949.

Dominguez-Sola, D., Ying, C. Y., Grandori, C., Ruggiero, L., Chen, B., Li, M., Galloway, D. A., Gu, W., Gautier, J., and Dalla-Favera, R. (2007). Non-transcriptional control of DNA replication by c-Myc. Nature *448*, 445-451.

Eggan, K., Baldwin, K., Tackett, M., Osborne, J., Gogos, J., Chess, A., Axel, R., and Jaenisch, R. (2004). Mice cloned from olfactory sensory neurons. Nature *428*, 44-49.

Evans, M. J., and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.

Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat Cell Biol *8*, 188-194.

Foster, K. W., Liu, Z., Nail, C. D., Li, X., Fitzgerald, T. J., Bailey, S. K., Frost, A. R., Louro, I. D., Townes, T. M., Paterson, A. J.*, et al.* (2005). Induction of KLF4 in basal keratinocytes blocks the proliferation-differentiation switch and initiates squamous epithelial dysplasia. Oncogene *24*, 1491-1500.

Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R., and Young, R. A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77-88.

Gurdon, J. B., and Byrne, J. A. (2003). The first half-century of nuclear transplantation. Proc Natl Acad Sci U S A *100*, 8048-8052.

Hanna, J., Markoulaki, S., Schorderet, P., Carey, B. W., Beard, C., Wernig, M., Creyghton, M. P., Steine, E. J., Cassady, J. P., Foreman, R.*, et al.* (2008). Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. Cell *133*, 250-264.

Hanna, J., Wernig, M., Markoulaki, S., Sun, C. W., Meissner, A., Cassady, J. P., Beard, C., Brambrink, T., Wu, L. C., Townes, T. M., and Jaenisch, R. (2007). Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. Science *318*, 1920-1923.

Harvey, M., McArthur, M. J., Montgomery, C. A., Jr., Butel, J. S., Bradley, A., and Donehower, L. A. (1993). Spontaneous and carcinogen-induced tumorigenesis in p53-deficient mice. Nat Genet *5*, 225-229.

Hill, J. R., Burghardt, R. C., Jones, K., Long, C. R., Looney, C. R., Shin, T., Spencer, T. E., Thompson, J. A., Winger, Q. A., and Westhusin, M. E. (2000). Evidence for placental abnormality as the major cause of mortality in first-trimester somatic cell cloned bovine fetuses. Biol Reprod *63*, 1787-1794.

Hochedlinger, K., and Jaenisch, R. (2002). Monoclonal mice generated by nuclear transfer from mature B and T donor cells. Nature *415*, 1035-1038.

Hochedlinger, K., and Jaenisch, R. (2003). Nuclear transplantation, embryonic stem cells, and the potential for cell therapy. N Engl J Med *349*, 275-286.

Hochedlinger, K., Yamada, Y., Beard, C., and Jaenisch, R. (2005). Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. Cell *121*, 465-477.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A. E., and Melton, D. A. (2008a). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol *26*, 795-797.

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D. A. (2008b). Induction of pluripotent stem cells

from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol *26*, 1269-1275.

Humpherys, D., Eggan, K., Akutsu, H., Friedman, A., Hochedlinger, K., Yanagimachi, R., Lander, E. S., Golub, T. R., and Jaenisch, R. (2002). Abnormal gene expression in cloned mice derived from embryonic stem cell and cumulus cell nuclei. Proc Natl Acad Sci U S A *99*, 12889-12894.

Jacks, T., Remington, L., Williams, B. O., Schmitt, E. M., Halachmi, S., Bronson, R. T., and Weinberg, R. A. (1994). Tumor spectrum analysis in p53-mutant mice. Curr Biol *4*, 1-7.

Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell *132*, 567-582.

Jiang, J., Chan, Y. S., Loh, Y. H., Cai, J., Tong, G. Q., Lim, C. A., Robson, P., Zhong, S., and Ng, H. H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat Cell Biol *10*, 353-360.

Kamachi, Y., Uchikawa, M., and Kondoh, H. (2000). Pairing SOX off: with partners in the regulation of embryonic development. Trends Genet *16*, 182-187.

Kato, Y., Tani, T., Sotomaru, Y., Kurokawa, K., Kato, J., Doguchi, H., Yasue, H., and Tsunoda, Y. (1998). Eight calves cloned from somatic cells of a single adult. Science *282*, 2095-2098.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S. H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049-1061.

Kimura, H., Tada, M., Nakatsuji, N., and Tada, T. (2004). Histone code modifications on pluripotential nuclei of reprogrammed somatic cells. Mol Cell Biol *24*, 5710-5720.

King, T. J., and Briggs, R. (1956). Serial transplantation of embryonic nuclei. Cold Spring Harb Symp Quant Biol *21*, 271-290.

Kunath, T., Saba-El-Leil, M. K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development *134*, 2895-2902.

Lee, T. I., Jenner, R. G., Boyer, L. A., Guenther, M. G., Levine, S. S., Kumar, R. M., Chevalier, B., Johnstone, S. E., Cole, M. F., Isono, K.*, et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Loh, Y. H., Wu, Q., Chew, J. L., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J.*, et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet *38*, 431-440.

Loh, Y. H., Zhang, W., Chen, X., George, J., and Ng, H. H. (2007). Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. Genes Dev *21*, 2545-2557.

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

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A *78*, 7634-7638.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A. A.*, et al.* (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol *9*, 625-635.

Matsumura, H., Tada, M., Otsuji, T., Yasuchika, K., Nakatsuji, N., Surani, A., and Tada, T. (2007). Targeted chromosome elimination from ES-somatic hybrid cells. Nat Methods *4*, 23-25.

Meissner, A., Wernig, M., and Jaenisch, R. (2007). Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. Nat Biotechnol *25*, 1177-1181.

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

Miller, R. A., and Ruddle, F. H. (1976). Pluripotent teratocarcinoma-thymus somatic cell hybrids. Cell *9*, 45-55.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.

Murry, C. E., and Keller, G. (2008). Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. Cell *132*, 661-680.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol *26*, 101-106.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev *12*, 2048-2060.

Niwa, H., Miyazaki, J., and Smith, A. G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet *24*, 372-376.

Ogawa, K., Nishinakamura, R., Iwamatsu, Y., Shimosato, D., and Niwa, H. (2006). Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. Biochem Biophys Res Commun *343*, 159-166.

Okamoto, K., Okazawa, H., Okuda, A., Sakai, M., Muramatsu, M., and Hamada, H. (1990). A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. Cell *60*, 461-472.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature *448*, 313-317.

Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., and Yamanaka, S. (2008). Generation of mouse induced pluripotent stem cells without viral vectors. Science *322*, 949-953.

Onishi, A., Iwamoto, M., Akita, T., Mikawa, S., Takeda, K., Awata, T., Hanada, H., and Perry, A. C. (2000). Pig cloning by microinjection of fetal fibroblast nuclei. Science *289*, 1188-1190.

Park, I. H., Zhao, R., West, J. A., Yabuuchi, A., Huo, H., Ince, T. A., Lerou, P. H., Lensch, M. W., and Daley, G. Q. (2008). Reprogramming of human somatic cells to pluripotency with defined factors. Nature *451*, 141-146.

Polejaeva, I. A., Chen, S. H., Vaught, T. D., Page, R. L., Mullins, J., Ball, S., Dai, Y., Boone, J., Walker, S., Ayares, D. L.*, et al.* (2000). Cloned pigs produced by nuclear transfer from adult somatic cells. Nature *407*, 86-90.

Rideout, W. M., 3rd, Hochedlinger, K., Kyba, M., Daley, G. Q., and Jaenisch, R. (2002). Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy. Cell *109*, 17-27.

Rosner, M. H., Vigano, M. A., Ozato, K., Timmons, P. M., Poirier, F., Rigby, P. W., and Staudt, L. M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature *345*, 686-692.

Rossant, J. (2008). Stem cells and early lineage development. Cell *132*, 527-531.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A. H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat Med *10*, 55-63.

Scholer, H. R., Ruppert, S., Suzuki, N., Chowdhury, K., and Gruss, P. (1990). New type of POU domain in germ line-specific protein Oct-4. Nature *344*, 435-439.

Shi, Y., Desponts, C., Do, J. T., Hahm, H. S., Scholer, H. R., and Ding, S. (2008a). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell Stem Cell *3*, 568-574.

Shi, Y., Do, J. T., Desponts, C., Hahm, H. S., Scholer, H. R., and Ding, S. (2008b). A combined chemical and genetic approach for the generation of induced pluripotent stem cells. Cell Stem Cell *2*, 525-528.

Silva, J., Chambers, I., Pollard, S., and Smith, A. (2006). Nanog promotes transfer of pluripotency after cell fusion. Nature *441*, 997-1001.

Silva, J., and Smith, A. (2008). Capturing pluripotency. Cell *132*, 532-536.

Singla, D. K., Schneider, D. J., LeWinter, M. M., and Sobel, B. E. (2006). wnt3a but not wnt11 supports self-renewal of embryonic stem cells. Biochem Biophys Res Commun *345*, 789-795.

Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688-690.

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

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

Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G., and Hochedlinger, K. (2008b). Induced pluripotent stem cells generated without viral integration. Science *322*, 945-949.

Strelchenko, N., Kukharenko, V., Shkumatov, A., Verlinsky, O., Kuliev, A., and Verlinsky, Y. (2006). Reprogramming of human somatic cells by embryonic stem cell cytoplast. Reprod Biomed Online *12*, 107-111.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr Biol *11*, 1553-1558.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tesar, P. J., Chenoweth, J. G., Brook, F. A., Davies, T. J., Evans, E. P., Mack, D. L., Gardner, R. L., and McKay, R. D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196-199.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., and Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Tokuzawa, Y., Kaiho, E., Maruyama, M., Takahashi, K., Mitsui, K., Maeda, M., Niwa, H., and Yamanaka, S. (2003). Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development. Mol Cell Biol *23*, 2699-2708.

Wakayama, S., Jakt, M. L., Suzuki, M., Araki, R., Hikichi, T., Kishigami, S., Ohta, H., Van Thuan, N., Mizutani, E., Sakaide, Y.*, et al.* (2006). Equivalency of nuclear transfer-derived embryonic stem cells to those derived from fertilized mouse blastocysts. Stem Cells *24*, 2023-2033.

Wakayama, T., Perry, A. C., Zuccotti, M., Johnson, K. R., and Yanagimachi, R. (1998). Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. Nature *394*, 369-374.

Wakayama, T., and Yanagimachi, R. (1999). Cloning of male mice from adult tail-tip cells. Nat Genet *22*, 127-128.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D. N., Theunissen, T. W., and Orkin, S. H. (2006). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364-368.

Welstead, G. G., Schorderet, P., and Boyer, L. A. (2008). The reprogramming language of pluripotency. Curr Opin Genet Dev *18*, 123-129.

Wernig, M., Meissner, A., Cassady, J. P., and Jaenisch, R. (2008a). c-Myc is dispensable for direct reprogramming of mouse fibroblasts. Cell Stem Cell *2*, 10-12.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B. E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.

Wernig, M., Zhao, J. P., Pruszak, J., Hedlund, E., Fu, D., Soldner, F., Broccoli, V., Constantine-Paton, M., Isacson, O., and Jaenisch, R. (2008b). Neurons derived from reprogrammed fibroblasts functionally integrate into the fetal brain and improve symptoms of rats with Parkinson's disease. Proc Natl Acad Sci U S A *105*, 5856-5861.

Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A., and Gough, N. M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature *336*, 684-687.

Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., and Campbell, K. H. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810-813.

Xie, H., Ye, M., Feng, R., and Graf, T. (2004). Stepwise reprogramming of B cells into macrophages. Cell *117*, 663-676.

Yamanaka, S. (2007). Strategies and new developments in the generation of patient-specific pluripotent stem cells. Cell Stem Cell *1*, 39-49.

Yang, X., Smith, S. L., Tian, X. C., Lewin, H. A., Renard, J. P., and Wakayama, T. (2007). Nuclear reprogramming of cloned embryos and its implications for therapeutic cloning. Nat Genet *39*, 295-302.

Ying, Q. L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281-292.

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

Young, L. E., Sinclair, K. D., and Wilmut, I. (1998). Large offspring syndrome in cattle and sheep. Rev Reprod *3*, 155-163.

Yu, J., and Thomson, J. A. (2008). Pluripotent stem cell lines. Genes Dev *22*, 1987-1997.

Yu, J., Vodyanik, M. A., He, P., Slukvin, II, and Thomson, J. A. (2006). Human embryonic stem cells reprogram myeloid precursors following cell-cell fusion. Stem Cells *24*, 168-176.

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R.*, et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-1920.

Acknowledgments

Todd Meyerrose, Caroline Beard and Rudolf Jaenisch provided helpful comments during the preparation of this introduction. I thank Megan Cole for providing Figure 3 to accompany the text. Figures 1, 2 and 5 are adapted from Jaenisch and Young, 2008 (Jaenisch and Young, 2008). Figure 4 is adapted from Welstead, et al., 2008 (Welstead et al., 2008).

Chapter 2

In vitro reprogramming of fibroblasts into a pluripotent ES cell-like state

Marius Wernig*, Alexander Meissner*, Ruth Foreman*, Tobias Brambrink*, Manching Ku*, Konrad Hochedlinger, Bradley E. Bernstein and Rudolf Jaenisch

* These authors contributed equally to this work

This chapter was initially published in Nature *448*, 318-324. Permission to publish it in this thesis has been granted by Nature Publishing Group.

Author contributions:

MW, AM and RJ conceived and designed the experiments and wrote the manuscript, MW derived all iPS lines, MW and AM performed the in vitro and in vivo characterization of the iPS lines (teratoma-, 2N-, 4N-injections and IHC) and the conditional Dnmt1 experiment, AM investigated the promoter and imprinting methylation, MK and BB performed and analyzed the real-time PCRs and ChIP experiments, RF and KH generated the selectable MEFs and TTFs, RF performed Western blot and PCR analysis, and TB the microarry analysis and the proviral integration Southern blots.

Summary:

Nuclear transplantation can reprogram a somatic genome back into an embryonic epigenetic state and the reprogrammed nucleus can create a cloned animal or produce pluripotent embryonic stem cells. One potential use of the nuclear cloning approach is the derivation of “customized” ES cells for patient specific cell treatment but technical and ethical considerations impede the therapeutic application of this technology. Reprogramming of fibroblasts to a pluripotent state can be induced *in vitro* through ectopic expression of the four transcription factors Oct4, Sox2, c-Myc and Klf4. Here we show that DNA methylation, gene expression and chromatin state of such induced reprogrammed cells are similar to those of ES cells. Importantly, the cells can form viable chimeras, can contribute to the germ line and can generate live late-term embryos when injected into tetraploid blastocysts. Our results show that the biological potency and epigenetic state of *in vitro* reprogrammed iPS cells are indistinguishable from that of ES cells.

Epigenetic reprogramming of somatic cells into embryonic stem (ES) cells has attracted much attention because of the potential for “customized” transplantation therapy as cellular derivatives of reprogrammed cells will not be rejected by the donor (Hochedlinger and Jaenisch, 2003; Yang et al., 2007). Thus far somatic cell nuclear transfer and fusion of fibroblasts with ES cells have been shown to promote the epigenetic reprogramming of the donor genome to an embryonic state(Cowan et al., 2005; Hochedlinger and Jaenisch, 2006; Tada et al., 2001). However, the therapeutic application of either approach has been hindered by technical complications as well as ethical objections (Jaenisch, 2004). Recently, a major breakthrough was reported by Takahashi and Yamanaka who demonstrated that expression of the transcription factors Oct4, Sox2, c-Myc and Klf4 induced fibroblasts to become pluripotent stem cells, designated as “induced pluripotent stem (“iPS”) cells, though with a low efficiency(Takahashi and Yamanaka, 2006). The iPS cells were isolated by selection for activation of *Fbx15*, which is a downstream gene of *Oct4*. This landmark study left a number of questions unresolved: (i) Although iPS cells were pluripotent they were not identical to ES cells; for example, iPS cells injected into blastocysts generated abnormal chimeric embryos that did not survive to term; (ii) gene expression profiling revealed major differences between iPS cells and ES cells; (iii) because the four transcription factors were transduced by constitutively expressed retroviral vectors it was unclear why the cells could be induced to differentiate and whether continuous vector expression was required for the maintenance of the pluripotent state; (iv) the epigenetic state of the endogenous pluripotency

genes *Oct4* and *Nanog* was incompletely reprogrammed raising questions about the stability of the pluripotent state.

Here we used the activation of the endogenous *Oct4* or *Nanog* gene as a more stringent selection strategy for the isolation for reprogrammed cells. We infected fibroblasts with retroviral vectors transducing the four factors and selected for the activation of the endogenous *Oct4* or *Nanog* gene. Positive colonies resembled ES cells and assumed an epigenetic state characteristic of ES cells. When injected into blastocysts the reprogrammed cells generated viable chimeras and contributed to the germ line. Our results establish that somatic cells can be reprogrammed to a pluripotent state that is similar if not identical to that of normal ES cells.

Selection of fibroblasts for *Oct4* or *Nanog* activation

Using homologous recombination in ES cells we generated mouse embryonic fibroblasts (MEFs) and tailtip fibroblasts (TTFs) that carried a neomycin resistance marker inserted into either the endogenous *Oct4* (Oct4-neo) or *Nanog* locus (Nanog-neo) (Fig. 1a). These cultures were sensitive to G418, indicating that the *Oct4* and *Nanog* loci were as expected silenced in somatic cells. These MEFs or TTFs were infected with Oct4, Sox2, c-Myc and Klf4 expressing retroviral vectors and G418 was added to the cultures 3, 6 or 9 days later. The number of drug resistant colonies increased substantially when analyzed at day 20 (Fig. 1i). Most colonies had a flat morphology (Fig.1 h right) and between 11 and 25% of the colonies were “ES-like” (Fig. 1h left) when selection was applied early (Fig.1k), a number which increased at later time points. At day 20, ES-like colonies were picked, dissociated and propagated in G418-containing media. They gave rise to ES-like cell lines (designated as Oct4-iPS or Nanog-iPS cells, respectively) which could be propagated without drug selection, displayed homogenous Nanog, SSEA1 and AP expression (Fig.1b-g and Fig. S1 and S5) and formed undifferentiated colonies when seeded at clonal density on gelatin-coated dishes (see inset in Fig.1b). Four of five analyzed lines had a normal karyotype (Table S1).

Although the timing and appearance of colonies was similar between the Oct4- and Nanog-selection we noticed pronounced quantitative differences between the two selection strategies: whereas Oct4-selected MEF cultures had 3- to 10-

fold fewer colonies, the fraction of “ES-like colonies” was 2 to 3 fold higher than in Nanog-selected cultures. Accordingly, ~4 times more Oct4-selected ES-like colonies gave rise to stable and homogenous iPS cell lines (Fig.1k). This suggests that although the *Nanog* locus was easier to activate a higher fraction of the drug resistant colonies in Oct4-neo cultures were reprogrammed to a pluripotent state. Therefore, the overall estimated efficiency of 0.05-.1% to establish iPS cell lines from MEFs was similar between Oct4- and Nanog-selection despite the larger number of total Nanog-neo resistant colonies (Fig.1k). Next we investigated the time course of reprogramming by studying the fraction of AP, SSEA1 and Nanog-positive cells in Oct4-selected MEF cultures. Fourteen days after infection some cells had already initiated AP activity and SSEA1 expression, but lacked detectable amounts of Nanog protein (Fig.1j) whereas by day 20 AP and SSEA1 expression had increased and ~8% of the cells were Nanog-positive. Thus, the four factor-induced reprogramming is a gradual and slow process.

Expression and DNA methylation

To characterize the reprogrammed cells on a molecular level we used quantitative RT-PCR (qRT-PCR) to measure expression of ES cell and fibroblast-specific genes. Fig. 2a shows that in Oct4-iPS cells the total level of Nanog and Oct4 was similar to that in ES cells but decreased upon differentiation to embryoid bodies (EB). MEFs did not express either gene. Using specific primers

for endogenous or total *Sox2* transcripts showed that the vast majority of *Sox2* transcripts originated from the endogenous locus rather than the viral vector (Fig. 2b). In contrast, *HoxA9* and *Zfpm2* were highly expressed in MEFs but at very low levels in iPS or ES cells (Fig. 2c). Western analysis showed that multiple iPS clones expressed Nanog and Oct4 protein at similar levels as ES cells (Fig. 2d). Finally, we used microarray technology to compare gene expression patterns on a global level. Fig. 2f shows that the iPS cells clustered with ES cells in contrast to wild type or donor MEFs.

To investigate the DNA methylation level of the *Oct4* and *Nanog* promoters we performed bisulfite sequencing and COBRA analysis with DNA isolated from ES cells, iPS cells and MEFs. As shown in Fig. 2g, both loci were demethylated in ES and iPS cells and fully methylated in MEFs. To assess whether the maintenance of genomic imprinting was compromised we assessed the methylation status of the four imprinted genes *H19, Peg1, Peg3* and *Snrpn*. As shown in Fig. 2e, bands corresponding to an unmethylated (U) and methylated (M) allele were detected for each gene in MEFs, iPS cells and tail tip fibroblasts. In contrast, embryonic germ (EG) cells, which have erased all imprints(Labosky et al., 1994), were unmethylated. Our results indicate that the epigenetic state of the *Oct4* and *Nanog* genes was reprogrammed from a transcriptionally repressed (somatic) to an active (embryonic) state and that the pattern of somatic imprinting was maintained in iPS cells. Also, the presence of imprints suggests a non-EG cell origin of iPS cells.

Chromatin modifications

Recently, downstream target genes of *Oct4, Nanog* and *Sox2* have been defined in ES cells by genome wide location analyses(Boyer et al., 2005; Loh et al., 2006). These targets include many important developmental regulators, a proportion of which are also bound and repressed by PcG complexes (Lee et al., 2006),(Boyer et al., 2006). Notably, the chromatin at many of these non-expressed target genes adopt a bivalent conformation in ES cells, carrying both the “active” histone H3 lysine 4 (H3K4) methylation mark and the “repressive” histone H3 lysine 27 (H3K27) methylation mark(Azuara et al., 2006; Bernstein et al., 2006). In differentiated cells, those genes tend instead to carry either H3K4 or H3K27 methylation depending on their expression state. We used chromatin immunoprecipitation (ChIP) and real-time PCR to quantify H3K4 and H3K27 methylation for a set of genes reported to be bivalent in pluripotent ES cells(Bernstein et al., 2006). Fig. 3a shows that the fibroblast specific genes *Zfpm2* and *HoxA9* carried stronger H3K4 than H3K27 methylation in the donor MEFs, whereas the silent genes *Nkx2.2, Sox1, Lbx1h and Pax5* primarily carried H3K27 methylation. In contrast, in the Oct4-iPS cells all of these genes showed comparable enrichment for both histone modifications, similar to normal ES cells (Fig. 3a). Identical results were obtained in Nanog-iPS clones selected from Nanog-neo MEFs (Fig. S2). These data suggest that the chromatin configuration of somatic cells is reset to one that is characteristic of ES cells.

iPS cells tolerate genomic demethylation

Tolerance of genomic demethylation is a unique property of ES cells in contrast to somatic cells that undergo rapid apoptosis upon loss of the DNA methyltransferase Dnmt1(Jackson-Grusby et al., 2001; Li et al., 1992; Meissner et al., 2005). We investigated whether iPS cells would be resistant to global demethylation after Dnmt1 inhibition and would be able to re-establish global methylation patterns after restoration of Dnmt1 activity. To this end, we utilized a conditional lentiviral vector harboring a Dnmt1-targeting shRNA and a GFP reporter gene (Fig. 3b(Ventura et al., 2004)). Infected iPS cells were plated at low density and GFP-positive colonies were picked and expanded. Southern analysis using HpaII digested genomic DNA showed that global demethylation of infected iPS cells (Fig. 3c, lanes 6, 7) was similar to Dnmt1$^{-/-}$ ES (lane 2). In contrast, uninfected iPS cells or MEFs (lanes 4, 5) displayed normal methylation levels. Morphologically, the GFP-positive cells were indistinguishable from the parental line or from uninfected sister subclones indicating that iPS cells tolerate global DNA demethylation. In a second step, the Dnmt1 shRNA was excised through Cre-mediated recombination and GFP negative clones were picked (Fig. 3d). The cells had excised the shRNA vector (Fig. 3e) and normal DNA methylation levels were restored (Fig. 3c, lane 8) as has been reported previously for ES cells(Holm et al., 2005). These observations imply that the *de novo* methyltransferases Dnmt3a/b were reactivated in iPS cells(Okano et al., 1999) leading to restoration of global methylation levels. As expected(Holm et al.,

2005), the imprinted genes *Snrpn* and *Peg3* were unmethylated and resistant to remethylation (Fig. 3f).

Maintenance of the pluripotent state

Southern analysis indicated that the Oct4-neo iPS clone 18 carried 4-6 copies of the *Oct4*, c-*Myc* and *Klf4* and only 1 copy of *Sox2* retroviral vectors (Fig. 4a). Because these four factors were under the control of the constitutively expressed retroviral LTR it was unclear in the previous study why iPS cells could be induced to differentiate(Takahashi and Yamanaka, 2006). To address this question, we designed primers specific for the 4 viral encoded factor transcripts and compared expression levels by qRT-PCR in MEFs 2 days after infection, in iPS cells, in embryoid bodies (EB) derived from iPS cells and in demethylated and remethylated iPS cells (Fig. 4b). Although the MEFs represented a heterogenous population composed of uninfected and infected cells, virally encoded RNA levels of *Oct4*, *Sox2*, and *Klf4* RNA were 5 fold and of *c-Myc* more than 10 fold lower in iPS cells than in the infected MEFs suggesting silencing of the viral LTR by *de novo* methylation upon reprogramming. Accordingly, the total *Sox2* and *Oct4* RNA levels in iPS cells were similar to that in wt ES cells and, moreover, the *Sox2* transcripts in iPS cells were mostly if not exclusively transcribed from the endogenous gene (compare Fig. 2b). Upon differentiation to EBs, both viral and endogenous transcripts were down regulated. All viral *Sox2*, *Oct4* and *Klf4* transcripts were about 2-fold upregulated in Dnmt1 knock down iPS cells and again downregulated upon restoration of Dnmt1 activity. This is consistent with

previous data that Moloney virus is efficiently *de novo* methylated and silenced in embryonic but not in somatic cells (Jähner et al., 1982; Stewart et al., 1982). Transcript levels of *c-Myc* were about 20 fold lower in iPS cells than in infected MEFs and did not change upon differentiation or demethylation.

To follow the kinetics of vector inactivation during the reprogramming process we isolated RNA from drug-resistant cell populations at different times after infection. Fig. 4c shows that the viral vector encoded transcripts were gradually silenced during the transition from MEFs to iPS cells with a time course that corresponded to the gradual appearance of pluripotency markers (compare Fig. 1j). Finally, to directly visualize Oct4 and Nanog expression during differentiation we injected Oct4-iPS cells into SCID mice to induce teratoma formation (Fig. 4d). Immunostaining revealed that Oct4 and Nanog were expressed in the centrally located undifferentiated cells but were silenced in the differentiated parts of the teratoma (Fig. 4e, f). Our results suggest that the retroviral vectors are subject to gradual silencing by *de novo* methylation during the reprogramming process. The maintenance of the pluripotent state and induction of differentiation strictly depends on the expression and normal regulation of the endogenous *Oct4* and *Nanog* genes.

Developmental potency

We determined the developmental potential of iPS cells by teratoma and chimera formation. Histological and immunohistochemical analysis of Oct4- or Nanog-iPS cell induced teratomas revealed that the cells had differentiated into cell types representing all three embryonic germ layers (Fig. S3 and S***4***). To more stringently assess their developmental potential various iPS cell-lines were injected into diploid (2N) or tetraploid (4N) blastocysts. Following injection into 2N blastocysts both Nanog- and Oct4-iPS clones derived from MEFs (Fig. 5a) or from tail tip fibroblasts (Fig. 5b,c) as well as iPS cells that had been subjected to a consecutive cycle of demethylation and remethylation (compare Fig. 3b, c) efficiently generated viable high-contribution chimeras (summarized in Table 1). To test for germ line transmission chimeras derived from two different iPS lines (Oct4-iPS O9 and O16) were mated with normal females, blastocysts were isolated and genotyped by 3 different PCR reactions for the presence of the multiple viral Oct4 and c-myc genes and for the single copy GFP-neo sequences inserted into the Oct4 locus of the donor cell (Fig. 1a). Fig. 5f shows that 9 of 16 embryos from 2 chimeras were positive for the viral copies. As expected, only half of the viral positive blastocysts contained the GFP-neo sequences (5/9 embryos, left panel). When E.10 embryos derived from an Oct4-iPS line O16 chimera were genotyped 3/8 tested embryos were transgenic (right panel). Finally, we injected iPS cells into 4N blastocysts as this represents the most rigorous test for developmental potency because the resulting embryos are composed only of the injected donor cells (“all ES embryo”). Fig. 5d,e show that

both Oct4- and Nanog-iPS cells could generate mid- and late-gestation “all iPS embryos” (summarized in Table 1). These findings indicate that iPS cells can contribute to all lineages of the embryo and thus have a similar developmental potential as ES cells.

Discussion

The results presented in this paper demonstrate that the four transcription factors *Oct4*, *Sox2*, *c-Myc* and *Klf4* can induce epigenetic reprogramming of a somatic genome to an embryonic pluripotent state. . In contrast to selection for Fbx15 activation(Takahashi and Yamanaka, 2006), fibroblasts that had reactivated the endogenous *Oct4* (Oct4-neo) or *Nanog* (Nanog-neo) loci grew feeder independently, expressed normal Oct4, Nanog and Sox2 RNA and protein levels, were epigenetically identical to ES cells by a number of criteria and were able to generate viable chimeras, contribute to the germ line and generate viable late-gestation embryos after injection into tetraploid blastocysts. Transduction of the 4 factors generated significantly more drug resistant cells from Nanog-neo than from Oct4-neo fibroblasts but a higher fraction of Oct4-selected cells had all characteristics of pluripotent ES cells suggesting that *Nanog* activation is a less stringent criterion for pluripotency than *Oct4* activation.

Our data suggest that the pluripotent state of Oct4- and Nanog-iPS cells is induced by the virally transduced factors but is largely maintained by the activity of the endogenous pluripotency factors including *Oct4*, *Nanog* and *Sox2* because

the viral controlled transcripts, though expressed highly in MEFs, become mostly silenced in iPS cells. The total levels of *Oct4*, *Nanog* and *Sox2* were similar in iPS and wt ES cells. Consistent with the conclusion that the pluripotent state is maintained by the endogenous pluripotency genes is that the *Oct4* and the *Nanog* genes became hypomethylated in iPS as in ES, and that the bivalent histone modifications of developmental regulators were re-established. Also, iPS cells were resistant to global demethylation induced by inactivation of Dnmt1 similar to ES cells and in contrast to somatic cells. Re-expression of Dnmt1 in the hypomethylated ES cells resulted in global remethylation indicating that the iPS cells had also reactivated the *de novo* methyltransferases Dnmt3a/b. All these observations are consistent with the conclusion that the iPS cells have gained an epigenetic state that is similar to that of normal ES cells. This conclusion is further supported by the recent observation that female iPS cells, similar to ES cells, reactivate the somatically silenced X chromosome (Maherali et al., 2007).

Expression of the four factors proved to be a robust method to induce reprogramming of somatic cells to a pluripotent state. However, the use of retrovirus-transduced oncogenes represents a serious barrier to the eventual use of reprogrammed cells for therapeutic application. Much work is needed to understand the molecular pathways of reprogramming and to eventually find small molecules that could achieve reprogramming without gene transfer of potentially harmful genes.

Methods summary

Cell culture, gene targeting and viral infections. ES and iPS cells were cultivated on irradiated MEFs. Using homologous recombination we generated ES cells carrying an IRES-GfpNeo fusion cassette downstream of Oct4 exon 5 (Fig. 1a). The Nanog gene was targeted as described (Mitsui et al., 2003). Transgenic MEFs were isolated and selected from E13.5 chimeric embryos following blastocyst injection of Oct4-IRES-GfpNeo or Nanog-neo targeted ES cells. MEFs were infected overnight with the Moloney-based retroviral vector pLIB (Clontech) containing the murine cDNAs of *Oct4*, *Sox2*, *Klf4* and *c-Myc*.

Southern blot, methylation and chromatin analysis. To assess the levels of DNA methylation, genomic DNA was digested with HpaII, and hybridized to a probe for the minor satellite repeats (Chapman et al., 1984), or with an IAP-probe (Walsh et al., 1998). Bisulfite treatment was performed with the Qiagen EpiTect Kit. For the methylation status of *Oct4* and *Nanog* promoters bisulfite sequencing analysis was performed as described previously (Blelloch et al., 2006). For imprinted genes, a COBRA assay was performed. PCR primers and conditions were as described previously (Lucifero et al., 2002). The status of bivalent domains was determined by chromatin immunopreciptation followed by quantitative PCR analysis as described before (Boyer et al., 2006)

Expression analysis. Total RNA was reverse transcribed and quantified using QuantTtect SYBR green RT-PCR Kit (Qiagen) on a 7000 ABI detection system. Western blot and immunofluorescence analysis was performed as described

(Hochedlinger et al., 2005; Wernig et al., 2004). Microarray targets from 2μg total RNA were synthesized and labeled using the Low RNA Input Linear Amp Kit (Agilent) and hybridized to Agilent whole mouse genome oligo arrays (G4122F) and analyzed as previously described (Brambrink et al., 2006).

References

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M.*, et al.* (2006). Chromatin signatures of pluripotent cell lines. Nat Cell Biol *8*, 532-538.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Blelloch, R., Wang, Z., Meissner, A., Pollard, S., Smith, A., and Jaenisch, R. (2006). Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. Stem Cells *24*, 2007-2013.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G.*, et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Brambrink, T., Hochedlinger, K., Bell, G., and Jaenisch, R. (2006). ES cells derived from cloned and fertilized blastocysts are transcriptionally and functionally indistinguishable. Proc Natl Acad Sci U S A *103*, 933-938.

Chapman, V., Forrester, L.,]Sanford, J., Hastie, N., and Rossant, J. (1984). Cell lineage specific undermethylation of mouse repetitive DNA. Nature, 284-286.

Cowan, C.A., Atienza, J., Melton, D.A., and Eggan, K. (2005). Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. Science *309*, 1369-1373.

Eads, C.A., and Laird, P.W. (2002). Combined bisulfite restriction analysis (COBRA). Methods Mol Biol *200*, 71-85.

Eggan, K., Akutsu, H., Loring, J., Jackson-Grusby, L., Klemm, M., Rideout, W.M., 3rd, Yanagimachi, R., and Jaenisch, R. (2001). Hybrid vigor, fetal overgrowth,

and viability of mice derived by nuclear cloning and tetraploid embryo complementation. Proc Natl Acad Sci U S A *98*, 6209-6214.

Hochedlinger, K., and Jaenisch, R. (2003). Nuclear transplantation, embryonic stem cells, and the potential for cell therapy. N Engl J Med *349*, 275-286.

Hochedlinger, K., and Jaenisch, R. (2006). Nuclear reprogramming and pluripotency. Nature *441*, 1061-1067.

Hochedlinger, K., Yamada, Y., Beard, C., and Jaenisch, R. (2005). Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. Cell *121*, 465-477.

Holm, T.M., Jackson-Grusby, L., Brambrink, T., Yamada, Y., Rideout, W.M., 3rd, and Jaenisch, R. (2005). Global loss of imprinting leads to widespread tumorigenesis in adult mice. Cancer Cell *8*, 275-285.

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

Jaenisch, R. (2004). Human cloning - the science and ethics of nuclear transplantation. N Engl J Med *351*, 2787-2791.

Jähner, D., Stuhlmann, H., Stewart, C.L., Harbers, K., Löhler, J., Simon, I., and Jaenisch, R. (1982). *De Novo* methylation and expression of retroviral genomes during mouse embryogenesis. Nature *298*, 623-628.

Labosky, P.A., Barlow, D.P., and Hogan, B.L. (1994). Mouse embryonic germ (EG) cell lines: transmission through the germline and differences in the methylation imprint of insulin-like growth factor 2 receptor (Igf2r) gene compared with embryonic stem (ES) cell lines. Development *120*, 3197-3204.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K.*, et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J.*, et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet *38*, 431-440.

Lucifero, D., Mertineit, C., Clarke, H.J., Bestor, T.H., and Trasler, J.M. (2002). Methylation dynamics of imprinted genes in mouse germ cells. Genomics *79*, 530-538.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, Y., Tchieu, J., Jaenisch, R*., et al.* (2007). Global epigenetic remodeling in directly reprogrammed fibroblasts. Cell StemCells, in press.

Meissner, A., Gnirke, A., Bell, G.W., Ramsahoye, B., Lander, E.S., and Jaenisch, R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res *33*, 5868-5877.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.

Naviaux, R.K., Costanzi, E., Haas, M., and Verma, I.M. (1996). The pCL vector system: rapid production of helper-free, high-titer, recombinant retroviruses. J Virol *70*, 5701-5705.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247-257.

