

# Reprogramming to pluripotency: stepwise resetting of the epigenetic landscape

Bernadett Papp<sup>1</sup>, Kathrin Plath<sup>1</sup>

<sup>1</sup>David Geffen School of Medicine, Department of Biological Chemistry, Jonsson Comprehensive Cancer Center, Molecular Biology Institute, Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California Los Angeles, 615 Charles E. Young Drive South, Los Angeles, CA, 90024, USA

In 2006, the “wall came down” that limited the experimental conversion of differentiated cells into the pluripotent state. In a landmark report, Shinya Yamanaka’s group described that a handful of transcription factors (Oct4, Sox2, Klf4 and c-Myc) can convert a differentiated cell back to pluripotency over the course of a few weeks, thus reprogramming them into induced pluripotent stem (iPS) cells. The birth of iPS cells started off a rush among researchers to increase the efficiency of the reprogramming process, to reveal the underlying mechanistic events, and allowed the generation of patient- and disease-specific human iPS cells, which have the potential to be converted into relevant specialized cell types for replacement therapies and disease modeling. This review addresses the steps involved in resetting the epigenetic landscape during reprogramming. Apparently, defined events occur during the course of the reprogramming process. Immediately, upon expression of the reprogramming factors, some cells start to divide faster and quickly begin to lose their differentiated cell characteristics with robust downregulation of somatic genes. Only a subset of cells continue to upregulate the embryonic expression program, and finally, pluripotency genes are upregulated establishing an embryonic stem cell-like transcriptome and epigenome with pluripotent capabilities. Understanding reprogramming to pluripotency will inform mechanistic studies of lineage switching, in which differentiated cells from one lineage can be directly reprogrammed into another without going through a pluripotent intermediate.

**Keywords:** reprogramming; pluripotency; epigenetic

*Cell Research* (2011) 21:486-501. doi:10.1038/cr.2011.28; published online 15 February 2011

## Introduction

Epigenetic mechanisms preside over our genetic information to enable development from the fertilized, totipotent oocyte to the adult body. The astonishing reprogramming experiment published by Shinya Yamanaka in 2006 demonstrates the profound flexibility of the mammalian epigenome: in less than one month’s time, a handful of transcription factors can reprogram differentiated mouse cells back to a pluripotent state, referred to as induced pluripotent stem (iPS) cell state [1]. Only one year, after the publication of this seminal study using mouse cells, human iPS cells were generated with very similar com-

binations of transcription factors [2-5]. Because human and mouse iPS cells represent an inexhaustible source of cells, highly similar to embryonic stem (ES) cells, the Yamanaka era of stem cell biology is driven by tremendous medical interest. Patient-specific pluripotent cells have already been created and will hopefully be used as substrates for modeling disease pathogenesis and provide immune-matched sources for cell or tissue grafts [2-12].

The Yamanaka screening strategy to find factors that can induce pluripotency is surprisingly simple and affordable [1]. The first reprogramming experiment involved retroviral-mediated overexpression of two dozen well-defined pluripotency regulators in mouse embryonic fibroblasts, and led to emergence of cells that morphologically resemble ES cells upon selection for expression of a resistance gene inserted into the *Fbx15* locus, which encodes an ES cell-specific gene. Subsequent experiments, in which factors were dropped from the original

Correspondence: Kathrin Plath  
Tel: +1-310-2068-688  
E-mail: kplath@mednet.ucla.edu

mix, showed that induction of pluripotency is more efficient when only four factors, Oct4, Sox2, Klf4 and c-Myc, are co-expressed in fibroblasts [1]. A characterization of the resulting iPS cell clones demonstrated, however, that not all of the genes typically expressed in ES cells were strongly upregulated. In agreement with this notion, these original iPS cells self-renewed and differentiated into diverse cell types of all three germ layers, but did not support adult chimerism upon blastocyst injection. Subsequent improvements of methods for the selection of faithfully reprogrammed cells allowed the derivation of iPS cells that are able to contribute to all three germ layers and the germline in mice [13-15], bringing them closer to the developmental potential of mouse ES cells. Some newer mouse iPS cell lines can even generate purely iPS cell-derived animals by tetraploid complementation, which is the most stringent pluripotency test available [16-20]. Many mouse and human iPS cell lines induced by overexpression of Oct4, Sox2, Klf4 and c-Myc were extensively characterized at the molecular level, and are similar to ES cells in their expression and chromatin signatures [15, 21-24]. Thus, reprogramming leads to the silencing of somatically expressed genes and upregulation of ES cell genes, concomitant with the resetting of chromatin structure.

To understand the reprogramming process, one could look at the role that Oct4, Sox2, Klf4 and c-Myc play in ES cells. These transcription factors are all important for the establishment and/or maintenance of pluripotent state during early embryonic development (see recent review [25] for further reading about their function). Importantly, Oct4, Sox2 and Klf4 are thought to maintain the pluripotent, self-renewing state of ES cells by co-occupying the promoter and enhancer regions of a large set of highly expressed ES cell-specific genes, often referred to as pluripotency genes [26-30]. Co-occupancy of Oct4, Sox2 and Klf4, is often predictive for co-occupancy by Nanog, another ES cell-specific transcription factor [21, 27, 29-31]. Thus, it has been suggested that Oct4, Sox2 and Klf4 cooperate over the course of reprogramming to establish functional enhancosomes required for upregulation of the ES cell-specific transcriptome. In contrast, solitary binding of these factors in ES cells is generally associated with transcriptional repression and this may explain how Oct4, Sox2 and Klf4 are able to silence somatic gene expression early in the course of reprogramming. In contrast, c-Myc, a well-known oncogene and cell cycle regulator, has a largely distinct set of target genes from Oct4, Sox2 and Klf4 in ES cells, including numerous cell cycle, metabolism genes etc., thus, forming a separate transcriptional network [28, 29, 32]. Though c-Myc can co-occupy some target genes with Oct4, Sox2 and Klf4,

it is believed that these transcription factors constitute two largely separate transcriptional networks in ES cells [32]. Interestingly, ectopic c-Myc is dispensable for the creation of iPS cells, but acts as an enhancer of kinetics and efficiency of reprogramming [33, 34], supporting the idea that pluripotency gene activation does not directly depend on c-Myc.

In this review, we will discuss the current knowledge of how the reprogramming factors accomplish the mammoth change in gene expression leading to iPS cell induction. Recent reprogramming reviews cover the historic events that led to the iPS cell reprogramming strategy, improved reprogramming methods and disease modeling with iPS cells in depth [35, 36]. Notably, new reprogramming methods that convert one differentiated cell into another, without establishing an intermediate pluripotent state (lineage conversion) [37-39], point us to alternative approaches of induced cell fate change that we will discuss at the end.

### **(Not) all roads lead to Rome**

A key characteristic of transcription factor-induced reprogramming to pluripotency is that the process is inefficient and slow, with only a few cells that express the reprogramming factors progressing to the pluripotent state, within one to two weeks. Optimizing delivery method of the transcription factors has been an attractive strategy to combat this inefficiency. Recent efforts have involved using non-integrating episomal plasmids [3] and viruses [40, 41], the use of cell membrane-penetrating proteins [42-44] and the direct transfection of RNA [45]. Though some of these methods have produced poorer efficiencies than conventional reprogramming with retroviral delivery, they represent a step closer to clinical application, considering that each random integration event of a retrovirus is a potential genomic hazard. The efficiency of reprogramming is typically calculated by determining the original starting cell number and the resulting iPS colony number. However, absolute reprogramming efficiency is, strictly speaking, lower than this metric, if taking into consideration that the starting cells divide several times before complete reprogramming is observed, and most dividing cells never reprogram and may even undergo apoptosis. The inefficiency of the process is a serious limitation to studies of the mechanism of reprogramming to the iPS cell state; the challenge being to molecularly detail the changes going on in a tiny fraction of cells camouflaged within a population of cells that will never reprogram to successful completion.

The fundamental question of whether terminally differentiated cells, or in fact any cell type, can be turned

into pluripotency, or whether iPS cells are derived from progenitor cells hidden in the differentiated cell population, dominated the iPS cell field for two years after the original Yamanaka publication. This controversy was laid to rest by sub-cloning pre-B cells designed to express the reprogramming factors from the same transgenic loci. Hanna and colleagues showed that in fact all individual cells can give rise to reprogrammed cells, if enough time i.e., up to 18 weeks, is given [52]. However, the majority of the daughter cells never reprogram, demonstrating that epigenetic and not genetic barriers limit reprogramming efficiency. The stochastic nature of the reprogramming process suggests that Oct4, Sox2, Klf4 and c-Myc must encounter epigenetic barriers that can be seen as roadblocks in the journey to pluripotency. Furthermore, many different cell types reprogram with similar kinetics under almost identical reprogramming conditions [1, 31, 46-51]. One might predict that the more characteristics the starting cell type and iPS cell end product share, the less roadblocks reprogramming faces and the more efficient the process will be. In agreement with this notion, it has become clear that the starting cell type influences reprogramming because the kinetics and efficiency of the process can dramatically differ amongst cell types. For example, neural progenitors expressing Sox2 and low levels of Klf4 can be reprogrammed more efficiently than fibroblasts and minimally require only ectopic Oct4 [50, 53-55].

