

REVIEW

# The roles of the reprogramming factors Oct4, Sox2 and Klf4 in resetting the somatic cell epigenome during induced pluripotent stem cell generation

Ryan Schmidt\* and Kathrin Plath\*

## Abstract

Somatic cell reprogramming to induced pluripotent stem (iPS) cells by defined factors is a form of engineered reverse development carried out *in vitro*. Recent investigation has begun to elucidate the molecular mechanisms whereby these factors function to reset the epigenome.

## Introduction

Current reprogramming technology, pioneered by Takahashi and Yamanaka [1], was built on several seminal advances in the field of developmental biology. First, nuclear transfer experiments demonstrated that a somatic cell nucleus could be epigenetically reset to an early developmental state [2]. Second, cell culture conditions were developed that allowed for the isolation and culture of pluripotent cells, termed embryonic stem (ES) cells, from the inner cell mass of the human and mouse blastocyst [3,4]. Finally, study of these cells and of early embryonic development led to the identification of factors that were ultimately able to reprogram mouse embryonic fibroblasts (MEFs) to the iPS cell state when ectopically expressed, albeit at low frequency [1].

Several groups rapidly followed up on the initial generation of iPS cells and demonstrated that these cells, in their ideal state, are functionally equivalent to ES cells in their ability to contribute to healthy adult mice and their offspring, in addition to forming teratomas when injected into athymic mice [5-10]. In accordance with these results, the gene expression and chromatin states of iPS cells were also found to be strikingly similar to their

ES cell counterparts, although subtle differences remain [10-12]. Tremendous innovation has occurred in the method of factor delivery and the type of somatic cells being reprogrammed. Initially, reprogramming factors were expressed from retroviral transgenes integrated into the genome. Subsequent advances have eliminated the requirement for genomic insertion and viral infection altogether (reviewed in [13]). Additionally, iPS cells have been generated from individuals with specific genetic lesions that can be used to model human diseases (reviewed in [14]). However, despite all of these advances, much remains to be learned about the reprogramming process itself. We believe that the MEF reprogramming paradigm still holds the most promise for future studies due to the ease of obtaining primary cells that are genetically tractable and easy to expand and reprogram, even though we acknowledge that additional lessons may be learned from the use of non-mesenchymal cells, such as hepatocytes or neural cells. The next frontier for the reprogramming field will be a complete mechanistic understanding of how the factors cooperate to reshape the epigenome and gene expression profile of the somatic cell.

## Enhancer and replacement factors

Reprogramming of somatic cells is a multistep process that culminates in the expression of pluripotency genes such as *Nanog*. Although morphological changes occur at early and intermediate stages of reprogramming, pluripotency gene expression is only induced during the late stage and indicates faithful reprogramming. The core reprogramming cocktail, consisting of the transcription factors Oct4, Sox2 and Klf4 (O, S and K), can be augmented by the addition of factors that enhance the efficiency of iPS cell generation, which is typically assessed by quantifying the number of *Nanog*-positive colonies in the culture (Figure 1a). The most well known of these enhancer factors is *c-Myc*, which was added alongside O, S and K in the original reprogramming experiment but later shown to be dispensable [1,5,9,10,15,16]. *c-myc* is a protooncogene that appears to

\*Correspondence: Kathrin Plath [kplath@mednet.ucla.edu](mailto:kplath@mednet.ucla.edu); Ryan Schmidt [schmidtr@ucla.edu](mailto:schmidtr@ucla.edu)

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, Los Angeles, CA 90095, USA

act early in reprogramming to promote an active chromatin environment, enhance cell proliferation, and may play a major role in enhancing the transition from transcriptional initiation to elongation [12,17]. In support of the notion that c-Myc acts mainly in early reprogramming stages, c-Myc greatly enhances the generation of partially reprogrammed cells, which have not turned on pluripotency genes, when combined with O, S and K [15,16]. It has been shown that the family members N-Myc and L-Myc can also enhance reprogramming [15] and that particularly L-Myc has little transforming potential, suggesting that reprogramming and transformation by Myc are distinct processes [18].