Peitz, M., Pfannkuche, K., Rajewsky, K., and Edenhofer, F. (2002). Ability of the hydrophobic FGF and basic TAT peptides to promote cellular uptake of recombinant Cre recombinase: a tool for efficient genetic engineering of mammalian genomes. Proc Natl Acad Sci U S A *99*, 4489-4494.

Stewart, C., Stuhlmann, H., Jähner, D., and Jaenisch, R. (1982). *De novo* methylation, expression, and infectivity of retroviral genomes introduced into embryonal carcinoma cells. Proc Natl Acad Sci USA *79*, 4098-4102.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr Biol *11*, 1553-1558.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Ventura, A., Meissner, A., Dillon, C.P., McManus, M., Sharp, P.A., Van Parijs, L., Jaenisch, R., and Jacks, T. (2004). Cre-lox-regulated conditional RNA interference from transgenes. Proc Natl Acad Sci U S A *101*, 10380-10385.

Walsh, C.P., Chaillet, J.R., and Bestor, T.H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation [letter]. Nat Genet *20*, p116-117.

Wernig, M., Benninger, F., Schmandt, T., Rade, M., Tucker, K.L., Bussow, H., Beck, H., and Brustle, O. (2004). Functional integration of embryonic stem cell-derived neurons in vivo. J Neurosci *24*, 5258-5268.

Yang, X., Smith, S.L., Tian, X.C., Lewin, H.A., Renard, J.P., and Wakayama, T. (2007). Nuclear reprogramming of cloned embryos and its implications for therapeutic cloning. Nat Genet *39*, 295-302.

Acknlowledgments

We thank H. Suh, D. Fu, J. Dausman for technical assistance, J. Love for help with the microarray analysis, S. Markoulaki for help with blastocyst injections, F. Edenhofer for a gift of tat-Cre and S. Yamanaka for the nanog-neo construct. We acknowledge L. Zagachin in the MGH Nucleic Acid Quantitation core for assistance with real-time PCR. We also thank C. Lengner, C. Beard and M. Creyghton for constructive criticism. MW was supported in part by fellowships from the Human Frontiers Science Organization Program and the Ellison Foundation, BB by grants from the Burroughs Wellcome Fund, the Harvard Stem Cell Institute and the NIH, and RJ by grants from the NIH.

Author Information

All microarray data from this study are available from Array Express at the EBI (http://www.ebi.ac.uk/arrayexpress) under the accession number E-MEXP-1037. The authors declare no competing financial interests.

Full Methods

Cell culture, MEF isolation, gene targeting and viral infections. ES and iPS cells were cultivated on irradiated MEFs in DME containing 15% fetal calf serum, Leukemia Inhibiting Factor (LIF), penicillin/streptomycin, L-glutamine, and non-essential amino acids. All cells were depleted of feeder cells for two passages on 0.2% gelatin before RNA, DNA or protein isolation. Transgenic MEFs were isolated and selected in 2µg/ml puromycin (Sigma) from E13.5 chimeric embryos following blastocyst injection of Oct4-inducible KH2 ES cells (Hochedlinger et al., 2005) which had been previously targeted with either Oct4-IRES-GfpNeo or Nanog-neo constructs (Fig. 1a, (Mitsui et al., 2003)). Using homologous recombination in ES cells, an IRES-GfpNeo fusion cassette was inserted into the BclI site downstream of Oct4 exon 5. Correctly targeted ES cell clones were screened by Southern analysis of NcoI digested DNA using a 5' external probe. The murine cDNAs for *Oct4*, *Sox2*, *Klf4* and *c-Myc* were PCR amplified from ES cell cDNA, sequence verified and cloned into the Moloney-based retroviral vector

pLIB (Clontech). $2x10^5$ MEFs or TTFs at passage 2-4 were infected overnight with pooled viral supernatant generated by transfection of $4x10^6$ HEK293T cells (Fugene, Roche) with 10µg of viral vectors and the packaging plasmid pCL-Eco in a 10 cm dish (Naviaux et al., 1996).

Blastocyst injection

Diploid or tetraploid blastocysts (94-98 hours post HCG injection) were placed in a drop of DMEM with 15% FCS under mineral oil. A flat tip microinjection pipette with an internal diameter of 12–15 mm was used for iPS cell injection (utilizing a Piezo micromanipulator (Eggan et al., 2001)). A controlled number of cells were injected into the blastocyst cavity. After injection, blastocysts were returned to KSOM media and placed at 37°C until transferred to recipient females.

Recipient females and caesarean sections

Ten to fifteen injected blastocysts were transferred to each uterine horn of 2.5 days post coitum pseudo pregnant B6D2F1 females. To recover full-term pups recipient mothers were sacrificed at 19.5 days postcoitum. Surviving pups were fostered to lactating BALB/c mothers.

Southern blot, methylation and chromatin analysis.

To assess the levels of DNA methylation, genomic DNA was digested with HpaII, and hybridized to pMR150 as a probe for the minor satellite repeats (Chapman et al., 1984), or with an IAP-probe (Walsh et al., 1998). Bisulfite treatment was performed with the Qiagen EpiTect Kit. For the methylation status of *Oct4* and

Nanog promoters bisulfite sequencing analysis was performed as described previously (Blelloch et al., 2006). 10-20 clones of each sample were sequenced in both directions. For imprinted genes, a COBRA assay was performed. PCR primers and conditions were as described previously (Lucifero et al., 2002). PCR products after bisulfite treatment and gel purification were digested with BstUI (CGCG; H19, Peg3 and Snrpn) or HpyCH4 IV (ACGT; Peg1) and resolved on a 2% agarose gel. Unmethylated CpGs in the recognition sequence will be converted to T and not cut. The status of bivalent domains was determined by chromatin immunopreciptation followed by quantitative PCR analysis as described before (Boyer et al., 2006)

Expression analysis.

50 ng of total RNA isolated using TRIzol reagent (Invitrogen) was reverse transcribed and quantified using QuantTtect SYBR green RT-PCR Kit (Qiagen) on a 7000 ABI detection system. Western blot and immunofluorescence analysis was performed as described (Hochedlinger et al., 2005; Wernig et al., 2004). Primary antibodies included Oct4 (monoclonal mouse, Santa Cruz), Nanog (polyclonal rabbit, Bethyl), actin (monoclonal mouse, Abcam), SSEA1 (monoclonal mouse, Developmental Studies Hybridoma Bank). Fluorophor-labeled secondary antibodies were purchased from Jackson Immunoresearch. Microarray targets from 2μg total RNA were synthesized and labeled using the Low RNA Input Linear Amp Kit (Agilent) and hybridized to Agilent whole mouse genome oligo arrays (G4122F). Arrays were scanned on an Agilent G2565B

scanner and signal intensities were calculated in Agilent FE software. Datasets were normalized using a R script (available at http://www.ebi.ac.uk/arrayexpress) and clustered as previously described (Brambrink et al., 2006).

Viral integrations

Genomic DNA was digested with *Spe*I overnight. Followed by electrophoresis and transfer, the blots were hybridized to the respective radioactively labeled cDNAs of the four transcription factors.

Genotyping

Blastocysts were lysed for 4 hours in 10µl 50mM Tris pH8.8 containing 1mM EDTA and 0.5% Tween20 and 200µg/ml Proteinase K. After heat inactivation for 15 min, PCR was performed with the following conditions: 95°C 30" (1 cycle); 95°C 10", 60°C 15", 72°C 15" (40 cycles); 72°C 5'.

Primer sequences for genotyping

GFP-F	TCCATGGCCAACACTAGTCA
GFP-R	TCCCAGAATGTTGCCATCTT
pLIB-FW1	CCCCCTTGAACCTCCTCGTTCGAC
Oct4R	GAGGTTCCCTCTGAGTTGCTTT
MycR	CGAATTTCTTCCAGATATCCTCAC

Primer sequences for viral-specific qRT-PCR

rtKlf4_virusF1	TCTCTAGGCGCCGGAATTC
rtKlf4_virusR1	CCATGTCAGACTCGCCAGGT
rtMyc_virusF1	CTTCTCTAGGCGCCGGAATT

rtMyc_virusR1 TGGTGAAGTTCACGTTGAGGG

rtOct4_virusF1 TACACCCTAAGCCTCCGCCT

rtOct4_virusR1 ATTCCGGCGCCTAGAGAAG

rtSox2_virusF1 TACACCCTAAGCCTCCGCCT

rtSox2_virusR1 ATTCCGGCGCCTAGAGAAG

DNMT1 hairpin target sequence DZ

GGAAAGAGATGGCTTAACA

Table 1

	2N injections				4N injections		
cell line	injected blast.	live chimeras	chimerism (%)	Germ line	injected blast.	dead embryos (arrested)	live embryos (analyzed
O6	-	-	-	-	13	0	2 (E12.5)
O9	30	5	30-70	yes	90	3 (E11-13.5)	12 (E10-12.5)
O16	15	3	10-30	yes	-	-	-
O18	95	8	5-50	no	134	7 (E9-11.5)	4* (E10-12.5)
O3-2	-	-	-	-	25	2 (E8,11.5)	0
O4-16	-	-	-	-	35	4 (E11-13.5)	3 (E14.5)
N7	30	1	30	-	-	-	-
N8	90	14	5-50	no	118	9 (E9-11.5)	1* (E12.5)
N14	30	5	5-20	-	46	2 (E8,11.5)	1 (E12.5)
TT-O25	50	2	30†	-	39	3 (E9.5)	0
O18 rem/3.1	25	1	30	-	-	-	-

*developmentally retarded or abnormal; † based on GFP fluorescence

Summery of 2N and 4N injections. The extent of chimerism was estimated based on coat color or EGFP expression (**†**); * developmentally retarded. – not determined. 4N injected blastocysts were analyzed between embryonic day E10.5 and E14.5. Analyzed = day of embryonic development analyzed; arrested = estimated stage of development of dead embryos.

Figure legends

Fig. 1

Generation of Oct4- and Nanog-selected iPS cells. a) Targeting strategy to generate an Oct4 IRES-GfpNeo allele. The resulting GfpNeo fusion protein has sufficient neomycin resistance activity in ES cells, GFP fluorescence however is not visible. b) Phasecontrast micrograph of Oct4-iPS cells (clone 18) grown on irradiated MEFs. Inset: ES cell-like colony 5 days after seeding in clonal density without feeder cells. iPS clone 18 cells exhibited strong alkaline phosphatase activity (c) and were homogenously labeled with antibodies against SSEA1 (d,e) and Nanog (f,g). h) One example of an ES-like colony 16 days after infection (left). The majority of G418-resistant colonies, however, consisted of flat non-ES-like cells (right). i) Gradual activation of the Nanog and Oct4-neo alleles. Shown are the total colony numbers of one experiment at day 20 after infection starting neo-selection at day 3, 6, and 9. j) Fraction of total selected cells expressing AP, SSEA1 and Nanog 0, 14, and 20 days after infection (counted were more than 10 visual fields containing n> 1000 total cells for every time point; s.d.). k) Estimated reprogramming efficiency of Oct4- and Nanog-selection (n=3 different experiments; +/- s.e.m.). Indicated are the total number of drug-resistant colonies per 100.000 plated MEF cells 20 days after infection; the fraction of “ES-like” colonies per total number of colonies; the fraction of iPS cell lines which could be established from picked “ES-like” colonies as defined by homogenous AP, SSEA1 and Nanog expression. After determining the fraction of Sox2 (83.4%), Oct4 (53.2%) and c-Myc (46.3%)-infected MEFs 2 days after infection by

immunofluorescence and assuming 50% were infected by Klf4 viruses we estimated the overall reprogramming efficiency as the ratio of quadruple infected cells and the extrapolated total number of iPS cell lines that could be established with G418 selection starting at day 6 post infection.

Fig. 2

Expression and promoter methylation analysis of iPS cells. a-c) qRT-PCR analysis (n = 3 independent PCR reactions; s.d.) of Oct4-iPS clone 18, subclone 18.1, 2 week-old embryoid bodies (EBs) derived from clone 18, V6.5 ES cells and Oct4-neo MEFs shows similar Nanog (red bars) and total Oct4 (blue bars) levels as in ES cells (a); slightly lower total *Sox2* levels (filled red bars) most of which was due to expression of endogenous *Sox2* transcripts (open red bars, b); strong downregulation of *HoxA9* (red) and *Zfpm2* (blue) transcripts in iPS cells (c). Transcript levels were normalized to GAPDH expression with expression levels in ES cells (a,b) and MEFs (c) set as “1”. d) Western blot analysis for Oct4 and Nanog expression of different Oct4-iPS clones (6,9,10,16,18) and a GFP labeled subclone of clone 18 (18.1). e) COBRA methylation analysis(Eads and Laird, 2002) of imprinted genes *H19* (maternally expressed), *Peg1* (paternally expressed), *Peg3* (paternally expressed) and *Snrpn* (paternally expressed). Upper band = unmethylated (U); lower bands = methylated (M). f) Unsupervised hierarchical clustering of averaged global transcriptional profiles obtained from Oct4-neo iPS clone 18, Nanog-neo iPS clone 8, genetically matched ES cells (V6.5;129/B6), Oct4-neo MEFs (O), Nanog-neo MEFs (N) and wildtype 129/B6

F1 MEFs (wt). g) Analysis of the methylation state of the Oct4 and Nanog promoter using bisulfite sequencing. Open circles indicate unmethylated and filled circles methylated CpG dinucleotides. Shown are 8 representative sequenced clones from ES cells (V6.5), Oct4-neo MEFs and Oct4-neo iPS clone 18.

Fig. 3

Reprogrammed MEFs acquire an ES cell-like epigenetic state. a) Real-time PCR after chromatin immunoprecipitation using antibodies against tri-methylated histone H3K4 and H3K27. Shown are the log2 enrichments for several previously reported 'bivalent' loci in ES cells (n=3 experiments; s.d.). *Zfpm2* and *HoxA9* show enrichment for the active (H3K4) mark in MEFs and are expressed (Fig. 1c and microarray data) whereas the other tested genes remain silent (microarray data). All loci tested in iPS clone O18 show enrichment for both H3K4 and H3K27 tri-methylation ('bivalent'), as seen in ES cells (V6.5, 129/B6F1). See Supplemental Fig. S2 for H3K4 and H3K27 tri-methylation analysis of a subclone (clone O18.1) and Nanog-neo iPS clone N8. b) Experimental design to de- and remethylate genomic DNA. Clone O18 was infected with the Dnmt1 hairpin containing lentiviral vector pSicoR-GFP. The shRNA and GFP marker in the pSicoR vector are flanked by LoxP sites(Ventura et al., 2004)). Green colonies were expanded and passaged 4 times. Tat-Cre protein transduction was used to remove the shRNA(Peitz et al., 2002). c) Southern blot analysis of the minor satellite repeats using a methylation sensitive restriction enzyme (HpaII) and its

methyl insensitive isoschizomer (MspI) as a control. Loss of methylation in two different clones (lanes 6 and 7) is comparable to Dnmt1 knockout ES cells (lane 2). After Cre-mediated recombination complete remethylation (lane 8) of the repeats is observed within 4 passages. d) Successful loop out after tat-Cre treatment was identified by disappearance of EGFP fluorescence (arrow) and verified by PCR analysis (e). f) COBRA assay of the imprinted genes Peg3 and Snrpn and a random intergenic region close to the Otx2 locus (Interg.) demonstrating the expected resistance to *de novo* methylation of imprinted genes in contrast to non-imprinted intergenic sequences. U=unmethylated band, M=methylated bands.

Fig. 4

Efficient silencing of retroviral transcripts in induced pluripotent cells. a) Southern blot analysis of proviral integrations in iPS clone O18 (left lanes) for the four retroviral vectors. Uninfected ES cells (right lanes) show only one or two bands corresponding to the endogenous gene (marked by *). b) Quantitative RT-PCR using primers specifically detecting the four viral transcripts. Shown are Oct4-neo iPS clone 18 and a GFP labeled subclone, Oct4-neo MEFs, 2-week old EB's generated from clone 18, two demethylated clones (18 dem/1 and 18 dem/3), a remethylated clone (18 rem/3.1) and Oct4-neo MEFs 2 days after infection with all 4 viruses but not selected with G418 (n = 3 independent experiments; s.d.). c) Viral transcript levels at various time points in cell populations after infection and Oct4-selection and in the two Oct4-iPS cell-lines

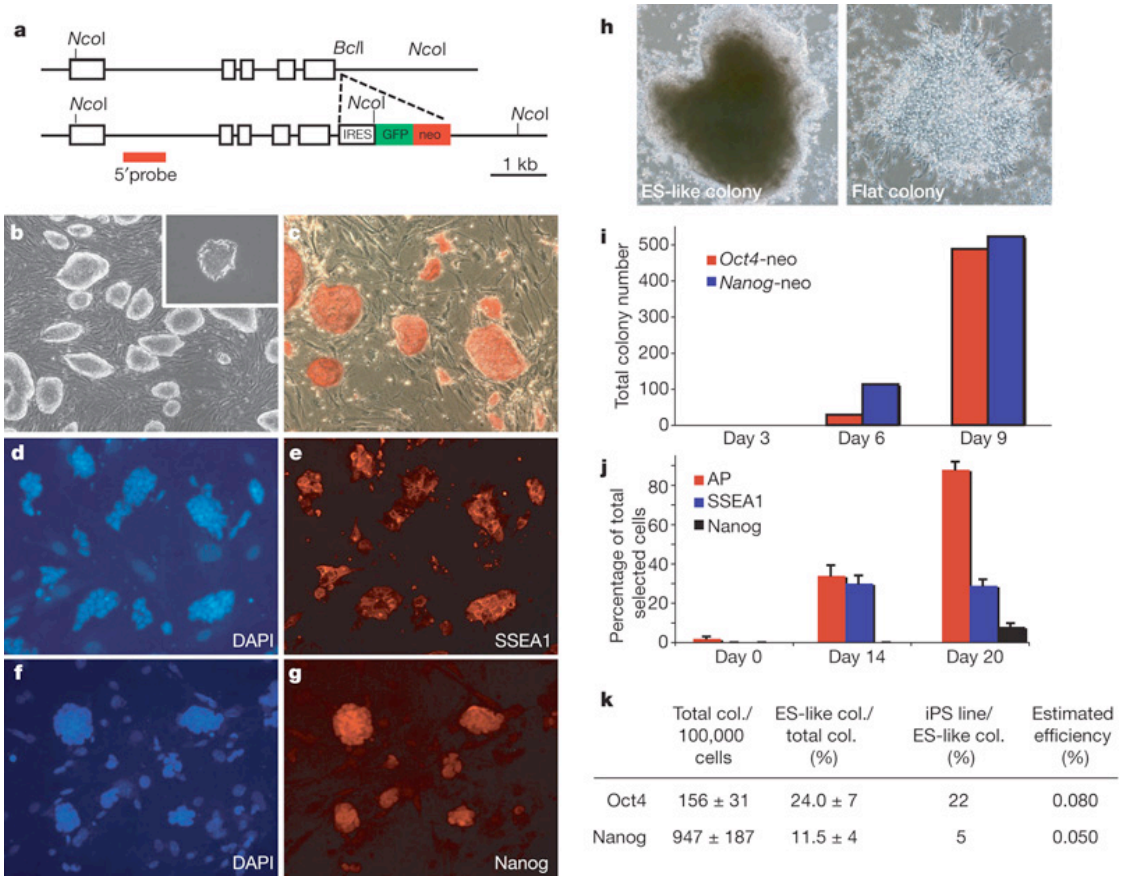
O1.3 and O9 (n = 3 independent experiments; s.d.). d-f) Paraffin sections of a teratoma 26 days after subcutaneous injection of Oct4-iPS clone 18 cells into SCID mice. Nanog (e) and Oct4 (f) expression was confined to undifferentiated cell types based on immunohistochemical analysis.

Fig. 5

Developmental pluripotency of reprogrammed fibroblasts. a) A 6 week old chimeric animal. Agouti coloured hairs originated from Oct4-iPS cell line O18.1. b) Two live pups after 2N blastocyst injection one of which shows high contribution (c) of the tailtip fibroblast-derived Oct4-iPS cell line TT-O25, which had been GFP labeled with a lentiviral Ubiquitin-EGFP vector. d) “all iPS cell embryos” were generated by injection of iPS cells into 4N blastocysts(Eggan et al., 2001). Live E12.5 embryos generated from Oct4-iPS line O6 (left embryo), from Nanog-iPS line N14 (middle embryo) and from V.6.5 ES cells (right embryo). e) A normally developed E14.5 embryo was derived from Oct4-iPS cell line O4-16 after tetraploid complementation which was isolated by screening MEFs for activation of GFP inserted into the Oct4 locus. f) Germline contribution of the Oct4-iPS clones O9 and O16. Genotyping of blastocysts from females mated with 3 chimeric males demonstrated the presence of Oct4- and c-Myc virus integrations and the Oct4-IRES-GfpNeo allele (left panel). Because of the multiple integrations (Fig. 4a) all embryos with iPS cell contribution are expected to be positive for proviral sequences in this assay. In contrast the single copy Oct4-IRESGfpNeo allele segregated into only 5 of the 9 virus positive embryos.

All six blastocysts from O9 chimera #1 were iPS-cell derived suggesting that this chimera was a pseudo-male. Additional genotyping identified 13/72 tested blastocysts derived from iPS line O9 and 4/13 blastocysts derived from iPS line O16 chimeras to carry the viral transgenes. The right panel shows that 3/8 tested E.10 mid-gestation embryos were sired by a chimera derived from the donor iPS line O16. + positive, - negative control samples.

Figure 1



k

	Total col./ 100,000 cells	ES-like col./ total col. (%)	iPS line/ ES-like col. (%)	Estimated efficiency (%)
Oct4	156 ± 31	24.0 ± 7	22	0.080
Nanog	947 ± 187	11.5 ± 4	5	0.050

Figure 2

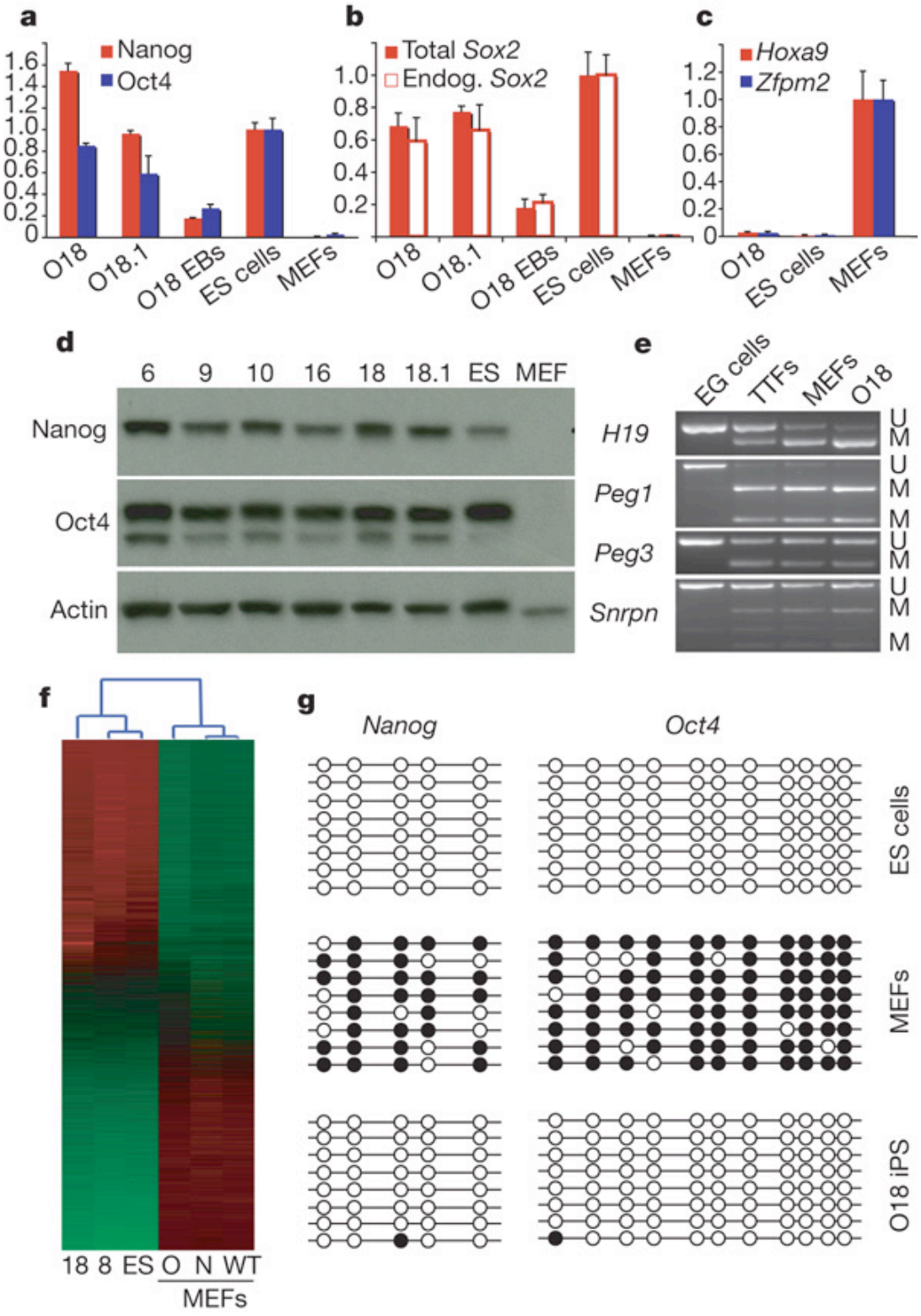

a
Nanog
Oct4
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
O18
O18.1
O18 EBs
ES cells
MEFs
b
Total Sox2
Endog. Sox2
1.0
0.8
0.6
0.4
0.2
O18
O18.1
O18 EBs
ES cells
MEFs
c
Hoxa9
Zfpm2
1.2
1.0
0.8
0.6
0.4
0.2
0
O18
ES cells
MEFs
d
6
9
10
16
18
18.1
ES
MEF
Nanog
Oct4
Actin
e
EG cells
TTFs
MEFs
O18
H19
Peg1
Peg3
Snrpn
U
M
f
18 8 ES O N WT
MEFs
g
Nanog
Oct4
ES cells
MEFs
O18 iPS

Figure 3

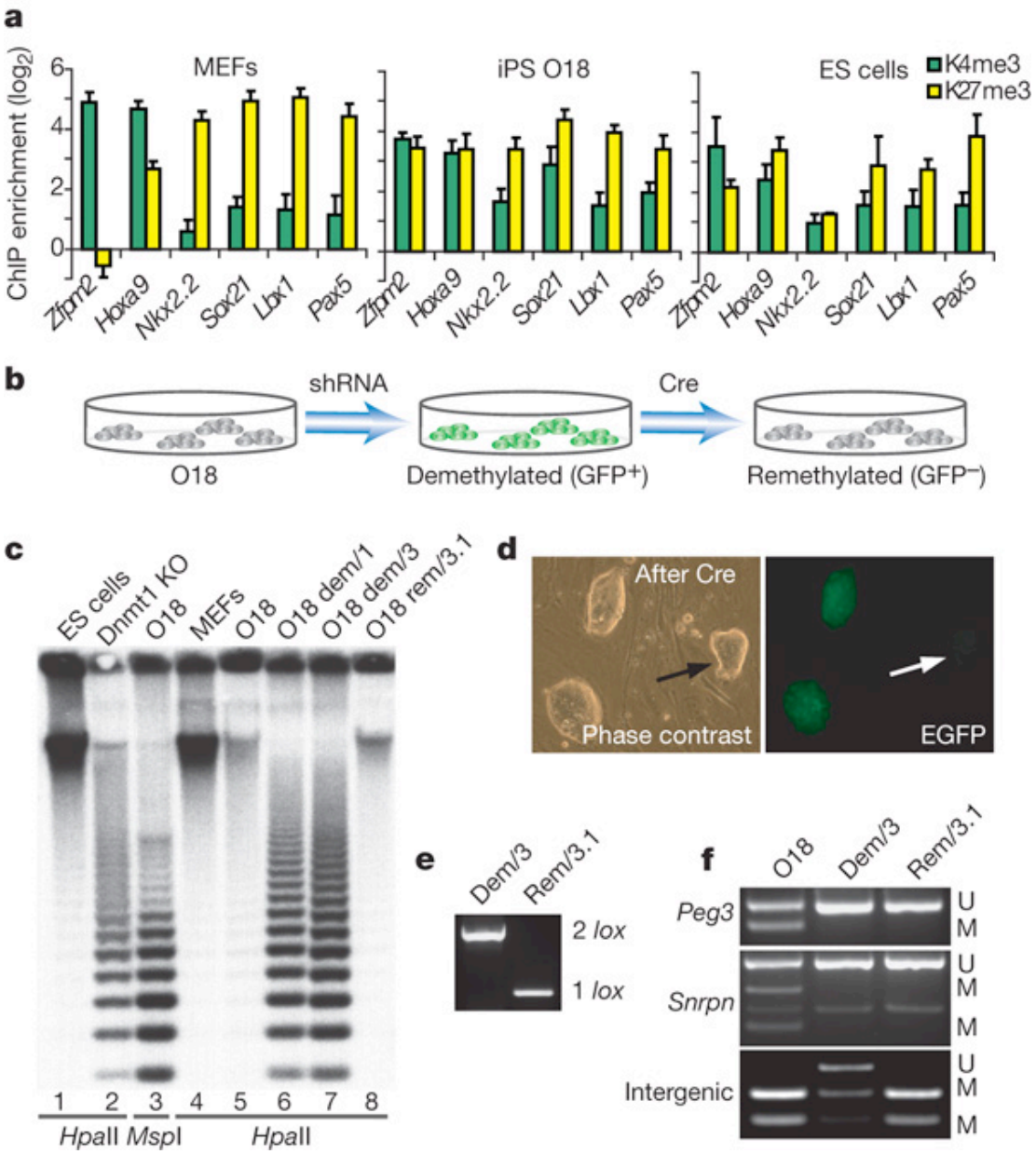

a
ChIP enrichment (log2)
MEFs
iPS O18
ES cells
K4me3
K27me3
Zfpm2
Hoxa9
Nkx2.2
Sox21
Lbx1
Pax5
b
shRNA
Cre
O18
Demethylated (GFP+)
Remethylated (GFP−)
c
ES cells
Dnmt1 KO
O18
MEFs
O18
O18 dem/1
O18 dem/3
O18 rem/3.1
HpaII MspI
HpaII
d
After Cre
Phase contrast
EGFP
e
Dem/3
Rem/3.1
2 lox
1 lox
f
O18
Dem/3
Rem/3.1
Peg3
Snrpn
Intergenic
U
M

Figure 4

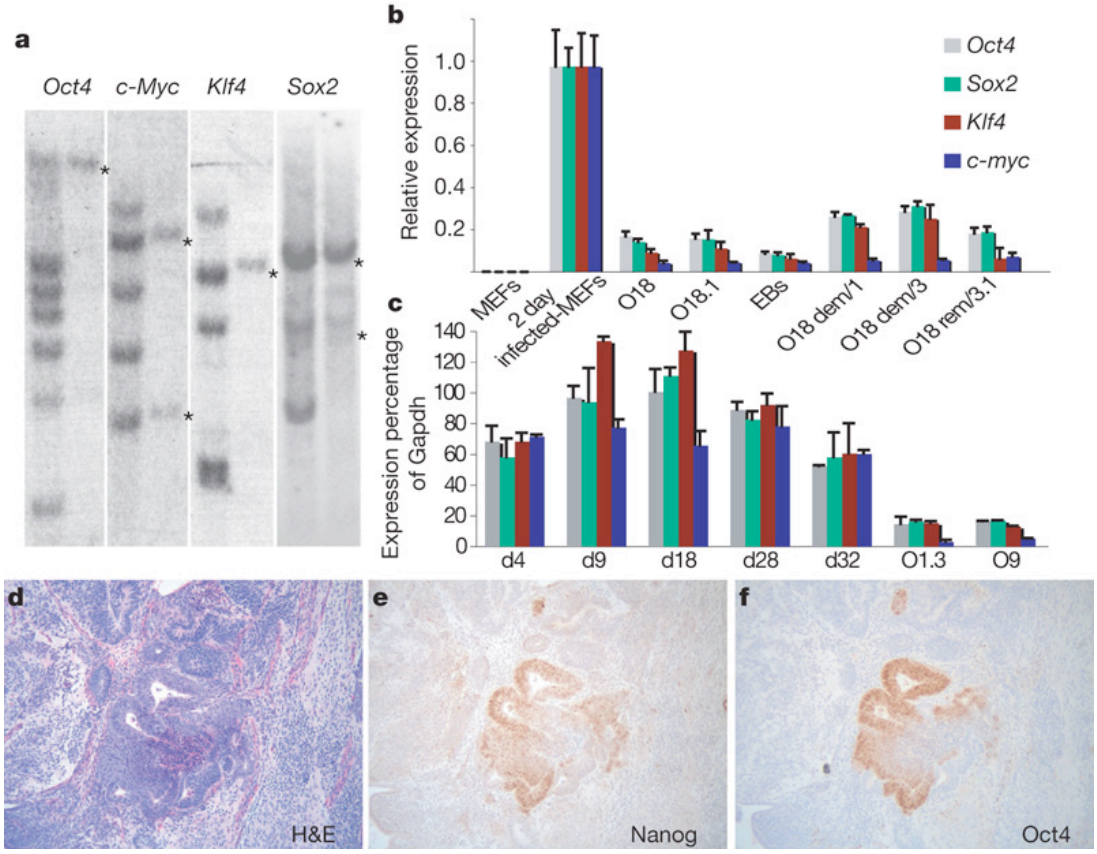
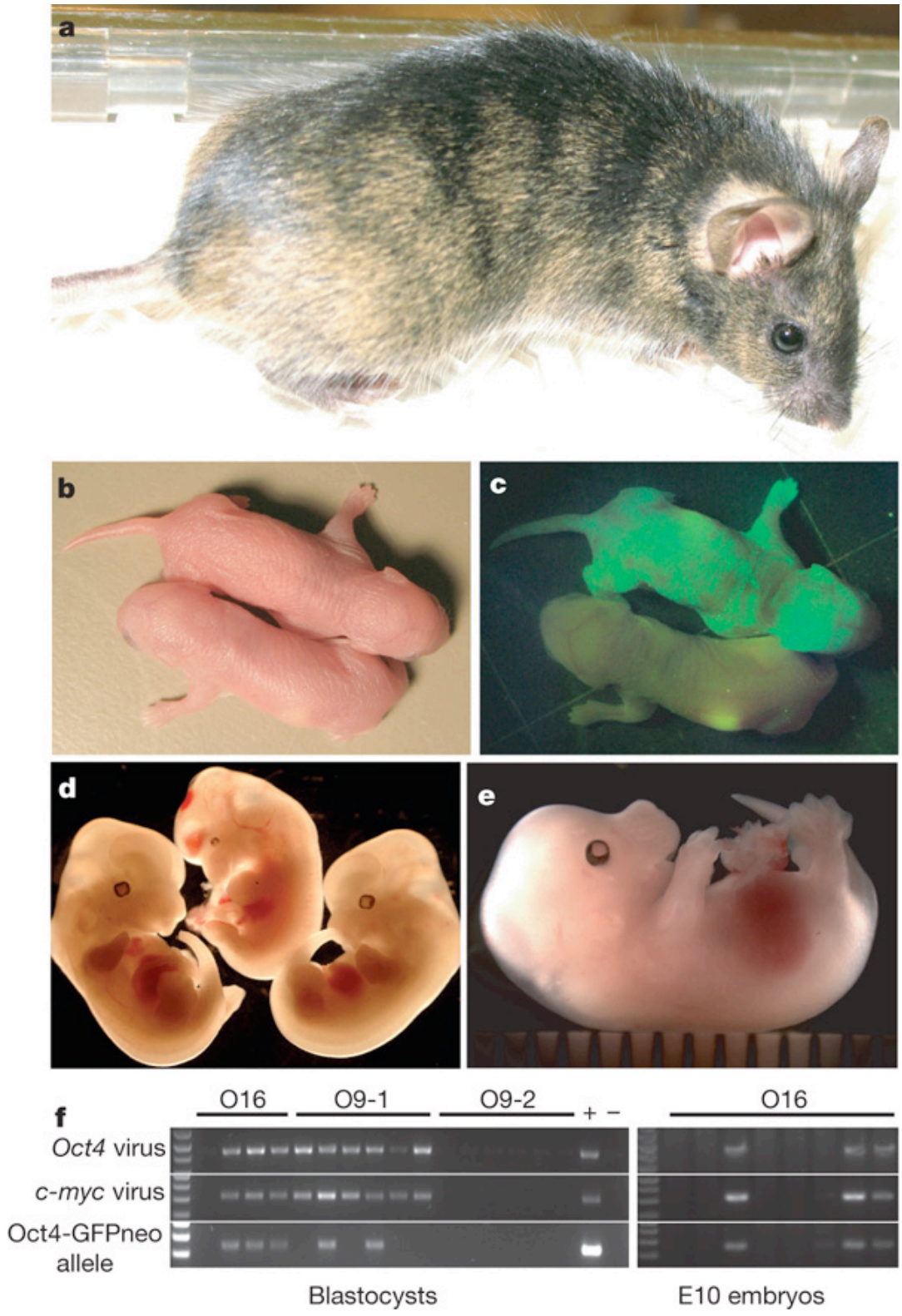

a
Oct4
c-Myc
Klf4
Sox2
b
Relative expression
1.0
0.8
0.6
0.4
0.2
Oct4
Sox2
Klf4
c-myc
MEFs
2 day infected-MEFs
O18
O18.1
EBs
O18 dem/1
O18 dem/3
O18 rem/3.1
c
Expression percentage of Gapdh
140
120
100
80
60
40
20
0
d4
d9
d18
d28
d32
O1.3
O9
d
H&E
e
Nanog
f
Oct4

Figure 5


a
b
c
d
e
f
O16
O9-1
O9-2
+ −
O16
Oct4 virus
c-myc virus
Oct4-GFPneo allele
Blastocysts
E10 embryos

Chapter 3

Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells

Tobias Brambrink[1], Ruth Foreman[1], G. Grant Welstead[1], Christopher J. Lengner, Marius Wernig, Heikyung Suh and Rudolf Jaenisch

[1]these authors contributed equally to this work, names are in alphabetical order

This chapter was initially published in Cell Stem Cell, *2*, 151-159. Permission to publish it in this thesis was granted by Elsevier.