Thus, less differentiated cells could reprogram more efficiently than differentiated cells. This idea was addressed by experiments in the hematopoietic lineage, with the reprogramming factors transgenically inserted and inducible from the same integration site in all cells in the population [51]. Hematopoietic stem and progenitor cells, both of which represent more immature cell types, are more amenable for reprogramming compared to terminally differentiated cell types, such as B and T cells, with up to 28% conversion efficiencies [51]. The reprogramming efficiency of these blood progenitors is the highest amongst any iPS cell experiment, suggesting that progenitors are the best starting material for reprogramming with fewest barriers to the reprogramming process. Importantly, this higher reprogramming efficiency of progenitors was not simply due to proliferation differences between the different cell types. Why are progenitors better substrates for reprogramming than differentiated cells? It is possible that in progenitor cells (1) chromatin is more accessible for the reprogramming factors, (2) the transcriptional network of progenitors is easier to disrupt than that of differentiated cells and/or (3) the progenitor transcriptome is more similar to that of ES cells such that more appropriate transcriptional enhancers or less somat-

ic cell stabilizers are present. The answer to this question is unclear at this point.

### Are we there yet? Are ES and iPS cells equivalent?

As introduced above, iPS cells are functionally and molecularly very similar to ES cells, but are these two pluripotent cell types in fact identical? We have to remember that ES cells are derived from the pre-implantation blastocyst and that these cells are inherently metastable. At any given time, ES cells fluctuate between at least two states: one biased to self-renew and the other biased towards differentiation. Heterogeneous expression of pluripotency genes is associated with these two states, and as example, ES cells interconvert between expressing low and high levels of the pluripotency transcription factor Nanog and those cells with lower levels will at that point in time be more receptive to differentiation-inducing signals [56, 57]. Other factors (such as Oct4), are present at similar levels in the entire ES cell population.

ES cells also possess a unique set of microRNAs that play an important role in regulating their gene expression [131]. While transcription factors act on the primary DNA sequence to regulate the expression of genes, miRNAs refine levels of transcribed gene products. It is conceivable that miRNAs can act on a pool of RNAs transcribed earlier, which is also relevant during cell fate changes, and thereby influence the switch to new gene expression programs. Importantly, miRNAs modulate differentiation by targeting pluripotency transcription factors [58, 59].

Global gene expression comparisons of various ES and iPS cell lines detected some differences between the two [60-66]. Importantly, besides protein-coding genes, miRNA expression profiles also differ between current iPS and ES cell clones [61, 67]. It should be noted though, that recent experiments profiling miRNA levels in large number of pluripotent cell lines (both human iPS and ES cells) suggest that differentially expressed miRNAs subdivide the pluripotent cells into two groups, irrespective of their induced pluripotent- or embryo-derived origin, but reflecting their p53 network status [68]. However, even though somatic genes and miRNAs are downregulated in iPS cells, in the direction of their respective levels in ES cells, levels for a handful remain statistically different. On the other hand, some ES cell-specific genes and miRNAs do not completely reach ES cell level in the iPS cell state. It is also observed that there are genes that are ectopically expressed, deviating from levels seen in either the starting cell type or in the ES cell lines. Some recently identified large intergenic

non-coding RNAs (lincRNAs) fall into this class and are elevated in iPS cells compared to ES cells and interestingly, some of these lincRNAs appear to be direct targets of Oct4, Sox2 and Nanog [23]. Overexpression of one of these RNAs, named lincRNA-RoR, leads to a slight increase in the reprogramming efficiency, indicating that high levels of these lincRNAs are advantageous for the generation of iPS cells [23]. These results argue that these statistical expression differences observed between ES and iPS cells can, at least to some extent, have functional consequences.

What is causing these transcriptional differences between ES and iPS cells? Is this perhaps a direct reflection in behavior of the reprogramming factors? Fully reprogrammed iPS cells are not dependent on the exogenous reprogramming factors, and when reprogramming is performed using the retroviral expression method, silencing of the reprogramming factor cassette occurs late in the reprogramming process when pluripotency is established [13-15, 69, 70]. However, some residual expression of exogenous reprogramming factors is often detected in iPS cell lines. Hence, the reprogramming factors are essential for inducing pluripotency, but leftover activity of the ectopic factors might also cause induction of inappropriate transcriptional changes at the end of reprogramming. While this may be true, the excision of reprogramming factor cassettes from the genome or the use of non-integrative reprogramming methods, such as RNA transfection or episomal transduction, still yields iPS cells with expression differences compared to the ES cell state, albeit at a much lower rate [15, 21-24, 61].

### An epigenetic memory for the starting cell type

Besides undergoing transcriptional changes, cells going through reprogramming also reset their pattern of DNA methylation and post-translational histone modifications to the ES cell-like state [15, 21, 22, 24, 61]. However, histone modifications, and DNA methylation levels and distribution also show some differences between ES and iPS cells, which may contribute to the observed gene expression differences [24, 71, 72]. For instance, repressive histone H3K9me3 mark in iPS cells is enriched in genes that are devoid of H3K9me3 in ES cells and among those are numerous differentially expressed genes [24]. The functional impact of these differences became apparent, when different adult cell types derived from the same mouse were reprogrammed and the DNA methylation profiles of resulting iPS cell lines compared to each other and to the cell type of origin [72, 73]. DNA methylation differences were indeed identified in these genetically identical iPS clones, in ways that reflect their cell

type of origin. An epigenetic memory of the starting cell types is further illustrated by the fact that each iPS cell line could be efficiently differentiated into the somatic cell type of origin, but showed reduced efficiency of differentiation into other lineages [72, 73]. The addition of the DNA methylation inhibitor 5-azacytidine (Aza-C) to established iPS cell lines increased their differentiation potential, making them more similar to ES cell lines [73], likely by wiping out improper DNA methylation events. Another feature of this “epigenetic memory” is the inappropriate silencing of imprinted genes, by DNA methylation and histone hypoacetylation, and this is shown to limit the developmental potential of iPS cell lines [20]. Specifically, there is a correlation of silencing of the imprinted region containing the *Dlk1* gene in iPS cells and the inability to support tetraploid complementation in mouse [20]. Thus, a fundamental difference between ES and iPS cells appears to be the presence of an epigenetic memory in iPS cells linked to inappropriate silencing or expression of genes.

So then how can we reach the epigenetic “blank page” of ES cells? Interestingly, prolonged passaging of iPS cell clones removes some of these transcriptional and chromatin differences [60, 61, 72]. It is unclear though whether a more completely reprogrammed cell is selected in this process, the reprogramming factors get silenced more efficiently, or whether epigenetic reprogramming continues in the culture, but it is possible that additional cell divisions help to passively erase somatic epigenetic marks. While functional and molecular differences are detectable between most iPS and ES cell lines, reprogramming of somatic nuclei by transfer into an enucleated oocyte (SCNT) appears to be a more reliable and faithful method to reset the somatic nucleus to the ES cell-like state, as SCNT-derived ES cell lines exhibit little chromatin and expression differences compared to traditional ES cells [73, 74]. Different reprogramming mechanisms might be in action during SCNT and a better understanding of them may lead to the development of reprogramming methods that are more efficient in resetting the epigenome to the ES cell stage.

Since translational applications of iPS cells require the efficient, safe, and reliable differentiation into specific lineages, a careful evaluation of each iPS cell line, potentially at different passages, is needed to choose the best line for a given differentiation protocol. To improve transcription factor-induced reprogramming, it is crucial to understand the temporal order of events leading to pluripotency and to find out how efficiently each step is reached, so that we can identify the steps blocking the reprogramming process or leading to the epigenetic memory in iPS cell clones.