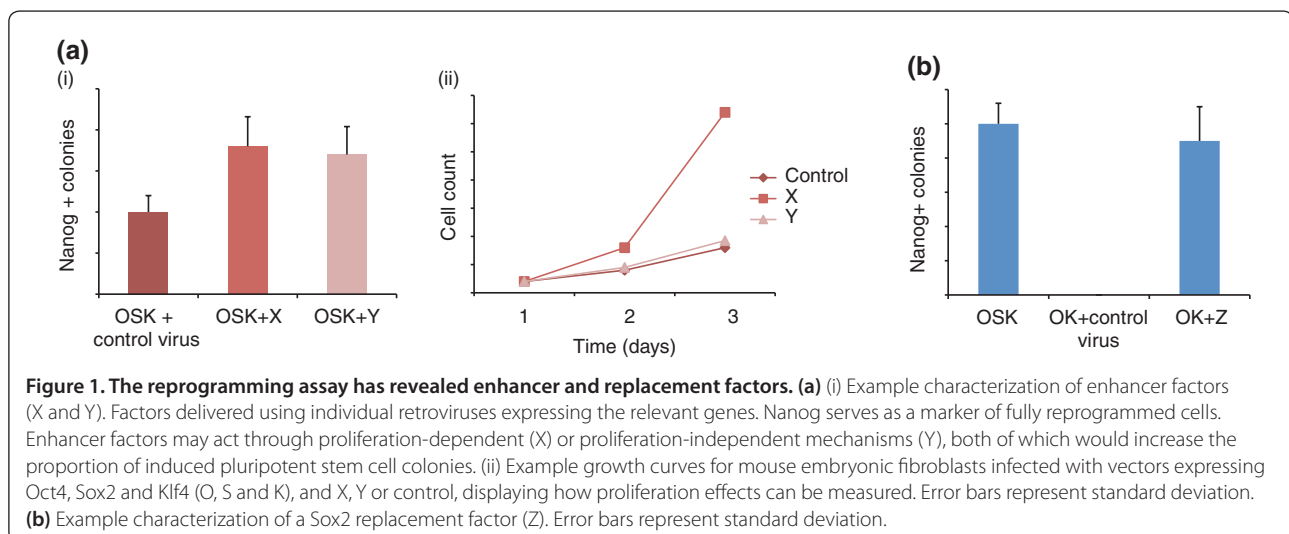
The frequency with which somatic cells convert to iPS cells is typically below 1%. Therefore, much effort has gone into improving reprogramming. Several transcription factors normally expressed in the early stages of embryonic development can enhance reprogramming when added ectopically to O, S and K treated MEFs. These include Glis1, Sall4 and Nanog [19-22]. This class of enhancer factors likely acts late in the reprogramming process to establish and stabilize the pluripotency transcription network. In contrast to c-Myc, Glis1 added to O, S and K enhances the generation of iPS cell colonies without producing Nanog-negative, partially reprogrammed colonies [20]. Remarkably, adding Glis1 and c-Myc together with O, S and K further enhances iPS cell colony formation without the presence of Nanog-negative colonies, suggesting that Glis1 is able to coerce them to the fully reprogrammed state. Forcing Nanog overexpression in partially reprogrammed cells leads to their conversion to iPS cells, demonstrating its late-stage reprogramming activity [22,23].

The ability of cells to pass through the cell cycle has also been shown to be an important determinant of reprogramming efficiency. Knockdown or gene deletion

of p53, p21 or proteins expressed from the *Ink4/Arf* locus allows cells undergoing reprogramming to avoid the activation of cell cycle checkpoints and cellular senescence, leading to greater iPS cell formation [21,24-27]. Consequently, it is likely that any manipulation that accelerates the cell cycle would enhance reprogramming. Thus, reprogramming cultures should be monitored for alterations in their proliferation rate to determine whether the action of an enhancer factor can be attributed to changes in the cell cycle (Figure 1a).

In summary, the induction of pluripotency by O, S and K is a multistep progression whose efficiency can be boosted by enhancer factors. Even though additional factors can positively influence reprogramming, the efficiency of reprogramming is typically still very low. The list of factors discussed above is a brief overview and is by no means exhaustive. Enhancer factors are not exclusively proteins and may consist of any manipulation, including small molecules, long non-coding RNAs and microRNAs, that improves reprogramming [28,29]. Their addition at different stages of the reprogramming process, the generation of partially reprogrammed cells, and the conversion of these cells to the fully reprogrammed state allows one to assay for enhancers of the early and late stages of reprogramming. It will be important to identify the subset of genes whose expression is changed by the introduction of each enhancer factor. Do these genes work alongside the core gene expression changes conferred by O, S and K, or do they simply amplify the magnitude and kinetics of these changes? Also, do known enhancer factors share common mechanisms of action?

Replacement factors possess the unique ability to substitute for O, S or K in reprogramming (Figure 1b). *Esrrb*, an orphan nuclear receptor that is expressed highly in ES cells, has been reported to replace *Klf4* [30].