Author contributions:
TB, RF, and GGW conceived of the experiments in this chapter and conducted the cell culture experiments described. RF generated the inducible contructs and performed the blastocyst injections. TB and GGW performed the majority of FACS analyses. TB prepared the manuscript.

SUMMARY

Pluripotency can be induced in differentiated murine and human cells by retroviral transduction of Oct4, Sox2, Klf4 and c-Myc. The resulting 'induced pluripotent stem' (iPS) cells resemble embryonic stem (ES) cells in their developmental potential and their epigenetic state. This direct reprogramming approach could potentially lead to the development of 'customized' pluripotent cell lines for human therapy and represents a valuable tool for basic research. Further advances in the field, such as the development of virus-free reprogramming strategies, would be greatly facilitated by a better understanding of the molecular mechanism governing the reprogramming process. To this end, we have devised a strategy in which the cDNAs of the four transcription factors are under the transcriptional control of a doxycycline (dox) dependent promoter in a lentiviral vector. We derived iPS cell lines from mouse fibroblasts using this inducible system and determined the timing of known pluripotency marker gene activation by FACS analysis. We found that alkaline phosphatase (AP) was activated first, followed by stage specific embryonic antigen 1 (SSEA1). Expression of Nanog and the endogenous Oct4 gene were only observed late in the process, marking fully reprogrammed cells. We also discovered that the virally transduced cDNAs needed to be expressed for at least 12 days in order to generate fully reprogrammed cells. In contrast to previously used retroviral constructs, which are efficiently silenced in iPS cells, the lentivirus-based vectors employed in this study allow for continued overexpression of the viral transcripts in iPS cells. We investigated the effect of sustained ectopic transgene expression on the differentiation potential of different iPS cell lines harboring inducible or

constitutively expressed viral transgenes and found that transgene silencing is a prerequisite for normal cell differentiation. Our results are a step towards an understanding of some of the molecular events underlying epigenetic reprogramming.

INTRODUCTION

The generation of pluripotent cell lines from somatic cells has great potential for basic research as well as clinical applications (Hochedlinger and Jaenisch, 2003; Rideout et al., 2002). Even though epigenetic reprogramming of a somatic genome to an embryonic state by nuclear transfer or fusion of ES cells with somatic cells has become a widely used procedure in various mammalian species over the last decade (Cowan et al., 2005; Kato et al., 1998; Polejaeva et al., 2000; Tada et al., 2001; Wakayama et al., 1998; Wilmut et al., 1997), the underlying molecular mechanism has not been identified.

It has been shown recently that retroviral transduction of mouse and human somatic cells with four transcription factors initiates the gradual conversion of a small subpopulation of the infected cells into a pluripotent, ES cell-like state (Maherali et al., 2007; Meissner et al., 2007; Okita et al., 2007; Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Wernig et al., 2007; Yu et al., 2007). Mouse fibroblasts infected with murine Moloney Leukemia Virus (MLV) derived vectors carrying the cDNAs of Oct4, Sox2, c-Myc, and Klf4 under the control of

the viral long terminal repeat (LTR) region gave rise to ES cell-like iPS cell colonies at a low frequency (Maherali et al., 2007; Meissner et al., 2007; Okita et al., 2007; Takahashi and Yamanaka, 2006; Wernig et al., 2007). These iPS cells were able to differentiate into cell types of all three germ layers in teratomas and contributed to the germline of chimeric mice (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007). DNA methylation and histone modification analyses revealed that the chromatin of iPS cells had been reprogrammed to an embryonic state. This included demethylation of the endogenous promoters controlling ES cell specific genes like Nanog and Oct4 (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007) and the simultaneous establishment of both H3K4 and H3K27 tri-methylation marks at loci previously identified to be 'bivalent' in ES cells (Bernstein et al., 2006; Maherali et al., 2007; Wernig et al., 2007). During the derivation process, iPS cells activated endogenous pluripotency markers, such as Nanog, Oct4, Sox2, SSEA1 and AP, and efficiently downregulated the transcriptional activity of the viral LTRs to basal levels (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007), suggesting that viral transgene expression becomes dispensable at a certain time point during the reprogramming process. These findings in mouse together with the recent derivation of human iPS cell lines let the development of customized human cell therapy approaches appear within reach (Takahashi et al., 2007; Yu et al., 2007). However, many problems remain unsolved. One major obstacle is that the stochastic activation of the viral transgenes has been linked to an elevated frequency of tumor formation in iPS cell-derived chimeric mice (Okita et

al., 2007). An in-depth understanding of the reprogramming process could accelerate the development of virus-free iPS cell derivation strategies, which is mandatory for clinical application of this technology.

In previous studies, iPS cell colonies were detected at widely differing time points after infection of somatic cells. Depending on the selection strategy employed, reprogrammed cells were identified as early as on day three and as late as several weeks after infection (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007). Our recently published data demonstrated that cells originating from the same infected parental cell activate the endogenous Oct4-locus at different time points, suggesting that the induction of pluripotency by viral transduction is a gradual process involving stochastic epigenetic events (Meissner et al., 2007). Viral transduction-mediated reprogramming involves the transcriptional activation of endogenous markers for pluripotency such as AP, SSEA1, Sox2, Oct4 and Nanog (Maherali et al., 2007; Meissner et al., 2007; Okita et al., 2007; Wernig et al., 2007). It is unclear, however, whether the reactivation of pluripotency-associated genes is a random, stochastic process or follows a specific set of sequential events and the minimal length of transgene expression required for successful reprogramming has not been determined. Understanding the timing of endogenous gene activation and transgene silencing is fundamental for the development of transient, non-viral reprogramming strategies.

In order to study the activation of endogenous genes during the reprogramming process and the time requirements for ectopic expression of the four factors, we have developed a lentiviral system in which the cDNAs of the four transcription factors are driven from a dox-inducible promoter. In contrast to previously employed retroviral systems (Maherali et al., 2007; Meissner et al., 2007; Okita et al., 2007; Wernig et al., 2007) or MLV vectors paired with the dox-inducible expression of Oct4 from a targeted allele (Maherali et al., 2007), this allows for controlled expression of all four transgenes independent of viral integration and cellular silencing of viral LTRs. We have derived iPS cell lines after infection of fibroblasts obtained from Oct4-GFP knock-in mice or Nanog-GFP knock-in mice expressing the M2rtTA from the endogenous Rosa26 (R26) promoter. By utilizing inducible viruses, we were able to activate or downregulate the expression of the four factors at different time points and to analyze the timing of the reprogramming process. Using FACS analysis, we assessed the activation of the four ES cell markers AP, SSEA1, Oct4-GFP and Nanog-GFP, and we observed that SSEA1 expression marks an intermediate state of reprogramming whereas Oct4 and Nanog expression was only detected in fully reprogrammed cells. We also identified the minimum time of viral transgene expression required for completion of the reprogramming process and determined the effect of continued expression of the four factors on the differentiation potential of several iPS cell lines.

RESULTS

A dox-inducible system for the derivation of pluripotent cell lines

A lentiviral backbone for dox-inducible transgene expression was constructed by replacing the human Ubiquitin-C promoter of the FUW plasmid (Lois et al., 2002) with a tetracycline operator and minimal CMV promoter. The cDNAs for Oct4, Sox2, Klf4, and c-Myc were subsequently cloned into this backbone. We generated mouse embryonic fibroblasts (MEFs) carrying the R26 promoter-driven M2rtTA (Beard et al., 2006) and a Green Fluorescent Protein (GFP) allele driven by either the endogenous Oct4 or Nanog promoter (Maherali et al., 2007; Meissner et al., 2007). *M2rtTA(+/-);Oct4GFP(+/-)* and *M2rtTA(+/-);NanogGFP(+/-)* MEFs were infected with inducible Oct4, Sox2, c-Myc, and Klf4 viruses and subsequently treated with dox (Figure 1A). To determine the expression levels of the four genes from the inducible viruses, we performed quantitative PCR analysis showing that ectopic transcript expression could be strongly induced in the presence of dox (Supplementary Figure 1).

To generate iPS cell lines, $2.5x10^5$ MEFs were infected in 10cm cell culture dishes. The cultures were split 1:5 at three days after infection. Dox was added to the cell culture medium one day following the split to induce transgene expression and initiate the reprogramming process. In the presence of dox, the

morphology of the infected MEFs changed within 3 days, with small, rounded cells forming in the culture. Similar to previous observations (Maherali et al., 2007; Meissner et al., 2007; Okita et al., 2007; Wernig et al., 2007) small colonies were observed by day 9 and distinctively ES-like colonies appeared by day 16, many of which had activated transcription of endogenous Oct4 or Nanog indicated by GFP expression. ES-like colonies were not observed in cultures that had not been treated with dox. We picked single GFP-positive colonies from the dox-treated plates at day 26 and expanded them on feeder MEFs in the absence of dox to derive iPS cell lines (Figure 1B). Viable chimeras were generated from both Oct4-GFP and Nanog-GFP iPS clones after injection into Balb/c blastocysts (Figure 1C and D).

Sequential activation of pluripotency markers

In order to provide a baseline measure of endogenous ES cell marker activation during the direct reprogramming of MEFs, we performed FACS analyses of AP, SSEA1, and Oct4- or Nanog-driven GFP. For each experiment, Oct4-GFP and Nanog-GFP cells were infected with the same batch of virus to reduce variability. Three days after infection, the MEFs were split onto gelatin-coated 10cm plates. The next day, dox was added to all plates (day 0). The expression of ES cell specific markers was analyzed at various time points after dox addition up to day 35. The average percentages of cells expressing AP, SSEA1, Oct4-GFP or Nanog-GFP from three independent experiments are displayed in Figure 2. Representative FACS plots for AP and SSEA1/Oct4-GFP or SSEA1/Nanog-GFP

analysis are displayed in Supplementary Figure 2. Since the GFP signal to background fluorescence ratio was low, we performed autofluorescence correction as outlined elsewhere (Alberti et al., 1987). Consistent with our previous observations (Wernig et al., 2007), AP was reactivated early after induction of the four factors and was detectable in a small percentage (~3-4%) of cells at day 3. SSEA1 was first expressed in a subpopulation of cells (~4%) at day 9 and GFP expression from either the Oct4 or Nanog endogenous promoters was first observed by FACS at day 16, with the percentage of GFP positive cells being below 1% in both cases. We performed quantitative RT-PCR assays to investigate whether we could detect transcriptional activation of endogenous Nanog or Sox2 loci before the detection of GFP fluorescence. We found that no appreciable levels of either transcript could be detected by this assay at early time points up to day 14 (Supplementary Figure 3). The portion of cells expressing AP, SSEA1 or GFP increased in both reporter lines over time, consistent with our previous results (Wernig et al., 2007).

SSEA1 activation marks an intermediate step of reprogramming

The timing of marker appearance observed by FACS suggests that the activation of AP, SSEA1, Oct4 and Nanog might be sequential events. This notion is supported by the observation that at day 9 after infection only about 7% of the AP positive cells also expressed SSEA1 and at day 16 only about 3% of the cells staining for SSEA1 also expressed Nanog or Oct4 as indicated by GFP fluorescence. In contrast, most if not all GFP positive cells also expressed

SSEA1. To test the hypothesis that SSEA1 and Oct4/Nanog are sequentially activated or that SSEA1 expression is required prior to Oct4 or Nanog activation, we sorted infected populations of Oct4-GFP and Nanog-GFP cells at day 21 of dox treatment (Figure 3A). Equal numbers of SSEA1+/GFP- cells were plated on feeders and cultured for three days either in the presence or in the absence of dox. After three days the medium was changed and the cells were incubated in dox-free medium for an additional seven days. GFP positive colonies were observed in both treated and untreated cultures of SSEA1+/GFP- cells (Figure 3B). The percentage of SSEA1+/GFP- cells that were able to generate GFP positive colonies was calculated. In both the Oct4-GFP and the Nanog-GFP cell populations ~0.02% of the SSEA1+/GFP- cells gave rise to GFP positive colonies without dox treatment. This percentage increased about ten-fold to ~0.2% when the cells were treated with dox, indicating that continuous expression of the transcription factors enhanced reprogramming in SSEA1+/GFP- cells (Figure 3C). We confirmed the existence of SSEA1+/GFP+ cells in the sorted populations by FACS analysis on day 11 after the initial sorting (Figure 3D). In contrast to SSEA1+/GFP- cells, SSEA1-/GFP- cells did not give rise to any colonies after ten days of culturing them with or without dox.

To investigate whether the SSEA1+/GFP- cells that did not give rise to iPS cell colonies after dox-withdrawal simply stopped dividing while maintaining SSEA1 expression or return to a MEF-like SSEA1-negative state, we repeated this experiment. This time, however, we sorted cells after only 9 days of dox

treatment - a time point well before the appearance of Nanog- or Oct4-GFP could be observed. The SSEA1+/GFP- cells were seeded onto plates with feeders and cultured for 20 days with or without dox. In addition, some of the SSEA1+/GFP- cells were also seeded onto plates without feeders to allow for the analysis of their morphology. At the end of the culture period, we assessed SSEA1 and GFP expression by FACS analysis (Figure 4A). In contrast to the SSEA1+/GFP- populations sorted at day 21, the populations sorted at day 9 did not give rise to any SSEA1+/GFP+ cells after culturing them without dox, indicating the complete absence of transgene-independent cells within the SSEA1+/GFP- populations sorted at this earlier time point. Furthermore, most if not all cells cultured without dox ceased to express SSEA1 and returned to a MEF-like morphology (Figure 4B). In contrast, when cultured in the presence of dox, the SSEA1+/GFP- cells from day 9 gave rise to iPS colonies and yielded SSEA1+/GFP+ iPS cells at the end of the culture period (Figure 4A and B).

Our results indicate that SSEA1 activation may represent an intermediate step that may have to precede activation of Oct4 or Nanog in order for reprogramming to be successful. The SSEA1-positive state is unstable and most SSEA1+/GFP- cells depend on continued transgene expression to complete the reprogramming process. When deprived of continued expression of the four transcription factors, most cells cannot proceed further towards an embryonic state, but return to a state resembling that of the donor cell. Our findings also indicate that reprogramming is a gradual, asynchronous process and that continuous vector

expression enhances reprogramming efficiency even at 21 days after the start of dox induction, a time point well beyond the initial appearance of GFP positive iPS cells.

Minimal time of transgene expression required for direct reprogramming

We investigated how long the exogenous expression of the four factors was required for successful reprogramming of fibroblasts to dox-independent, self-renewing iPS cells. For this, cells were infected with the inducible lentiviruses, and split onto six gelatin-coated 10cm plates three days later as outlined in Figure 5A. One day after splitting the cells, dox was added to the culture medium (day 0). The drug was withdrawn from individual cultures at various time points (day 12, 14, 16, 19, 22 and 26) and the dishes were fixed and stained for AP on day 35 (Figure 5A). We found that day 14 was the earliest time point of dox-withdrawal that was permissive to the establishment of dox-independent colonies that stained positive for AP (Figure 5A). No dox-independent, AP positive colonies were observed on plates that had dox withdrawn on day 12 (Figure 5A), even though cell colonies were visible on those plates at the time of dox-withdrawal. These initial colonies did not proliferate after dox-withdrawal and did not stain for AP at day 35. In the dishes where dox was withdrawn at later time points, the number of dox-independent, AP positive colonies increased with the time that the cells were exposed to dox and therefore with the length of transgene expression (Figure 5A).

To determine the time point of dox-withdrawal at which fully reprogrammed iPS cells could reproducibly be derived, we repeated this experiment in triplicate for both Nanog-GFP and Oct4-GFP cells. This time we withdrew dox on days 9, 12, 14, 16, 21 and 24, respectively, and examined the plates for GFP expressing iPS colonies on day 35 (Figure 5B). No marked differences in colony counts were detected between Oct4-GFP and Nanog-GFP cells and the data obtained from both donor cell lines were combined. We observed that dox-independent, GFP expressing iPS cell colonies could not reliably be detected on plates that had dox removed on or prior to day 12, with four of the six “day 12” plates containing no GFP-positive colonies, one plate containing a single colony and another plate containing two GFP-positive colonies. Notably, a number of small colonies were visible on these early withdrawal plates at the time of dox removal (Supplementary Figure 4A). However, these did not resemble ES cell colonies and did not continue to proliferate once dox was withdrawn (Supplementary Figure 4B). Dox-independent, GFP positive colonies consistently appeared on plates treated with dox for 16 days or longer suggesting that some cells on those plates could proliferate and self-renew independently of viral transgene expression. Consistent with our results from the AP stainings displayed in Figure 5A, continued dox-treatment allowed more cells to reach this transgene-independent state of the reprogramming process (Figure 5B). Together, these results suggest that transgene expression is critical for the formation of iPS cells at least up to day 12-16, and that the continued expression of the transgenes beyond this time point increases the number of iPS cell colonies. Shorter

transgene expression yields cells that are AP and SSEA1 positive, as evident by our FACS sorting experiments, but still depend on transgene expression for their progress towards a fully reprogrammed iPS cell state.

Differentiation of iPS cells requires downregulation of transgene expression

To test whether continuous transgene expression affected the differentiation potential of the derived iPS lines, we compared three iPS cell lines derived using inducible lentiviruses to iPS cell lines obtained after transduction of MEFS with the four cDNAs cloned into the constitutively expressed FUW lentiviral vector (Lois et al., 2002). In this vector, the cDNAs are expressed from the human Ubiquitin C promoter that is ubiquitously expressed in mouse cells (Lois et al., 2002). We injected $5x10^5$ cells of three iPS lines harboring inducible vectors (NS12, NS14, and NS16) and of five lines harboring constitutive vectors subcutaneously into SCID mice. We used wt ES cells and ES cells harboring the R26-M2rtTA allele as controls. Figure 6 shows representative sections of ES-cell and iPS-derived tumors that were stained for Oct4 expression. Whereas the tumors obtained from the control ES cells and the inducible iPS cells displayed similar states of differentiation, the tumors generated from iPS cell lines harboring constitutively expressing vectors showed virtually no differentiated cells. Most if not all cells in the FUW-derived iPS cell tumors stained strongly for Oct4. This was in contrast to the controls and the tumors obtained after injection of iPS cells harboring inducible vectors, where Oct4 was only expressed in very

few cells throughout the tumor. These results suggest that continued ectopic expression of the four transcription factors ablates the differentiation capacity of iPS cell lines.

DISCUSSION

In this study, we used an inducible, lentiviral system to generate GFP positive, pluripotent iPS cells from MEFs derived from Oct4-GFP/R26-M2rtTA and Nanog-GFP/R26-M2rtTA knock-in mice. We established iPS cell lines that were able to differentiate in teratoma assays and could produce viable chimeras. This is in contrast to iPS cells harboring constitutively expressing lentiviral constructs that are not efficiently silenced in iPS cells. The inducible system allowed us to investigate the timing and sequence of ES cell marker gene activation using FACS analysis. We found direct in vitro reprogramming to be a gradual process encompassing the sequential activation of four pluripotency marker genes, with AP being expressed first on day 3 followed by SSEA1 on day 9. GFP driven off of the endogenous promoters of Oct4 or Nanog was first detectable on day 16 (Figure 7). SSEA1-/GFP- cells were not able to activate GFP expression in the same time frame as SSEA1+/GFP- cells and the progress from the SSEA1+/GFP- state to the dox-independent, fully reprogrammed SSEA1+/GFP+ state was greatly enhanced by continued transgene expression. Together, these results indicate that SSEA1+/GFP- cells might represent a transgene-dependent

intermediate state of reprogramming. Most of these partially reprogrammed cells require continued transgene expression in order to progress towards a fully reprogrammed state, whereas transgene downregulation causes them to stop proliferating and to return to a fibroblast-like, SSEA1-negative state. Finally, we demonstrated that the generation of iPS cells requires the ectopic expression of the four transcription factors for a minimum of 12-16 days (Figure 7). The removal of dox prior or on day 12 resulted in the return of virtually all colony-forming cells to a MEF-like state. The independence from transgene expression closely correlates with the reactivation of the endogenous Oct4 and Nanog loci (Figure 7). We postulate, therefore, that the onset of endogenous Oct4 or Nanog expression might be a marker for fully reprogrammed, transgene-independent iPS cells.

The sequential activation of pluripotency marker genes found in this study is consistent with our previous observations from MLV-based reprogramming strategies where cells at day 14 after MLV-infection expressed AP and SSEA1, but not Nanog, whereas cells had activated all three markers at day 20 (Wernig et al., 2007). Intriguingly, Nanog-neo MEFs in that study yielded neomycin resistant colonies even when the selection process was initiated at day 6 after infection, a time point well before Nanog expression could be detected by immunostaining (Wernig et al., 2007). Similar results were reported in Okita et al., 2007, and Maherali et al., 2007, where puromycin resistant colonies were derived from Nanog-GFP-IRES-puro MEFs by starting puromycin selection as

early as 3 days after infection. This discrepancy in timing between antibiotic resistance and protein detection could be explained by the hypothesis that the activation of pluripotency marker genes might be a drawn-out, gradual process, with initial expression levels sufficient to render the respective, partially reprogrammed cells resistant to antibiotics, but not high enough to warrant detection by less sensitive methods like immunostaining, GFP assays or whole population QPCR. This gradual activation seems to occur at different times for specific markers, with AP and SSEA1 being activated earlier than Oct4 or Nanog (Figure 7). It is also possible that the full activation of 'early' markers, such as AP or SSEA1, is a prerequisite for the activation of 'late' genes, such as Oct4 and Nanog. Both notions would be consistent with our finding that cells expressing SSEA1, but not Oct4- or Nanog-promoter driven GFP may represent an intermediate state of reprogramming.

We found the time span at which transgene expression becomes dispensable for iPS cell derivation (12-16 days) to precede the time point at which GFP expression indicating the activation of the endogenous Oct4 and Nanog loci could be observed (day 16). Since GFP detection by FACS requires a significant level of protein expression, it is likely that this method slightly overestimates the minimum time span required for Nanog and Oct4 activation. This slightly delayed detection could also explain our finding that SSEA1+/GFP- cell populations sorted at day 21, but not those sorted at day 9, contained a small fraction of transgene-independent cells. These cells might have activated the Oct4 or

Nanog loci but did not display GFP expression levels high enough for FACS detection. Sustained transgene expression beyond the minimal time requirement increased the number of cells activating endogenous Nanog and Oct4 expression, supporting the idea of stochastic epigenetic events playing a role in four-factor reprogramming. This is consistent with our previous observation that infected cell populations continue to generate iPS colonies over a drawn-out time window (Meissner et al., 2007). Our results suggest that individual cells either enter the reprogramming process at different time points after transgene induction or take different times to go through the reprogramming sequence. Longer transgene expression, therefore, would give more cells the chance to undergo the required stochastic epigenetic changes and, consequentially, reach a state of transgene independence.

In previous reports on iPS cell derivation from somatic cells, transgene expression was driven by the LTRs of MLV based vectors, which were shown to be efficiently silenced in iPS cells (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007). When we used constitutively expressed lentiviral vectors to generate iPS cell lines, we found these cells to be poorly capable to differentiate normally in teratoma assays. This result supports the notion that efficient transgene silencing is essential for the derivation of truly pluripotent iPS cell lines. In a recent publication, iPS cell lines generated using constitutive lentiviral vectors were reported to differentiate in teratoma assays and to contribute to various tissues of mid-gestation chimeric fetuses (Blelloch et al., 2007). The

cDNAs in that study were driven from the CMV promoter, which has been shown to undergo methylation-mediated silencing in embryonic stem cells (Hong et al., 2007; Xia et al., 2007), in contrast to the Ubiquitin C promoter employed in our constitutively expressing viral constructs (Lois et al., 2002). It can be argued, therefore, that the differentiation capabilities observed in iPS cells harboring CMV-driven cDNAs could be a result of at least partial silencing of the viral transgenes. Notably, no viable chimeras have been reported from iPS cell lines derived using constitutive lentiviral constructs so far (Blelloch et al., 2007).

Very recently, it has been reported that mouse and human iPS cells can be generated without the use of a c-Myc transgene (#add citations to Endnote#Takahashi et al., 2007: Wernig et al., 2008; Yu et al., 2007). However, the forced expression of c-Myc, as well as other factors such as Nanog and Lin28, has been found to have a substantial effect on the efficiency and the timing of the reprogramming process. While these reports are significant steps towards reducing the tumorigenic potential of iPS cells and their derivatives, the final solution to this problem will be the generation of transgene-free iPS cells. The ability to quantify reprogramming dynamics in a controlled system as presented here will be an invaluable tool for the development of transient reprogramming strategies.

In conclusion, the results presented in this report clarify steps involved in the generation of iPS cells by the expression of the four transcription factors Oct4,

Sox2, Klf4 and c-Myc. The information on pluripotency marker activation and transgene expression provided here can be utilized as a benchmark for further analyses of the reprogramming process and should allow for the identification of factors that positively affect epigenetic reprogramming. When compared to nuclear transfer, it is obvious that reprogramming by viral transduction requires a longer period of time, and it will be important to understand the molecular basis for this difference. The determination of the minimum length of transgene expression has implications for the development of non-retroviral delivery methods of these four factors to derive genetically unmodified iPS cells. Specifically, any transient expression strategy for iPS cell generation, such as protein transduction, will need to provide protein expression at sufficient levels for a minimum of 12-16 days. The generation of transient, non-viral approaches in conjunction with a better understanding of the reprogramming process will be an important step in the development of stem cell based therapies.

MATERIALS AND METHODS

Reporter cells

MEFs used in the infections with inducible lentiviruses were harvested at 13.5dpc from F1 matings between R26-M2rtTA mice (Beard et al., 2006)and either Oct4-GFP (Meissner et al., 2007) or Nanog-GFP mice (Maherali et al., 2007). MEFs

used in the infections with constitutive lentiviruses were derived from matings of mice carrying an Oct4-neomycin reporter allele (Wernig et al., 2007) and expressed a p53 hairpin (Ventura et al., 2004).

Viral Constructs

To generate the tetracycline inducible viral backbone the human Ubiquitin C promoter and intron were excised from FUW (Lois et al., 2002) by first digesting with PacI, followed by 3' overhang removal with Klenow, and subsequent digesting with BamHI. The Tet Operator/CMV promoter was excised from pRevTRE (Clontech) by digesting first with BglII followed by fill-in followed by BamHI digest. The TetO/CMV promoter and FUW backbone were then ligated. The cDNAs for Oct4, Klf-4, Sox2 and c-Myc were subsequently cloned into the EcoRI sites of the resulting vector. For the constitutive constructs, cDNAs were cloned into the EcoRI site of FUW.

Cell Culture

Cells were cultured in standard ES-cell culture conditions in DMEM supplemented with 10% FBS and LIF as described previously (Beard et al., 2006).

Infections

Virus was prepared as previously described. Briefly, 293 cells were transfected with a mixture of viral plasmid and packaging constructs expressing the viral

packaging functions and the VSV-G protein. Medium was replaced 24hrs after transfection and viral supernatants were collected at 48hrs and 72hrs. After filtration, supernatants were pooled and MEFs were incubated with viral supernatants and fresh media at a ratio of 1:1 for 24 hours and subsequently cultured in ES-cell medium.

Blastocyst Injections

Injections of iPS cells into Balb/c host blastocysts were carried out as previously described (Beard et al., 2006).

Antibodies

For FACS analysis we used an APC conjugated anti-mouse SSEA1 (R&D systems, Minneapolis, MN) and an alkaline phosphatase substrate kit: Vector Red substrate kit (Vector Laboratories, Burlingame, CA).

FACS

Cells were trypsinized, washed once in PBS and resuspended in FACS buffer (PBS+5%CCS). 10^6 cells were stained with 10ul of anti-SSEA1 antibody for 30 minutes and for the stains that included the AP substrate, cells were then washed once in PBS and fixed/permeabilized using an intracellular staining kit (R&D systems, Minneapolis, MN). After permeabilization, cells were treated with 500ul of AP substrate solution (Vector Laboratories, Burlingame, CA) for 20 minutes. Cells were then washed once with wash buffer and resuspended in

FACS buffer for analysis on a FACS-calibur cell sorter. For live sorting, cells were stained with an APC-conjugated antibody against SSEA1 (R&D systems, Minneapolis, MN) and sorted on a FACS-Aria cell sorter.

Teratoma assay

Cells were trypsinized and $5x10^5$ cells were injected subcutaneously into SCID mice. After 14-21 days, teratomas were dissected, fixed in 10% phosphate-buffered formalin overnight and subsequently embedded in paraffin wax using a Tissue-Tek VIP embedding machine (Miles Scientific, Naperville, IL) and a Thermo Shandon Histocenter 2 (Thermo Fisher Scientific, Waltham, MA). Sections were cut at a thickness of 2 μm using a Leica RM2065 (Leica, Wetzlar, Germany) and stained with hematoxylin and eosin or antibodies against Oct4 as previously described(Hochedlinger et al., 2005).

Quantitative RT-PCR

Total RNA was isolated using Trizol reagent (Invitrogen, Carlsbad, CA). Five micrograms of total RNA was treated with DNase I to remove potential contamination of genomic DNA using a DNA Free RNA kit (Zymo Research, Orange, CA). One microgram of DNase I-treated RNA was reverse transcribed using a First Strand Synthesis kit (Invitrogen) and ultimately resuspended in 100 ul of water. Quantitative PCR analysis was performed in triplicate using 1/50 of the reverse transcription reaction in an ABI Prism 7000 (Applied Biosystems, Foster City, CA) with Platinum SYBR green qPCR SuperMix-UDG with ROX

(Invitrogen). Primers used for amplification were as follows: Oct4 F, 5_-ACATCGCCAATCAGCTTGG-3_ and R, 5_-AGAACCATACTCGAACCACATCC-3_; c-myc F, 5'-CCACCAGCAGCGACTCTGA-3' and R, 5'-TGCCTCTTCTCCACAGACACC-3'; Klf4 F, 5'-GCACACCTGCGAACTCACAC-3' and R, 5'-CCGTCCCAGTCACAGTGGTAA-3'; Sox2 F, 5'-ACAGATGCAACCGATGCACC-3' and R, 5'-TGGAGTTGTACTGCAGGGCG-3'. To ensure equal loading of cDNA into RT reactions, GAPDH mRNA was amplified using the following: F, 5_-TTCACCACCATGGAGAAGGC-3_; and R, 5_-CCCTTTTGGCTCCACCCT-3_. Data were extracted from the linear range of amplification. All graphs of qRT-PCR data shown represent samples of RNA that were DNase treated, reverse transcribed, and amplified in parallel to avoid variation inherent in these procedures.

ACKNOWLEDGEMENTS

The authors thank Glenn Paradis for critical evaluation of the FACS data and help with the autofluorescence correction, Jessica Dausman and Ruth Flannery for their help with the generation of chimeras, Dongdong Fu for the immunostaining of the teratomas, Caroline Beard, Laurie Boyer, Alexander Meissner, Krishanu Saha, Frank Soeldner and Betty Zhou for critical review of the manuscript. Rudolf Jaenisch was supported by NIH grants RO1-HD045022 and R37-CA084198 and a grant from the Ellison Foundation.

REFERENCES

Alberti, S., Parks, D. R., and Herzenberg, L. A. (1987). A single laser method for subtraction of cell autofluorescence in flow cytometry. Cytometry *8*, 114-119.

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

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Blelloch, R., Venere, M., Yen, J., and Ramalho-Santos, M. (2007). Generation of Induced Pluripotent Stem Cells in the Absence of Drug Selection. Cell Stem Cell *1*, 245-247.

Cowan, C. A., Atienza, J., Melton, D. A., and Eggan, K. (2005). Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. Science *309*, 1369-1373.

Hochedlinger, K., and Jaenisch, R. (2003). Nuclear transplantation, embryonic stem cells, and the potential for cell therapy. N Engl J Med *349*, 275-286.

Hochedlinger, K., Yamada, Y., Beard, C., and Jaenisch, R. (2005). Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. Cell *121*, 465-477.

Hong, S., Hwang, D. Y., Yoon, S., Isacson, O., Ramezani, A., Hawley, R. G., and Kim, K. S. (2007). Functional analysis of various promoters in lentiviral vectors at different stages of in vitro differentiation of mouse embryonic stem cells. Mol Ther *15*, 1630-1639.

Kato, Y., Tani, T., Sotomaru, Y., Kurokawa, K., Kato, J., Doguchi, H., Yasue, H., and Tsunoda, Y. (1998). Eight calves cloned from somatic cells of a single adult. Science *282*, 2095-2098.

Lois, C., Hong, E. J., Pease, S., Brown, E. J., and Baltimore, D. (2002). Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. Science *295*, 868-872.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R.*, et al.* (2007). Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. Cell Stem Cell *1*, 55-70.

Meissner, A., Wernig, M., and Jaenisch, R. (2007). Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. Nat Biotechnol *25*, 1177-1181.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature *448*, 313-317.

Polejaeva, I. A., Chen, S. H., Vaught, T. D., Page, R. L., Mullins, J., Ball, S., Dai, Y., Boone, J., Walker, S., Ayares, D. L.*, et al.* (2000). Cloned pigs produced by nuclear transfer from adult somatic cells. Nature *407*, 86-90.

Rideout, W. M., 3rd, Hochedlinger, K., Kyba, M., Daley, G. Q., and Jaenisch, R. (2002). Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy. Cell *109*, 17-27.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr Biol *11*, 1553-1558.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

#Takahashi et al, c-Myc paper citation needs to be added to Endnote

Ventura, A., Meissner, A., Dillon, C. P., McManus, M., Sharp, P. A., Van Parijs, L., Jaenisch, R., and Jacks, T. (2004). Cre-lox-regulated conditional RNA interference from transgenes. Proc Natl Acad Sci U S A *101*, 10380-10385.

Wakayama, T., Perry, A. C., Zuccotti, M., Johnson, K. R., and Yanagimachi, R. (1998). Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. Nature *394*, 369-374.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B. E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.

#Wernig et al., c-Myc paper needs to be added to Endnote.

Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., and Campbell, K. H. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810-813.

Xia, X., Zhang, Y., Zieth, C. R., and Zhang, S. C. (2007). Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner. Stem Cells Dev *16*, 167-176.

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R.*, et al.* (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. Science.

FIGURE LEGENDS

Figure 1: Derivation of iPS cells using an inducible lentiviral system.

Somatic cells harboring a GFP reporter driven by the endogenous Oct4 or Nanog promoters were infected with tet-inducible lentiviral vectors carrying the cDNAs of Oct4, Sox2, Klf4 and c-Myc (A). The addition of doxycycline after infection induced lentiviral expression and subsequently led to reprogrammed, GFP positive iPS cell colonies (B). The iPS cells were pluripotent and contributed to viable chimeras after injection into Balb/c host blastocysts as indicated by coat

color (C, D). White mice are non-chimeric Balb/c animals, whereas mixed coat color mice are chimeric.

Figure 2: Timing of pluripotency marker reactivation during reprogramming.

FACS analysis of AP, SSEA1 and GFP reactivation was performed on NanogGFP/M2rtTA MEFs and Oct4GFP/M2rtTA MEFs at different times after the induction of reprogramming. Cells were harvested at various time points after the addition of dox to the ES cell medium and incubated with an APC labeled anti-SSEA1 antibody and a fluorescent substrate that is activated in the presence of AP activity. The data from three independent experiments is summarized. Columns display average percentages of cells expressing AP (yellow columns), SSEA1 (blue columns) and Nanog-GFP (black columns) or Oct4-GFP (white columns). AP and SSEA1 values represent averages of Nanog-GFP and Oct4-GFP experiments. Error bars indicate standard deviations.

Figure 3: SSEA1 expression marks an intermediate state of reprogramming.

SSEA1+/GFP- cells were sorted with an APC conjugated SSEA1-antibody after 21 days of transgene expression and seeded on feeder MEFs (A). GFP positive ES-like cell colonies were observed ten days after sorting (B). Supplementation of the media with dox for an additional three days after sorting increased the percentage of SSEA1+/GFP- cells that gave rise to GFP positive iPS cell

colonies (C). Confirmation of GFP-positive cells in the sorted SSEA1+/GFP- populations eleven days after initial sorting and seeding on feeder MEFs (D).