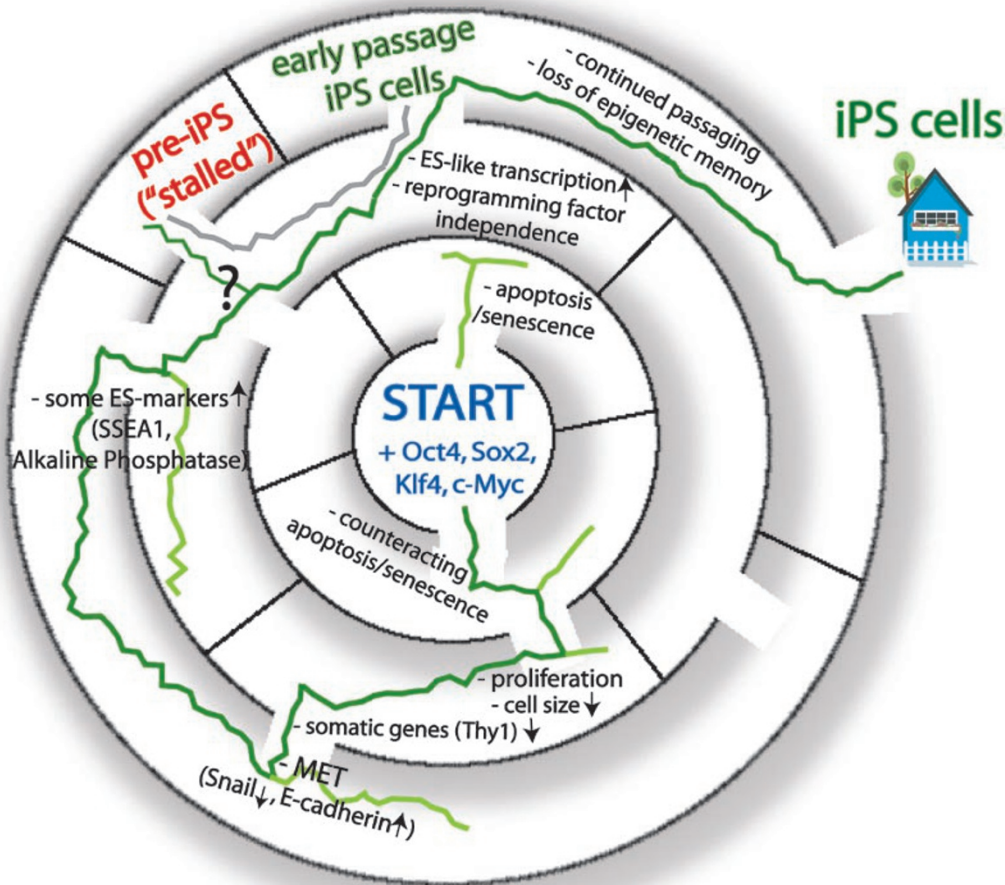
## Steps leading to pluripotency

Is reprogramming to pluripotency a stepwise process with common intermediate stages? Recent live imaging of mouse fibroblasts undergoing reprogramming supports the idea that reprogramming proceeds through highly synchronized progressive events, with first events initiating almost immediately after induction of the reprogramming factors [74]. Even though reprogramming events are initiated early, the presence of reprogramming factors is required until the end, but most cells in which these initial reprogramming events occur do not complete reprogramming [69, 70, 75]. Figure 1 accompanies the discussion below and represents the complex reprogramming process in a labyrinth, with several possible ways, which do not always lead to the pluripotency exit.

In the last couple of years, mainly from studies on mouse and human fibroblasts, the following steps of transcription factor-induced reprogramming have been defined. Initial steps lead to loss of differentiated cell characteristics; then a pre-pluripotent state is acquired, characterized by upregulation of some ES cell markers, which leads to the emergence of the self-sustained pluripotent state with all key pluripotency genes, including Nanog, Esrrb and the endogenously encoded Oct4, expressed.

### Early events toward pluripotency

The first characterized event during reprogramming of fibroblasts is the increase of cell cycle rate. iPS cell colonies are descendents of cells that increase their division



**Figure 1** The “labyrinth to pluripotency” represents the transcriptional and morphological changes during reprogramming to the iPS cell state. (Center) Reprogramming starts from somatic cells induced to ectopically express Oct4, Sox2, Klf4 and c-Myc. The green line leads via indicated cornerstones to the faithfully reprogrammed iPS cell state. Many cells do not succeed in reprogramming as indicated by lines ending in the labyrinth at different steps. The gray line, parallel to the green, shows that pre-iPS cells, a stalled reprogramming intermediate, can be converted to the iPS stage by diverse treatments.

rates from the fibroblast cell cycle length of about 22 hours toward that of cycling ES cells with 11-12 hours after a day or two of the induction of the reprogramming factors [74]. Already, only a minority of fibroblasts starts to divide faster, and the majority of Oct4-, Sox2-, Klf4- and c-Myc-overexpressing fibroblasts retain their slow-dividing nature and fail to reprogram, and often undergo apoptosis or senescence [74]. Accordingly, it has been shown that suppression of apoptosis and senescence is helpful for successful reprogramming [76-81]. As successful daughter cells retain the shorter cell cycle rate, the event must be epigenetic in nature though a molecular explanation remains elusive. Morphology changes as emerging rapid cycling cells get smaller over time and continue to grow as monolayer [74]. At 4 to 8 days later, some of the small cycling cells form compact colonies, concurrent with a mesenchymal-to-epithelial transitions (MET) [74, 75, 82]. ES cells and their *in vivo* counterparts, the epiblast progenitor cells of the pre-implantation blastocyst, are epithelial in nature, meaning they have close cell-cell contact, are highly proliferative with an extremely short G1 cell cycle phase, and have a large nucleus to cytoplasm ratio. However, fibroblasts, the cell type most often used as starting cells for reprogramming studies, are of mesenchymal origin and subject to contact inhibition. Importantly, MET is exactly the opposite of the epithelial-mesenchymal transitioning of pluripotent cells to give rise to tissues during development.

The combined action of the reprogramming factors in these first reprogramming steps must therefore block EMT, and induce the downregulation of mesenchymal genes and the upregulation of metabolic, cell cycle and epithelial genes, respectively [21, 22, 75, 82]. These events are exemplified by the loss of the fibroblast cell surface marker Thy1, downregulation of the mesenchymal transcription factor Snail, and upregulation of proliferation genes, followed by the appearance of the E-cadherin [22, 69, 75, 82]. Importantly, E-cadherin-mediated cell-cell contacts are required for reprogramming [82, 83], thus the completion of these early reprogramming events is necessary for iPS cell generation, but not yet sufficient, as not all cells proceed from here to the iPS state and later steps form additional barriers.

### Late events towards pluripotency

In the following intermediate steps after MET, ES cell markers like alkaline phosphatase (AP) and the surface marker SSEA1 are upregulated [69, 70]. Importantly, again, only a subset of cells is able to transition from one expression state to the next. While most of the starting cells downregulate the somatic marker Thy1, only

few induce SSEA1 and of those even fewer complete the final pluripotency program [69]. Importantly, it appears that only Thy1-negative and SSEA1-positive cells have the potential to transition toward pluripotency by inducing the expression of Nanog, Esrrb and other key pluripotency genes, but until then, these late intermediate cells revert back to a fibroblast-like morphology and transcriptome if ectopic reprogramming factors are removed [69, 75, 70]. However, after upregulating the complete pluripotency program, the ectopic reprogramming factors are not essential anymore, in agreement with the notion that the reprogrammed pluripotent state is self-sustained and that a heritable change in cell identity has occurred. Thus, based on the reprogramming experiments with fibroblasts, it seems that a sequential and fairly synchronous order of events is initiated upon expression of the reprogramming factors, and the continuous expression of reprogramming factors is required to overcome roadblocks toward pluripotency.

### Epigenetic roadblocks

What are the epigenetic mechanisms that suppress the transition from one step to the next? While it is clear that the chromatin signature gets reset to an ES cell-like pattern during reprogramming, the identity of the major chromatin-modifying and chromatin-binding factors involved in this process is not yet known. Blunt treatment with histone deacetylase inhibitors (TSA, VPA, SAHA, butyrate) results in an enhancement of the reprogramming process, likely by raising the global levels of histone acetylation, although secondary effects of acetylation levels on other proteins that, in their acetylated state, could enhance the efficiency of reprogramming cannot be excluded [84-87]. Treatment with VPA can even replace the function of c-Myc in reprogramming experiments with mouse cells and of c-Myc and Klf4 with human cells [84, 85]. In ES cells, Myc interacts with a histone acetyltransferase protein complex, the NuA4 histone acetyltransferase, and its target promoters have abundant histone acetylation levels [32]. In somatic cells, c-Myc is known to recruit p300 [88], another histone acetyltransferase; however, in ES cells, p300 does not appear to get recruited to c-Myc targets, rather p300 is targeted to enhancers with the help of Oct4-Sox2-Nanog [29]. Thus, specific reprogramming factors may exert some of their functions by modulating histone acetylation levels and are required less when acetylation levels are raised in cells by other means.

Higher histone acetylation levels are generally associated with elevated gene expression and more open chromatin structure. Accordingly, global gene expres-

sion analysis on early reprogramming steps confirmed that treatment with the deacetylase inhibitor butyrate enhances the upregulation of genes typically expressed at higher levels in ES cells, but only when c-Myc is part of the reprogramming-inducing cocktail [86]. When butyrate is added to fibroblasts, transduced with only Oct4, Sox2, and Klf4, reprogramming becomes less efficient and gene expression changes towards the ES cell stage are less dramatic [86]. Furthermore, butyrate seems to be advantageous only in the beginning of reprogramming since its addition at later time points does not yield any detectable enhancement of the reprogramming process, in line with the finding that ectopic c-Myc expression is only enhancing (but not essential for) reprogramming [33, 34] and acts during the early steps [21]. Systematic screens are needed to find the specific acetyltransferases involved in the modulation of the reprogramming process and to reveal the mechanism underlying this enhancement.