**Figure 1. The reprogramming assay has revealed enhancer and replacement factors. (a)** (i) Example characterization of enhancer factors (X and Y). Factors delivered using individual retroviruses expressing the relevant genes. Nanog serves as a marker of fully reprogrammed cells. Enhancer factors may act through proliferation-dependent (X) or proliferation-independent mechanisms (Y), both of which would increase the proportion of induced pluripotent stem cell colonies. (ii) Example growth curves for mouse embryonic fibroblasts infected with vectors expressing Oct4, Sox2 and Klf4 (O, S and K), and X, Y or control, displaying how proliferation effects can be measured. Error bars represent standard deviation. **(b)** Example characterization of a Sox2 replacement factor (Z). Error bars represent standard deviation.

Additionally, p53 knockdown has been shown to permit reprogramming in the absence of Klf4 [31]. High-throughput screens have been used successfully to identify small molecule replacement factors. Treatment of cells with kenpaullone allows reprogramming to occur without Klf4, albeit with slightly lower efficiency [32], and several distinct classes of small molecules contribute to iPS cell generation in the absence of Sox2 [33-35]. Reprogramming enhancer and replacement factors are not necessarily mutually exclusive. Nr5a2, for instance, is capable of both enhancing reprogramming and replacing Oct4 [36]. In the human reprogramming system, Lin28 and Nanog, mentioned above as enhancer factors, combine to replace Klf4 [37].

Replacement factors, despite their substantial molecular and functional divergence, may provide important insights into the mechanism whereby O, S and K function in reprogramming. Future work will demonstrate whether these factors regulate the same key genes and pathways as the reprogramming factors that they replace or whether they help achieve the iPS cell state via different means.

### Gene expression changes during reprogramming

Even though causal events are difficult to pinpoint during reprogramming due to the inefficiency of the process, important changes have nonetheless been identified through global expression profiling [11,12,38]. The introduction of O, S and K brings about a dramatic change in the MEF transcriptional profile that eventually leads to induced pluripotency. Of the genes examined by Sridharan *et al.* [12] (GEO:GSE14012) using expression microarrays, more than 6,000 change their expression by more than twofold between MEFs and iPS cells (Figure 2a). The expression changes in response to reprogramming factors begin immediately; however, the pluripotent state is not achieved until several days later [11,38,39]. Hierarchical clustering of data obtained from a reprogramming time course has suggested that reprogramming can be separated into three distinct gene expression phases [38].

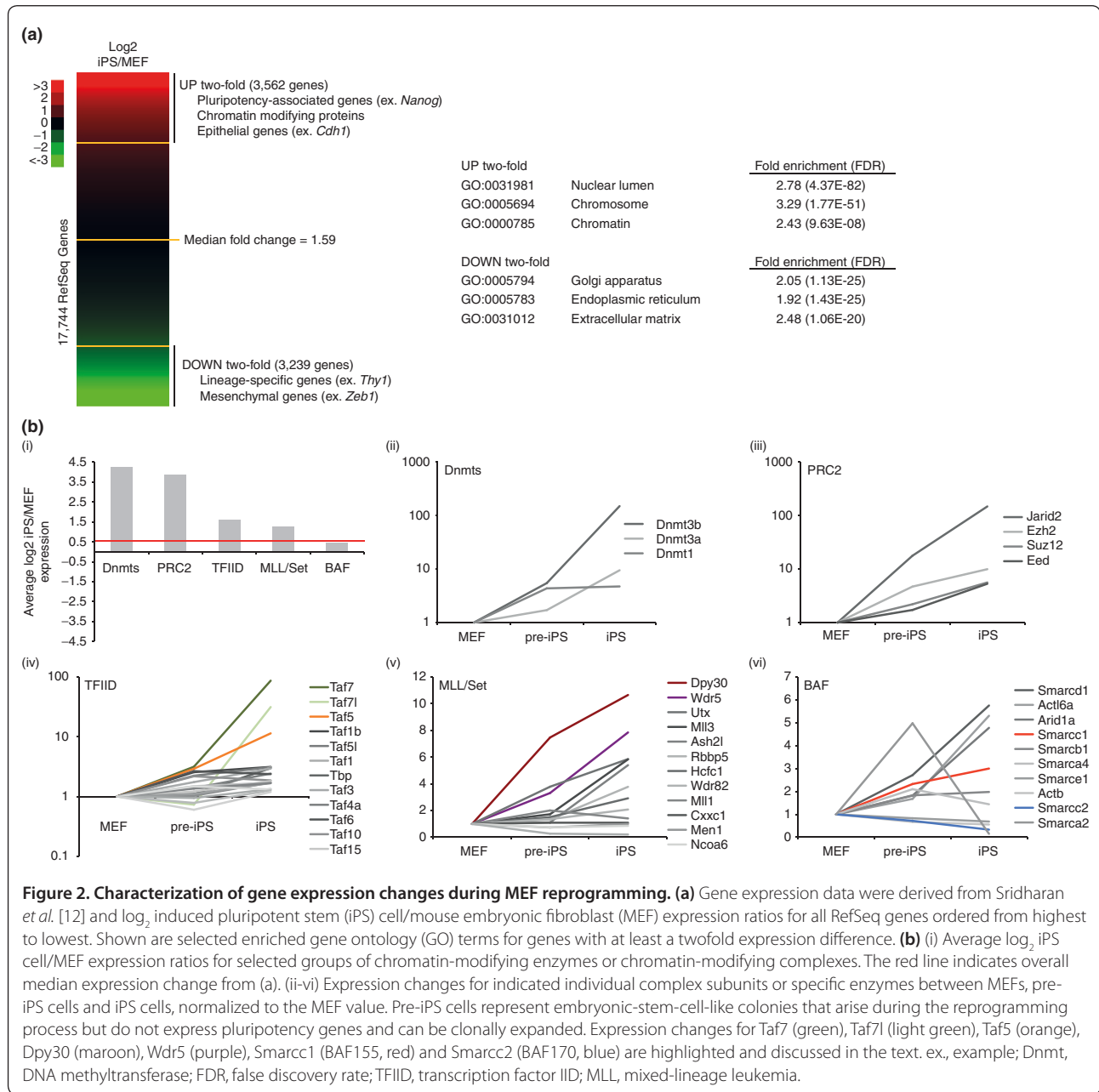
The first of these phases includes downregulation of lineage-specific genes and activation of a genetic program that radically alters cell morphology [38]. This change, known as mesenchymal-to-epithelial transition (MET), is activated by BMP/Smad signaling and inhibited by activation of the TGF- $\beta$  pathway [34,38,40]. The difference in morphology that results from MET is not simply cosmetic. For example, knockdown of *Cdh1*, which encodes the epithelial cell adhesion protein E-cadherin, significantly reduces reprogramming efficiency [40]. Additionally, reduction in cell size has been shown to be an important early event that occurs in cells that go on to reach the pluripotent state [41].