Figure 4: SSEA1+/GFP- cells return to a MEF-like state upon dox withdrawal

Oct4GFP/M2rTTA MEFs undergoing reprogramming were FACS sorted for SSEA1 expression after 9 days of transgene expression (A). The cells were cultured either in the presence or absence of dox for 20 days after the intial sorting and then re-analyzed for SSEA1 and GFP expression. While plates that were treaed with dox contained SSEA1+/GFP+ and SSEA1+/GFP- cells, virtually all cells on untreated plates had lost their SSEA1 expression (A). After culture, Culture in the presence of dox yielded ES-like colonies, but in the absence of dox the SSEA1+/GFP- cells returned to a MEF like morphology (B).

Figure 5: Time requirement of transgene expression for iPS derivation.

Infected MEFs were split into 6 dishes and cultured in the presence of dox for a different number of days as indicated by numbers next to the red lines (A). Dox was withdrawn at different time points as indicated by red bars and the cells were subsequently cultured in the absence of the drug until day 35. On day 35 all plates were stained for AP at ay 35 of reprogramming to determine how many transgene-independent cells existed at the time point of dox withdrawal. No AP positive colonies were observed when dox was withdrawn before day 12 and the

number of AP positive colonies increased with the time the cells were exposed to the drug (A). Similarly, infected MEFs were grown in the presence of dox and dox was removed at different time points from the culture medium. GFP expressing iPS colonies were counted on day 35. Data for Nanog-GFP and Oct4-GFP experiments were averaged and the number of colonies detected in six wells (3 for Oct4-GFP and 3 for Nanog-GFP cells) for each time point is shown in the bar graph (B). Error bars indicate standard deviation.

Figure 6: Constitutive expression of the four factors prevents differentiation of iPS cells in teratomas.

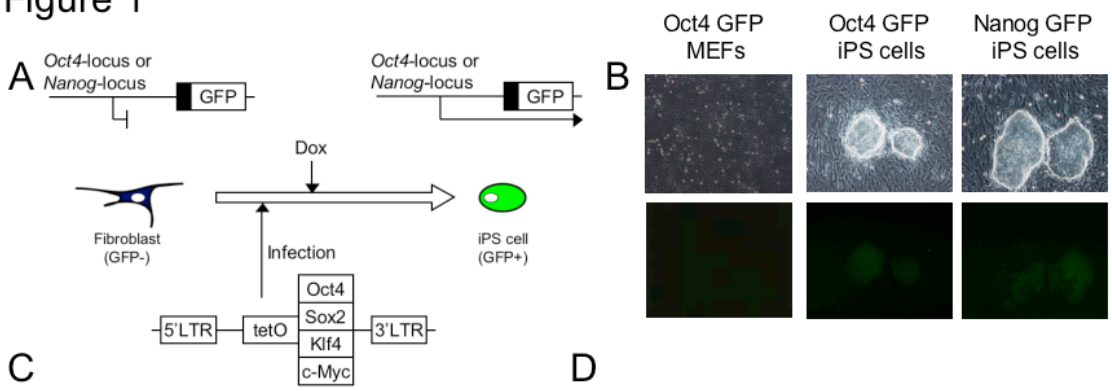
Differentiation of iPS cells in teratomas was ablated by continuous expression of transgenes in iPS cells. Tumors derived from iPS cells expressing the four transgenes under the control of the constitutive Ubiquitin C promoter stained homogeneously positive for Oct4 and did not show normal cell differentiation (A). IPS cells harboring the inducible viral backbones differentiated readily after downregulation of the transgenes in the absence of dox (B). The tumors derived from these cells were similar in size, fraction of cells staining positive for Oct4 and differentiation to tumors derived from wildtype control ES cell lines (C) or ES cells harboring the R26-M2rtTA allele (D).

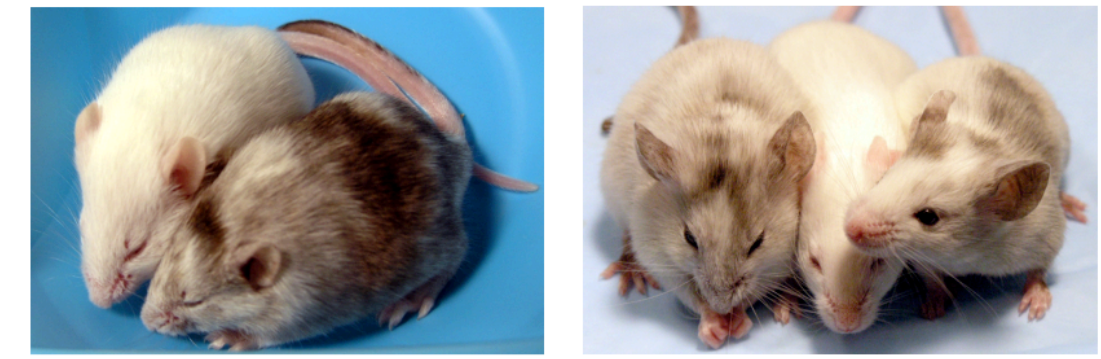
Figure 7: A timeline of reprogramming dynamics.

The black arrow represents the time starting from infection (above) or dox-induction (below) in one-day increments. For MLV infection, the earliest time

points for two antibiotic selection strategies are indicated by blue bars and the expression of Nanog-GFP and Oct4-GFP is indicated by green bars. For dox-inducible transgenes, the minimum time of dox-induced transgene expression for successful iPS cell derivation is displayed as a black bar. The expression timing of four pluripotency marker genes is indicated by red (AP and SSEA1) and green bars (Nanog-GFP and Oct4-GFP)

Figure 1





Nanog-GFP iPS Chimera

Oct4-GFP iPS Chimeras

Figure 2

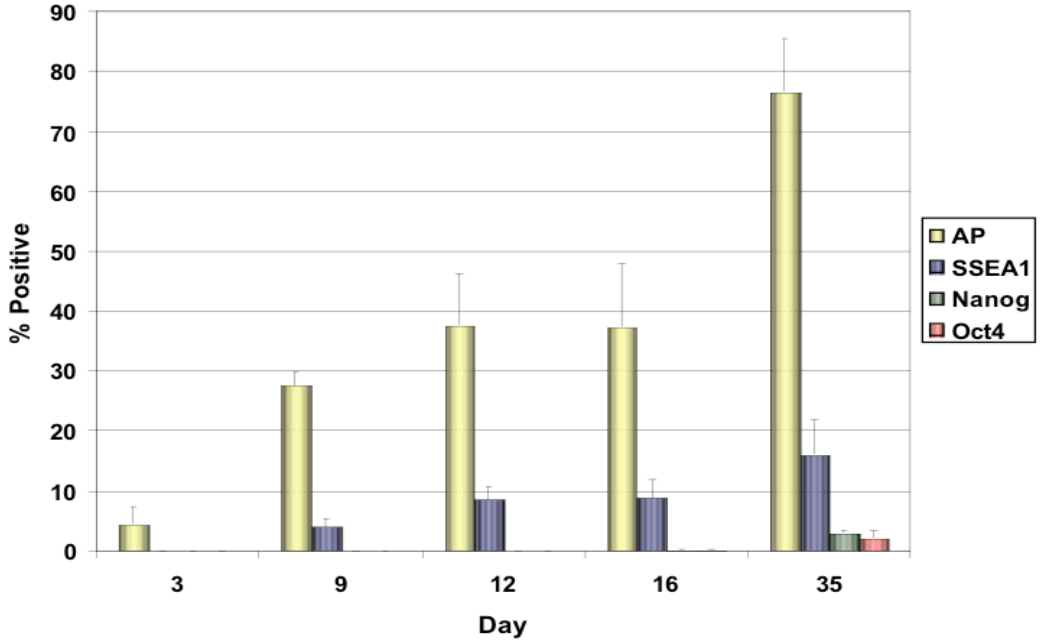

90
80
70
60
50
40
30
20
10
0
% Positive
3
9
12
16
35
Day
AP
SSEA1
Nanog
Oct4

Figure 3

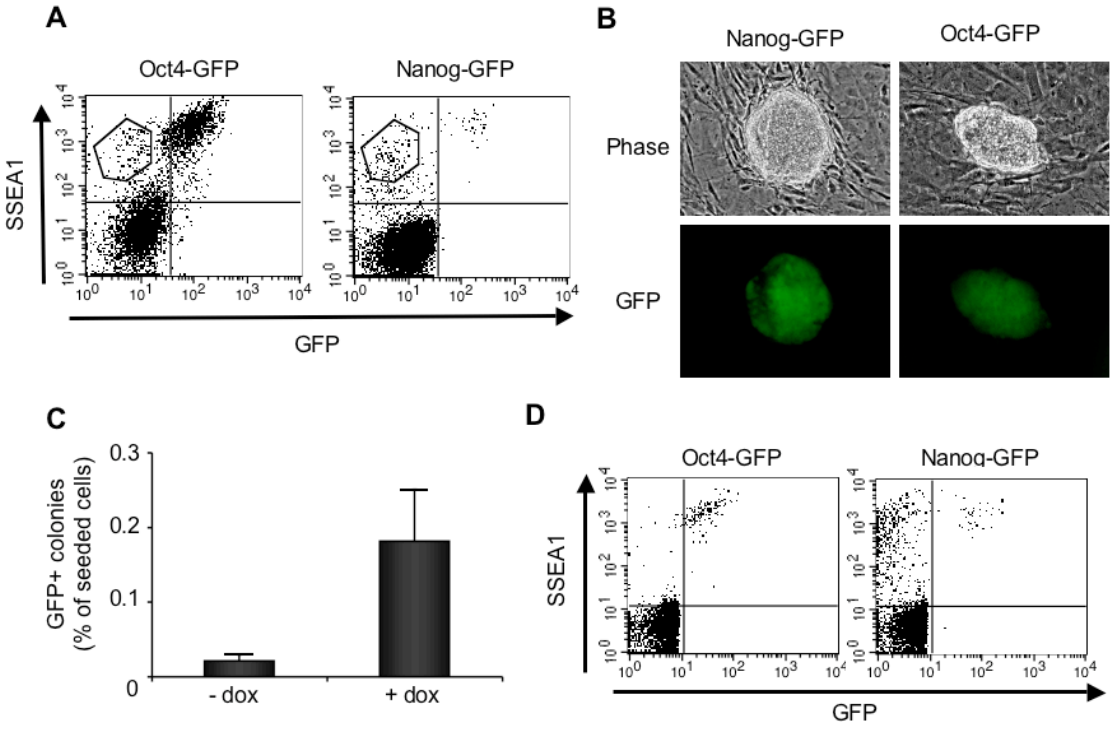

A
Oct4-GFP
Nanog-GFP
SSEA1
GFP
B
Nanog-GFP
Oct4-GFP
Phase
GFP
C
GFP+ colonies (% of seeded cells)
0.3
0.2
0.1
0
- dox
+ dox
D
Oct4-GFP
Nanog-GFP
SSEA1
GFP

Figure 4

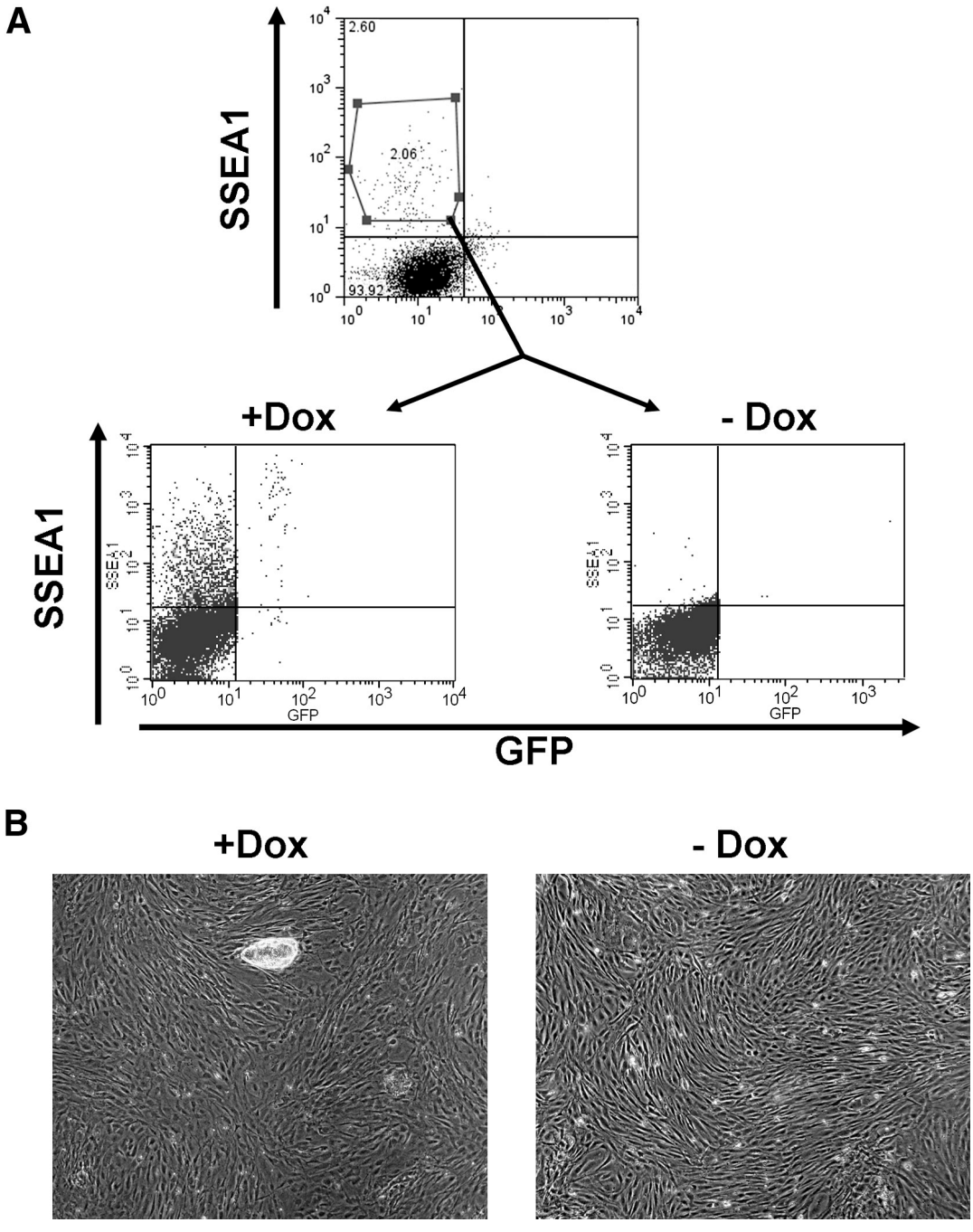


Figure 5

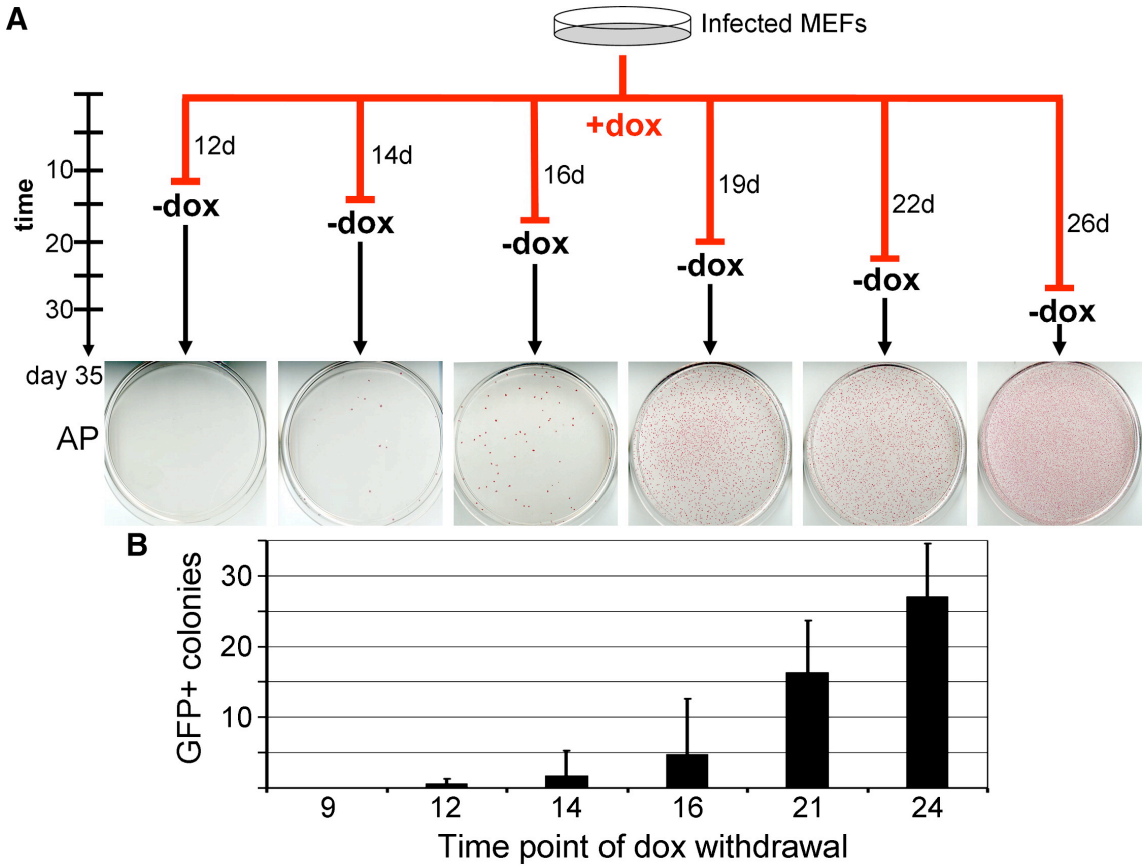

A
Infected MEFs
+dox
12d
14d
16d
19d
22d
26d
-dox
time
10
20
30
day 35
AP
B
GFP+ colonies
30
20
10
9
12
14
16
21
24
Time point of dox withdrawal

Figure 6

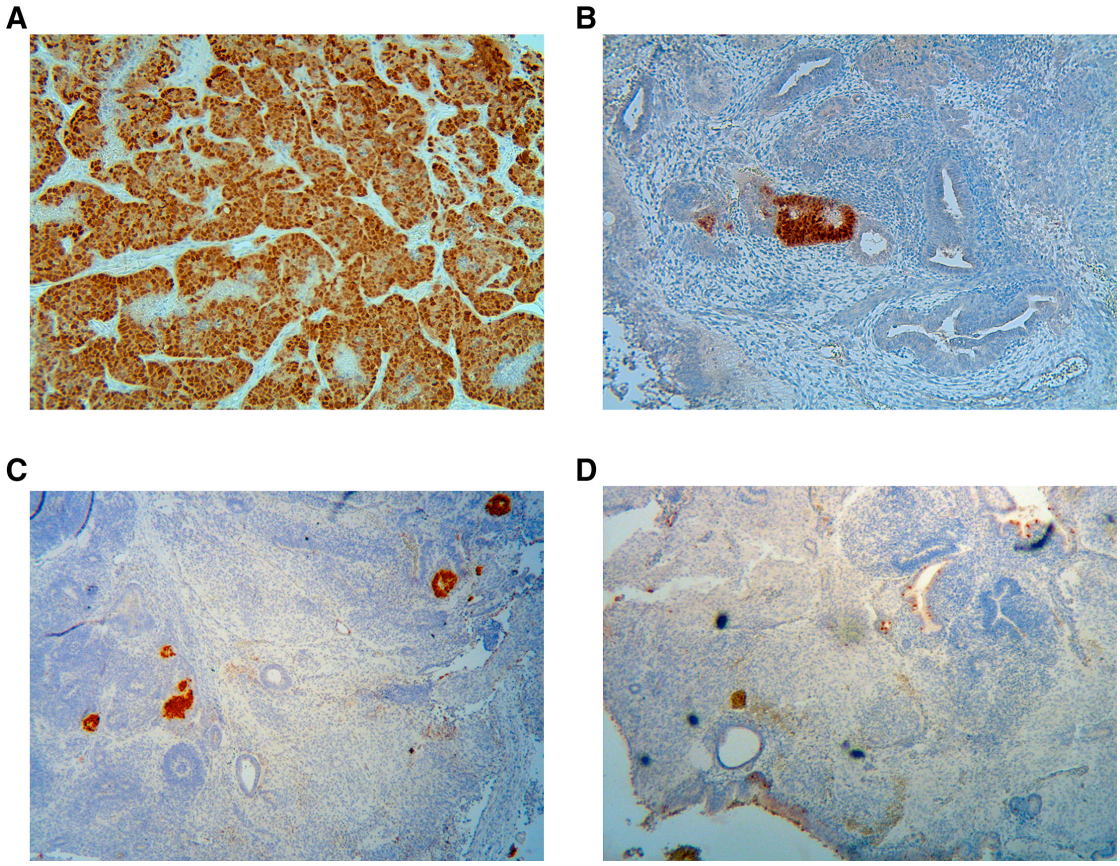
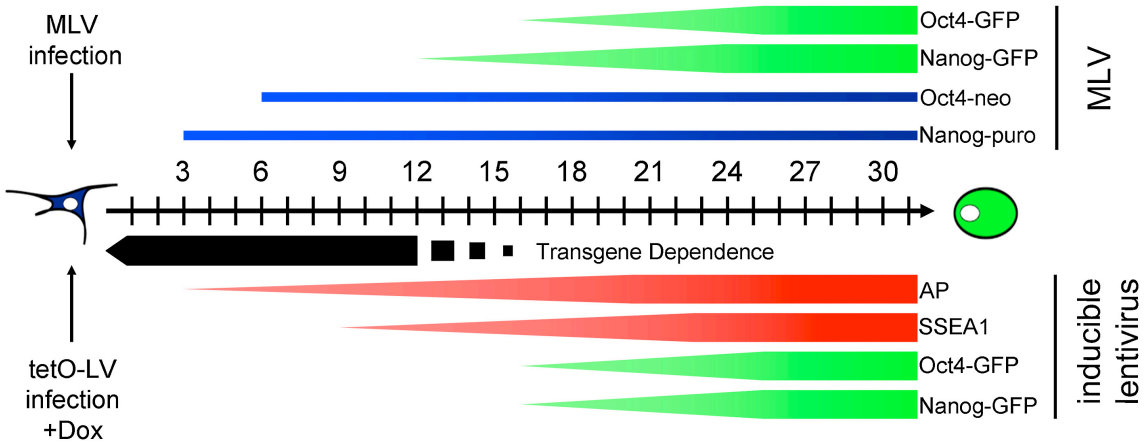


Figure 7


MLV
infection
Oct4-GFP
Nanog-GFP
Oct4-neo
Nanog-puro
MLV
3
6
9
12
15
18
21
24
27
30
Transgene Dependence
AP
SSEA1
Oct4-GFP
Nanog-GFP
inducible
lentivirus
tetO-LV
infection
+Dox

Chapter 4

Wnt signaling promotes reprogramming of somatic cells to pluripotency

Alexander Marson*, Ruth Foreman*, Brett Chevalier*, Steve Bilodeau, Michael Kahn, Richard A. Young‡, Rudolf Jaenisch‡

**These authors contributed equally to this work*

‡Corresponding authors

This chapter was initially published in Cell Stem Cell, *3* 132-135. Permission to publish it in this thesis was granted by Elsevier.

Author contributions: RY suggested that AM examine the role of Wnt signaling in iPS cell generation given the work going on in his lab at the time (and eventually published in Cole et al., 2008). AM approached RF about a collaboration. AM, RF and BC conducted the cell culture experiments. RF performed blastocyst injections.

Somatic cells can be reprogrammed to induced pluripotent stem (iPS) cells by retroviral transduction of four transcription factors, Oct4, Sox2, Klf4 and c-Myc (Takahashi et al., 2006; Takahashi et al., 2007; reviewed in Jaenisch and Young, 2008). While the reprogrammed pluripotent cells are thought to have great potential for regenerative medicine (Hanna et al., 2007; Wernig et al., 2008b), genomic integrations of the retroviruses, especially c-Myc, increase the risk of tumorigenesis (Okita et al., 2007). Recently, iPS cells have been generated without c-Myc retrovirus ($Myc^{[-]}$), but in the absence of exogenous c-Myc the efficiency and kinetics of reprogramming are significantly reduced (Nakagawa et al., 2008; Wernig et al., 2008a). We report here that soluble Wnt3a promotes the generation of iPS cells in the absence of c-Myc retrovirus. These data demonstrate that signal transduction pathways and transcription factors can act coordinately to reprogram differentiated cells to a pluripotent state.

Naturally occurring signaling molecules that modulate the expression of endogenous ES cell transcription factors are promising candidates for soluble factors that enhance reprogramming. The Wnt signaling pathway contributes to the maintenance of pluripotency in mouse and human ES cells (Sato et al., 2004; Ogawa et al., 2006; Singla et al., 2006; Cai et al., 2007), as well as the self-renewal of undifferentiated adult stem cells in multiple tissues (Reya and Clevers, 2005). In addition, initial studies of iPS cell generation suggested that constitutively active b-catenin, another downstream component of the Wnt pathway, might promote reprogramming of fibroblasts to pluripotency (Takahashi

and Yamanaka, 2006). Studies of Tcf3, one of the key transcriptional regulators downstream of the Wnt pathway in embryonic stem cells, have revealed that this factor co-occupies almost all promoter regions occupied by ES cell specific transcription factors, including Oct4 and Nanog, and can regulate the expression of key ES cell transcription factors (Cole et al, 2008; Tam et al., 2008; Yi et al., 2008).

Since the Wnt pathway is intimately connected to the core circuitry of pluripotency, we hypothesized that the stimulation of the pathway using soluble factors could modulate the efficiency of inducing pluripotency in somatic cells. Here we examined the influence of Wnt stimulation on the reprogramming of murine fibroblasts in the absence of c-Myc retrovirus. For this, cells with GFP driven by the endogenous *Oct4* promoter (Meissner et al., 2007) were infected with doxycycline (DOX)-inducible lentiviruses encoding Oct4, Sox2 and Klf4. Oct4/Sox2/Klf4 infected cells were either cultured in standard ES cell medium or in Wnt3a-conditioned medium (Wnt3a-CM) and were analyzed for GFP expression by flow cytometry at days 10, 15 and 20 after DOX induction. No GFP positive cells were present with or without Wnt3a-CM treatment on day 10 or day 15. By day 20 a small population of GFP expressing cells was detected only in the cells cultured in Wnt3a-CM (**Figure 1a**). The Wnt3a-CM exposed cultures formed GFP expressing colonies with morphology typical for ES or iPS cells (**Figure 1b**). However, unlike four-factor-transduced cells, which usually form a highly heterogeneous population of cells when propagated without selection, the Oct4/Sox2/Klf4/Wnt3a-CM colonies appeared homogenously ES-like, similar to previously reported Myc$^{[-]}$ iPS clones (Nakagawa et al., 2008).

Several assays were performed to characterize the developmental potential

of Myc$^{[-]}$ iPS cells derived with Wnt3a-CM treatment. Immunocytochemistry confirmed the expression of markers of pluripotency, including the nuclear factor Nanog and the surface glycoprotein SSEA1 (**Figure 1b**). Functional assays confirmed that, like ES cells, these iPS cells were pluripotent. When injected into SCID mice subcutaneously, the Myc$^{[-]}$ iPS cells gave rise to teratomas with histological evidence of cells differentiating into all three germ layers (**Figure 1c**). More importantly, Myc$^{[-]}$ iPS cells derived with Wnt3a-CM treatment contributed to the formation of somatic tissues as well as the germline in chimeric mice (**Figures 1d and 1e and Supplemental Table S1**). These results indicate that Wnt3a-CM treated Myc$^{[-]}$ clones are pluripotent cells that are morphologically and functionally indistinguishable from ES cells.

To quantify the effects of Wnt3a-CM, triplicate experiments were performed on Oct4/Sox2/Klf4-inducible MEFs carrying a G418 resistance cassette downstream of the *Oct4* promoter (**Figure 2a)**. G418 was added to the cultures at 15 days after infection to select for cells that had reactivated the Oct4 locus. When scored on day 28 after infection, only a few Myc$^{[-]}$ G418 resistant colonies (between 0-3 colonies forming on each ten centimetre plate) were detected in standard ES cell culture conditions. In contrast, ~20 fold more drug resistant colonies formed when G418 selection was initiated on Wnt3a-CM-treated cells, consistent with the conclusion that activation of the Wnt pathway enhances reprogramming. It should be noted that conditioned medium from control fibroblasts (L cells, ATCC) lacking Wnt3a over-expression also caused a moderate increase in the number of G418-resistant colonies relative to standard ES medium, suggesting that normal fibroblasts may secrete factors, perhaps including Wnt3a, that promote reprogramming. So far, we have been unable to

recapitulate the effects of Wnt3a-CM on reprogramming with small molecule inhibitors of glycogen synthase kinase-3 (Gsk3) that modulate the Wnt pathway and promote the self-renewal of ES cells (Sato et al., 2004; Ying et al., 2008) (**Supplementary Figure S1**). While there are many plausible explanations for this negative result, one possibility is that the half-life of these chemicals is too short to provide sustained stimulation of the Wnt-pathway required to enhance reprogramming.

To independently assess the effect of Wnt3a on reprogramming, we cultured cells in the presence of ICG-001 (Teo et al., 2005; McMillan and Kahn, 2005), an inhibitor of the Wnt/ß-catenin pathway. **Figure 2a** (right columns) shows that 4μM ICG-001 strongly inhibited the effect of Wnt3a-CM on Myc$^{[-]}$ iPS formation. The effects of Wnt3a-CM and ICG-001 were also examined in MEFs over-expressing all four reprogramming factors, including c-Myc (**Figure 2b**). High numbers of G418 resistant colonies were observed in both standard ES cell media (66 colonies/32cm^2) and Wnt3a-CM in four-factor-reprogrammed cells, with only a subtle increase in the number of colonies with Wnt3a-CM (80 colonies/32cm^2). In contrast to the dramatic effect of ICG-001 on Myc$^{[-]}$ cells, at the same dose, the compound had only a subtle effect on the number of G418 colonies in c-Myc transduced cells, and a relatively high number of resistant colonies was observed under these conditions (51 colonies/32cm^2) (**Figure 2b**). When the experiment was repeated with higher doses of ICG-001, iPS colony numbers were further reduced, but even at 25μM multiple Oct4/Sox2/Klf4/c-Myc iPS colonies were observed (mean= 8 colonies/32cm^2) (**Supplemental Figure S2**). These data suggest that Wnt signaling and *c-Myc* could have overlapping, but not completely redundant roles in the induction of pluripotency.

Recent reports have demonstrated that Myc is dispensable for inducing pluripotency in MEFs (Nakagawa et al., 2008; Wernig et al., 2008a), establishing Oct4, Sox2 and Klf4 as a core set of reprogramming factors, although a slightly different combination of factors also functions in human cells (Yu et al., 2007). Epigenetic reprogramming of MEFs with defined factors is a gradual process, where transcription factors progressively re-establish the core circuitry of pluripotency over the course of weeks (Brambrink et al, 2008; Stadtfeld et al., 2008). Myc accelerates this process (Nakagawa et al., 2008; Wernig et al., 2008a), conceivably because it enhances cell proliferation allowing the changes to unfold more rapidly. Additionally, Myc is suspected to have widespread effects on chromatin state (Knoepfler, 2008; Kim et al., 2008) and could facilitate the productive binding of the core reprogramming factors to their appropriate genomic targets (Jaenisch and Young, 2008).

The Wnt signaling pathway has been shown to connect directly to the core transcriptional regulatory circuitry of ES cells, suggesting a mechanism by which this pathway could directly promote the induction of pluripotency in the absence of c-Myc transduction (**Figure 2c)**. In ES cells, Tcf3 occupies and regulates the promoters of *Oct4*, *Sox2* and *Nanog* (Cole et al., 2008; Tam et al., 2008; Yi et al., 2008). In MEFs, these endogenous pluripotency transcription factors are silenced. During reprogramming, as exogenous Oct4, Sox2 and Klf4 contribute to the reactivation of the endogenous pluripotency factors (Jaenisch and Young, 2008), Wnt signaling could directly potentiate the effect of these transcription factors, as it does in ES cells (Cole et al., 2008).

Additionally, Wnt could serve to activate endogenous *c-Myc* directly, thereby substituting for exogenous c-Myc (**Figure 2c)**. Indeed, *c-Myc* is a well-

established target of the Wnt pathway in colorectal cancer cells (He et al., 1998). In ES cells, Tcf3 occupies the *c-Myc* promoter, and Wnt3a positively contributes to expression of the gene (Cole et al., 2008). However, quantitative PCR studies did not reveal Wnt3a-CM-dependent activation of *c-Myc* during the first 48 hours after induction of Oct4/Sox2/Klf4 in MEFs (**Supplementary Figure S3**). Although we cannot exclude that Wnt3a enhances *c-Myc* expression in the rare cells that will eventually become iPS cells, the expression data suggest that the role of Wnt in reprogramming could be independent of *c-Myc* induction, consistent with recent findings in ES cells (Ying et al., 2008). Wnt signaling and c-Myc could play partially redundant functional roles in iPS cell generation. For example, Wnt-induced effects on cell proliferation could help to accelerate the sequence of events that lead to the generation of $Myc^{[-]}$ iPS colonies.

A major goal of current research is to identify transient cues that can reprogram somatic cells, eliminating the need for retroviruses. The studies described here establish that Wnt stimulation can be used to enhance the efficiency of reprogramming in combination with nuclear factors, Oct4, Sox2 and Klf4. By enhancing the efficiency of reprogramming in the absence of c-Myc retrovirus, soluble Wnt, or small molecules that modulate the Wnt signaling pathway, will likely prove useful in combination with other transient cues, yet to be identified, that can replace the remaining retroviruses.

Acknowledgements

We thank members of the Jaenisch, Young and Kahn laboratories, especially Jamie Newman, Megan Cole and Tony Lee for discussions and critical review of the manuscript. We are grateful to Jamie Newman, Megan Cole, and Sarah Johnstone for sharing data prior to publication and for Wnt3a conditioned media, and to Seshamma Reddy, Jessica Dausman, Ruth Flannery and Betty Zhou for technical assistance. We also thank Tom DiCesare for graphic assistance. This work was supported by NIH grants 5-RO1-HDO45022, 5-R37-CA084198, and 5-RO1- CA087869 to RJ and by NIH grant HG002668 and a grant from the Whitehead Institute to RY. RJ is an advisor to Stemgent which has obtained a license from MIT to distribute some of the reagents used in this paper.

References

Brambrink, T., Foreman, R., Welstead, G. G., Lengner, C. J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell *2*, 151-159.

Cai, L., Ye, Z., Zhou, B. Y., Mali, P., Zhou, C., and Cheng, L. (2007). Promoting human embryonic stem cell renewal or differentiation by modulating Wnt signal and culture conditions. Cell Res *17*, 62-72.

Cole, M. F., Johnstone, S. E., Newman, J. J., Kagey, M. H., and Young, R. A. (2008). Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes and Development *in press*.

Hanna, J., Wernig, M., Markoulaki, S., Sun, C. W., Meissner, A., Cassady, J. P., Beard, C., Brambrink, T., Wu, L. C., Townes, T. M., and Jaenisch, R. (2007). Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. Science *318*, 1920-1923.

He, T. C., Sparks, A. B., Rago, C., Hermeking, H., Zawel, L., da Costa, L. T., Morin, P. J., Vogelstein, B., and Kinzler, K. W. (1998). Identification of c-MYC as a target of the APC pathway. Science *281*, 1509-1512.

Jaenisch, R., and Young, R. A. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell *132*, 567-582.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S. H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049-1061.

Knoepfler, P. S. (2008). Why Myc? An unexpected ingredient in the stem cell cocktail. Cell Stem Cell *2*, 18-21.

McMillan, M., and Kahn, M. (2005). Investigating Wnt signaling: a chemogenomic safari. Drug Discov Today *10*, 1467-1474.

Meissner, A., Wernig, M., and Jaenisch, R. (2007). Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. Nat Biotechnol *25*, 1177-1181.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol *26*, 101-106.

Ogawa, K., Nishinakamura, R., Iwamatsu, Y., Shimosato, D., and Niwa, H. (2006). Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. Biochem Biophys Res Commun *343*, 159-166.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature *448*, 313-317.

Reya, T., and Clevers, H. (2005). Wnt signalling in stem cells and cancer. Nature *434*, 843-850.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A. H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat Med *10*, 55-63.

Singla, D. K., Schneider, D. J., LeWinter, M. M., and Sobel, B. E. (2006). wnt3a but not wnt11 supports self-renewal of embryonic stem cells. Biochem Biophys Res Commun *345*, 789-795.

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

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tam, W. L., Lim, C. Y., Han, J., Zhang, J., Ang, Y. S., Ng, H. H., Yang, H., and Lim, B. (2008). Tcf3 Regulates Embryonic Stem Cell Pluripotency and Self-Renewal by the Transcriptional Control of Multiple Lineage Pathways. Stem Cells.

Teo, J. L., Ma, H., Nguyen, C., Lam, C., and Kahn, M. (2005). Specific inhibition of CBP/beta-catenin interaction rescues defects in neuronal differentiation caused by a presenilin-1 mutation. Proc Natl Acad Sci U S A *102*, 12171-12176.