A few repressive chromatin modifiers have been identified as limiters of reprogramming. Treatment of cultures with the small molecule named BIX-01294, thought to target the repressive histone H3K9 methyltransferase G9a, or the DNA methyltransferase inhibitor Aza-C, enhances the reprogramming process [22, 89, 90]. The success with these inhibitors of transcriptional repressors point to overabundant repressive modifications in cells undergoing reprogramming, however, we do not know their key targets during reprogramming. Another major repressive chromatin regulator proteins are Polycomb group (PcG) proteins form versatile repressive chromatin-modifying multiprotein complexes, which co-occupy hundreds of target genes and inhibit chromatin remodeling to maintain silencing of their targets, as summarized in recent reviews [91]. Simultaneous loss of both of the major polycomb complexes, PRC1 and PRC2 abrogates ES cell differentiation [92]. It will be interesting to determine whether overexpression of PcG proteins affects reprogramming as PcG proteins maintain the silencing of key senescence regulating genes, such as Ink4/ARF, whose depletion enhances reprogramming [76, 77, 79, 81, 93, 94]. One could predict that overexpression of PcG proteins will lead to reduction of senescence/apoptosis thereby enhancing reprogramming. To this end, the activity of several PcG proteins is necessary for inducing pluripotency in differentiated cells in ES/somatic cell fusion experiments [95]. Surprisingly, in this instance, lack of PcG proteins in the ES cell counterpart dominantly represses reprogramming, leading to the theory that PcGs are required to repress an inhibitor of the reprogramming process. Further studies are needed to identify the specific reprogramming events that require PcG proteins as

well as to find the implicated downstream targets.

### Partially reprogrammed cells

In addition to the iPS cell colonies, colonies with an ES cell-like morphology, which do not express endogenous pluripotency factors like Nanog or Esrrb, start to appear in the reprogramming culture after 4-7 days post induction of the reprogramming factors [96]. When one attempts to clonally expand these colonies, some of them regress or apoptose, but a few can be maintained as stable lines and are referred to as partially reprogrammed cells or “pre-iPS” cell lines, as they share some of the characteristics of iPS cells. So far, all pre-iPS cell lines, even if derived from fibroblasts or B cells, appear to be stalled at a similar stage, suggesting a common barrier for reprogramming [22]. All of these lines express high levels of ectopic reprogramming factors from retroviral vectors and this likely is needed for their stable propagation [21]. Typically, the somatic transcriptome is efficiently downregulated in pre-iPS cells, but most of the key pluripotency genes are not upregulated [21, 22, 55]. Pre-iPS cells have been shown to retain fibroblast-like hypermethylated regions (at the DNA level), most notably the Nanog and Oct4 loci, which are typically robustly demethylated in iPS/ES cells [22]. Given that the gene expression state of pre-iPS cells has all the characteristics of a late intermediate of the reprogramming process, we and others have argued that these cells can be used to gain a deeper understanding of late events in the reprogramming process. This is particularly important because pure populations of late intermediates of faithful reprogramming events cannot be isolated yet due to the lack of predictive markers. In agreement with the notion that pre-iPS cells are valuable tools for studying reprogramming, these cells can be converted to the pluripotent iPS cell state by overexpressing pluripotency transcription factors, adding ascorbic acid, or modulating specific signaling pathways [21, 22, 55, 97, 98]. Furthermore, specific transcription factors implicated in regulating diverse lineages in development are often ectopically induced during reprogramming and present in pre-iPS cells and reducing their levels by knockdown leads to an enhancement of conversion to the fully reprogrammed state [22]. Interestingly, not all pre-iPS cell lines can be triggered to reach the faithfully reprogrammed state and various lines react differently to stimuli. For instance, TGF- $\beta$  inhibition leads to Nanog upregulation and pluripotency induction only in a subset of pre-iPS lines and similarly, DNA methylation inhibitor (Aza-C) converts only a fraction of lines to iPS cells [98]. Correlating the epigenetic state of pre-iPS cells with the ability to respond or not

to respond to diverse perturbations should be a powerful tool to identify key epigenetic mechanisms regulating the establishment of pluripotency.

### **Lack of chromatin engagement of the reprogramming factors in pre-iPS cells**

As mentioned above, the analysis of the epigenetic and transcriptional profile of pre-iPS cells can help to understand late steps of reprogramming. A deeper appreciation of the function of the reprogramming factors has been gained from chromatin immunoprecipitation followed by microarray (ChIP-on-CHIP) experiments comparing the binding targets of Oct4, Sox2, Klf4 and c-Myc in ES, iPS and pre-iPS cells [21]. This study revealed that c-Myc already binds many of its ES/iPS cell targets in pre-iPS cells, and that Oct4, Sox2 and Klf4 are properly recruited to many ES/iPS cell target genes at which each of them binds alone or with c-Myc [21]. However, Oct4, Sox2 and Klf4 are not generally recruited to pluripotency genes in pre-iPS cells, many of which are co-occupied by these three transcription factors in ES/iPS cells [21]. These findings indicate that, in pre-iPS cells, the c-Myc transcriptional network appears to be already in a more ES cell-like state, establishing the more undifferentiated cell metabolism and cell cycle state, while the Oct4-Sox2-Klf4-based transcriptional network, mainly regulating pluripotency gene expression in ES cells, is not established properly yet in these cells.

### **Missing cofactors and repressive chromatin state as a barrier to reprogramming factor binding**

These results of pre-iPS point to a special recruitment mechanism for Oct4, Sox2 and Klf4, when their co-binding occurs in the absence of c-Myc. Perhaps specific partners of Oct4, Sox2 and Klf4, which are required for their efficient recruitment, are not yet expressed at the pre-iPS cell stage, or inhibitory chromatin features at their target genes interfere with the access of the reprogramming factors. A missing cooperative factor for Oct4, Sox2 and Klf4 co-binding could be the pluripotency transcription factor Nanog, which is essential for the generation of ES cells from the blastocyst but not for their maintenance [57, 99]. Nanog has a large protein-interactome in ES cells that includes Oct4, Sall4, Esrrb and other well-known pluripotency regulators and co-binds many of the genes in ES cells with Oct4, Sox2, Klf4 and Esrrb [27-30, 100]. Nanog overexpression enhances the generation of iPS cells from pre-B cells in a cell division rate-independent manner, indicating that Nanog alters cell-intrinsic parameters to enhance the re-

programming process [52]. Nanog expression is also essential for reprogramming, as its deletion throughout the early stages does not affect the reprogramming process, but it does during late steps [101]. In agreement with this result, the Nanog protein is not present in pre-iPS cells and its ectopic expression enhances the transition of pre-iPS cells to the iPS state [21, 101].

Is there a unique combination of chromatin modifications at the pluripotency gene targets of Oct4, Sox2 and Klf4, which could explain the lack of their binding at the pre-iPS cell state? Studies comparing chromatin modification patterns of pre-iPS cells to those of the starting cells and iPS/ES cells find fibroblast-like or intermediate chromatin-modification patterns on not properly expressed/bound genes in pre-iPS cells, but no single mark was revealed, which could explain the lack of binding [21, 22]. Like the starting fibroblasts, pre-iPS cells retain DNA methylation at the promoters of key pluripotency genes, such as Oct4, Nanog, Utf1 and Dppa5, in line with the lack of transcriptional activation of these genes [22]. As discussed above, treating specific pre-iPS cell lines with Aza-C or reducing the level of the maintenance DNA methyltransferase Dnmt1, enhances their transition into the iPS cell stage, potentially by erasing the repressive DNA methylation mark from these key promoters [22, 98]. The exact mechanism remains unclear though, since these treatments could also affect modifications on a more global level rather than targeting specific promoters directly, and it has not been studied whether such treatment alters reprogramming factor binding in pre-iPS cells.

ChIP-on-CHIP studies of histone modifications in pre-iPS cells focused on two N-terminal histone methylation marks, a modification associated with transcriptional activation, H3K4me3, and are associated with transcriptional repression, H3K27me3, mediated by PcG proteins [21]. Focusing on pluripotency genes that show lacking of binding by Oct4, Sox2 and Klf4 in pre-iPS cells, it was demonstrated that a subset of these genes, namely those that undergo a dramatic change in expression level from fibroblasts to iPS cells, only carry the H3K27me3 mark in fibroblasts, while they solely carry H3K4me3 upon their transcriptional activation in ES/iPS cells. In pre-iPS cells, an intermediate combination of these modifications was found at these genes with the H3K4me3 mark not efficiently elevated to the ES cell level and H3K27me3 not completely depleted. It is unclear whether this pre-iPS cell chromatin state directly prevents binding of Oct4, Sox2 and Klf4. A correlation of the genome-wide locations of reprogramming factors at the pre-iPS cell stage with the absence or presence of a wide range of chromatin marks and nucleosomal posi-



tioning will further our understanding of how the reprogramming factors engage chromatin at key pluripotency genes. Such studies might identify interfering epigenetic marks associated with or even functionally responsible for the lack of pluripotency.

### Status of the X chromosome inactivation in female cell reprogramming

Female mammals silence one of their two X chromosomes in a process called X chromosome inactivation (XCI) during early embryonic development as a mechanism to equalize X-linked gene dose between the two sexes [102]. XCI is a random process such that either the maternally or the paternally inherited X chromosome becomes inactivated leading to female mosaicism of X-linked gene expression. The process of silencing is initiated when pluripotent cells, i.e., female ES cells or their *in vivo* equivalent are induced to differentiate. The first step in XCI is the upregulation of the non-coding RNA *Xist* on the future inactive X chromosome (Xi), which immediately leads to exclusion of RNA polymerase II and transcriptional silencing [103, 104]. Subsequently, specific and step-wise changes in chromatin structure occurs leading to the accumulation of repressive chromatin marks, such as H3K27me<sub>3</sub>, and the exclusion of active chromatin marks along the entire Xi. This Xi, whether maternal or paternal, is then stably propagated to all somatic daughter cells.