Wernig, M., Meissner, A., Cassady, J. P., and Jaenisch, R. (2008a). c-Myc is dispensable for direct reprogramming of mouse fibroblasts. Cell Stem Cell *2*, 10-12.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B. E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.

Wernig, M., Zhao, J. P., Pruszak, J., Hedlund, E., Fu, D., Soldner, F., Broccoli, V., Constantine-Paton, M., Isacson, O., and Jaenisch, R. (2008b). Neurons derived from reprogrammed fibroblasts functionally integrate into the fetal brain and improve symptoms of rats with Parkinson's disease. Proc Natl Acad Sci U S A *105*, 5856-5861.

Yi, F., Pereira, L., and Merrill, B. J. (2008). Tcf3 Functions as a Steady State Limiter of Transcriptional Programs of Mouse Embryonic Stem Cell Self Renewal. Stem Cells.

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

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R.*, et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-1920.

Figure Legends

Figure 1 Wnt3a promotes reprogramming of somatic cells to pluripotency. **a**. Scatter plots comparing GFP intensity to autofluoresence, using flow cytometry, in *Oct4*-GFP cells on day 20 post-induction of Oct4/Sox2/Klf4, reveal a GFP expressing population of cells (indicated with an arrow) only with Wnt3a-CM treatment (right), not in the control (Ctrl) Oct4/Sox2/Klf4-infected cells cultured in standard ES cell medium (left). **b**. Phase-contrast micrographs (Phase, left) of GFP expressing $Myc^{[-]}$ cells derived with Wnt3a-CM treatment and without any genetic selection**.** Immunostaining reveals induction of pluripotency markers, Nanog (upper) and SSEA-1(lower) in Wnt3a-CM treated $Myc^{[-]}$ cells. **c**. Wnt3a-CM treated $Myc^{[-]}$ lines formed teratomas when injected into SCID mice subcutaneously. Teratomas from Oct4/Sox2/Klf4/Wnt3a-CM iPS lines showed evidence of differentiated cells of three germ layers similar to teratomas formed from V6.5 mES injections. Arrows indicated neural tissue in (upper), cartilage in (middle), and endodermal cells in (lower), **d**. Oct4/Sox2/Klf4/Wnt3aCM iPS lines derived without selection gave rise to chimeric mice (as shown on the left) with agouti coat color and pigmented eyes (in contrast to wild type Balb/c mouse, right) providing evidence of contribution to somatic cells. **e**. Agouti coat color of offspring (left) of chimeric mouse (right) confirmed that the Oct4/Sox2/Klf4/Wnt3a-CM iPS line generated here is germline-competent.

Figure 2 Wnt/ß-catenin stimulation enahances iPS colony formation in absence of c-Myc retrovirus. a. Counts are shown for G418-resistant colonies in Oct4/Sox2/Klf4 over-expressing MEFs cultured in ES cell media, control L cell conditioned media (ATCC), Wnt3a over-expressing conditioned media (ATCC), and Wnt3a over-expressing conditioned media with ICG-001 (4μM). Selection was initiated on day 15 post-induction, and colonies were assessed on day 28. Wnt3a-CM treatment was maintained for 6-9 days after selection was initiated.

Mean number of counts from triplicate experiments is displayed with error bars indicating S.D. **b**. Counts are shown for G418-resistant colonies (in a 32cm^2 area) in Oct4/Sox2/Klf4/c-Myc over-expressing MEFs cultured in ES cell media, Wnt3a over-expressing conditioned media (ATCC), and Wnt3a over-expressing conditioned media with ICG-001 (4μM). Selection was initiated on day 10 post-induction, Wnt3a-CM was maintained for the first 6-9 days of selection, and colonies were assessed on day 20. **c**. Wnt stimulation promotes the formaton of iPS cells in the absence of c-Myc transduction. This could be due to: i) direct regulation by the Wnt pathway of key endogenous pluripotency factors, such as *Oct4*, *Sox2* and *Nanog* as suggested by genomic studies in ES cells (Cole et al., 2008), ii) Wnt pathway-induced activation of endogenous *Myc* (He et al., 1998; Cole et al., 2008), or other cell proliferation genes, accelerating the sequential process of forming iPS colonies.

Figure 1

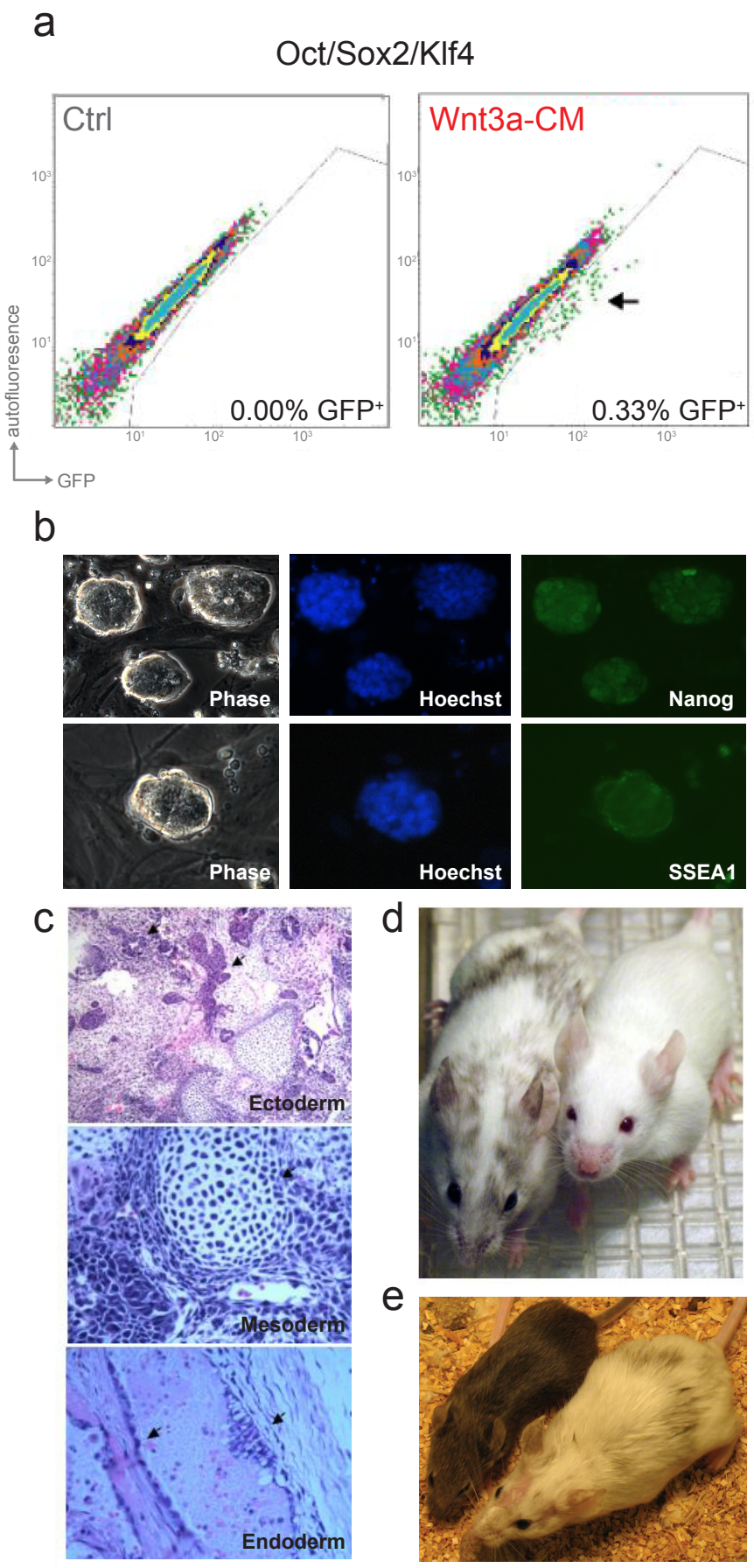
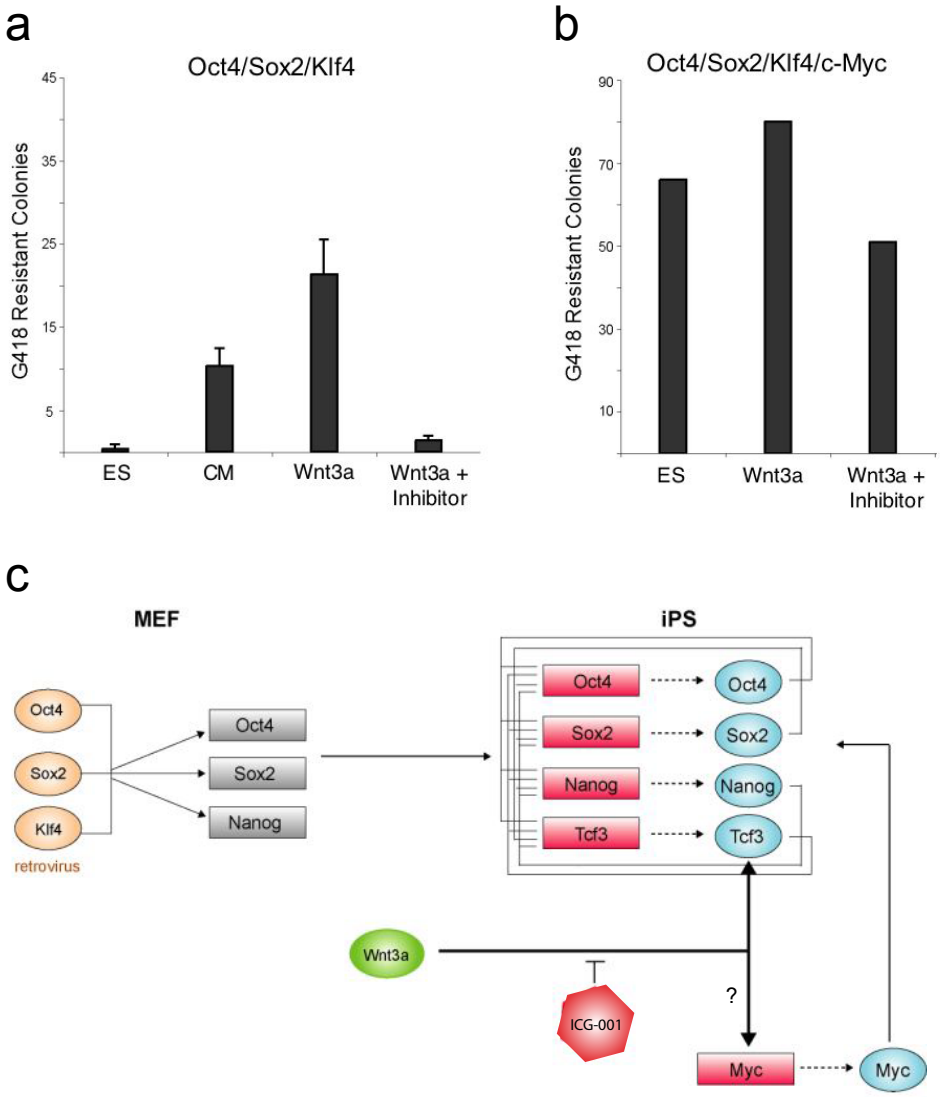

a
Oct/Sox2/Klf4
Ctrl
Wnt3a-CM
0.00% GFP+
0.33% GFP+
autofluoresence
GFP
b
Phase
Hoechst
Nanog
Phase
Hoechst
SSEA1
c
Ectoderm
Mesoderm
Endoderm
d
e

Figure 2


a
Oct4/Sox2/Klf4
G418 Resistant Colonies
45
35
25
15
5
ES
CM
Wnt3a
Wnt3a +
Inhibitor
b
Oct4/Sox2/Klf4/c-Myc
G418 Resistant Colonies
90
70
50
30
10
ES
Wnt3a
Wnt3a +
Inhibitor
c
MEF
iPS
Oct4
Sox2
Klf4
retrovirus
Oct4
Sox2
Nanog
Oct4
Sox2
Nanog
Tcf3
Oct4
Sox2
Nanog
Tcf3
Wnt3a
ICG-001
?
Myc
Myc

Chapter 5

Reprogramming of murine fibroblasts to iPS cells: Chemical Complementation of Klf4

Costas A. Lyssiotis[1], Ruth K. Foreman[1], Judith Staerk[1], Michael Garcia, Divya Mathur, Styliani Markoulaki, Jacob Hanna, Luke L. Lairson, Bradley D. Charette, Laure C. Bouchez, Michael Bollong, Conrad Kunick, Achim Brinker, Charles Cho, Peter G. Schultz[2] and Rudolf Jaenisch[2]

[1]These authors contributed equally to this work.
[2]Corresponding authors

Author contributions: CAL, RF and JS contributed to the concept, designed and performed the experiments, analyzed the data, and wrote the manuscript. MG helped in design of and ran the primary small molecule screen. DM, SM, JH, LLL, BDC, LCB and MB contributed reagents and/or provided technical assistance. CK contributed analytical tools. AB and CC contributed to the design of the screening platform. PGS and RJ provided the concept, assisted in experimental design and analysis, and revised the manuscript.

Summary

Ectopic expression of defined transcription factors can reprogram somatic cells to induced pluripotent stem (iPS) cells, but the utility of iPS cells is hampered by the use of viral delivery systems. Small molecules offer an alternative to replace virally transduced transcription factors with chemical signaling cues responsible for reprogramming. In this study we describe a small molecule screening platform that was applied to identify compounds that functionally replace the reprogramming factor Klf4. A series of small molecule scaffolds were identified that activate Nanog expression in mouse fibroblasts transduced with a subset of reprogramming factors lacking Klf4. Furthermore, application of one such molecule, kenpaullone, in lieu of Klf4, gave rise to iPS cells that are indistinguishable from murine embryonic stem cells. This experimental platform can be used to screen large chemical libraries in search of novel compounds to replace the reprogramming factors that induce pluripotency. Ultimately, such compounds may provide mechanistic insight into the process of reprogramming.

Recently it has been demonstrated that forced expression of combinations of four transcription factors – Oct4, Klf4, Sox2 and c-Myc or Oct4, Sox2, Nanog and Lin28 – can reprogram somatic cells into induced pluripotent stem (iPS) cells that closely resemble embryonic stem (ES) cells (Maherali et al., 2007; Okita et al., 2007; Takahashi and Yamanaka, 2006; Wernig et al., 2007). The induction of pluripotency in somatic cells by ectopic expression of four transcription factors has created new opportunities in the stem cell field. For example, patient-derived pluripotent cells may ultimately represent a personalized source of tissue to replace cells lost to degenerative or age associated diseases (Jaenisch and Young, 2008). However, the clinical use of iPS cells is mainly hindered by the use of virally transduced transcription factors and the use of proto-oncogenes.

Optimization of the original reprogramming method has achieved reprogramming of both mouse and human somatic cells with the three-factor combination of Oct4, Sox2 and Klf4 in the absence of c-Myc (Hockemeyer et al., 2008; Nakagawa et al., 2007; Wernig et al., 2008b). Additional methods to reduce the number of reprogramming factors have taken advantage of endogenously expressed reprogramming factors, thereby replacing the need for ectopic expression of these factors – e.g., Sox2 (Aasen et al., 2008; Aoi et al., 2008; Eminli et al., 2008; Kim et al., 2009; Kim et al., 2008b). More recently, iPS cells have been generated by the use of excisable vectors (Kaji et al., 2009; Soldner et al., 2009; Woltjen et al., 2009), non-integrating vectors (Stadtfeld et al., 2008; Yu et al., 2009) or transient transfection approaches (Okita et al., 2008). While these methods have solved some of the limitations posed by the presence of multiple

proviruses in the genome of iPS cells, they either remain inefficient or have failed to further our understanding of the mechanistic details involved epigenomic reprogramming.

Small molecules that induce epigenetic changes have also been used to increase the efficiency of reprogramming and/or replace reprogramming factors (Huangfu et al., 2008a; Huangfu et al., 2008b; Mikkelsen et al., 2008; Shi et al., 2008a; Shi et al., 2008b; Wernig et al., 2008a). Recently, the histone deacetylase inhibitor valproic acid enabled two-factor reprogramming (Oct4 and Sox2) of human fibroblasts in the absence of Klf4 and c-Myc (Huangfu et al., 2008a; Huangfu et al., 2008b). This method relies on the use of known molecules that target known mechanisms. The identification of novel small-molecules that induce pluripotency could provide useful chemical tools to unravel and study the molecular mechanisms governing this process. Furthermore, application of such molecules may ultimately lead to useful new therapeutics for the treatment of disease. Toward this end, we have developed a high-throughput screening strategy to identify small molecules that can functionally replace reprogramming transcription factors. In this paper we describe the identification of a series of compounds that can substitute for the reprogramming factor Klf4 in the formation of iPS cells. Our results provide proof-of-principle that high-throughput screening is a robust vehicle to identify small-molecules that can replace the transcription factors used to reprogram somatic cells.

Results

Nanog-luciferase reporter mouse strain. The induction of pluripotency in fibroblasts by the ectopic expression of Oct4, Sox2, Klf4 and c-Myc requires a minimum of 8-12 days and occurs at low frequency with <0.1% of the somatic cells giving rise to iPS cells (Brambrink et al., 2008; Wernig et al., 2007). The slow kinetics of reprogramming impedes the development of robust assays to screen large chemical libraries with a reliable readout that can be captured on a well-to-well basis in a miniaturized format. In order to overcome these limitations, we relied on the high sensitivity and quantitative readout capabilities of luciferase to develop a screening platform. To establish a sensitive readout for iPS cell formation, the firefly luciferase gene was inserted into the *Nanog* locus by homologous recombination. *Nanog* was chosen because it plays important roles in the maintenance of the undifferentiated state (Chambers et al., 2003; Chambers et al., 2007; Mitsui et al., 2003), is completely inactivated in somatic lineages, and iPS cells selected for *Nanog* reactivation demonstrate complete developmental potential (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007).

A targeting vector was constructed using 1.2kb of the *Nanog* promoter as a 5' arm and 1.4kb of *Nanog* intron 2 as a 3' arm. The firefly luciferase gene from pGL3-basic and a floxed pgk-neo cassette were inserted between the targeting arms (Fig. 1A). Nanog-luciferase (NL) ES cell clones that carried the targeting construct in the *Nanog* locus, as confirmed by Southern blot analysis (Fig. 1B), gave rise to a signal 10-30 times greater than wild type V6.5 ES cells (Fig. 1C). The NL ES cells were introduced into blastocysts to generate chimeric mice, which were bred to generate both transgenic

mice and mouse embryonic fibroblasts (MEFs). NL-MEFs produced negligible amounts of luciferase (Fig. 1C) and were used for subsequent experiments.

High-throughput chemical screening strategy: Klf4 replacement. In this report, we customized our screening strategy to identify compounds that would replace Klf4. It has been suggested that Klf4 plays a major role in chromatin remodeling and recent evidence demonstrated that Klf4 is an important mediator of the undifferentiated of ES cell state (Jiang et al., 2008; Kim et al., 2008a). In addition, over-expression of Klf4 can promote breast cancer (Foster et al., 2000) or squamous cell carcinomas (Foster et al., 2005) ascribing a potent oncogenic role to this transcription factor. Thus, replacement of Klf4 is a critical step for the eventual application of iPS cells as a therapeutically viable strategy.

In order to identify a chemical replacement for Klf4, NL-MEFs were transduced with a subset of reprogramming factors consisting of Oct4, Sox2 and c-Myc (OSM). To obtain the approximately 10^9 cells needed to screen large chemical libraries required that OSM transduced NL-MEFs be expanded for three passages prior to screening. At passage 4, OSM transduced NL-MEFs were seeded into 1536-well plates at 500 cells/well in mouse ES cell growth media supplemented with LIF (20 ng/mL) and screened against a large (500,000 compounds), chemically diverse small molecule library (Plouffe et al., 2008). Compounds were added at a final concentration of 2.2 μM (20 nL of 1 mM stock solutions in 9 μL of media) immediately after plating, and Nanog-driven luciferase expression was assayed 10 days later (Fig. 1D). Analysis of a focused library in pilot experiments demonstrated a high degree of homogeneity in luciferase

response from well to well (10^4 molecules in triplicate; SI Fig. 1). Furthermore, normalization of these data using the 'jackknife' Z score method – which accounts for plate to plate variation and does not rely on positive controls readings – and subsequent analysis using the one-sample Kolmogorov-Smirnov test (P = 0.54) indicated that there was little difference between the empirical and theoretical inverse gamma distribution. These statistical analyses confirmed both that the screening platform was robust and that the output data were highly reliable (Malo et al., 2006).

Nanog activity from OSM transduced NL-MEFs was one order of magnitude greater than that in non-transduced cells and 4-fold less than the activity in pluripotent NL reporter cells (Fig. 1D). From the primary screen, compounds that gave rise to a signal >2.2-fold (~5K compounds) over the plate average –assuming that most compounds are inactive and can serve as controls (Malo et al., 2006) – were selected for reanalysis in triplicate. Measurements with these putative hit compounds were collected, averaged and those that activated the luciferase signal 2.5-fold over DMSO treated controls (~1K compounds) were counter-screened in a cell based SV40-driven firefly luciferase assay to rule out false positives that directly and non-specifically activate luciferase reporters (Auld et al., 2009). Compounds that reproducibly hit in the Nanog assay that did not hit in the SV40 assay were further analyzed in a dose response format to identify the optimal concentration for activity (Fig. 2A). Collectively, this strategy further narrowed our hit compounds to 135 small molecules (~0.03% of the initial set), which grouped into three lead structural classes – flavones, lysergamides, and paullones (Fig. 2B).

Activity of lead compounds. To distinguish compounds that replace Klf4 from those that simply activate Nanog expression, colony formation was used as a secondary assay. Representative compounds from each of the lead structural classes were added to OSM transduced MEFs and analyzed for the presence of colonies 10 days later. Members of each structural class induced the OSM-MEFs to form colonies (Fig. 2B); however, colonies that stained positive for AP were observed only in the presence of kenpaullone and lysergic acid ethylamide (data not shown). Kenpaullone proved to be the most robust reprogramming molecule based on this series of assays and was studied in more detail.

While AP positive colony formation is consistent with the conversion of somatic to iPS cells, it does not prove establishment of a pluripotent state. To determine whether iPS cells derived with kenpaullone were fully reprogrammed, we transduced MEFs carrying the gene encoding neomycin resistance in the *Oct4* locus – ONR-MEF (Wernig et al., 2007) – with OSM vectors and treated these cells with retrovirally delivered Klf4, vehicle control or kenpaullone 2 days later (Fig. 2C). Nanog expression was analyzed by semi-quantitative RT-PCR at three-day intervals to confirm the re-activation observed using the NL reporter line. Consistent with observations using NL-MEFs, Nanog expression was activated in kenpaullone treated OSM-transduced MEFs in a time dependent manner (Fig. 2C and SI Fig. 2).

After 20 days of concurrent kenpaullone treatment, colonies expressing Oct4 from their endogenous locus were selected upon supplementation of the culture media with neomycin (Fig. 2C). Kenpaullone was able to replace Klf4, as demonstrated by the appearance of Oct4-neo colonies, though reprogramming was less efficient than after

transduction with a Klf4 vector. Likewise, OSM transduction and kenpaullone treatment of MEFs carrying a green fluorescent protein (GFP) reporter in the *Oct4* locus – O4G-MEF (Lengner et al., 2007) – gave rise to GFP positive colonies (Fig. 3A) with similar efficiency and kinetics as observed in the ONR-MEFs (Fig. 3B). DMSO treated OSM-transduced O4G-MEFs grown did not display GFP positive cells or colony formation even after >40 days in culture (data not shown).

Characterization of iPS cells. iPS cells generated by OSM and kenpaullone treatment were indistinguishable from ES cells by morphological criteria and expressed the pluripotency associated markers Oct4 and Nanog from the endogenous loci (Fig. 3A). However, the reprogramming efficiency of kenpaullone-treated OSM transduced O4G-MEFs was 10-fold lower than that of Klf4-infected controls (Fig. 3B). To confirm the activity of kenpaullone in an independent culture system, we used “secondary” MEFs that carry doxycycline (DOX)-inducible proviruses encoding Oct4, Sox2 and c-Myc (Markoulaki et al., 2009). The results shown in Figure 3C indicate that kenpaullone induces reprogramming of secondary OSM-MEFs with a similar efficiency as primary infected OSM-MEFs. The reprogramming kinetics of kenpaullone treated iPS cell formation was, however, delayed as DOX independent iPS cells only appeared after 25-30 days, rather than after 15 days observed for 4-factor control iPS cells. The slight kinetic delay in reprogramming suggests that kenpaullone is less efficient in epigenetic reprogramming than Klf4.

Once established, Klf4-free iPS cells grew in the absence of DOX and kenpaullone while maintaining ES cell morphology and expression of Nanog and Oct4. To test for

pluripotency, the cells were injected into blastocysts and found to generate germline competent chimeras (Fig. 3A and data not shown).

It has been shown that mouse neural progenitor cells (NPC) that express endogenous Sox2 can be reprogrammed to a pluripotent state with the addition of Oct4 and Klf4 +/- c-Myc (Eminli et al., 2008; Kim et al., 2008b) or simply Oct4 (Kim et al., 2009), albeit with delayed kinetics and low efficiency. To determine the functionality of kenpaullone in the reprogramming of neural cells, NPCs were transduced with all combinations of Oct4, Klf4 and c-Myc and treated with 5 μM kenpaullone or vehicle control. Eight days later, discernable AP positive colonies had formed in Oct4/c-Myc kenpaullone-treated wells that, when passaged, appeared morphologically indistinguishable from ES cells and stained positive for the pluripotency-associated markers Nanog, Sox2 and AP (Fig. 3D). Furthermore, these cells no longer expressed GFP from the virally delivered c-Myc vector (pCMV-Myc-IRES-GFP; SI Fig. 3), consistent with retroviral silencing observed in reprogrammed cells (Wernig et al., 2007). The efficiency of iPS cell formation from NPCs was lower with kenpaullone than with Klf4 gene transduction, though not as low as that observed in MEFs (Fig. 3B versus 3E). It is also worth noting that neither MEFs nor NPCs reprogrammed in the absence of c-Myc when kenpaullone was substituted for Klf4. Collectively, these results suggest that kenpaullone requires transduction of c-Myc for reprogramming and emphasizes that kenpaullone may not completely recapitulate the role of Klf4.

Biological Activity of Kenpaullone. Previous work with kenpaullone has demonstrated a wide range of biological utility extending from maintenance of

pancreatic beta cell survival and proliferation (Stukenbrock et al., 2008) to the induction of apoptosis in cancer cells (Zaharevitz et al., 1999). Kenpaullone's diverse range of activity is a direct result of its kinase inhibition promiscuity; *i.e.*, kenpaullone is a potent inhibitor of GSK-3b (23 nM), CDK1/cyclin B (400 nM), CDK2/cyclin A (680 nM), CDK5/p35 (850 nM), and exhibits inhibitory activity toward a variety of other kinases at higher concentrations (Knockaert et al., 2002; Zaharevitz et al., 1999).

To determine if the reprogramming activity of kenpaullone results from its well established role as a GSK-3β or CDK inhibitor, a highly selective small molecule inhibitor of GSK-3β (CHIR99021), a promiscuous inhibitor of CDKs (purvalanol A) or both were applied to OSM-transduced ONR-MEFs and compared against kenpaullone. As shown in Figure 3F, known inhibitors of these kinases are unable to replace Klf4 in the reprogramming of murine fibroblasts and have negligible activity in the NL reporter MEF line (SI Fig. 4). It is interesting to note that kenpaullone, like CHIR99021, increases the efficiency of 4-factor reprogramming (Fig. 3F). This observation is consistent with reports indicating that GSK-3β inhibitors can facilitate, in part, mouse ES cell self-renewal (Ying et al., 2008) and the reprogramming of mouse NPCs back to the pluripotent state (Silva et al., 2008). However, as other more specific GSK-3β inhibitors are not able to activate Nanog expression (SI Fig. 4) or to replace Klf4 (Fig. 3F), this activity is likely independent of the reprogramming activity of kenpaullone.

In addition to small molecule mediated inhibition, short hairpin RNA (shRNA)-targeted ablation of each of the canonical kenpaullone kinase targets – individually, combinatorially or in a pooled format – did not result in colony formation or an increase in Nanog driven luciferase expression in OSM-transduced MEFs (SI Fig. 5).

Collectively, these data strongly suggest that the activity of kenpaullone does not result from its well-documented role as a GSK-3β or cell cycle inhibitor. Lastly, kenpaullone does not directly activate Klf4 expression at the mRNA or protein level (SI Fig. 6) suggesting that it is functioning through an entirely novel mechanism.

Discussion

In this work we describe the development and application of an unbiased, high-throughput (>500K compounds) cell based screening technology that can be applied to identify small molecules to replace virally transduced reprogramming transcription factors. In a proof-of-concept application, we have identified three classes of molecules that activate Nanog expression in murine fibroblasts during direct reprogramming in the absence of Klf4, and have characterized one such molecule, kenpaullone, in detail. We showed that kenpaullone is able to replace Klf4 in the reprogramming of primary and secondary fibroblasts and NPCs. OSM iPS cells derived in the presence of kenpaullone display the characteristics of pluripotent ES cells, including the ability to contribute to chimeric mice.

Previous studies with small molecules that induce epigenetic changes suggest that chromatin remodeling is a rate-limiting step in the conversion from a somatic to a pluripotent epigenetic state (Huangfu et al., 2008a; Huangfu et al., 2008b; Mikkelsen et al., 2008; Shi et al., 2008a; Shi et al., 2008b; Wernig et al., 2008a). For example, treatment of MEFs with a small molecule inhibitor of DNA methyltransferase (5-azacytidine) facilitates reprogramming during a brief temporal window by removing demethylation marks, thereby lowering the kinetic barrier to transition to pluripotency

(Mikkelsen et al., 2008; Wernig et al., 2008a). Furthermore, histone deacetylase inhibitors and histone methyl transferase inhibitors have also been shown to increase the efficiency of reprogramming (Huangfu et al., 2008a; Huangfu et al., 2008b; Shi et al., 2008a; Shi et al., 2008b). In particular, the histone deacetylase inhibitor valproic acid dramatically increased three-factor reprogramming efficiency in the absence of c-Myc in both mouse and human cells and enabled two-factor reprogramming (Oct4 and Sox2) of human fibroblasts in the absence of Klf4 and c-Myc (Huangfu et al., 2008a; Huangfu et al., 2008b).

While modulation of known epigenetic control elements has provided significant progress toward reprogramming not requiring genetic manipulations, such techniques rely on specific compounds to target known mechanisms. As the study of induced pluripotency is still in its infancy, the mechanisms governing reprogramming remain largely unknown. It is likely that the activation of different alternative pathways by small molecules will enhance the efficiency of vector-free reprogramming. Furthermore, the identification of chemicals with novel mechanisms of action to modulate this process is of great interest. This work demonstrates that unbiased, cell-based screens can identify small molecules from large chemical libraries to replace transcription factors for reprogramming. Moreover, the identification of these and other molecules provide novel tools to study the molecular mechanisms at play during epigenome overhaul and may ultimately help to bring iPS cell technology one step closer to clinical application.

Materials and Methods

Chemicals. Kenpaullone, purvalanol A and 7-hydroxyflavone were purchased from Sigma, CHIR99021 from Axon MedChem and lysergic acid ethylamide was acquired from a commercial chemical vendor. The chemical library was described previously (Plouffe et al., 2008).

Gene Targeting and Blastocyst Injection. Generation of the Nanog-luciferase targeting vector is described in detail in the SI methods. For targeting, the Nanog-luciferase vector was linearized with XhoI and electroporated into ES cells. G418 selection (350 μg/ml) was started 24 hours after electroporation and continued for 10-12 days. Resistant clones were picked and expanded for screening by Southern blot analysis. Injections of Nanog-luciferase ES or iPS cells into either BDF2 or BALB/C host blastocysts were carried out as previously reported (Beard et al., 2006).

Cell Culture and Viral Infections. Transgenic MEFs were isolated from Nanog-luciferase, secondary OSM (Markoulaki et al., 2009), ONR (4) or OGF (Lengner et al., 2007) mice, selected from E13.5 embryos and cultivated in DMEM (Invitrogen) containing 15% FBS, L-glutamine, beta-mercaptoethanol and nonessential amino acids. ES and iPS cells were cultivated on irradiated MEFs in the aforementioned media plus 20 ng/mL LIF (Chemicon).

MEFs at passage 1-2 were infected with pooled viral supernatant generated by transfection of HEK293T cells (Fugene, Roche) with doxycycline inducible lentiviral vectors (Brambrink et al., 2008) or the maloney retroviral vectors (pMXs, Addgene; generously deposited by S. Yamanaka) containing the cDNAs of *Oct4*, *Sox2* and *c-Myc +/- Klf4*, and pseudo-typed using with a plasmid encoding the VSV-G envelope protein.

NPCs were derived from ES cells (Bibel et al., 2007) and expanded 4-5 passages in DMEM/F-12 (Invitrogen) supplemented with B27 (1X; Invitrogen), EGF (20 ng/mL; R&D Systems) and FGF-2 (20 ng/mL; Invitrogen). NPCs (10^5 per well of a 6-well dish) were infected overnight, followed by transfer to ES cell growth media. For all experiments, compound(s) or vehicle control (DMSO 0.05%), doxycycline (2 μg/mL) and media were refreshed every 2-3 days. The experimental conditions concerning the generation of iPS cells from NPCs, ONR-, OG4- and secondary-MEFs are described in detail in the SI Methods section accompanying this paper.

Chemical Screening and Nanog Assay. Nanog luciferase MEFs (~10^8 cells) were transduced at passage 1 with concentrated VSV-G pseudotyped *Oct4*, *Sox2* and *c-Myc* expressing lentivirus in 15cm dishes. Immunofluorescent staining confirmed that >90% of NL-MEFs expressed Oct4 (polyclonal mouse, Santa Cruz) and Sox2 (monoclonal mouse, Santa Cruz). OSM transduced NL-MEFs were expanded four passages (~2 weeks) prior to screening. For large-scale screens, the OSM NL-MEFs were plated into 1536-well plates at 500 cells/well in mouse ES cell growth media supplemented with LIF (20 ng/mL). Compounds were added at 2.2 mM final concentration (20 nL of 1 mM stock solutions in 9 mL of media) immediately after plating. Nanog-driven luciferase expression was assayed 10 days later by adding 2 μL of Bright-Glo (Promega). Primary hits were re-screened in triplicate in 1536-well format. Dose response curves were generated by screening OSM NL-MEFs in 8-pt dilution (1 mM stock, half log serial dilutions) in 384-well format. 'Jackknife' Z score values, the one-sample K-S test value (P = 0.54) and the empirical and theoretical inverse gamma distribution were calculated according to Malo, *et al.* (Malo et al., 2006) and references therein.

A HEK293T counter-screening cell line was generated by co-transfection of the pGL3 SV40-luciferase vector (Promega, # E1761) with pcDNA3.1+ (Invitrogen, V790-20) encoding the gene for neomycin resistance. Neomycin selections (400 μg/mL) were carried out for 7 days. The SV40-luciferase HEK293T line was passaged in DMEM plus 10% FBS. Hits that reconfirmed in triplicate in NL-MEFs were screened in SV40-luc 293T cells to filter out false positives (Auld et al., 2009).

Counter-screening, dose response experiments, structure and activity relationship analysis and the timecourse experiments were all run in 384-well plates (Greiner). Cells were plated at 1K cells/well in ES cell growth media and assayed 10 days later, unless otherwise specified. All luciferase readings were acquired following addition of Bright-Glo reagent (Promega). All experiments were run in duplicate and repeated at least three times.

Semiquantitative RT-PCR Analysis. Total RNA was obtained and reverse transcribed, as described previously (Lyssiotis et al., 2007). PCR was performed using the Phusion polymerase (New England Biosciences) and gene-specific primers (SI Table 4). Cycle parameters were 10s at 98°C, 15s at 63°C (Nanog) or 67°C (Nat1), and 30s at 72°C for 30 cycles. The amount of total RNA and PCR conditions were optimized so that amplification of *Nat1* and *Nanog* were in the exponential phase.

Immunocytochemistry and Antibodies. Immunofluorescence analysis was performed as described previously (Lyssiotis et al., 2007). Antibodies against Sox2 (polyclonal mouse, Santa Cruz), Oct4 (monoclonal mouse, Santa Cruz) and Nanog (polyclonal rabbit, Santa Cruz) were incubated with fixed cells, washed and labeled with fluorophore appropriate secondary antibodies (Jackson Immunoresearch). Alkaline

phosphatase staining was performed according to the manufacturer's recommendations (Millipore). Images were captured using an Eclipse TE2000-U microscope (Nikon) with _40 or _100 magnification.

Acknowledgements

We would like to thank the Novartis Institute for Functional Genomics chemical screening facility for technical assistance and acknowledge Dr. Anthony Orth and Chang Liu for helpful discussions. This work is supported by a National Science Foundation Predoctoral Fellowship (C.A.L.), a Human Frontier Science Program Long-term Fellowship (J.S.), a Canadian Institute for Health Research Post-Doctoral Fellowship (L.L.L.), a grant from the Skaggs Institute of Chemical Biology of The Scripps Research Institute (P.G.S), the Novartis Research Foundation (P.G.S), and the National Institute of Health (RO1-HD045022 and R37-CA084198; R.J.).

References

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilic, J., Pekarik, V., Tiscornia, G.*, et al.* (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat Biotechnol.

Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T., and Yamanaka, S. (2008). Generation of pluripotent stem cells from adult mouse liver and stomach cells. Science *321*, 699-702.

Auld, D.S., Thorne, N., Maguire, W.F., and Inglese, J. (2009). Mechanism of PTC124 activity in cell-based luciferase assays of nonsense codon suppression. Proc Natl Acad Sci U S A *106*, 3585-3590.

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

Bibel, M., Richter, J., Lacroix, E., and Barde, Y.A. (2007). Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. Nat Protoc *2*, 1034-1043.

Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential Expression of Pluripotency Markers during Direct Reprogramming of Mouse Somatic Cells. Cell Stem Cell *2*, 151-159.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643-655.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230-1234.

Eminli, S., Utikal, J., Arnold, K., Jaenisch, R., and Hochedlinger, K. (2008). Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous sox2 expression. Stem Cells *26*, 2467-2474.

Foster, K.W., Frost, A.R., McKie-Bell, P., Lin, C.Y., Engler, J.A., Grizzle, W.E., and Ruppert, J.M. (2000). Increase of GKLF messenger RNA and protein expression during progression of breast cancer. Cancer Res *60*, 6488-6495.

Foster, K.W., Liu, Z., Nail, C.D., Li, X., Fitzgerald, T.J., Bailey, S.K., Frost, A.R., Louro, I.D., Townes, T.M., Paterson, A.J.*, et al.* (2005). Induction of KLF4 in basal keratinocytes blocks the proliferation-differentiation switch and initiates squamous epithelial dysplasia. Oncogene *24*, 1491-1500.

Hockemeyer, D., Soldner, F., Cook, E.G., Gao, Q., Mitalipova, M., and Jaenisch, R. (2008). A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. Cell Stem Cell *3*, 346-353.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008a). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol *26*, 795-797.

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D.A. (2008b). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol *26*, 1269-1275.

Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell *132*, 567-582.

Jiang, J., Chan, Y.-S., Loh, Y.-H., Cai, J., Tong, G.-Q., Lim, C.-A., Robson, P., Zhong, S., and Ng, H.-H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat Cell Biol *10*, 353-360.

Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P., and Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008a). An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. Cell *132*, 1049-1061.

Kim, J.B., Sebastiano, V., Wu, G., Arauzo-Bravo, M.J., Sasse, P., Gentile, L., Ko, K., Ruau, D., Ehrich, M., van den Boom, D.*, et al.* (2009). Oct4-induced pluripotency in adult neural stem cells. Cell *136*, 411-419.

Kim, J.B., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V., Arauzo-Bravo, M.J., Ruau, D., Han, D.W., Zenke, M.*, et al.* (2008b). Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. Nature *454*, 646-650.

Knockaert, M., Wieking, K., Schmitt, S., Leost, M., Grant, K.M., Mottram, J.C., Kunick, C., and Meijer, L. (2002). Intracellular Targets of Paullones. Identification following affinity purification on immobilized inhibitor. J Biol Chem *277*, 25493-25501.

Lengner, C.J., Camargo, F.D., Hochedlinger, K., Welstead, G.G., Zaidi, S., Gokhale, S., Scholer, H.R., Tomilin, A., and Jaenisch, R. (2007). Oct4 expression is not required for mouse somatic stem cell self-renewal. Cell Stem Cell *1*, 403-415.

Lyssiotis, C.A., Walker, J., Wu, C., Kondo, T., Schultz, P.G., and Wu, X. (2007). Inhibition of histone deacetylase activity induces developmental plasticity in oligodendrocyte precursor cells. Proc Natl Acad Sci U S A *104*, 14982-14987.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R.*, et al.* (2007). Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. Cell Stem Cell *1*, 55-70.

Malo, N., Hanley, J.A., Cerquozzi, S., Pelletier, J., and Nadon, R. (2006). Statistical practice in high-throughput screening data analysis. Nat Biotechnol *24*, 167-175.

Markoulaki, S., Hanna, J., Beard, C., Carey, B.W., Cheng, A.W., Lengner, C.J., Dausman, J.A., Fu, D., Gao, Q., Wu, S.*, et al.* (2009). Transgenic mice with defined combinations of drug-inducible reprogramming factors. Nat Biotechnol *27*, 169-171.

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

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2007). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature *448*, 313-317.

Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., and Yamanaka, S. (2008). Generation of Mouse Induced Pluripotent Stem Cells Without Viral Vectors. Science.

Plouffe, D., Brinker, A., McNamara, C., Henson, K., Kato, N., Kuhen, K., Nagle, A., Adrian, F., Matzen, J.T., Anderson, P.*, et al.* (2008). From the Cover: In silico activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen. Proc Natl Acad Sci U S A *105*, 9059-9064.

Shi, Y., Desponts, C., Do, J.T., Hahm, H.S., Scholer, H.R., and Ding, S. (2008a). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell Stem Cell *3*, 568-574.

Shi, Y., Do, J.T., Desponts, C., Hahm, H.S., Scholer, H.R., and Ding, S. (2008b). A combined chemical and genetic approach for the generation of induced pluripotent stem cells. Cell Stem Cell *2*, 525-528.

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

Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G.W., Cook, E.G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M.*, et al.* (2009). Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. Cell *136*, 964-977.

Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G., and Hochedlinger, K. (2008). Induced pluripotent stem cells generated without viral integration. Science *322*, 945-949.

Stukenbrock, H., Mussmann, R., Geese, M., Ferandin, Y., Lozach, O., Lemcke, T., Kegel, S., Lomow, A., Burk, U., Dohrmann, C.*, et al.* (2008). 9-cyano-1-azapaullone (cazpaullone), a glycogen synthase kinase-3 (GSK-3) inhibitor activating pancreatic beta cell protection and replication. J Med Chem *51*, 2196-2207.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Wernig, M., Lengner, C.J., Hanna, J., Lodato, M.A., Steine, E., Foreman, R., Staerk, J., Markoulaki, S., and Jaenisch, R. (2008a). A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. Nat Biotechnol *26*, 916-924.

Wernig, M., Meissner, A., Cassady, J.P., and Jaenisch, R. (2008b). C-Myc is Dispensible for Direct Reprogramming of Mouse Fibroblasts. Cell Stem Cell *2*, 1-3.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.

Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hamalainen, R., Cowling, R., Wang, W., Liu, P., Gertsenstein, M.*, et al.* (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature.

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

Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, II, and Thomson, J.A. (2009). Human Induced Pluripotent Stem Cells Free of Vector and Transgene Sequences. Science.

Zaharevitz, D.W., Gussio, R., Leost, M., Senderowicz, A.M., Lahusen, T., Kunick, C., Meijer, L., and Sausville, E.A. (1999). Discovery and initial characterization of the paullones, a novel class of small-molecule inhibitors of cyclin-dependent kinases. Cancer Res *59*, 2566-2569.

Fig. 1 Nanog luciferase screening platform. **(A)** Scheme for targeting the firefly (FF) luciferase gene to the *Nanog* locus. **(B)** Southern blot hybridization analysis of the Nanog-luciferase targeted allele. The targeted allele-specific 5.2kb and 4.7 kb PvuII fragments are detected with the 5' internal and 3' external probes, respectively, shown in A. **(C)** Nanog-luciferase activity of correctly targeted ES cell clones (NL2 and NL5) versus Nanog-luciferase MEFs. Differentiated Nanog-luciferase tissue shows a significant reduction in luciferase expression. V6.5 ES cells and MEFs are provided as a

luciferase negative control. **(D)** Small molecule screening strategy. Nanog-luciferase MEFs were transduced with a reduced reprogramming cocktail (Oct4, Sox2 and c-Myc; OSM), expanded for 2 weeks, plated into 1536-well plates, treated with compound and assayed for luciferase expression 10 days after plating. The initial library of 500,000 compounds was narrowed down to ~1,000 compounds that reproducibly activated the Nanog reporter, 135 of which were confirmed in dose-response format. Representative members of each chemical class activated the Nanog reporter approximately 3-fold over untreated OSM-MEFs and near equivalent to that observed in pluripotent cells. Flavones (7-hydroxyflavone, 20 μM), Ergolines (lysergic acid ethylamide, 2 μM) and Paullones (kenpaullone, 5 μM). Nanog activity is reported in RLU and was read in 96- **(B)** or 384- **(D)** well format. Error bars for all experiments represent the standard deviation (n = 3).

Fig. 2 Activity of lead chemical scaffolds. **(A)** Dose-dependent Nanog activation of a lead compound from each chemical class: lysergic acid ethylamide (Ergolines), 7-hydroxyflavone (Flavones), kenpaullone (Paullones). Above 10 μM, kenpaullone and lysergic acid ethylamide are toxic. **(B)** OSM-MEFs took on colony morphology upon application lead compounds. **i** – DMSO; **ii** – KLF4; **iii**, 7-hydroxyflavone, 20 μM; **iv**, lysergic acid ethylamide; 2 μM; **v**, kenpaullone, 5 μM. **(C)** ONR-MEFs were transduced with OSM (day -2) and treated with DMSO, retrovirally-delivered Klf4 or kenpaullone 2 days later (day 0). Semi-quantitative RT-PCR of samples collected at days 3, 6, 9 and 12 confirmed the re-activation of Nanog observed using NL-MEFs. At day 20, kenpaullone and DMSO were removed, and neomycin was added to the culture media

to initiate selection of iPS cells. Neomycin-resistant colonies – indicating re-activation of endogenous Oct4 – were counted at day 25 (per fifty thousand MEFs). Error bars, standard deviation (n = 4 wells of a 6-well plate). Dose response curves are representative of at least 6 independent experiments and were read in 384-well format. Images were captured at _40 magnification.

Fig. 3 Characterization of iPS cells derived with kenpaullone. **(A)** iPS cells generated with OSM and 5 μM kenpaullone express the pluripotency associated markers Oct4 and Nanog from their endogenous loci and are capable of contributing to chimeric mice**. (B)** Chemical complementation of Klf4 with kenpaullone is less efficient than retroviral delivery of Klf4, as scored by the number of colonies (per ten thousand MEFs) that express GFP from the endogenous *Oct4* locus. Error bars, standard deviation (n = 3). **(C)** Secondary MEFs harboring proviral copies of OSM were treated with kenpaullone (5 μM) or retrovirally delivered Klf4. Colonies were scored based on expression of OCT4. **(D)** Kenpaullone-treated Oct4/c-Myc-NPCs gave rise to iPS cells that expressed pluripotency associated proteins (Nanog, Sox2 and AP) from endogenous loci. **(E)** NPCs were transduced with Oct4 +/- Klf4 and c-Myc and treated with kenpaullone (5 μM) or vehicle control. Data represent colonies that expressed AP after 8 days (per one hundred thousand NPCs). Error bars, standard deviation (n = 3). **(F)** ONR-MEFs were transduced with OSM +/- Klf4 and treated with a GSK-3β specific inhibitor (GSKi; CHIR99021, 3 μM), a general CDK inhibitor (CDKi; purvalanol A, 3 μM), both the GSKi and the CDKi, or kenpaullone. Colonies were scored for neomycin-resistance at day 25 (per fifty thousand MEFs). Error bars, standard deviation (n = 4).

Fig. 1

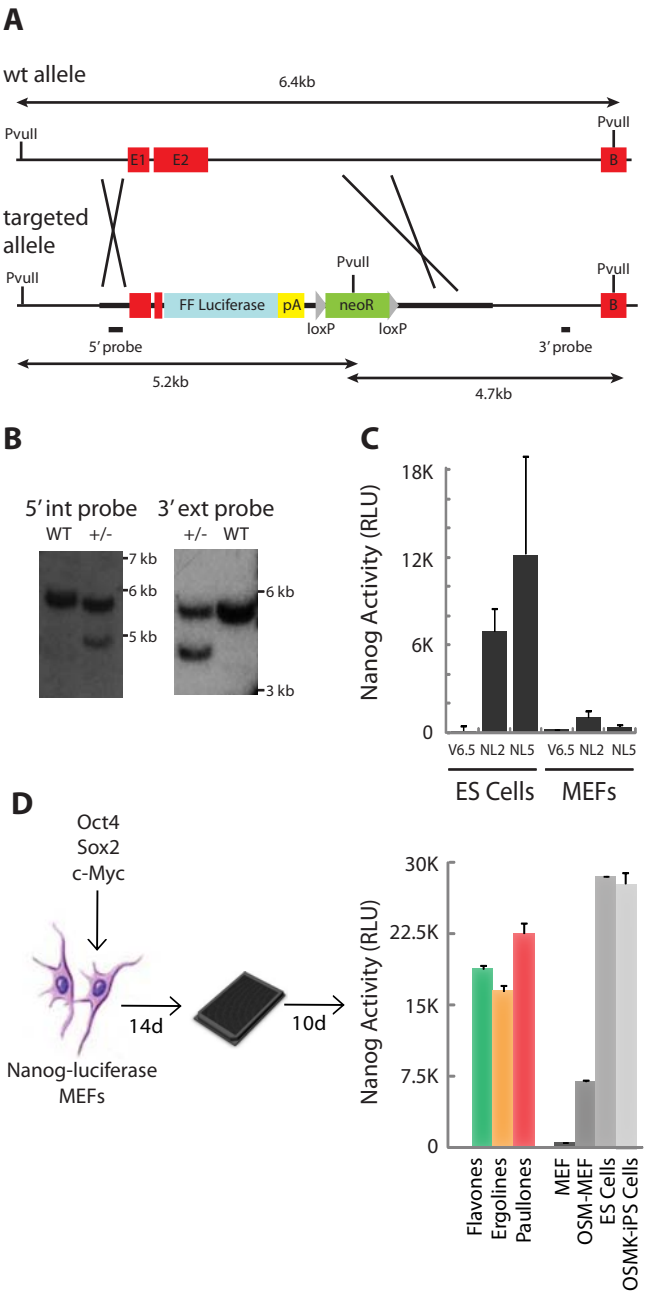

A
wt allele
6.4kb
PvuII
PvuII
E1
E2
B
targeted allele
PvuII
PvuII
PvuII
FF Luciferase
pA
neoR
loxP
loxP
B
5′ probe
3′ probe
5.2kb
4.7kb
B
5′ int probe
3′ ext probe
WT +/-
+/- WT
7 kb
6 kb
5 kb
6 kb
3 kb
C
Nanog Activity (RLU)
18K
12K
6K
0
V6.5 NL2 NL5 V6.5 NL2 NL5
ES Cells
MEFs
D
Oct4
Sox2
c-Myc
14d
10d
Nanog-luciferase MEFs
Nanog Activity (RLU)
30K
22.5K
15K
7.5K
0
Flavones
Ergolines
Paullones
MEF
OSM-MEF
ES Cells
OSMK-iPS Cells

Fig. 2

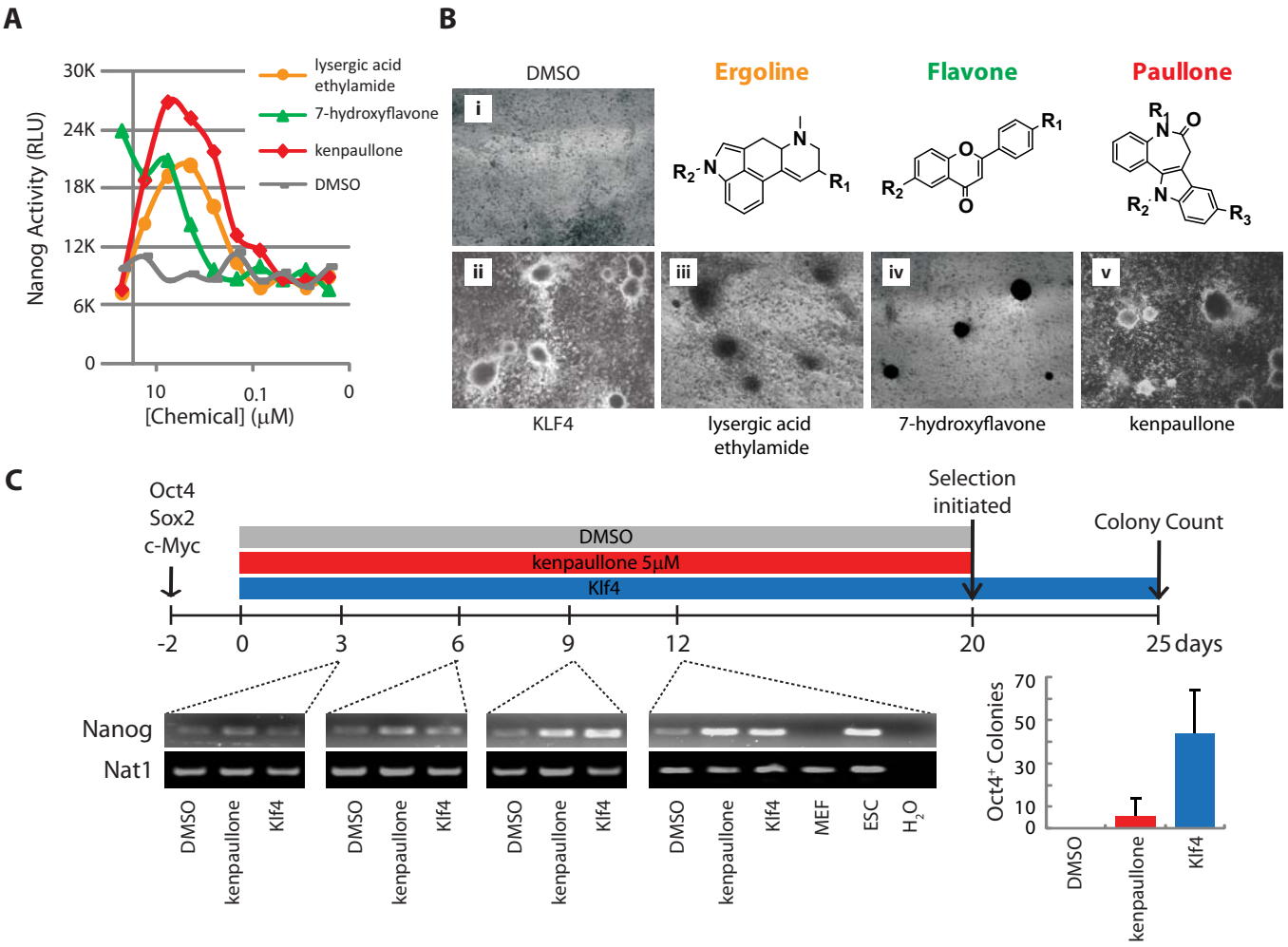
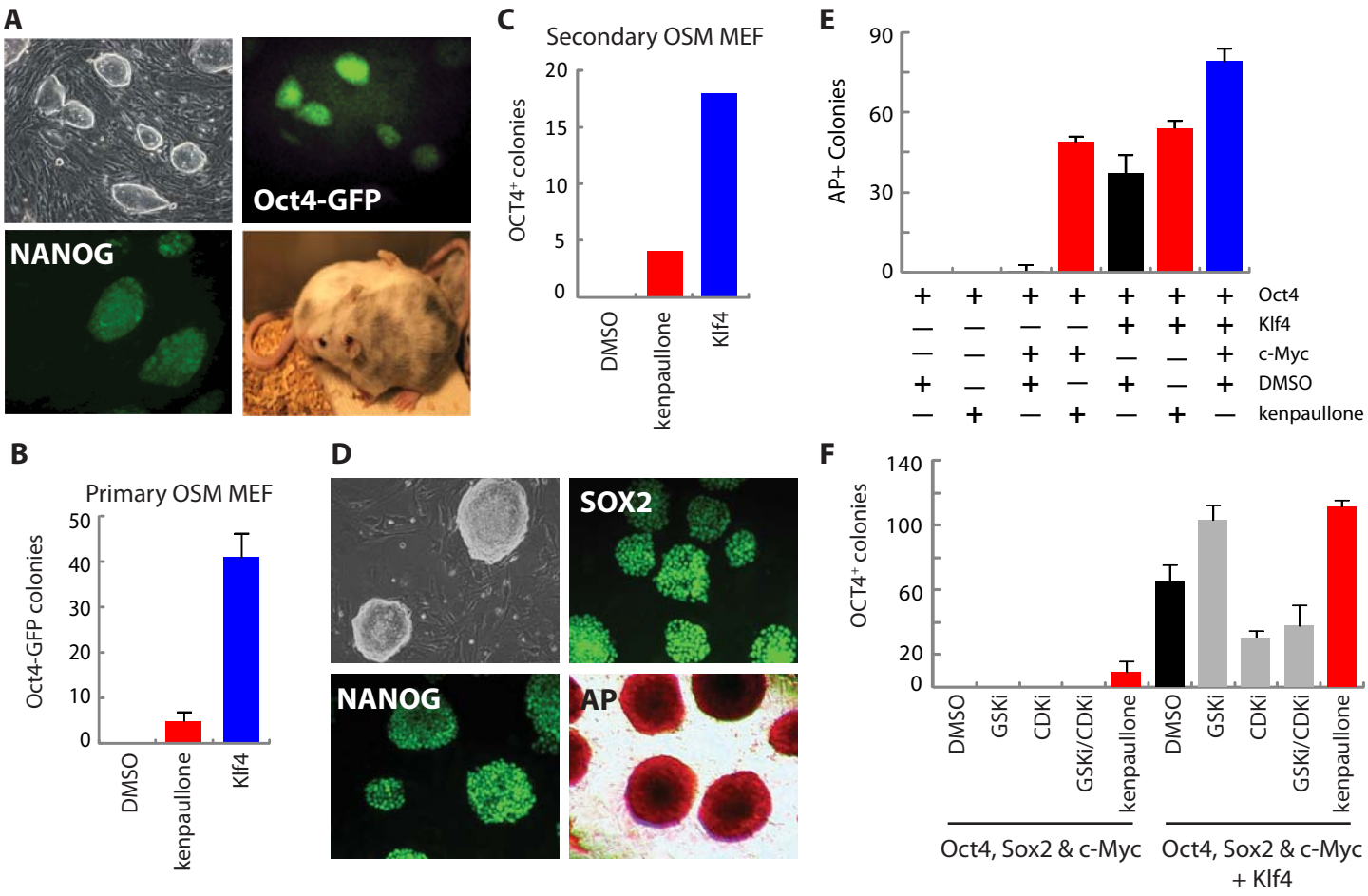

A
30K
24K
18K
12K
6K
0
Nanog Activity (RLU)
lysergic acid ethylamide
7-hydroxyflavone
kenpaullone
DMSO
10
0.1
0
[Chemical] (μM)
B
DMSO
Ergoline
Flavone
Paullone
i
ii
iii
iv
v
KLF4
lysergic acid ethylamide
7-hydroxyflavone
kenpaullone
C
Oct4
Sox2
c-Myc
DMSO
kenpaullone 5μM
Klf4
Selection initiated
Colony Count
-2
0
3
6
9
12
20
25 days
Nanog
Nat1
DMSO
kenpaullone
Klf4
MEF
ESC
H_2O
Oct4⁺ Colonies
70
50
30
10
0

Fig. 3


A
Oct4-GFP
NANOG
B
Primary OSM MEF
Oct4-GFP colonies
50
40
30
20
10
0
DMSO
kenpaullone
Klf4
C
Secondary OSM MEF
OCT4+ colonies
20
15
10
5
0
DMSO
kenpaullone
Klf4
D
SOX2
NANOG
AP
E
AP+ Colonies
90
60
30
0
Oct4
Klf4
c-Myc
DMSO
kenpaullone
F
OCT4+ colonies
140
100
60
20
0
DMSO
GSKi
CDKi
GSKi/CDKi
kenpaullone
Oct4, Sox2 & c-Myc
Oct4, Sox2 & c-Myc + Klf4

Chapter 6

Perspectives

Embryonic stem cells have significant therapeutic potential for use in regenerative medicine given their pluripotency, or their ability to differentiate into tissues of all three germ layers. Nuclear reprogramming of somatic cells offers the possibility of generating patient-specific stem cell lines, removing some of the obstacles hindering the therapeutic use of ES cells, including the need for immunosuppression after tissue transplants. A breakthrough in the field of reprogramming occurred when Takahashi and Yamanaka reported that retroviral transduction of the four transcription factors Oct4, Sox2, Klf4 and c-Myc induces a pluripotent state in somatic cells, generating iPS (induced pluripotent stem) cells (Takahashi and Yamanaka, 2006). The work in this thesis focuses on reprogramming using these defined factors.

Takahashi and Yamanaka's initial study did not generate fully reprogrammed iPS cells, but subsequent work, including that detailed in Chapter 2, describes the use of more stringent selection markers Oct4 and Nanog to isolate fully reprogramming iPS cells from four-factor-transduced fibroblasts. Once it was established that viral transduction of the four reprogramming factors indeed induced a pluripotent state in somatic cells, we began to explore the mechanism underlying nuclear reprogramming, providing an analysis of the re-activation kinetics of key pluripotency markers. We found, as detailed in Chapter 3, that reprogramming events occur gradually and sequentially over time. Using doxycycline-inducible viruses, this study also identified a minimum time requirement for transgene expression during the reprogramming process. This

timeline was then used as a benchmark for subsequent reprogramming experiments.

One main obstacle to realizing the therapeutic potential of iPS cells is the current reliance on viral transduction of the reprogramming factors and the multiple genomic integration sites iPS cells can contain. Not only can integrated viruses activate endogenous genes, including oncogenes, but reactivation of the viral transgenes can also occur. Roughly 50% of chimeric mice generated from iPS cells were prone to tumorigenesis due to c-Myc re-activation (Okita et al., 2007). Recent studies have begun to address viral integrations in reprogramming. One group was able to reprogram hepatocytes using non-integrating adenoviral vectors encoding the reprogramming factors, while another used transient transfection of a plasmid containing Oct4, Sox2 and Klf4 in a single expression vector (Okita et al., 2008; Stadtfeld et al., 2008). While iPS cells from both studies contain no viral integrants, reprogramming was highly inefficient, hindering its widespread use. More recently iPS cells were generated by genomic integration of the four reprogramming factors using plasmids, lentiviruses, or transposons, followed by excision of the viral transgenes using Cre-recombinase or transposase expression (Kaji et al., 2009; Soldner et al., 2009; Woltjen et al., 2009). While the majority of the transgene is excised in these studies, small portions still remain in the iPS cells, which could cause safety concerns if they are used for therapeutic purposes.

The ability to identify viral alternatives remains a key focus of scientific research. Work in this thesis, specifically in Chapters 4 and 5, details the identification of transient cues as replacements for the viral reprogramming factors. One method employed a candidate approach, applying the knowledge that Wnt signaling feeds into the core regulatory circuitry of ES cells (Cole et al., 2008). We show in Chapter 4 that stimulation of the Wnt pathway can promote iPS cell formation in the absence of c-Myc. In addition to a candidate approach, we developed a high-throughput chemical screening platform described in Chapter 5 to identify small molecules that replace the viral factors essential to reprogramming. One compound in particular, Kenpaullone (Kp), was identified that functionally replaces Klf4. This screening platform could be extrapolated and used to isolate replacements for the other reprogramming factors. High-throughput screening also offers the possibility of identifying novel compounds that provide mechanistic insight into the process of reprogramming, which is poorly understood.

The identification of chemical combinations that can achieve reprogramming in somatic cells may require sequential rounds of screening, using the previously identified compounds in the subsequent screening steps, rather than simply combining compounds identified from two independent experiments. This is highlighted in our work with the compound Kp. Although c-Myc has been shown to be dispensible for reprogramming in the presence of Oct4, Sox2 and Klf4 (Nakagawa et al., 2008; Wernig et al., 2008b), the replacement of Klf4 with Kp was dependent on c-Myc. Furthermore, the combination of Kp and Wnt3a was

unable to replace viral Klf4 and c-Myc, although independently they can replace the viral factors. Sequential high-throughput screening could also be combined with a low-throughput, candidate approach to comprehensively identify combinations of compounds that could be used to reprogram somatic cells. Chemical compounds could also be used to improve the efficiency of reprogramming, which becomes more critical as we replace the viral reprogramming factors with less efficient alternatives.

Published chemical screens to date have used primary infected cells to screen for compounds. Inherent in the use of primary infected cells is the genetic heterogeneity of the transduced cell population and the repeated need for high titers of retroviruses. One solution to the heterogeneity of transduced cell populations is the use of “secondary” cells. “Secondary” cells carry doxycycline-inducible proviruses that do not require new virus infection to induce reprogramming and can reprogram with higher efficiency than primary infected cells (Wernig et al., 2008a). More recently, improved secondary lines were isolated with single proviral copies of the reprogramming factors (Markoulaki et al., 2009). Some of these lines were missing one or more of the four transcription factors and were able to reprogram only if the missing factor(s) was virally transduced. These secondary lines would be ideal for use in a chemical screen to identify small molecules to replace the reprogramming transcription factors.

Inherent in the generation of chemically reprogrammed cells is the careful examination that chemical treatment does not cause genetic abnormalities or genomic instability in the resulting iPS cells. Genomic sequencing is one way in which this could be achieved, especially in cells used therapeutically. An additional method that could be used to replace viral reprogramming is transient protein transfection of the reprogramming factors. A recent publication details the generation of iPS cells using media supplemented with engineered proteins that penetrate the cell membrane of somatic cells and translocate to the nuclei (Zhou et al., 2009). While this method was highly inefficient and could only reprogram cells in the presence of the chemical valproic acid, it represents a viable alternative to viral reprogramming and one that does not introduce any genetic material into the somatic cells.

Although iPS cells could be used therapeutically as a donor source for cell transplantation, the ability to isolate patient-specific iPS lines opens new possibilities for human disease models (Dimos et al., 2008; Park et al., 2008; Soldner et al., 2009). The *in vitro* differentiation of patient-specific iPS cells offers the opportunity to model human disease with the exact genetic mutations that cause that disease rather than using mouse models, which can be artificial representations of disease processes. Inherent in their use in *in vitro* modeling systems, is the ability to differentiate iPS cells into multiple tissue lineages. This is another area where chemical screening platforms can identify molecules that will drive stem cell differentiation down specific developmental pathways. Not

only will this be important for *in vitro* modeling of diseases, but also in the differentiation of iPS cells for tissue transplantation and regenerative therapy.

I have discussed the use of chemical screens to isolate compounds to replace the reprogramming transcription factors, but one use of iPS cells could be in therapeutic drug testing (Rubin, 2008). Patient-specific iPS cells could be *in vitro* differentiated into the cell type affected by a particular disease and drug screens could then be performed to identify pharmaceuticals that alleviate specific cellular defects caused by that disease. Furthermore, iPS cells could revolutionize toxicity testing of pharmaceuticals. Rather than first identifying toxic side effects in clinical trials, an *in vitro* panel of tissue types, including hepatocytes, differentiatied from iPS cells could be analyzed before Phase I clinical trials. These examples suggest that the true benefit of iPS cells to human disease may be as an investigative, laboratory tool rather than as a direct source of transplantable tissue.

While reprogramming to a pluripotent state has broad clinical applications, the ability to convert one somatic cell type to another without going through a pluripotent intermediate, or “transdifferentiation,” also has therapeutic potential. By not going through a pluripotent intermediate, “transdifferentiated” cells may be less likely to form teratomas upon tissue transplantation and may therefore be more desirable to use therapeutically. A recent study on the transcription factor-induced reprogramming of pancreatic exocrine, acinar cells to endocrine, insulin-

producing β cells is an important advance in the understanding and feasibility of "transdifferentiation" (Zhou et al., 2008). It remains to be seen whether any cell type can be derived from a skin cell or whether there are epigenetic barriers to crossing germ layers or cell lineages. A more detailed knowledge of cellular signaling programs and transcription factor-induced cell fate changes will inform the future study of transdifferentiation.

In summary, the induction of pluripotency by defined transcription factors has broadened the field of stem cell research. The work in this thesis focused on the process of reprogramming, its mechanisms and how to make it therapeutically applicable. We are just beginning to understand the reprogramming process and the multitude of applications for which it can be used. We stand at the forefront of a new era in the understanding not only of pluripotency, but of human disease processes and the determination of cellular fate.

References:

Cole, M. F., Johnstone, S. E., Newman, J. J., Kagey, M. H., and Young, R. A. (2008). Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev *22*, 746-755.

Dimos, J. T., Rodolfa, K. T., Niakan, K. K., Weisenthal, L. M., Mitsumoto, H., Chung, W., Croft, G. F., Saphier, G., Leibel, R., Goland, R.*, et al.* (2008). Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. Science *321*, 1218-1221.

Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P., and Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature *458*, 771-775.

Markoulaki, S., Hanna, J., Beard, C., Carey, B. W., Cheng, A. W., Lengner, C. J., Dausman, J. A., Fu, D., Gao, Q., Wu, S.*, et al.* (2009). Transgenic mice with defined combinations of drug-inducible reprogramming factors. Nat Biotechnol *27*, 169-171.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol *26*, 101-106.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature *448*, 313-317.

Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., and Yamanaka, S. (2008). Generation of mouse induced pluripotent stem cells without viral vectors. Science *322*, 949-953.

Park, I. H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M. W., Cowan, C., Hochedlinger, K., and Daley, G. Q. (2008). Disease-specific induced pluripotent stem cells. Cell *134*, 877-886.

Rubin, L. L. (2008). Stem cells and drug discovery: the beginning of a new era? Cell *132*, 549-552.

Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G. W., Cook, E. G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M.*, et al.* (2009). Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. Cell *136*, 964-977.

Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G., and Hochedlinger, K. (2008). Induced pluripotent stem cells generated without viral integration. Science *322*, 945-949.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Wernig, M., Lengner, C. J., Hanna, J., Lodato, M. A., Steine, E., Foreman, R., Staerk, J., Markoulaki, S., and Jaenisch, R. (2008a). A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. Nat Biotechnol *26*, 916-924.

Wernig, M., Meissner, A., Cassady, J. P., and Jaenisch, R. (2008b). c-Myc is dispensable for direct reprogramming of mouse fibroblasts. Cell Stem Cell *2*, 10-12.

Woltjen, K., Michael, I. P., Mohseni, P., Desai, R., Mileikovsky, M., Hamalainen, R., Cowling, R., Wang, W., Liu, P., Gertsenstein, M.*, et al.* (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature *458*, 766-770.

Zhou, H., Wu, S., Joo, J. Y., Zhu, S., Han, D. W., Lin, T., Trauger, S., Bien, G., Yao, S., Zhu, Y.*, et al.* (2009). Generation of Induced Pluripotent Stem Cells Using Recombinant Proteins. Cell Stem Cell.

Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J., and Melton, D. A. (2008). In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. Nature *455*, 627-632.

Appendix A: Supplemental Data for Chapter 2

Figure S1

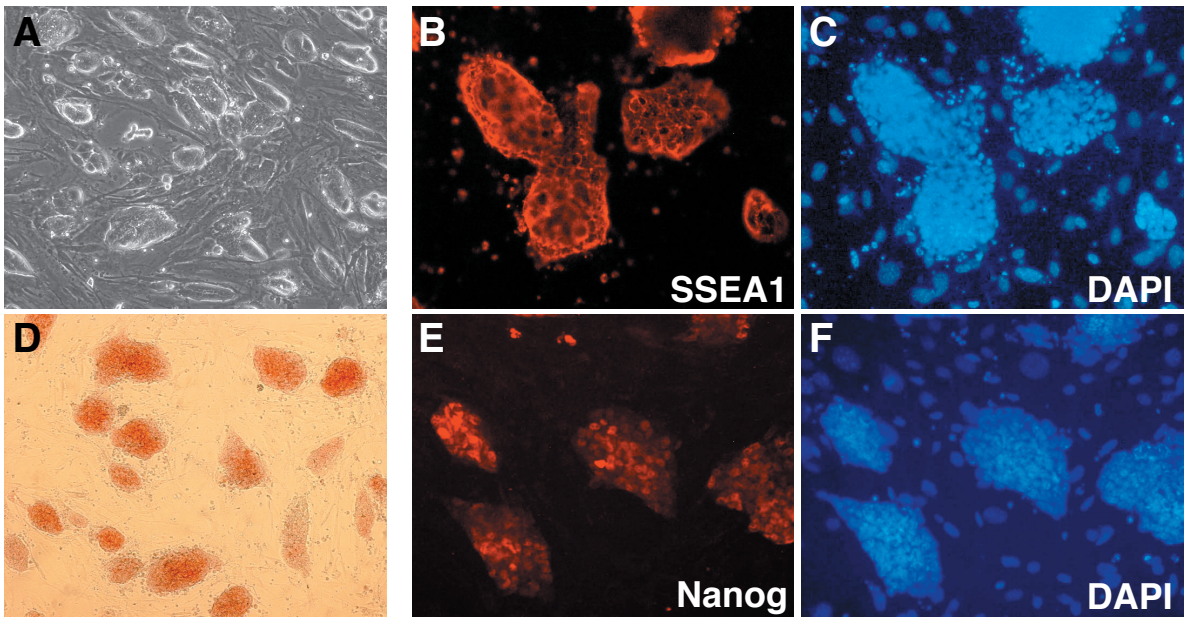


Figure S1: *In Vitro* Characterization of a Nanog-neo iPS cell line (clone 8). Nanog-neo resistant colonies closely resembling ES cells were selected to establish iPS cell lines. Nanog-neo iPS clone 8 grew as ES cell-like colonies on irradiated MEFs (A), was labeled homogenously with antibodies against SSEA1 (B, DAPI in C). Most colonies had a strong alkaline phosphatase activity (D) and a large fraction of the cells contained nuclear Nanog protein (E), DAPI (F).