XCI represents a developmental program that is genetically amenable and allows single cell-resolution tracking of the interconversion between heterochromatin and euchromatin *in vitro*. Since XCI is one of the most dramatic forms of heterochromatin formation associated with differentiation of pluripotent cells, an interesting question has been whether the Xi reactivates during reprogramming to the iPS cell state. Using mouse fibroblasts, it was found that reactivation of the Xi occurs as a late step in reprogramming that roughly coincides with the reactivation of the endogenous *Nanog* and *Oct4* loci [15, 69]. Thus, Xi reactivation during mouse reprogramming appears to be tightly coupled to the gain of pluripotency. In support of this finding, *Xist* RNA and the repressive H3K27me<sub>3</sub> mark are still enriched on the Xi in pre-iPS cells, indicating that reactivation does not occur at this stage [21, 55]. It will be interesting to see if molecules involved in Xi reactivation are generally effectors of reprogramming to pluripotency as these processes are so tightly coupled developmentally, at least in mouse.

A long standing puzzling question has been why most human female ES cell lines carry an Xi, as opposed to female mouse ES cells that always have two active X

chromosomes [105-107]. Strikingly, in stark contrast to mouse iPS cell lines, currently all available (FGF4-dependent) human female iPS cells do not reactivate the Xi [108], and the ramifications of this Xi retention are proposed to affect studies of X-linked diseases [108]. The difference in X chromosome status between mouse and human iPS cell lines may relate to a difference in their developmental state described in the next section.

### Primed versus naive pluripotency in reprogramming

Mouse epiblast stem cells (EpiSCs) are derived from the post-implantation epiblast of day 5.5 embryos, depend on the FGF4 signaling pathway, and, for female cells, have an Xi. In contrast, mouse ES cells are obtained from epiblast progenitors of the earlier blastocyst (day 3.5), require LIF signaling, and female ES cells have two active X chromosomes. EpiSCs are able to differentiate *in vitro* into the three germ layers similar to ES cells and therefore considered pluripotent, but opposed to ES cells, EpiSCs are almost unable to contribute to chimera. Therefore, EpiSCs are commonly referred to as “primed” pluripotent cells as opposed to the “naïve” pluripotency of mouse ES/iPS cells [109-111]. Mouse EpiSCs express many of the same genes as iPS/ES cells, including *Nanog*, *Oct4* and *Sox2*, and can be induced to revert back to the ES cell-like naive state, when culturing conditions are changed, and/or transcription factors such as *Klf4*, *c-Myc* or *Nanog* are overexpressed [112-114]. It also has been shown that STAT3 activation, downstream of the LIF signaling pathway, is limiting for reprogramming to naive pluripotency, and that STAT3 overexpression enhances the process [115]. Importantly, reprogramming is still not very efficient from EpiSCs to the ES cell state, and it is not understood why [114]. As we will discuss later, a possible reason for this observation is that EpiSCs and ES cells differ in their global nuclear organization [116].

The most stringent pluripotency tests cannot be performed on human ES/iPS cell lines. As a proxy, comparisons of mouse ES cells and EpiSCs with human ES cells have led to a new understanding on the nature of human cell pluripotency *ex vivo*. Some readily apparent differences between human and mouse ES/iPS cells are morphology, the aforementioned Xi status and the requirement of different culturing conditions (FGF/activin versus LIF/STAT) for their maintenance, which place human ES/iPS cells closer to mouse EpiSCs than mouse ES cells [63, 109, 110].

The application of methods to human ES or iPS cell cultures, that convert the mouse EpiSCs back to the na-

ive ES cell-like state [112, 114, 117], leads to the establishment of human pluripotent cells with mouse ES cell characteristics, including LIF dependence, morphology, two active X chromosomes and appropriate global gene expression [118]. The established naïve human pluripotent state appears to be not very robust, as continued Klf4 or Klf4/Oct4 overexpression is required for its maintenance [118]. Understanding why the reprogramming factors cannot easily induce this naïve pluripotent state in the human cell system should reveal important epigenetic differences between primed and naïve pluripotency. A similar observation has been made in mouse reprogramming experiments in which the naïve pluripotent state in certain genetic backgrounds is metastable [112].

Interestingly, pre-iPS cells appear to be intermediates between ES cells and fibroblasts, also in their cell cycle profile. A pre-iPS cell culture seems to have about equal percent of cells in G1 and S phase, and interestingly, this cell cycle profile is rather similar to EpiSCs, whereas the vast majority of fibroblast and ES cells are in G1 or S phase, respectively [116]. What is different in pre-iPSCs/EpiSCs as opposed to ES/iPS cells in terms of cell cycle regulation and how is that related to pluripotency? Perhaps by inducing cell cycle changes in pre-iPS cells to convert them to more ES-like with the extremely short G1 phase, we could understand how G1 cell cycle length is connected to pluripotency.

### Resetting of DNA replication

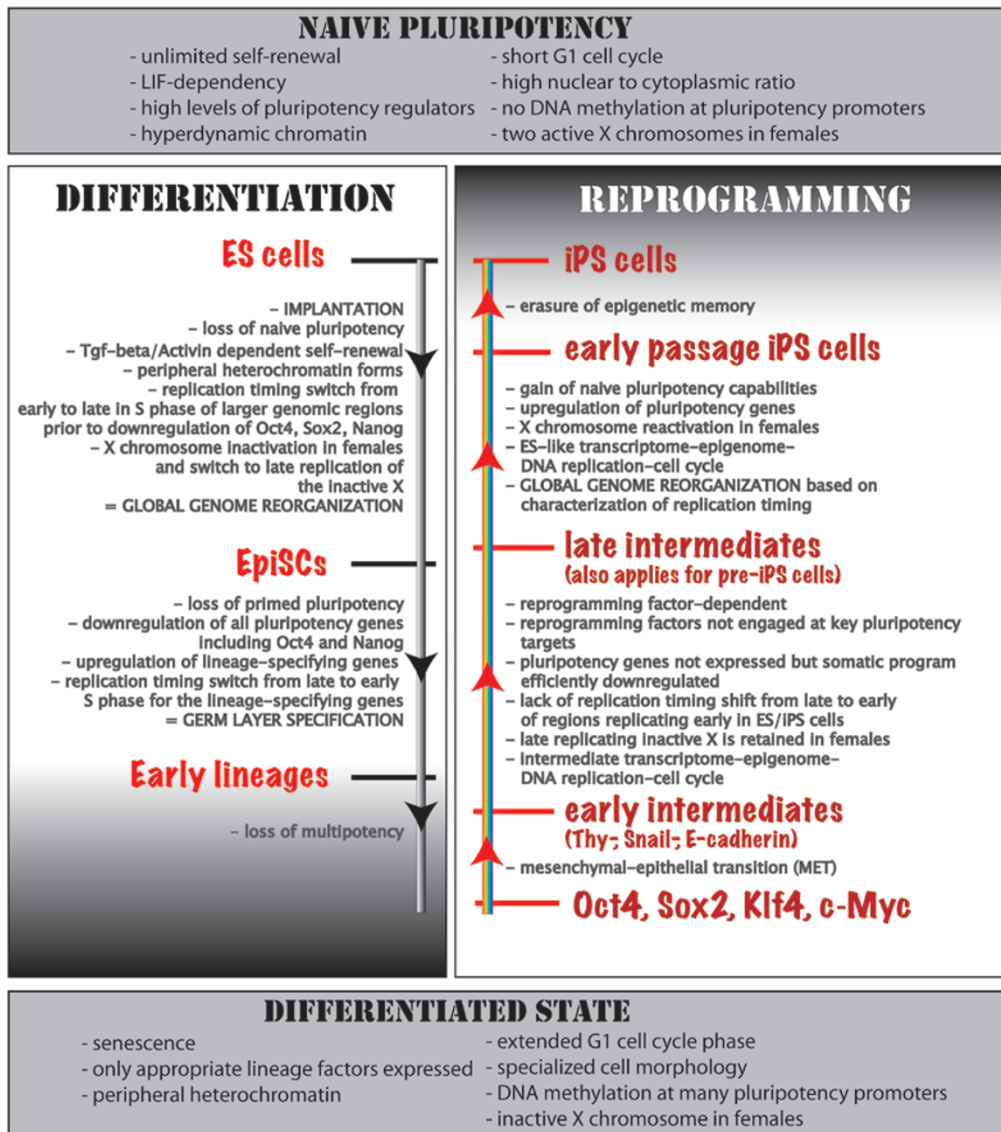
Not only genetic information but also epigenetic information is replicated through each cell cycle and each genomic region is replicated at a specific time in S phase. Replication timing correlates positively with transcription, and replication timing switches are coordinated with transcriptional changes and accompanied by sub-nuclear repositioning [119]. Genome-wide replication timing studies revealed the existence of large genomic domains with similar replication timing and showed that about 20% of the genome changes its replication timing during differentiation of mouse ES cells into neuronal lineages [119]. In differentiated cells, these large replication domains tend to align with static genomic features such as GC content. Strikingly, smaller replication domains and no strong relationship between replication timing and GC content characterize not only ES cells but also iPS cells, indicating that the pluripotent state is characterized by a unique DNA replication timing control [119].

Detailed genome-wide DNA replication studies of ES cells, their differentiation and EpiSCs, revealed that differentiation is associated with global nuclear reorganization events and replication timing changes that occur in

a sequential manner [116]. First, shortly after induction of differentiation of ES cells, replication timing switches occur mainly from early to late in S phase, concomitant with the downregulation of only very few pluripotency genes such as *Dppa2* and *Zfp42*, and a switch of the *Xi* in female cells from early-to-late replication [116]. Thus, during *in vitro* differentiation, before the EpiSC-like stage is reached, a global epigenetic reorganization has occurred, prior to the silencing of most pluripotency genes, including *Oct4* and *Nanog* [116]. The G1 cell cycle length also increases at this time, suggesting that the two events are linked and that in the *in vivo* equivalent stage they might happen approximately in one day at around the implantation of the blastocyst [116]. Subsequently, depending on the specific differentiation path, additional changes in replication timing are as follows: coupled to the downregulation of most pluripotency regulators (e.g., *Nanog*, *Oct4*) more early-to-late S phase changes occur and the transcriptional upregulation of lineage-specific regulators is accompanied by late-to-early replication timing switches [116].

The analysis of replication timing in fibroblasts, pre-iPS and iPS cells indicated that during reprogramming, replication timing reaches the ES cell-like pattern [116, 119]. Even though pre-iPS cells achieve proper ES/iPS cell-like replication timing for the majority of their genome, they retain the somatic replication timing for a subset of genomic regions [116]. Strikingly, many of the early-to-late switching regions, which are the first to be changed during differentiation of ES cells, are among the regions that have not reached the ES/iPS cell-like replication timing state in pre-iPS cells and failed to switch back to early S-phase replication [116]. Thus, not only the upregulation of pluripotency genes occurs late in the reprogramming process but also their replication timing change, suggesting that a global reorganization is required at the end of reprogramming, which could present a major epigenetic barrier to reprogramming. Intriguingly, this idea may explain why the transition of EpiSCs to ES cells is so inefficient. Figure 2 illustrates this genome reorganizing event and the similarity between differentiation and reprogramming.

Nuclear position of chromosomes and interactions between genomic subdomains are non-random in cells, as also revealed by a recent “Hi-C” study that measured the proximity of genomic regions in human cell lines [120]. The Hi-C study proposed that the genome is divided into two distinct compartments, an open compartment characterized by euchromatic features, such as histone 3 lysine 4 (H3K4) trimethylation, and a closed compartment characterized by heterochromatic features, such as H3K9 methylation [120]. Importantly, DNA replication



**Figure 2** A comparison of genome reorganization in differentiation and reprogramming in the mouse system. Top and bottom: main characteristics of the starting/end point cell types of the indicated processes are shown. (Left) events occurring during differentiation of ES cells and in early mouse embryonic development are given. Note that it reads top to bottom. (Right) steps of factor-induced reprogramming to pluripotency, reading from bottom to top. Note that some of the steps are similar between the differentiation and reprogramming processes, just happening in reverse order. See main text for more details.

correlates to 3-D genome organization even more so than to transcription, with early replicating regions being mostly in the euchromatic compartment and late replicating ones in the heterochromatic one [116, 121]. Thus, lack of global reorganization of DNA replication in pre-iPS cells might reflect a not properly reorganized global 3-D nuclear architecture, suggesting that nuclear architecture might be another barrier for reprogramming that is closely associated with replication timing control.

### Converting a compact to an open hyperdynamic chromatin

ES cell chromatin is hyperdynamic, through a combination of loose association of histones and chromatin binding proteins with DNA and rapid turnover of chromatin-binding proteins and histones; heterochromatin binding proteins such as HP1 and histones associate less tightly with chromatin in ES cells than in cells undergoing differentiation [122]. Dynamic chromatin is essential

for pluripotency since restricting the exchange of linker histones leads to differentiation arrest of embryonic stem cells [122]. The absence of DNA synthesis-independent nucleosome assembly factor, HirA, largely elevates the level of soluble core histones in ES cells, and leads to accelerated embryoid body differentiation [122]. This unique chromatin structure of ES cells appears to be actively maintained as the downregulation of the chromatin remodeler Chd1 in ES cells leads to an accumulation of heterochromatin and loss of pluripotency [123]. Perhaps not surprisingly, manipulation at this level of chromatin organization has effects on cellular reprogramming as Chd1 is essential for iPS cell generation and overexpression of ES cell-specific components of the ATP-dependent chromatin remodeling complex (named ES-BAF) enhances reprogramming [123, 124].

Electron spectroscopy imaging also showed that the epiblast progenitor cells of the mouse blastocyst and mouse ES cells have highly dispersed global chromatin architecture, which is distinct from the more compact chromatin state of differentiated cells, with specialized silencing compartments formed [125]. In contrast, mouse EpiSCs, like differentiated cells, contain an area of heterochromatin at their nuclear periphery, thus have already formed specialized nuclear compartments [116]. Thus, not only establishing the ES cell-like hyperdynamic chromatin state during reprogramming is important, but also destabilizing the more repressive compartmentalized chromatin structure of differentiated cells.

In summary, reprogramming to pluripotency goes through specific epigenetic events in a sequential order. Some of these steps appear to be the reversion of processes happening *in vivo* during differentiation and typically represent barriers of the process. As discussed earlier, the mesenchymal-to-epithelial transition happens in the beginning of reprogramming, which is the reverse of EMT events happening around implantation [126]. Furthermore, completion of reprogramming requires upregulation of pluripotency genes and concomitant resetting of their DNA replication profile and global nuclear and genome reorganization, which normally happens in early development during implantation, between the EpiSC and ES cell states.

Currently, the hope is that the development of improved reprogramming strategies will make the conversion of differentiated into pluripotent cells faster and more efficient, perhaps in a more targeted manner using small molecule inhibitors targeting particular factors, rather than modulating the epigenome at a global level. Alternatively, a high-yield isolation of specific cell types may be best accomplished by forgoing the return to pluripotency altogether as described next.

## Shortcut to new lineages without pluripotency intermediates

Intriguingly, efficient conversions of somatic cell types into different somatic cell types without going through the pluripotent state have been achieved, by overexpressing specific sets of lineage-specific transcription factors in combination with appropriate culture conditions. First of these lineage-switching experiments was the conversion of fibroblasts to myoblasts by MyoD overexpression [127] and B cells to macrophages with C/EBP overexpression acting through Pax5 (a B cell-specific transcription factor) inhibition [128]. By Pax5 deletion, B-cells dedifferentiated into progenitor-like cells, which then differentiated into T-lymphocytes [129]. Importantly, even reprogramming of adult pancreatic exocrine cells to insulin-producing  $\beta$ -cells has been achieved *in vivo* through overexpressing Ngn3, Pdx1 and Mafa [130].

More recently, direct reprogramming of fibroblasts to functional cardiomyocytes was accomplished, without any detectable intermediate progenitors, by overexpression of three transcription factors Gata4, MEf2c and Tbx5, which normally function in early heart development [38]. The impressive kinetics, where cardiomyocytes were detected after only 3 days with a very high efficiency of 20% contrasts transcription factor-induced reprogramming to pluripotency. Similarly, a mesodermal cell (fibroblast) can be converted directly into ectodermal cell (neuron). Overexpression of a panel of neuronal master regulators (e.g., Brn2, Myt11, Zic1, Olig2 and Ascl1) leads to mitotic arrest of treated fibroblasts within one day and emergence of immature neuron-like cell morphology within three days. Subsequently functional neurons form, as determined by synapse formations and action potential measurements [39]. The functionality of the neurons improved in culture with time, suggesting that as with reprogramming to pluripotency, time in culture is an essential aspect of successful cell fate changes. Intriguingly, this cell fate change occurs independently of cell division, as opposed to the generation of iPS cells, which appears to require cell division [52]. A last example of note, and maybe the most surprising one, is the conversion of human fibroblasts to multipotent hematopoietic progenitors and mature cells of hematopoietic fate with sole overexpression of Oct4 and modification of culturing protocols [37]. Consistent with the observed bypassing of the pluripotent state to generate blood fate, gene regulatory programs, specific for the adult hematopoietic state were activated, distinct from the embryonic programs involved in the generation of blood cells from pluripotent stem cells. Astoundingly, today, in 2011, it appears that somatic cells can be made

to order and it should very soon become clear whether these cell types are of clinical quality and use. Hopefully, in the future, studies of the mechanism of iPS cell induction will inform studies of direct lineage conversion and *vice versa*.

### Concluding remarks

The ability to generate patient-specific cell types has tremendous implications for disease studies and cell-replacement approaches. While reprogramming to pluripotency may generate an unlimited pool of cells for such studies, the method of direct reprogramming from one to another differentiated cell type may have the advantage that cells are less prone to tumorigenesis. However, cell types generated by lineage conversion still need to be tested more extensively for their functional attributes and have to be compared to cells generated via transition through the pluripotent stage. Our mechanistic understanding of the epigenetic processes leading to cell fate changes is still limited, but the breathtaking speed of new discoveries in the field of reprogramming will surely fill this gap fast and reveal the epigenetic tools that maintain the differentiated state and establish the pluripotent state.