Figure S2

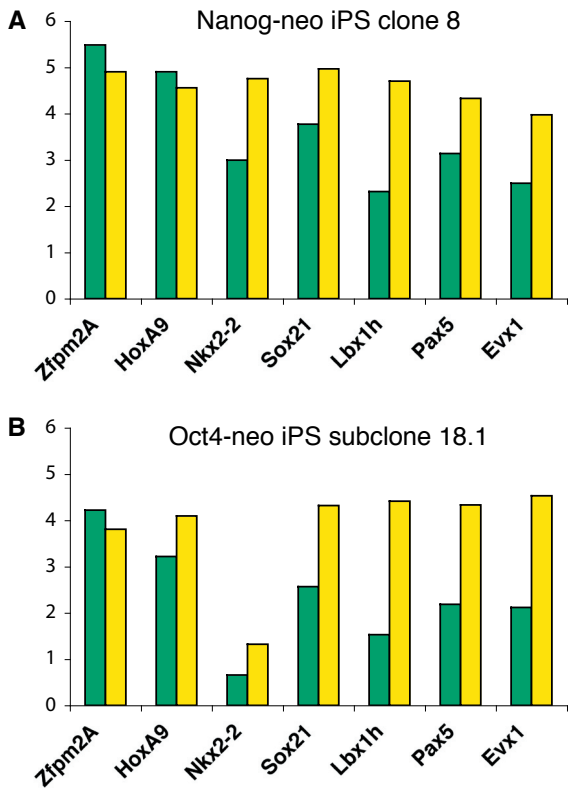


Figure S2: "Bivalent domains" in Nanog-neo iPS clone 8 (A) and subclone 18.1 of Oct4-neo iPS clone 18 (B). Shown are the log2 enrichments of chromatin immunoprecipitated material with antibodies specific for tri-methylated histone H3K4 and H3K27 as detected by quantitative RT-PCR. Green bars represent the "active" H3K4 mark, yellow bars the "repressive" H3K27 mark.

Figure S3

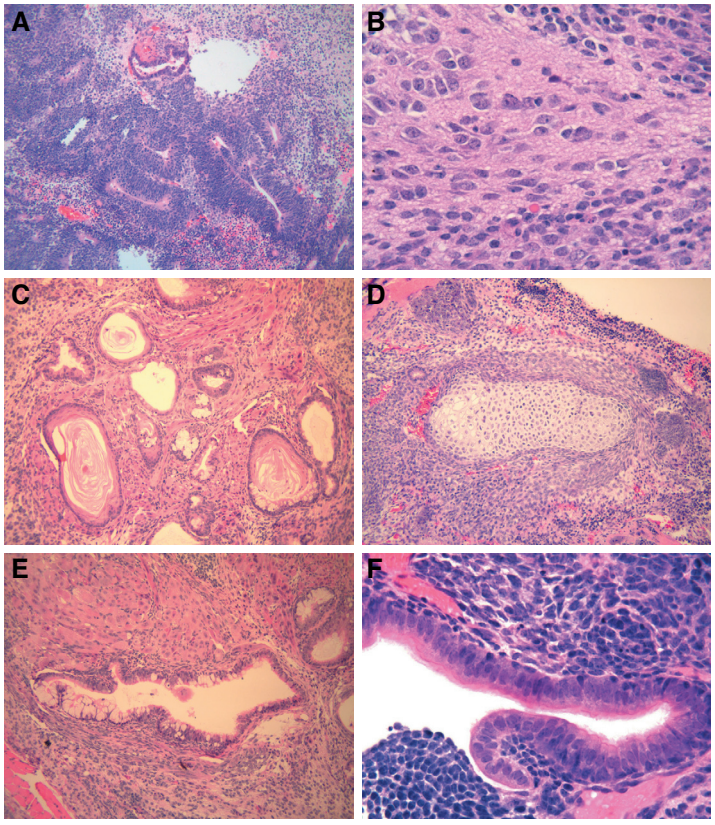


Figure S3 Multilineage differentiation of Oct4-neo iPS cells following injection into SCID mice. Sixteen to 26 days after subcutaneous injection of iPS cells resulting tumors were fixed, embedded in paraffin, sectioned and stained with Hematoxylin and Eosin. Somatic cell types found included proliferative neuroectoderm (A), mature central nervous system (CNS) tissue (B), keratinizing squamous epithelia (C), cartilage (D), cuboidal epithelia containing goblet cells (E) and ciliated cuboidal epithelia (F).

Figure S4

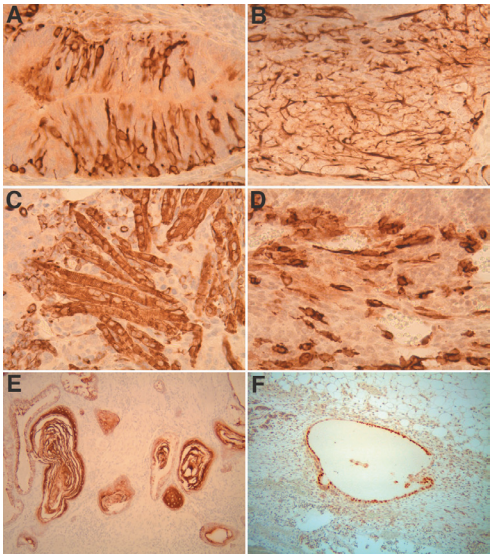


Figure S4: Immunohistochemical characterization of teratomas. A) Postmitotic neurons within a neural tube-like structure were labeled with antibodies against beta-III-tubulin, B) Glial fibrillary acidic protein (GFAP)-positive astrocytes in another part of the tumor resembling mature CNS tissue C) Desmin-positive striated muscle-fibers, D) Smooth muscle actin (SMA) expressing cells) E) Cytokeratin expression in epidermis-like epithelial cells, F) GATA4-positive endothelium.

Figure S5

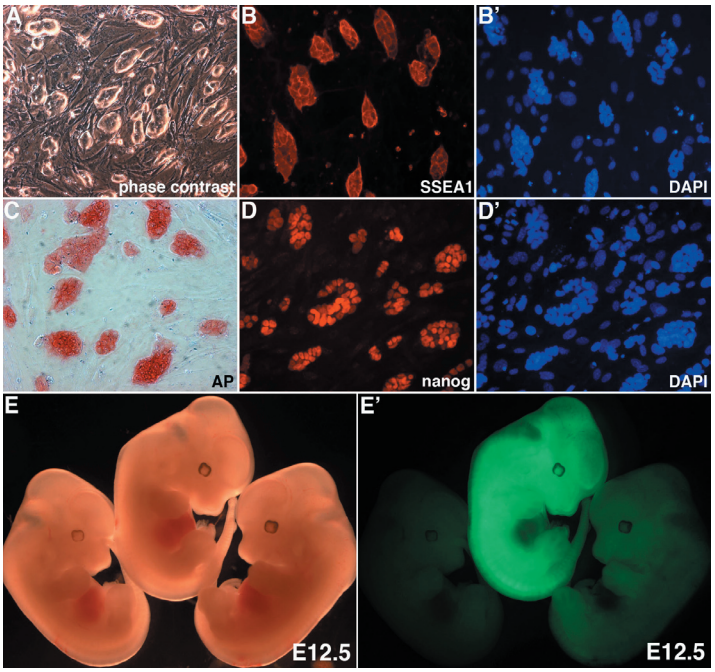


Figure S5: Characterization of tailtip fibroblast-derived iPS cells. iPS cell lines were established from tailtip fibroblasts utilizing Nanog or Oct4 selection. A-D) In vitro characterization of Oct4-iPS clone TT-O16 displaying ES cell-like morphology (A), homogenous staining for SSEA1 (B, corresponding DAPI image B'), alkaline phosphatase (C), Nanog (D, corresponding DAPI image D'). Oct4-iPS Clone TT-O25 was injected into diploid blastocysts resulting in high contribution mid-gestation embryos (E bright field, E' EGFP fluorescence).

Table S1

cell line	karyotype (20 cells analyzed per cell line)
O3-2	40, XY (20)
O9	40, XY (16); 39, X0 (1); various losses or gains of other chromosomes (3)
O18	41, XY+12 (17); 42, XY+10+12 (3)
N8	40, XY (19); 39, X0 (1)
N14.2	40, XY (20)

Appendix B: Supplemental Material for Chapter 3

Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells

Tobias Brambrink, Ruth Foreman, G. Grant Welstead, Christopher J. Lengner, Marius Wernig, Heikyung Suh and Rudolf Jaenisch

Figure S1: Quantitative RT-PCR assays of viral trangene expression in infected MEFs.

MEFs were infected and grown in the absence or the presence of dox. Average relative expression levels and standard deviations from two RT-PCR reactions are shown for all four factors. ES cell values are given as a positive control.

Figure S2: FACS analyses of AP, SSEA1 and GFP expression in infected MEFs and control ES cells.

FACS analysis of AP, SSEA1 and GFP reactivation was performed on NanogGFP/M2rtTA MEFs and Oct4GFP/M2rtTA MEFs at different times after the induction of reprogramming. Cells were harvested at various time points after the addition of dox to the medium and stained with an APC labeled anti-SSEA1 antibody and a fluorescent substrate detecting AP activity. Representative FACS plots of three independent experiments are shown. The fraction of AP positive cells increases with time after transgene induction. In the histogram plots of AP staining, the red line represents the negative control (infected MEFs cultured without dox) and the solid purple represents induced MEFs analyzed at the

specified time (A). The numbers displayed for each plot are the number of cells in the M1 gate which was set so that less than 1% of cells in the negative control were in the M1 population. Analysis of SSEA1 and GFP expression (B). Dot plots of GFP and SSEA1 signals are displayed and the percentages of cells for the SSEA1+/GFP- and the SSEA1+/GFP+ populations are shown. AP positive cells were first observed on day 3, while SSEA1 positive cells appeared on Day 9. Cells positive for Nanog-GFP or Oct4-GFP were first detected on day 16 and most of these cells also stained positive for SSEA1.

Figure S3: Morphology of non-reprogrammed cell colonies.

Pictures of cell colonies on induced plates at the time specified (A). Plates were re-assessed at day 35 of the experiment and pictures of cell colonies that were detectable are shown (B).

Figure S1

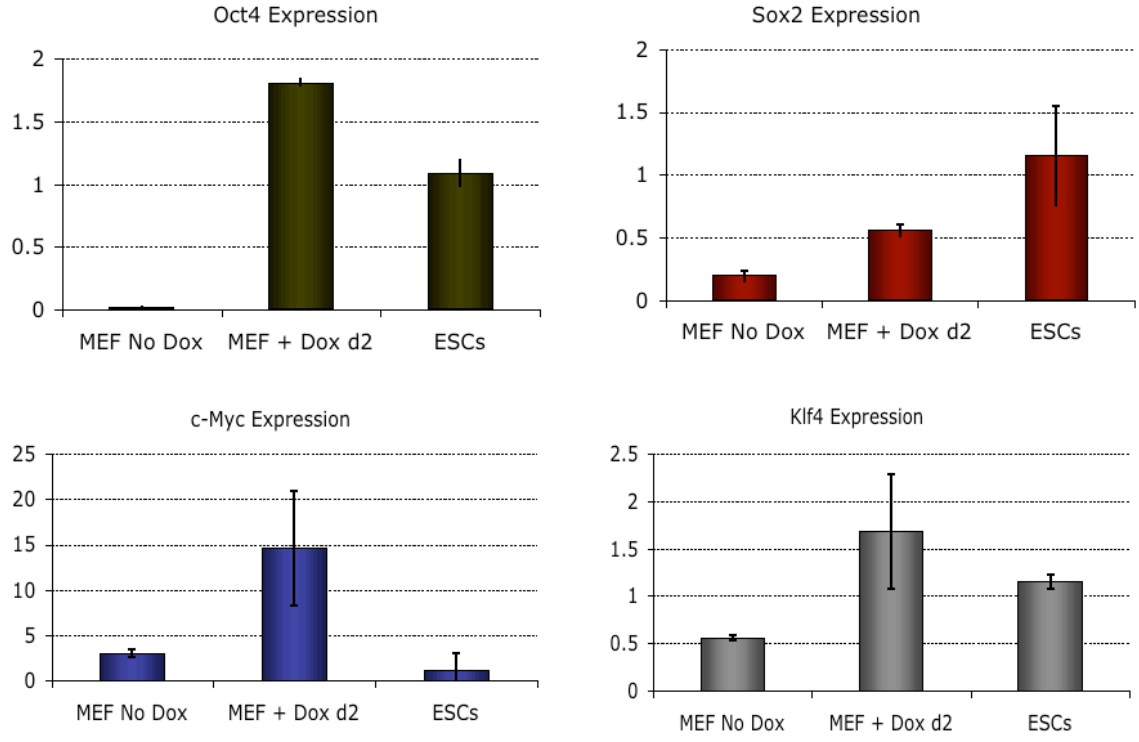

Oct4 Expression
2
1.5
1
0.5
0
MEF No Dox
MEF + Dox d2
ESCs
Sox2 Expression
2
1.5
1
0.5
0
MEF No Dox
MEF + Dox d2
ESCs
c-Myc Expression
25
20
15
10
5
0
MEF No Dox
MEF + Dox d2
ESCs
Klf4 Expression
2.5
2
1.5
1
0.5
0
MEF No Dox
MEF + Dox d2
ESCs

Figure S2

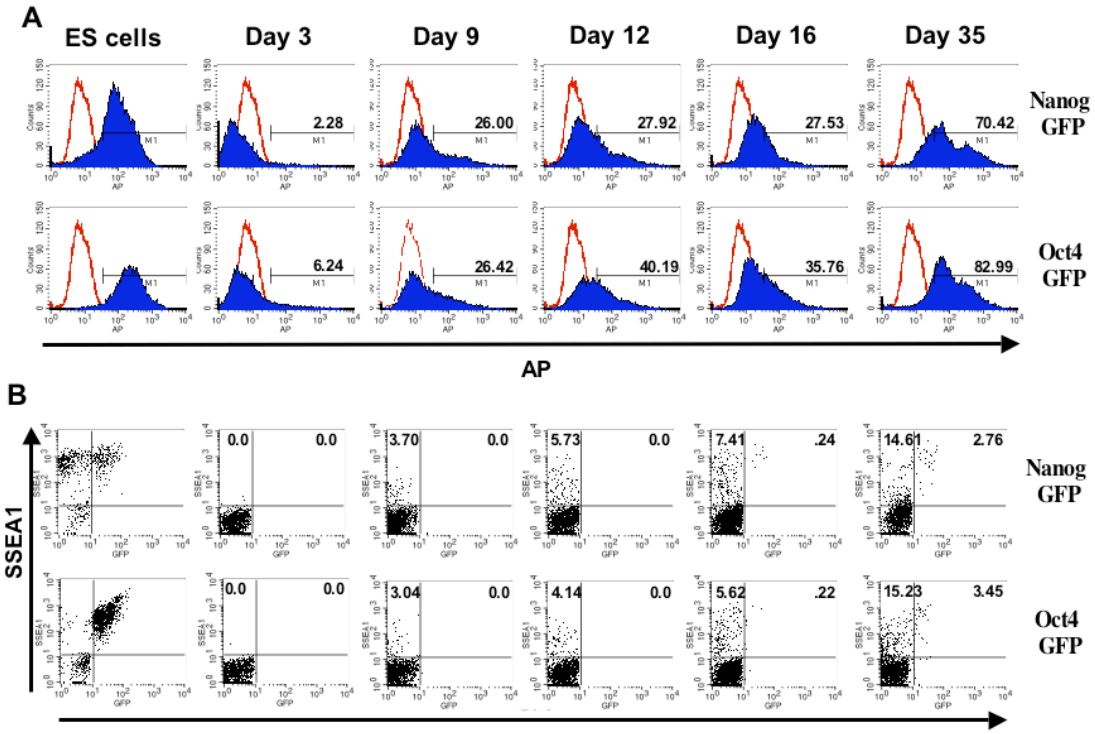
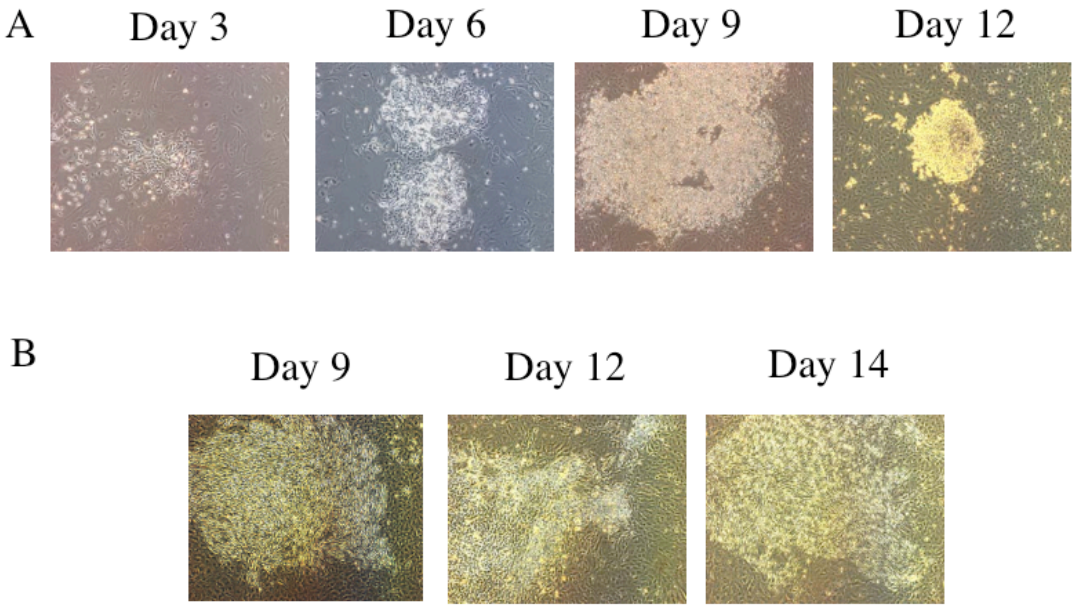

A
ES cells
Day 3
Day 9
Day 12
Day 16
Day 35
Nanog GFP
Oct4 GFP
2.28
26.00
27.92
27.53
70.42
6.24
26.42
40.19
35.76
82.99
AP
B
SSEA1
GFP
0.0
0.0
3.70
0.0
5.73
0.0
7.41
.24
14.61
2.76
0.0
0.0
3.04
0.0
4.14
0.0
5.62
.22
15.23
3.45
Nanog GFP
Oct4 GFP

Figure S3



Appendix C: Supplemental Data for Chapter 4

Wnt signaling promotes reprogramming of somatic cells to pluripotency

Alexander Marson, Ruth Foreman, Brett Chevalier, Steve Bilodeau, Michael Kahn, Richard A. Young, and Rudolf Jaenisch

Supplemental Experimental Procedures

Mice, cell lines, retroviral vectors, cell culture, flow cytometry, immunostainings, teratoma formation and blastocyst injections were previously described in detail (Wernig et al., 2007; Meissner et al., 2007; Brambrink et al., 2008). Wnt3a conditioned media (Willert et al., 2003) and control conditioned media from parental L cells were generated according to standard protocols (ATCC), and used in a 1:1 dilution with standard ES cell medium. It has been previously shown that the conditioned media from Wnt3a-overexpressing cells has Wnt signaling activity as determined by TOP flash assays (Ogawa et al., 2006) and that functional Wnt3a protein can be purified from this media (Willert et al., 2003). Furthermore, secretion of Wnt3a protein can be detected by immunoblot of the conditioned media from Wnt3a-overexpressing cells (Shibamoto et al., 1998). ICG-001 was dissolved in DMSO to a stock concentration of 0.1M and used at a final, working concentration of 4uM (or as indicated in Supplemental Figure S2).

References

Brambrink, T., Foreman, R., Welstead, G. G., Lengner, C. J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell *2*, 151-159.

Meissner, A., Wernig, M., and Jaenisch, R. (2007). Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. Nat Biotechnol *25*, 1177-1181.

Ogawa, K., Nishinakamura, R., Iwamatsu, Y., Shimosato, D., and Niwa, H. (2006). Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. Biochem Biophys Res Commun *343*, 159-166.

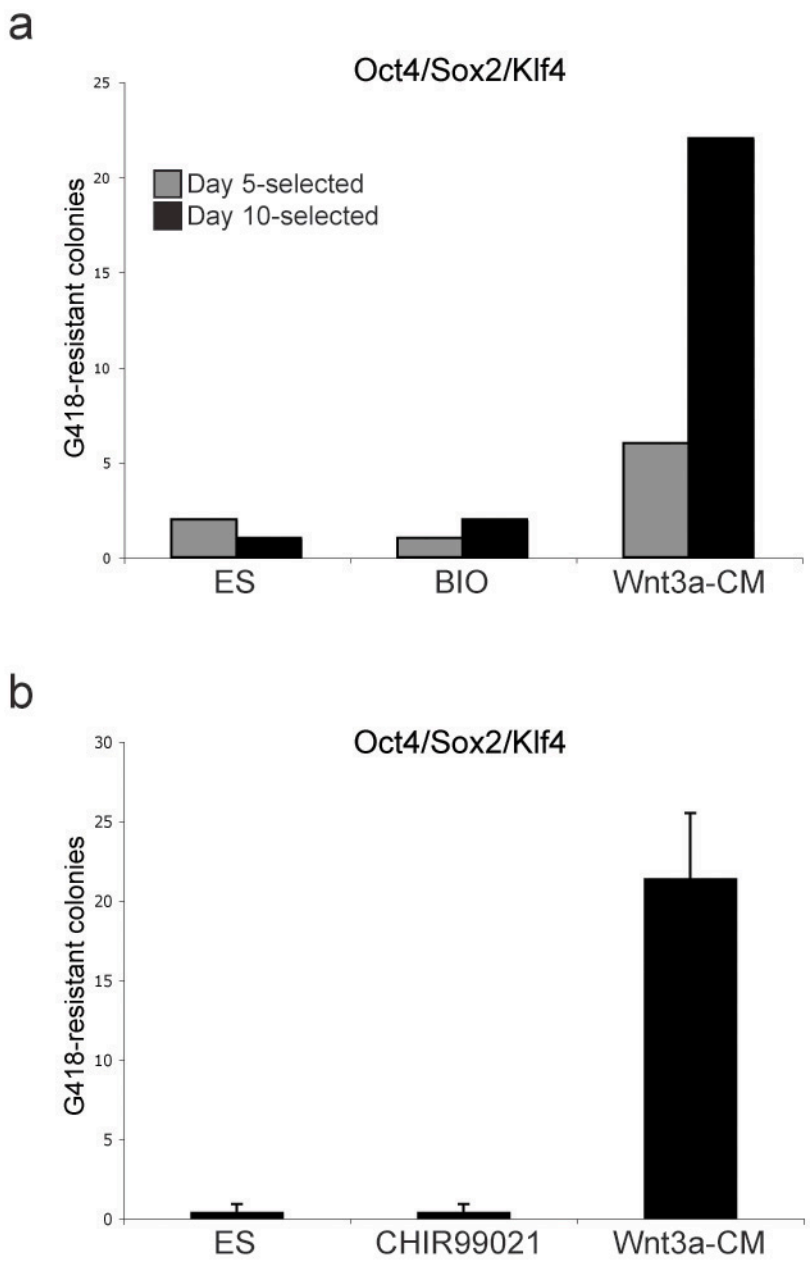
Shibamoto, S., Higano, K., Takada, R., Ito, F., Takeichi, M., and Takada, S. (1998). Cytoskeletal reorganization by soluble Wnt-3a protein signalling. Genes Cells. *3,* 659-670.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B. E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.

Willert, K., Brown, J. D., Danenberg, E., Duncan, A. W., Weissman, I. L., Reya, T., Yates, J. R., 3rd, and Nusse, R. (2003). Wnt proteins are lipid-modified and can act as stem cell growth factors. Nature *423*, 448-452.

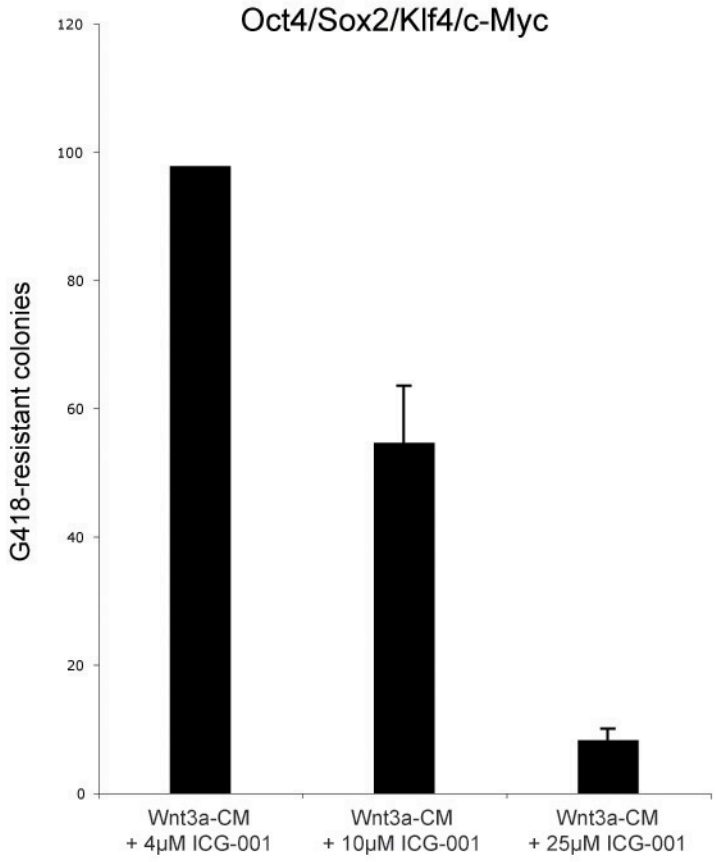
Cell Line	Injected Blastocysts	Live Chimeras/Live Births	Germline Transmission
KSOW49	21	1/2	Yes
KSOW52	37	1/7	Yes

Supplemental Table S1. Two clonal cell lines (KSOW49 and KSOW52) derived from Oct4-GFP MEFS overexpressing Oct4/Sox2/Klf4 in the presence of Wnt3a-CM were injected into blastocysts. Total number of blastocysts, live births and live births with chimeric contribution from iPS cells (as determined by coat color) are shown. Both lines were competent to contribute to the germline of chimeric mice (shown here in and in Figures 1E and 1F).

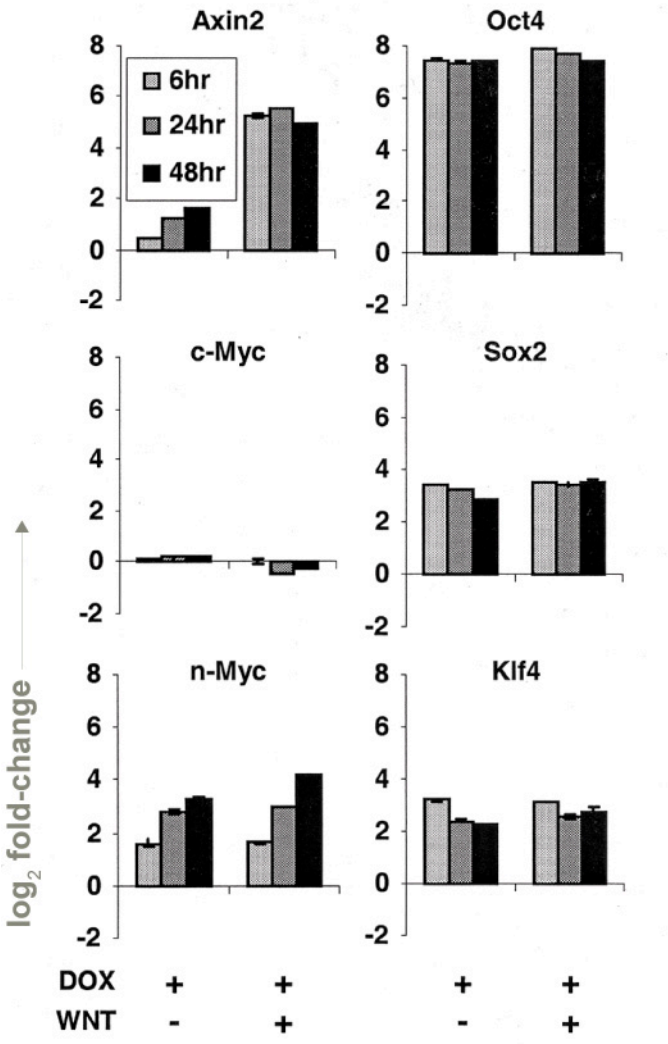


Supplemental Figure S1 Gsk3 inhibitors and reprogramming in the absence of c-Myc retrovirus. a. MEFs with doxycycline (DOX)-inducible Oct4/Sox2/Klf4 were cultured in ES cell medium (ES), ES cell medium supplemented with the glycogen synthase kinase-3 (Gsk-3) inhibitor 6-bromoindirubin-3'-oxime (BIO,Sigma), or Wnt3a over-expressing conditioned media (Wnt3a-CM). G418 selection was initiated on Day 5 (gray bars) or Day 10 (black bars) post-

doxycycline induction as indicated. Counts are shown for G418-resistant colonies. **b.** MEFs with doxycycline (DOX)-inducible Oct4/Sox2/Klf4 were cultured in ES cell medium (ES), ES cell medium supplemented with the Gsk-3 inhibitor CHIR99021 (Ying et al., 2008), or Wnt3a over-expressing conditioned media (Wnt3a-CM). Mean number of G418-resistant colony counts from triplicate experiments is displayed with error bars indicating S.D. Although neither BIO nor CHIR99021 promoted reprogramming in these experiments, these small molecules could prove useful at different doses or in combination with other chemicals.



Supplemental Figure S2 Effect of ICG-001 on reprogramming with four factor-induced MEFs. MEFs with doxycycline (DOX)-inducible Oct4/Sox2/Klf4/c-Myc were cultured in Wnt3a over-expressing conditioned media (Wnt3a-CM) with 4μM, 10μM or 25μM ICG-001. G418 selection was initiated on Day 15 post-doxycycline induction. Counts (in a 32cm^2 area) are shown for G418-resistant colonies. Mean number of G418-resistant counts from triplicate experiments is displayed with error bars indicating S.D. Because G418-resistant colonies on one plate of cells treated with Wnt3a-CM and 4μM ICG-001 were too dense to count, data shown for this concentration of ICG-001 only reflect with mean of duplicate data.



Supplemental Figure S3 Absence of endogenous *Myc* induction with Wnt3a-CM treatment. Quantitative RT-PCR analysis of RNA extracted (RNeasy, Qiagen) from *Oct4*-inducible, *Oct4*-G418 selectable cells transduced with Klf4 and Sox2. Cells were either treated with doxycycline (DOX), or doxycycline and Wnt3a-CM (WNT), as indicated. Transcript levels, assessed with Taqman probes (Applied Biosystems), were normalized to Gapdh levels. Values are reported as fold-change relative to cells not treated with DOX or WNT. Error bars indicate standard deviation derived from triplicate PCR reactions. No Wnt-dependent activation of *c-Myc* was observed in the population of cells examined. However,

we cannot exclude that Wnt3a enhances *c-Myc* expression in the rare cells that will eventually become iPS cells.

Appendix D: Supplemental Data for Chapter 5

Reprogramming of murine fibroblasts to iPS cells: Chemical Complementation of Klf4

Costas A. Lyssiotis, Ruth K. Foreman, Judith Staerk, Michael Garcia, Divya Mathur, Styliani Markoulaki, Jacob Hanna, Luke L. Lairson, Bradley D. Charette, Laure C. Bouchez, Michael Bollong, Conrad Kunick, Achim Brinker, Charles Cho, Peter G. Schultz and Rudolf Jaenisch

Supplemental Text

Structure and activity relationship (SAR) analysis of a 100 member paullone collection (1, 2) in the Nanog-luciferase assay demonstrated that the core paullone scaffold is sensitive to modification and that kenpaullone is the most active member (SI Fig. 5 and SI Table 1). Interestingly, however, a variety of analogs that have no kinase inhibitory activity against canonical kenpaullone kinase targets (i.e., CDK1, CDK5, GSK-3β) were still active in the Nanog reporter assay (SI Fig. 5). One analog (Kun152) that was active in the Nanog assay (2.4-fold relative to DMSO-treated controls), ineffectual toward the canonical paullone kinases (>10 μM) and did not activate Wnt signaling (SI Fig. 5F) was selected for further kinase profiling. Kun152 demonstrated negligible activity against a 57 member tyrosine kinase profile at 5 μM in vitro (SI Table 2), suggesting that it is not a non-selective kinase inhibitor. Collectively these data, together with the data Figure 3F and SI Figure 4, suggest that the Nanog activating ability of the paullone scaffold does not result from inhibition of the canonical kinase targets nor via general tyrosine kinase inhibition.

Supplemental Methods

Characterization of kenpaullone (3). 9-Bromo-7,12-dihydro-indolo[3,2-*d*][1]benzazepin-6(*5H*)-one, colorless crystalline solid, $C_{16}H_{11}BrN_2O$. Purity of purchased compound was checked by NMR and MS respectively. ^{1}H NMR (400MHz, DMSO-d_6): _ = 11.81 (*br*.s, 1H, NH), 10.13 (*br*.s, 1H, NH), 7.92 (d, *J* = 2.0 Hz, 1H), 7.73 (dd, *J* = 8.2, 1.4 Hz, 1H), 7.42-7.37 (m, 2H), 7.32-7.24 (m, 3H), 3.35 (s, 2H). ^{13}C NMR (100MHz, DMSO-d_6): _ = 171.4, 136.0, 135.6, 134.0, 128.5, 128.4, 128.3, 127.0, 124.5, 123.7, 122.3, 120.4, 113.4, 111.7, 107.2, 31.3. MS (EI): m/z: 327, 328, 329, 349, 351.

Chemicals. Synthesis of the 100 member paullone collection was described previously (2, 4, 5). CDK and GSK-3β inhibitors were purchased from Calbiochem with the exception of kenpaullone and purvalanol A, which were purchased from Sigma and CHIR99021 (Axon MedChem).

Gene Targeting and Blastocyst Injection. The Nanog-luciferase targeting vector was generated by first subcloning a PGK-neomycin cassette flanked by 2 loxP sites into a pSP72 vector with BamHI and XhoI. The 5' targeting arm was inserted into this BamHI site and a ClaI site in the pSP72 vector. The 5' arm spanned 1188bp of the *Nanog* promoter and was flanked by a ClaI site at the 5' end and a BamHI site at the 3' end. These restriction sites were added to the PCR primers that were used to amplify the 5' arm from a BAC containing this region of the *Nanog* promoter. The primers for PCR of the 5' arm were as follows: Forward (ATCGATCTGGGTTAGAGTGTCTTTCACTCAC) and Reverse (GGATCCGTCAGTGTGATGGCGAGGGAAGGG). The 3' arm spanned 1.4kb of *Nanog* intron 2 and was amplified from a BAC using the following PCR primers: Forward (GCGGCCGCGTAAGGAATTCAGTCCCCGAA) and Reverse (GCGGCCGCCTCGAGGCCCTCTTCTGGAGTGTCTGAAGAC). The 3'arm was cloned into the XhoI site of the vector. A 1.9kb NcoI/BamHI fragment of pGL3-basic (Promega)

containing the firefly luciferase gene and a SV40 pA signal was blunt-end ligated into the vector.

For targeting, the Nanog-luciferase vector was linearized with XhoI and electroporated into ES cells. G418 selection (350 μg/ml) was started 24 hours after electroporation and continued for 10-12 days. Resistant clones were picked and expanded for screening by Southern blot analysis. DNA from resistant clones was digested with PvuII and run on a 0.8% agarose gel. 3' end homologous recombination was examined by a 415bp external probe that was PCR-amplified from *Nanog* intron 2 using primers: Forward (GCTACCTGAGACCCTATCCCTTAG) and Reverse (CATCTCACCAGCCCTACATACAGTG). This 3' external probe identified the *Nanog* wild-type allele 6.4kb PvuII fragment and the Nanog-luciferase targeted allele 4.7kb PvuII fragment. 5' end homologous recombination of resistant clones was analyzed using a 340bp XbaI/HindIII fragment of the Nanog-luciferase targeting vector. This internal probe identified a wild-type 6.4kb PvuII fragment and a targeted 5.2kb PvuII fragment.

Cell Culture and Viral Infections. ONR-MEFs (at passage 1-2; 5 _ 10^4 cells/well of a 6-well dish) were infected overnight with concentrated viral supernatant. Virus was generated by transfection (FugeneHD, Roche) of GP2-293 packaging cells (Clontech) with moloney based retroviral vector pMXs containing the cDNA of Oct4, Sox2 and c-Myc +/-Klf4 (Addgene, deposited by S. Yamanaka) together with the VSV-G packaging plasmid. Media and compound (or vehicle control; DMSO 0.05%) were changed every 2-3 days. At day 20, selection was initiated by adding G418 (300 μg/mL) to the culture media. Colonies were counted – based on survival – or passaged onto irradiated MEFs in ES cell growth media at day 25.