### Acknowledgments

We are grateful to Alissa Minkovsky for critical reading of the manuscript. KP is supported by the NIH Director's Young Innovator Award (DP2OD001686), a CIRM Young Investigator Award (RN1-00564), and by the Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research at UCLA.

### References

- 1 Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006; **126**:663-676.
- 2 Takahashi K, Tanabe K, Ohnuki M, *et al*. Induction of pluripotent stem cells from adult human fibroblasts by defined Factors. *Cell* 2007; **131**:861-872.
- 3 Yu J, Hu K, Smuga-Otto K, *et al*. Human induced pluripotent stem cells free of vector and transgene sequences. *Science* 2009; **324**:797-801.
- 4 Lowry WE, Richter L, Yachechko R, *et al*. Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc Natl Acad Sci USA* 2008; **105**:2883-2888.
- 5 Park IH, Zhao R, West JA, *et al*. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* 2008; **451**:141-146.
- 6 Park IH, Arora N, Huo H, *et al*. Disease-specific induced pluripotent stem cells. *Cell* 2008; **134**:877-886.
- 7 Soldner F, Hockemeyer D, Beard C, *et al*. Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell* 2009; **136**:964-977.
- 8 Carvajal-Vergara X, Sevilla A, D'Souza SL, *et al*. Patient-specific induced pluripotent stem-cell-derived models of LEOPARD syndrome. *Nature* 2010; **465**:808-812.
- 9 Raya A, Rodríguez-Pizà I, Guenechea G, *et al*. Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature* 2009; **460**:53-59.
- 10 Lee G, Papapetrou EP, Kim H, *et al*. Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature* 2009; **461**:402-406.
- 11 Ebert AD, Yu J, Rose FF Jr, *et al*. Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature* 2009; **457**:277-280.
- 12 Marchetto MC, Carromeu C, Acab A, *et al*. A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* 2010; **143**:527-539.
- 13 Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature* 2007; **448**:313-317.
- 14 Wernig M, Meissner A, Foreman R, *et al*. *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 2007; **448**:318-324.
- 15 Maherali N, Sridharan R, Xie W, *et al*. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* 2007; **1**:55-70.
- 16 Han J, Yuan P, Yang H, *et al*. Tbx3 improves the germ-line competency of induced pluripotent stem cells. *Nature* 2010; **463**:1096-1100.
- 17 Zhao X, Li W, Lv Z, *et al*. iPS cells produce viable mice through tetraploid complementation. *Nature* 2009; **461**:86-90.
- 18 Boland MJ, Hazen JL, Nazor KL, *et al*. Adult mice generated from induced pluripotent stem cells. *Nature* 2009; **461**:91-94.
- 19 Kang L, Wang J, Zhang Y, Kou Z, Gao S. iPS cells can support full-term development of tetraploid blastocyst-complemented embryos. *Cell Stem Cell* 2009; **5**:135-138.
- 20 Stadtfeld M, Apostolou E, Akutsu H, *et al*. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* 2010; **465**:175-181.
- 21 Sridharan R, Tchieu J, Mason MJ, *et al*. Role of the murine reprogramming factors in the induction of pluripotency. *Cell* 2009; **136**:364-377.
- 22 Mikkelsen TS, Hanna J, Zhang X, *et al*. Dissecting direct reprogramming through integrative genomic analysis. *Nature* 2008; **454**:49-55.
- 23 Loewer S, Cabili MN, Guttman M, *et al*. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat Genet* 2010; **42**:1113-1117.
- 24 Hawkins RD, Hon GC, Lee LK, *et al*. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* 2010; **6**:479-491.
- 25 Pei D. Regulation of pluripotency and reprogramming by transcription factors. *J Biol Chem* 2009; **284**:3365-3369.
- 26 Jiang J, Chan YS, Loh YH, *et al*. A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat Cell Biol* 2008; **10**:353-360.
- 27 Loh Y, Wu Q, Chew JL, *et al*. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* 2006; **38**:431-440.

- 28 Kim J, Chu J, Shen X, Wang J, Orkin SH. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* 2008; **132**:1049-1061.
- 29 Chen X, Xu H, Yuan P, *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 2008; **133**:1106-1117.
- 30 Boyer LA, Lee TI, Cole MF, *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 2005; **122**:947-956.
- 31 Kim JB, Zaehres H, Wu G, *et al.* Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* 2008; **454**:646-650.
- 32 Kim J, Woo AJ, Chu J, *et al.* A myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell* 2010; **143**:313-324.
- 33 Wernig M, Meissner A, Cassady JP, Jaenisch R. c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* 2008; **2**:10-12.
- 34 Nakagawa M, Koyanagi M, Tanabe K, *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotech* 2008; **26**:101-106.
- 35 Ho R, Chronis C, Plath K. Mechanistic insights into reprogramming to induced pluripotency. *J Cell Physiol* 2010 Oct 13. doi:10.1002/jcp.22450
- 36 Stadtfeld M, Hochedlinger K. Induced pluripotency: history, mechanisms, and applications. *Genes Dev* 2010; **24**:2239-2263.
- 37 Szabo E, Rampalli S, Risueño RM, *et al.* Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature* 2010; **468**:521-526.
- 38 Ieda M, Fu JD, Delgado-Olguin P, *et al.* Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* 2010; **142**:375-386.
- 39 Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Südhof TC, Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 2010; **463**:1035-1041.
- 40 Stadtfeld M, Nagaya M, Utikal J, Weir G, Hochedlinger K. Induced pluripotent stem cells generated without viral integration. *Science* 2008; **322**:945-949.
- 41 Zhou W, Freed CR. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. *Stem Cells* 2009; **27**:2667-2674.
- 42 Zhou H, Wu S, Joo JY, *et al.* Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* 2009; **4**:381-384.
- 43 Kim D, Kim CH, Moon JI, *et al.* Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* 2009; **4**:472-476.
- 44 Cho HJ, Lee CS, Kwon YW, *et al.* Induction of pluripotent stem cells from adult somatic cells by protein-based reprogramming without genetic manipulation. *Blood* 2010; **116**:386-395.
- 45 Warren L, Manos PD, Ahfeldt T, *et al.* Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* 2010; **7**:618-630.
- 46 Aoi T, Yae K, Nakagawa M, *et al.* Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* 2008; **321**:699-702.
- 47 Aasen T, Raya A, Barrero MJ, *et al.* Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotech* 2008; **26**:1276-1284.
- 48 Hanna J, Markoulaki S, Schorderet P, *et al.* Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* 2008; **133**:250-264.
- 49 Stadtfeld M, Brennand K, Hochedlinger K. Reprogramming of pancreatic [beta] cells into induced pluripotent stem cells. *Curr Biol* 2008; **18**:890-894.
- 50 Eminli S, Utikal J, Arnold K, Jaenisch R, Hochedlinger K. Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. *Stem Cells* 2008; **26**:2467-2474.
- 51 Eminli S, Foudi A, Stadtfeld M, *et al.* Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nat Genet* 2009; **41**:968-976.
- 52 Hanna J, Saha K, Pando B, *et al.* Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* 2009; **462**:595-601.
- 53 Kim JB, Greber B, Araúzo-Bravo MJ, *et al.* Direct reprogramming of human neural stem cells by OCT4. *Nature* 2009; **461**:649-643.
- 54 Kim JB, Sebastiano V, Wu G, *et al.* Oct4-induced pluripotency in adult neural stem cells. *Cell* 2009; **136**:411-419.
- 55 Silva J, Barrandon O, Nichols J, Kawaguchi J, Theunissen TW, Smith A. Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol* 2008; **6**:e253.
- 56 Singh AM, Hamazaki T, Hankowski KE, Terada N. A heterogeneous expression pattern for Nanog in embryonic stem cells. *Stem Cells* 2007; **25**:2534-2542.
- 57 Chambers I, Silva J, Colby D, *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* 2007; **450**:1230-1234.
- 58 Tay YM, Tam WL, Ang YS, *et al.* MicroRNA-134 modulates the differentiation of mouse embryonic stem cells, where it causes post-transcriptional attenuation of Nanog and LRH1. *Stem Cells* 2008; **26**:17-29.
- 59 Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS. MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell* 2009; **137**:647-658.
- 60 Chin MH, Pellegrini M, Plath K, Lowry WE. Molecular analysis of human induced pluripotent stem cells and embryonic stem cells. *Cell Stem Cell* 2010; **7**:263-269.
- 61 Chin MH, Mason MJ, Xie W, *et al.* Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 2009; **5**:111-123.
- 62 Guenther MG, Frampton GM, Soldner F, *et al.* Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell* 2010; **7**:249-257.
- 63 Greber B, Wu G, Bernemann C, *et al.* Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. *Cell Stem Cell* 2010; **6**:215-226.
- 64 Ghosh Z, Wilson KD, Wu Y, Hu S, Quertermous T, Wu JC. Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. *PLoS ONE* 2010; **5**:e8975.

- 65 Marchetto MC, Yeo GW, Kainohana O, Marsala M, Gage FH, Muotri AR. Transcriptional signature and memory retention of human-induced pluripotent stem cells. *PLoS ONE* 2009; **4**:e7076.
- 66 Newman AM, Cooper JB. Lab-specific gene expression signatures in pluripotent stem cells. *Cell Stem Cell* 2010; **7**:258-262.
- 67 Wilson KD, Venkatasubrahmanyam S, Jia F, Sun N, Butte AJ, Wu JC. MicroRNA profiling of human-induced pluripotent stem cells. *Stem Cells Dev* 2009; **18**:749-758.
- 68 Neveu P, Kye MJ, Qi S, et al. MicroRNA profiling reveals two distinct p53-related human pluripotent stem cell states. *Cell Stem Cell* 2010; **7**:671-681.
- 69 Stadtfeld M, Maherali N, Breault DT, Hochedlinger K. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* 2008; **2**:230-240.
- 70 Brambrink T, Foreman R, Welstead GG, et al. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* 2008; **2**:151-159.
- 71 Doi A, Park IH, Wen B, et al. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet* 2009; **41**:1350-1353.
- 72 Polo JM, Liu S, Figueroa ME, et al. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat Biotech* 2010; **28**:848-855.
- 73 Kim K, Doi A, Wen B, et al. Epigenetic memory in induced pluripotent stem cells. *Nature* 2010; **467**:285-290.
- 74 Smith ZD, Nachman I, Regev A, Meissner A. Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nat Biotech* 2010; **28**:521-526.
- 75 Samavarchi-Tehrani P, Golipour A, David L, et al. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* 2010; **7**:64-77.
- 76 Hong H, Takahashi K, Ichisaka T, et al. Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. *Nature* 2009; **460**:1132-1135.
- 77 Kawamura T, Suzuki J, Wang YV, et al. Linking the p53 tumor suppressor pathway to somatic cell reprogramming. *Nature* 2009; **460**:1140-1144.
- 78 Li H, Collado M, Villasante A, et al. The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature* 2009; **460**:1136-1139.
- 79 Marión RM, Strati K, Li H, et al. A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* 2009; **460**:1149-1153.
- 80 Utikal J, Polo JM, Stadtfeld M, et al. Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. *Nature* 2009; **460**:1145-1148.
- 81 Banito A, Rashid ST, Acosta JC, et al. Senescence impairs successful reprogramming to pluripotent stem cells. *Genes Dev* 2009; **23**:2134-2139.
- 82 Li R, Liang J, Ni S, et al. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* 2010; **7**:51-63.
- 83 Chen T, Yuan D, Wei B, et al. E-Cadherin-mediated cell-cell contact is critical for induced pluripotent stem cell generation. *Stem Cells* 2010; **28**:1315-1325.
- 84 Huangfu D, Maehr R, Guo W, et al. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat Biotech* 2008; **26**:795-797.
- 85 Huangfu D, Osafune K, Maehr R, et al. Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotech* 2008; **26**:1269-1275.
- 86 Liang G, Taranova O, Xia K, Zhang Y. Butyrate promotes induced pluripotent stem cell generation. *J Biol Chem* 2010; **285**:25516-25521.
- 87 Mali P, Chou BK, Yen J, et al. Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. *Stem Cells* 2010; **28**:713-720.
- 88 Faiola F, Liu X, Lo S, et al. Dual regulation of c-Myc by p300 via acetylation-dependent control of Myc protein turnover and coactivation of Myc-induced transcription. *Mol Cell Biol* 2005; **25**:10220-10234.
- 89 Shi Y, Do JT, Despons C, Hahm HS, Schöler HR, Ding S. A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell* 2008; **2**:525-528.
- 90 Shi Y, Despons C, Do JT, et al. Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* 2008; **3**:568-574.
- 91 Morey L, Helin K. Polycomb group protein-mediated repression of transcription. *Trends Biochem Sci* 2010; **35**:323-332.
- 92 Leeb M, Pasini D, Novatchkova M, Jaritz M, Helin K, Wutz A. Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev* 2010; **24**:265-276.
- 93 Li H, Collado M, Villasante A, et al. The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature* 2009; **460**:1136-1139.
- 94 Utikal J, Polo JM, Stadtfeld M, et al. Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. *Nature* 2009; **460**:1145-1148.
- 95 Pereira CF, Piccolo FM, Tsubouchi T, et al. ESCs require PRC2 to direct the successful reprogramming of differentiated cells toward pluripotency. *Cell Stem Cell* 2010; **6**:547-556.
- 96 Meissner A, Wernig M, Jaenisch R. Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nat Biotech* 2007; **25**:1177-1181.
- 97 Esteban MA, Wang T, Qin B, et al. Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* 2010; **6**:71-79.
- 98 Ichida JK, Blanchard J, Lam K, et al. A small-molecule inhibitor of Tgf-Beta signaling replaces Sox2 in reprogramming by inducing Nanog. *Cell Stem Cell* 2009; **5**:491-503.
- 99 Mitsui K, Tokuzawa Y, Itoh H, et al. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* 2003; **113**:631-642.
- 100 Wang J, Rao S, Chu J, et al. A protein interaction network for pluripotency of embryonic stem cells. *Nature* 2006; **444**:364-368.
- 101 Silva J, Nichols J, Theunissen TW, et al. Nanog is the gateway to the pluripotent ground state. *Cell* 2009; **138**:722-737.
- 102 Payer B, Lee JT. X chromosome dosage compensation: how mammals keep the balance. *Genetics* 2008; **42**:733-772.

- 103 Marahrens Y, Panning B, Dausman J, Strauss W, Jaenisch R. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev* 1997; **11**:156-166.
- 104 Wutz A, Jaenisch R. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol Cell* 2000; **5**:695-705.
- 105 Lengner CJ, Gimelbrant AA, Erwin JA, *et al.* Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. *Cell* 2010; **141**:872-883.
- 106 Silva SS, Rowntree RK, Mekhoubad S, Lee JT. X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. *Proc Natl Acad Sci USA* 2008; **105**:4820-4825.
- 107 Shen Y, Matsuno Y, Fouse SD, *et al.* X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. *Proc Natl Acad Sci USA* 2008; **105**:4709-4714.
- 108 Tchieu J, Kuoy E, Chin MH, *et al.* Female human iPSCs retain an inactive X chromosome. *Cell Stem Cell* 2010; **7**:329-342.
- 109 Tesar PJ, Chenoweth JG, Brook FA, *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 2007; **448**:196-199.
- 110 Brons IG, Smithers LE, Trotter MW, *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 2007; **448**:191-195.
- 111 Nichols J, Smith A. Naive and primed pluripotent states. *Cell Stem Cell* 2009; **4**:487-492.
- 112 Hanna J, Markoulaki S, Mitalipova M, *et al.* Metastable pluripotent states in NOD-mouse-derived ESCs. *Cell Stem Cell* 2009; **4**:513-524.
- 113 Bao S, Tang F, Li X, *et al.* Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* 2009; **461**:1292-1295.
- 114 Guo G, Yang J, Nichols J, *et al.* Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* 2009; **136**:1063-1069.
- 115 Yang J, van Oosten AL, Theunissen TW, Guo G, Silva JC, Smith A. Stat3 activation is limiting for reprogramming to ground state pluripotency. *Cell Stem Cell* 2010; **7**:319-328.
- 116 Hiratani I, Ryba T, Itoh M, *et al.* Genome-wide dynamics of replication timing revealed by *in vitro* models of mouse embryogenesis. *Genome Res* 2010; **20**:155-169.
- 117 Bao S, Tang F, Li X, *et al.* Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* 2009; **461**:1292-1295.
- 118 Hanna J, Cheng AW, Saha K, *et al.* Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc Natl Acad Sci USA* 2010; **107**:9222-9227.
- 119 Hiratani I, Ryba T, Itoh M, *et al.* Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol* 2008; **6**:e245.
- 120 Lieberman-Aiden E, van Berkum NL, Williams L, *et al.* Comprehensive mapping of long range interactions reveals folding principles of the human genome. *Science* 2009; **326**:289-293.
- 121 Ryba T, Hiratani I, Lu J, *et al.* Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res* 2010; **20**:761-770.
- 122 Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* 2006; **10**:105-116.
- 123 Gaspar-Maia A, Alajem A, Polesso F, *et al.* Chd1 regulates open chromatin and pluripotency of embryonic stem cells. *Nature* 2009; **460**:863-868.
- 124 Singhal N, Graumann J, Wu G, *et al.* Chromatin-remodeling components of the BAF complex facilitate reprogramming. *Cell* 2010; **141**:943-955.
- 125 Ahmed K, Dehghani H, Rugg-Gunn P, Fussner E, Rossant J, Bazett-Jones DP. Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. *PLoS One* **5**:e10531.
- 126 Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. *Cell* 2009; **139**:871-890.
- 127 Davis R, Weintraub H, Lassar A. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 1987; **51**:987-1000.
- 128 Xie H, Ye M, Feng R, Graf T. Stepwise Reprogramming of B cells into macrophages. *Cell* 2004; **117**:663-676.
- 129 Cobaleda C, Jochum W, Busslinger M. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature* 2007; **449**:473-477.
- 130 Zhou Q, Brown J, Kanarek A, Rajagopal J, Melton DA. *In vivo* reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature* 2008; **455**:627-632.
- 131 Martinez NJ, Gregory RI. MicroRNA gene regulatory pathways in the establishment and maintenance of ESC identity. *Cell Stem Cell* 2010; **7**:31-35.