OG4-MEFs (at passage 1-2; 10^4 cells/well of a 6-well) were infected with pooled viral supernatant generated by transfection of HEK293T cells (Fugene, Roche) with doxycycline inducible lentiviral vectors (6) containing the cDNAs of Oct4, Sox2 and c-Myc +/- Klf4, together with a plasmid encoding the VSV-G envelope protein. Media was supplemented the following day with doxycycline (2 μg/mL) and kenpaullone or DMSO. Green colonies were selected two weeks later and expanded under ES cell growth conditions.

NPCs (at passage 4-5; 10^5 cells/well of a 6-well dish) were infected overnight with concentrated viral supernatant generated with the following plasmids; pMXs-Oct4 +/- pMXs-Klf4 and pCIGAR-Myc (an MSCV-based, bicistronic retroviral vector modified to permit Gateway-mediated insertional recombination of full-length or ORF transgenes immediately upstream of an IRES-eGFP marker). Compounds were added 12 hours later and changed every two days thereafter. At day 8, colonies were either fixed and stained for AP or passaged onto irradiated MEFs in ES cell growth media.

Colony forming assays with secondary OSM-MEFs were performed as described previously (7) and analyzed for Oct4 expression 25 days later.

Timecourse Analysis. OSM-transduced NL-MEFs were plated at 1K cells/well in ES cell growth media, treated with kenpaullone (5 μM) and assayed at the indicated time points. All luciferase readings were acquired following addition of Bright-Glo reagent (Promega). Cell number was normalized at each time point by assaying parallel plates

for total ATP levels using the cell titer-glo reagent (Promega). The assay was run in 384-well plates (Greiner).

Quantitative and Semi-Quantitative RT-PCR. Total RNA was obtained and reverse transcribed, as described previously (8). TaqMan real-time PCR was performed using the TaqMan Gene Expression Master Mix (Applied Biosystem) and ABI Prism 7900HT fast real time PCR system with Klf4 probes (Applied Biosystems: Mm00516105_g1, MGB probe) and normalized to GAPDH (4326317E, VIC probe) as endogenous controls.

For semi-quantitative analyses, PCR was performed using the Phusion polymerase (New England Biosciences) and gene-specific primers (SI Table 4). Cycle parameters were 10s at 98°C, 15s at 51°C (CDK1), 54°C (CDK2), 63°C (GSK-3β), 61°C (CDK5) or 67°C (Nat1), and 30s at 72°C for 30 cycles. The amount of total RNA and PCR conditions were optimized so that amplification of both *Nat1* and the cDNAs of interest were in the exponential phase.

Western Blotting. Protein extraction, preparation, gel electrophoresis and transfer were performed as described previously (9). Membranes were blocked (5% non-fat dry milk in Tris buffered saline with 0.5% Tween-20), washed and incubated with primary antibodies (Klf4, Santa Cruz; γ-tubulin, Sigma; β-actin,) in blocking solution for 2 hr at room temperature or 4˚C overnight, washed and incubated with the corresponding horse radish peroxidase-conjugated secondary antibodies at room temperature for 1 hr. Finally, membranes were washed and subjected to enhanced chemi-luminescence detection.

Short-Hairpin Knock-downs. Lentiviral particles containing short hairpin (sh)RNAs were generated by transfecting HEK293T cells (FugeneHD, Roche) with pooled pLL3.7 vectors encoding the desired knock-down (SI Table 5), packaging vectors and VSV-G. 72 hours later, viral supernatant was collected, concentrated (Amicon) and applied to OSM-transduced NL-MEFs. Individual cell lines were made for each kinase knock-down. 2 days later, cells were sorted for pLL3.7-driven GFP; flow cytometric analysis indicated that infectivity was >90%. Combinatorial knock-down was achieved by treating single knock-down lines with lentiviral particles targeting additional kinases.

ONR-MEFs at passage 1-2 were infected with pooled viral supernatant containing shRNAs targeting CDK1, CDK2, CDK5 and GSK-3β. 2 weeks later, cells were fixed and AP stained.

SI Fig. 1 Nanog-luciferase validation assay. OSM-transduced NL-MEFs were treated with a known drug compound plate in 1536-well format and analyzed 10 days later for luciferase activity; values represent the mean of three independent runs. Nanog activity is reported in relative light units (RLU).

SI Fig. 2 Timecourse of Nanog activation. OSM-transduced NL-MEFs were treated with kenpaullone (5 μM) or DMSO (0.05%) and analyzed for luciferase activity as a function of time. Nanog expression is reported in RLU and is normalized for cell number (obtained via analysis of ATP concentration; read in parallel plates). Error bars, standard deviation (n=3).

SI Fig. 3 Retroviral vectors are silenced in iPS cells derived from NPCs with Oct4/c-Myc and kenpaullone. NPCs were transduced with Oct4 and a c-Myc vector harboring an IRES-GFP cassette. **(A)** After eight days of kenpaullone treatment (5 μM), definitive colonies **(B)** expressing GFP were visible. **(C)** Colonies passaged under mouse ES cell conditions resembled murine ES cells morphologically and **(D)** did not express GFP, consistent with the silencing of retroviral elements observed in reprogrammed cells.

SI Fig. 4 Dual inhibition of CDKs and GSK-3β fail to recapitulate the activity of kenpaullone. **(A)** OSM-transduced NL-MEFs were treated in a dose dependent fashion with CDK or **(B)** GSK-3β inhibitors in 384-well format and analyzed for luciferase activity 10 days later. **(C)** OSM-transduced NL-MEFs were plated in screening media containing a potent, selective GSK-3β inhibitor (CHIR99021, 3 μM) or **(D)** a general CDK inhibitor (purvalanol A, 3 μM) and treated in a dose dependent manner with CDK or GSK-3β inhibitors, respectively. 10 days after treatment, the cells were analyzed for luciferase expression. Data points represent the mean of triplicate runs and are representative of at least 3 independent experiments. Values that drop off sharply at high concentration indicate cellular toxicity. Kinase inhibition data are presented in SI Table 3.

SI Fig. 5 Paullone induced activation of Nanog activity does not result from inhibition of the canonical paullone kinase targets. **(A)** Paullone scaffold. Atom labeling corresponds to the SAR data in SI Table 1. **(B)** Structure of kenpaullone and the paullone analog, Kun152, profiled for kinase activity. Kinase inhibition data is detailed in SI Table 2. **(C)** Plots of Nanog activity versus inhibition of CDK1, **(D)** CDK5 or **(E)** GSK-3_. Each point on the graph represents a different member of the paullone collection. Kenpaullone is depicted in red. Kinase inhibition data was obtained previously (1), and is plotted as the IC50 value in μM. **(F)** Nanog activity was plotted against TCF/LEF-driven luciferase activity – representative of GSK-3β inhibition (10) – for the 100-member paullone library. Kun152 is depicted in orange. **(G)** OSM-transduced NL-MEFs were treated with lentiviral particles containing short hairpin (sh)RNAs targeting the canonical paullone kinase targets, either individually or combinatorially, and analyzed for luciferase expression 10 days later. Kenpaullone (5μM) was included as a positive control. **(H)** Lentiviral particles containing shRNAs targeting Cdk1, Cdk2, Cdk5 and Gsk-3_ were pooled and applied to OSM-MEFs. Cells were AP stained two weeks later. Klf4 and kenpaullone (5μM) were included as positive controls. **(I)** RT-PCR analysis of OSM-

transduced NL-MEFs for the indicated knock-down(s). Data points presented in **A-D** represent the mean of triplicate runs and are representative of at least 3 independent experiments. Nanog and Wnt activity for the paullone analogs were plotted at the most active, non-toxic compound concentration. Error bars in **G**, standard deviation (n=5).

SI Fig. 6 Klf4 is not activated by kenpaullone. **(A)** OSM-transduced MEFs were treated with DMSO, retrovirally-delivered Klf4 or kenpaullone (5 μM). Klf4 expression was analyzed at 3 day intervals by quantitative RT-PCR and normalized to GAPDH. Error bars, standard deviation (n=3). **(B)** Secondary MEFs that carry doxycycline (DOX)-inducible proviruses encoding Oct4, Sox2 and c-Myc (7) were treated with kenpaullone (5 μM) or vehicle control and western blotted for Klf4 expression at days 4 and 7. **(C)** Secondary MEFs carrying a DOX-inducible proviral copy of Klf4 (TetO-Klf4) were treated with kenpaullone (5 M), vehicle control or DOX (positive control). Wild type (wt) MEFs were treated with kenpaullone (5 μM), a GSK-3β specific inhibitor (CHIR99021, 3 μM) or vehicle control. Samples were analyzed by western blot 92 hours later. Data are representative of three independent experiments.

SI Table 1 Paullone SAR. Effects of substitution to the paullone scaffold (SI Fig. 5A) as determined by Nanog activity. Bolded data represent modification to the paullone carbon backbone.

Position	Substitution
1	**C = N > S**
2	H > OH > OCH_3 > short alkyl chain >>> bulky substituents (not tolerated)
2	**C >> N (nitrogen abolishes activity)**
3	H > OCH_3 > OH
4	H >> OH > OCH_3
4	**C > N**
5	H > CH_3 > hydroxy-amidine > p/Cl-Bz >> bulky substituents
9	Br = I = CH_3 > CF_3 > OCH_3 > Cl >>> bulky and (F,CN,OH,H,CO_2H, NO_2) not tolerated
11	**C > N**
12	**N >> O**
12	H > CH_3 >>> bulky substituents (not tolerated)

SI Table 2 Tyrosine kinase panel. Data is represented as percent kinase inhibition (% i) in the presence of 5 μM Kun 152. Kun152 increases the Nanog luciferase signal 2.4-fold, relative to DMSO-treated controls.

Kinase	% i
Abl	7
Abl (T315I)	12
Akt1	4
AMPK	10
AurB	15
AXL	11
BMX	-8
BTK	-7
CaMK2a	31
CDK2/cycA	19
CHK2	12
CK1a	5
cKit	-2
c-RAF	6
CSK	3
DYRK1a	4
EGFR	-2
EphA3	7
EphB3	2

Kinase	% i
FAK2	7
FGFR3	1
Flt3	15
FMS/CSF1R	31
Fyn	5
GSK3b	52
IGF1R	-12
IKKb	1
InsR	-29
IRAK4	1
JAK2	-1
JNK2a2	5
KDR	14
Lck	8
Lyn	23
MAPK1	3
MAPK13	-7
MAPK14	-6
MAPKAPK2	7

Kinase	% i
MET	5
NEK2	7
p70S6K	4
PAK2	1
PDGFRa	-6
PDK1	6
Pim2	3
PKA	-1
PKCa	-2
PLK1	6
RET	1
ROCK2	5
RSK1	3
SGK1	8
Src	11
Syk	1
TrkA	10
TTK	8
ZAP70	-14

SI Table 3. Kinase inhibition profile for the small molecule inhibitors used in this study. Class indicates whether the small molecule is a CDK inhibitor (CDKi), GSK-3β inhibitor (GSKi) or an inhibitor of both CDKs and GSK-3β (dual). Data is reported as the IC50 in μM for *in vitro* kinase inhibition. n/a, data not available.

Compound	Class	Cdk1	Cdk2/Cyclin A	Cdk2/Cyclin E	Cdk5	GSK-3β
Kenpaullone (11, 12)	Dual	0.4	0.68	7.5	0.85	0.23
Flavopiridol (13)	Dual	0.4	n/a	n/a	n/a	0.45
Aloisine (14)	Dual	0.7	n/a	n/a	1.5	0.92
Aloisine A (14)	Dual	0.15	0.12	0.4	0.2	1.5
BIO (13, 15)	Dual	0.32	0.3	n/a	3	0.005
Hymenialdisine (13, 16)	Dual	0.022	0.07	0.04	0.028	0.01
Purvalanol A (17)	CDKi	0.004	0.07	0.035	0.075	>10
Roscovatine (18)	CDKi	0.65	0.7	0.7	0.2	n/a
NU6102 (19)	CDKi	0.01	0.0054	n/a	n/a	n/a
AR-A014418 (13, 20)	GSKi	n/a	>100	>100	>100	0.104
CHIR99021 (13, 21)	GSKi	8.8	n/a	n/a	n/a	0.007

SI Table 4. List of primers used for semi-quantitative RT-PCR.

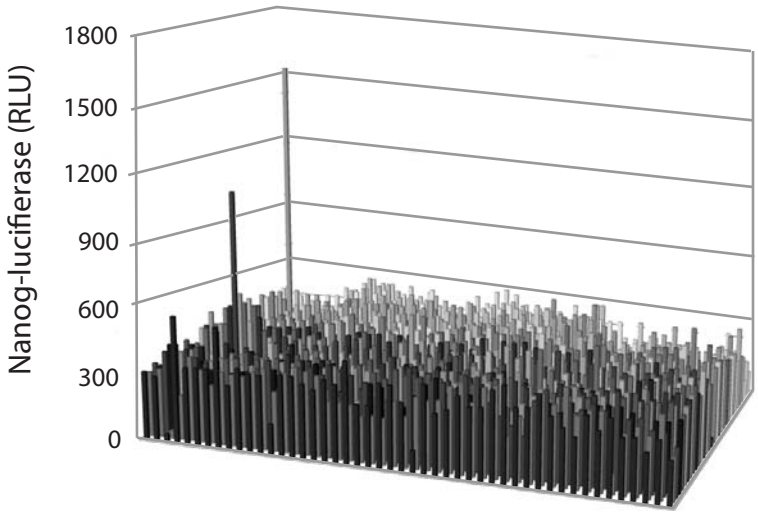
Gene (NA #)	Forward Primer	Reverse Primer
Nanog (NM_028016.2)	AGGGTCTGCTACTGAGATGCTCTG	CAACCACTGGTTTTTCTGCCACCG
Nat1 (NM_008673.1)	ATTCTTCGTTGTCAAGCCGCCAAAGTGGAG	AGTTGTTTGCTGCGGAGTTGTCATCTCGTC
Cdc2 (NM_007659)	GTCCGTCGTAACCTGTTGAG	CATGTTAATTAAGTTATACGTAGG
Cdk2 (NM_183417)	GGAGAACTTTCAAAAGGTGG	CCGTGAGAGCAGAGGCATCCATG
Gsk-3β (NM_019827.6)	CAGAGTCGCCAGACACTATAGTCGA	GCCGGAAGACCTTTGTCCAA
Cdk5 (NM_007668.3)	TCTGAAGCGTGTCAGGCTGG	ACAAGCTGTGGGCCACATCA

SI Table 5. List of short hairpin (sh)RNA sequences used for knock-down analysis.

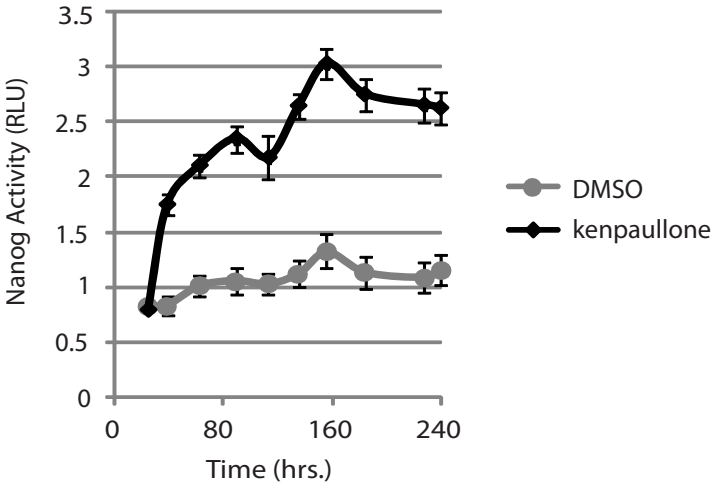
Gene	19mer Short Hairpin RNA Sequences
Glycogen Synthase Kinase-3β (GSK-3β)	GTTCTACAGGACAAGCGAT
	GTGATTGGAAATGGATCAT
	GTTGTATATGTATCAGCTG
	GATGAGGTCTACCTTAACC
	GCATGAAAGTTAGCAGAGA
Cyclin Depdendent Kinase 2 (CDK2)	GGGCCCTATTCCCTGGAGA
	GCCTGATTATAAGCCAAGT
	GCAGCCCTGGCTCACCCTT
	GCTGCTCCAGGGCCTGGCT
	GGTGTACCCAGTACTGCCA
Cyclin Depdendent Kinase 5 (CDK5)	GGCTTCATGATGTCCTGCA
	GCTGCTGAAAGGCCTGGGA
	GACTATAAGCCCTACCCAA
	GTGGTCAGCCGGCTGCATC
Cell division cycle 2 homolog A (CDC2A) Also known as CDK1	GCACCCGTACTTTGATGAC
	GGAGTGCCCAGTACTGCAA
	GGGACCATATTTGCAGAAC
	GAACACCTTTCCCAAGTGG

1. Leost M, Schultz C, Link A, Wu YZ, Biernat J, Mandelkow EM, Bibb JA, Snyder GL, Greengard P, Zaharevitz DW*, et al.* (2000) *Eur J Biochem* **267,** 5983-5994.
2. Schultz C, Link A, Leost M, Zaharevitz DW, Gussio R, Sausville EA, Meijer L, & Kunick C (1999) *J Med Chem* **42,** 2909-2919.
3. Kunick C (1992) *Arch. Pharm.* **325,** 297-299.
4. Kunick C, Lauenroth K, Leost M, Meijer L, & Lemcke T (2004) *Bioorg Med Chem Lett* **14,** 413-416.
5. Kunick C, Zeng Z, Gussio R, Zaharevitz D, Leost M, Totzke F, Schachtele C, Kubbutat MH, Meijer L, & Lemcke T (2005) *Chembiochem* **6,** 541-549.
6. Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, Suh H, & Jaenisch R (2008) *Cell Stem Cell* **2,** 151-159.
7. Markoulaki S, Hanna J, Beard C, Carey BW, Cheng AW, Lengner CJ, Dausman JA, Fu D, Gao Q, Wu S*, et al.* (2009) *Nat Biotechnol* **27,** 169-171.
8. Lyssiotis CA, Walker J, Wu C, Kondo T, Schultz PG, & Wu X (2007) *Proc Natl Acad Sci U S A* **104,** 14982-14987.
9. Zhu S, Wurdak H, Wang J, Lyssiotis CA, Peters EC, Cho C, Xu W, & Schultz PG (2009) *Cell Stem Cell* **In Press**.
10. Liu J, Wu X, Mitchell B, Kintner C, Ding S, & Schultz PG (2005) *Angew Chem Int Ed Engl* **44,** 1987-1990.
11. Knockaert M, Wieking K, Schmitt S, Leost M, Grant KM, Mottram JC, Kunick C, & Meijer L (2002) *J Biol Chem* **277,** 25493-25501.
12. Zaharevitz DW, Gussio R, Leost M, Senderowicz AM, Lahusen T, Kunick C, Meijer L, & Sausville EA (1999) *Cancer Res* **59,** 2566-2569.
13. Meijer L, Flajolet M, & Greengard P (2004) *Trends Pharmacol Sci* **25,** 471-480.
14. Mettey Y, Gompel M, Thomas V, Garnier M, Leost M, Ceballos-Picot I, Noble M, Endicott J, Vierfond JM, & Meijer L (2003) *J Med Chem* **46,** 222-236.
15. Meijer L, Skaltsounis AL, Magiatis P, Polychronopoulos P, Knockaert M, Leost M, Ryan XP, Vonica CA, Brivanlou A, Dajani R*, et al.* (2003) *Chem Biol* **10,** 1255-1266.
16. Wan Y, Hur W, Cho CY, Liu Y, Adrian FJ, Lozach O, Bach S, Mayer T, Fabbro D, Meijer L*, et al.* (2004) *Chem Biol* **11,** 247-259.
17. Gray NS, Wodicka L, Thunnissen AM, Norman TC, Kwon S, Espinoza FH, Morgan DO, Barnes G, LeClerc S, Meijer L*, et al.* (1998) *Science* **281,** 533-538.
18. Meijer L, Borgne A, Mulner O, Chong JP, Blow JJ, Inagaki N, Inagaki M, Delcros JG, & Moulinoux JP (1997) *Eur J Biochem* **243,** 527-536.
19. Davies TG, Bentley J, Arris CE, Boyle FT, Curtin NJ, Endicott JA, Gibson AE, Golding BT, Griffin RJ, Hardcastle IR*, et al.* (2002) *Nat Struct Biol* **9,** 745-749.
20. Bhat R, Xue Y, Berg S, Hellberg S, Ormo M, Nilsson Y, Radesater AC, Jerning E, Markgren PO, Borgegard T*, et al.* (2003) *J Biol Chem* **278,** 45937-45945.
21. Ring DB, Johnson KW, Henriksen EJ, Nuss JM, Goff D, Kinnick TR, Ma ST, Reeder JW, Samuels I, Slabiak T*, et al.* (2003) *Diabetes* **52,** 588-595.

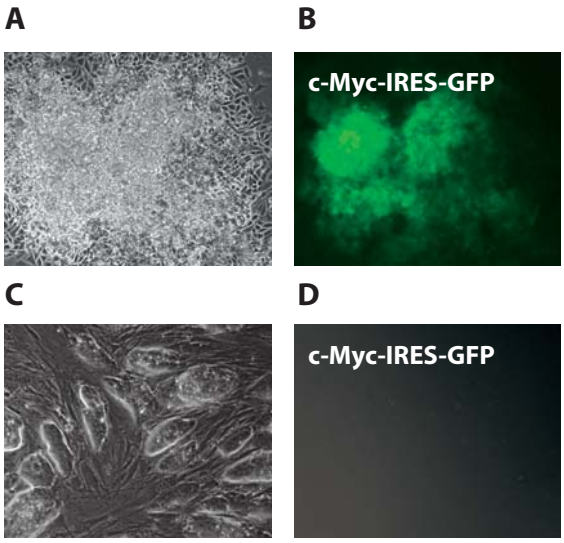
SI Fig. 1



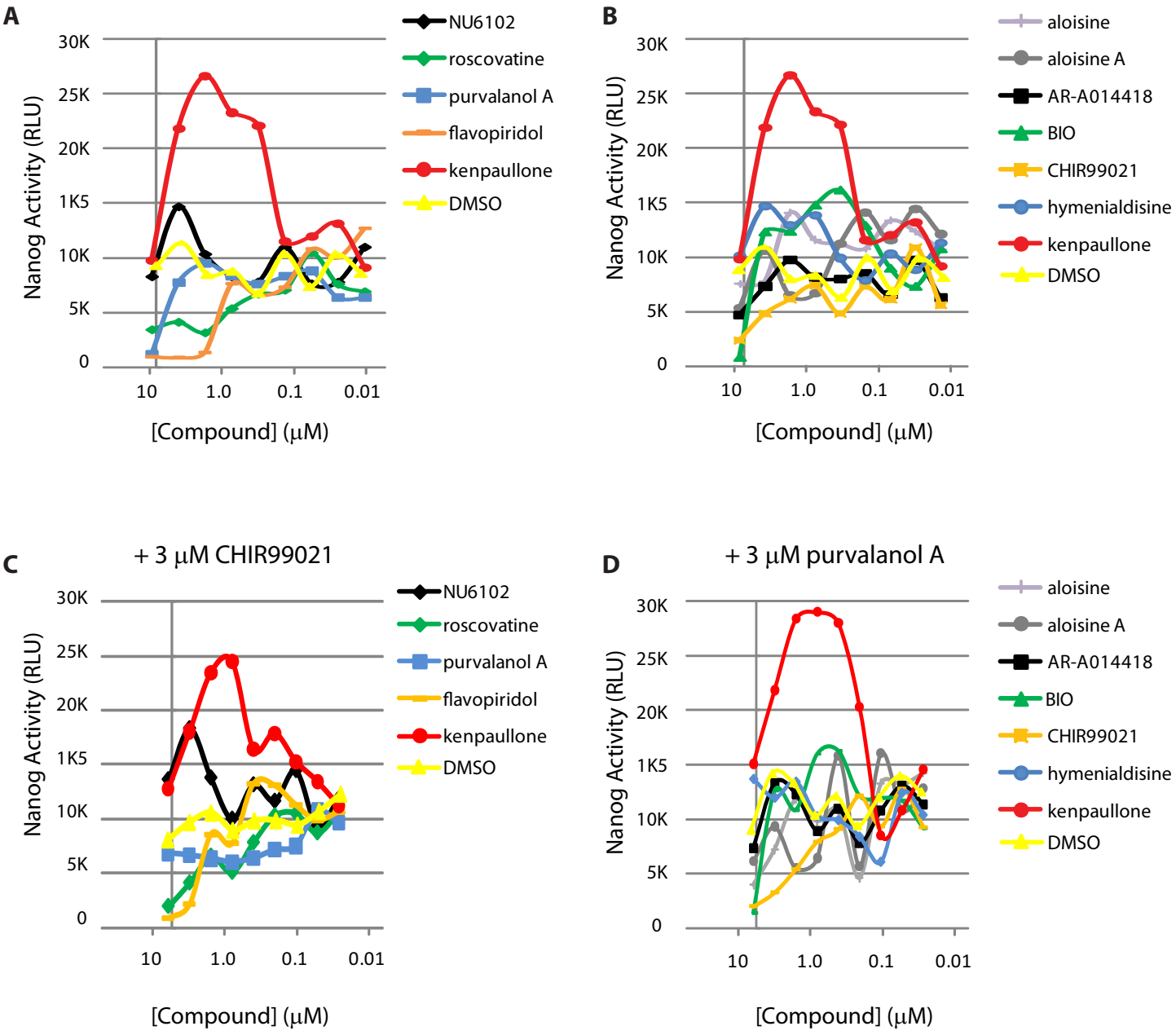
SI Fig. 2


3.5
3
2.5
2
1.5
1
0.5
0
Nanog Activity (RLU)
0
80
160
240
Time (hrs.)
DMSO
kenpaullone

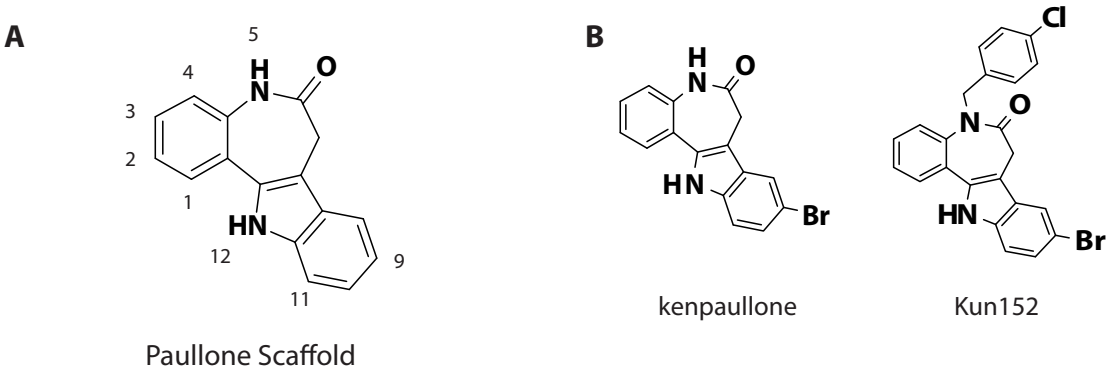
SI Fig. 3

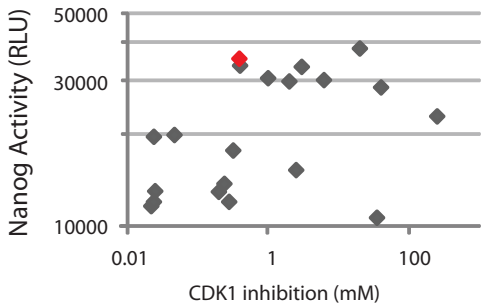


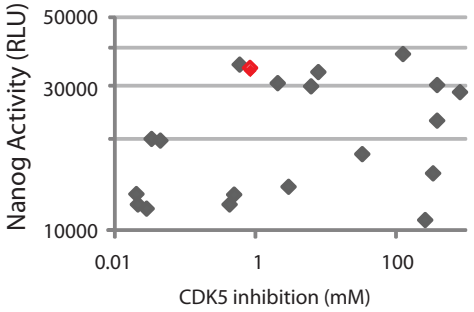
SI Fig. 4

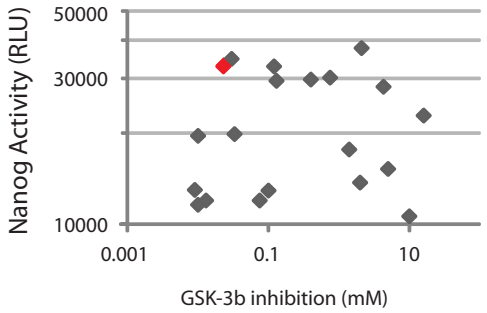

A
Nanog Activity (RLU)
30K
25K
20K
1K5
10K
5K
0
10
1.0
0.1
0.01
[Compound] (μM)
NU6102
roscovatine
purvalanol A
flavopiridol
kenpaullone
DMSO
B
Nanog Activity (RLU)
[Compound] (μM)
aloisine
aloisine A
AR-A014418
BIO
CHIR99021
hymenialdisine
kenpaullone
DMSO
C
+ 3 μM CHIR99021
Nanog Activity (RLU)
[Compound] (μM)
NU6102
roscovatine
purvalanol A
flavopiridol
kenpaullone
DMSO
D
+ 3 μM purvalanol A
Nanog Activity (RLU)
[Compound] (μM)
aloisine
aloisine A
AR-A014418
BIO
CHIR99021
hymenialdisine
kenpaullone
DMSO

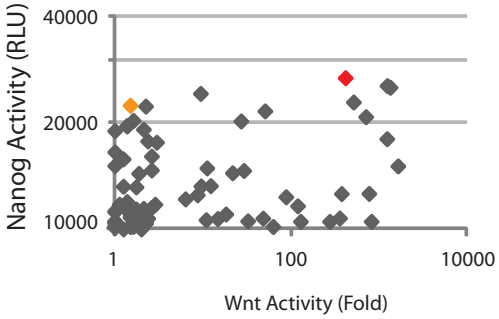
SI Fig. 5

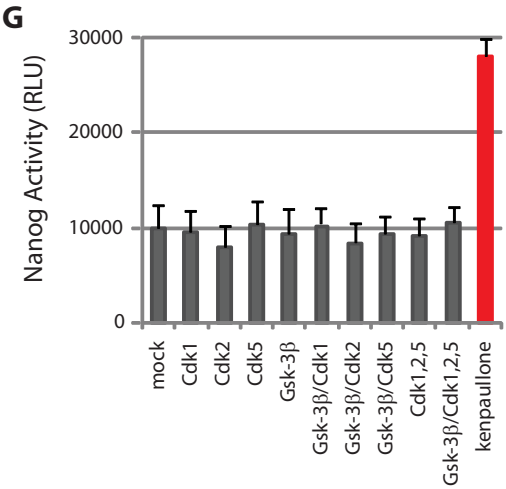

A
Paullone Scaffold
B
kenpaullone
Kun152

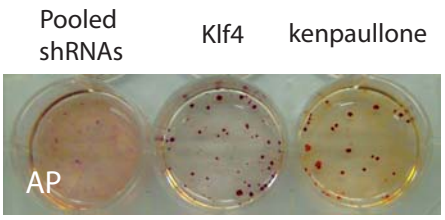

C
Nanog Activity (RLU)
CDK1 inhibition (mM)

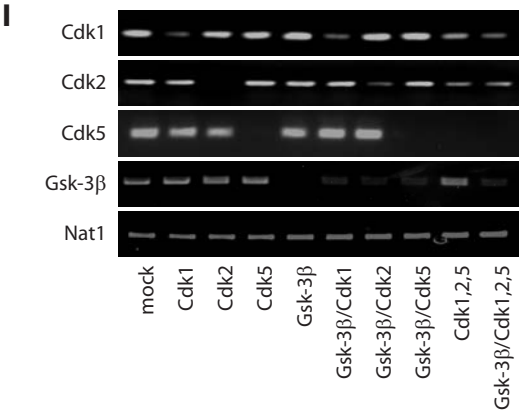

D
Nanog Activity (RLU)
CDK5 inhibition (mM)


E
Nanog Activity (RLU)
GSK-3b inhibition (mM)

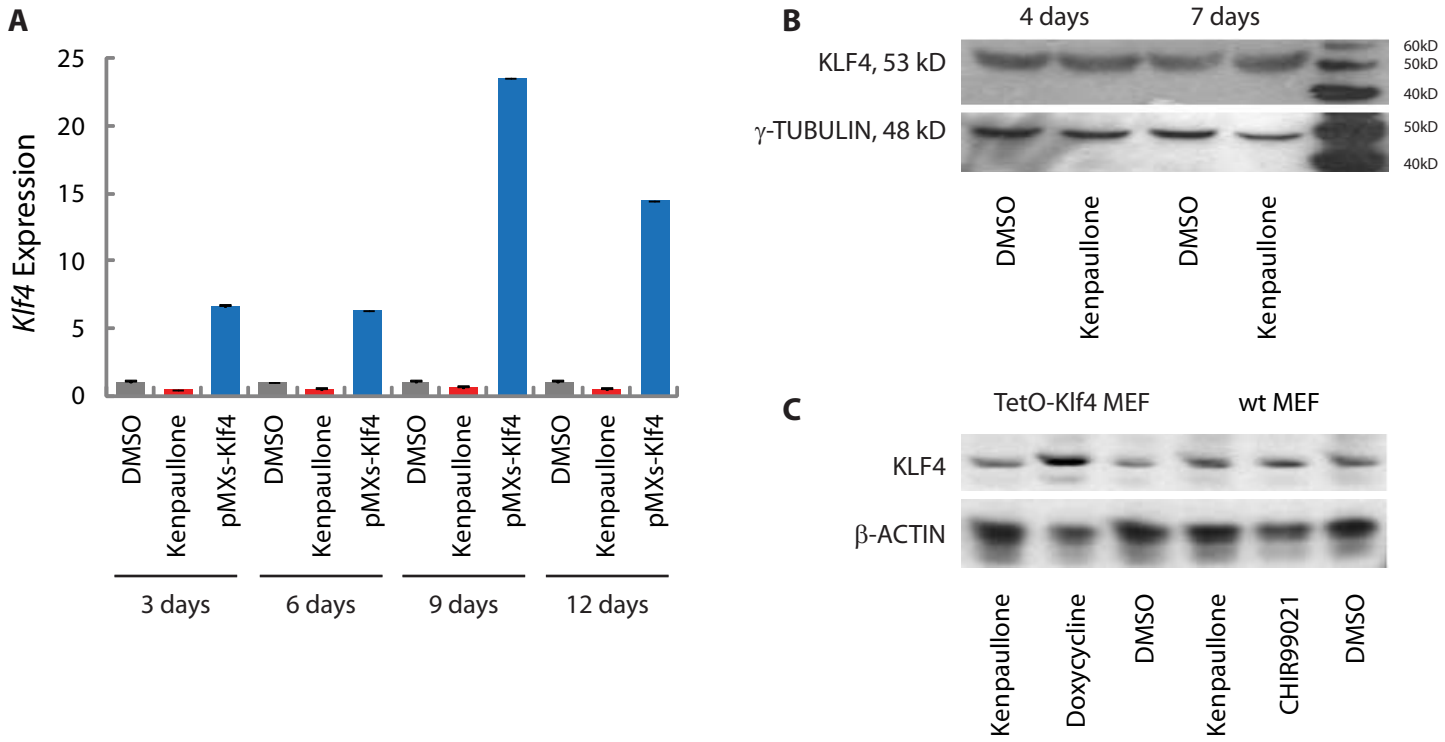

F
Nanog Activity (RLU)
Wnt Activity (Fold)


G
Nanog Activity (RLU)
mock
Cdk1
Cdk2
Cdk5
Gsk-3β
Gsk-3β/Cdk1
Gsk-3β/Cdk2
Gsk-3β/Cdk5
Cdk1,2,5
Gsk-3β/Cdk1,2,5
kenpaullone


H
Pooled shRNAs
Klf4
kenpaullone
AP


I
Cdk1
Cdk2
Cdk5
Gsk-3β
Nat1
mock
Cdk1
Cdk2
Cdk5
Gsk-3β
Gsk-3β/Cdk1
Gsk-3β/Cdk2
Gsk-3β/Cdk5
Cdk1,2,5
Gsk-3β/Cdk1,2,5
OSM, NL-MEF Lines

SI Fig. 6


A
Klf4 Expression
25
20
15
10
5
0
DMSO
Kenpaullone
pMXs-Klf4
3 days
6 days
9 days
12 days
B
4 days
7 days
KLF4, 53 kD
γ-TUBULIN, 48 kD
60kD
50kD
40kD
50kD
40kD
DMSO
Kenpaullone
DMSO
Kenpaullone
C
TetO-Klf4 MEF
wt MEF
KLF4
β-ACTIN
Kenpaullone
Doxycycline
DMSO
Kenpaullone
CHIR99021
DMSO