The intermediates generated in a reprogramming culture do not appear to be stable when factor expression is turned off before pluripotency is achieved [38,42,43]. In this instance, cells revert back to a MEF-like gene expression pattern. In agreement with this notion, stable reprogramming intermediates isolated in the form of pre-iPS cells with an ES-cell-like morphology retain high levels of ectopic O, S, K and c-Myc [11,12]. These cells have successfully downregulated fibroblast genes and initiated MET, but have not activated the self-reinforcing network of transcription that characterizes the ES/iPS state [11,12,44,45].

Fully reprogrammed cells arise with low frequency in reprogramming cultures. These cells exhibit indefinite self-renewal and possess the capacity to differentiate into any of the cell types that make up the developing organism. These unique properties are governed by a complex transcriptional program involving many transcription factors, including the reprogramming factors O, S and K, now expressed from their endogenous loci, and additional genes such as *Nanog*, *Esrrb*, *Smad* family members and *Stat* family members [44,45]. Transcription factors within the pluripotency network appear to work cooperatively to regulate genes. Genome-wide chromatin immunoprecipitation (ChIP) experiments demonstrate co-binding among these factors at levels well beyond what would be expected by chance [12,44,45]. Additionally, the presence of multiple factors at a given locus is associated with increased levels of ES/iPS cell-specific gene expression [12,44,45].

In ES cells, which are viewed as a proxy for iPS cells due to their high level of functional similarity, knockdown of any one of a number of transcription factors leads to loss of the pluripotent state, indicating the interconnected nature of the transcriptional network [46]. However, one factor – Nanog – seems to be of special importance. Overproduction of Nanog was able to rescue several of the aforementioned loss-of-function effects and allow ES cells to maintain pluripotency in the absence of the growth factor LIF [46-48]. Furthermore, reprogramming of Nanog-deficient cells proceeds to a partially reprogrammed state that cannot transition to the iPS cell state due to impaired upregulation of the pluripotency network [22,23]. These data illustrate the central role of Nanog in the establishment and maintenance of pluripotency and are consistent with its role as a late-stage enhancer of reprogramming.

Now that transcription factors within the pluripotency network have been largely identified, future research can determine their relative importance by performing similar gain-of-function and loss-of-function assays to those described above involving Nanog. Are all pluripotency-associated factors capable of acting as enhancers of reprogramming? Does their abrogation block reprogramming? Why or why not?



**Figure 2. Characterization of gene expression changes during MEF reprogramming.** (a) Gene expression data were derived from Sridharan *et al.* [12] and log<sub>2</sub> induced pluripotent stem (iPS) cell/mouse embryonic fibroblast (MEF) expression ratios for all RefSeq genes ordered from highest to lowest. Shown are selected enriched gene ontology (GO) terms for genes with at least a twofold expression difference. (b) (i) Average log<sub>2</sub> iPS cell/MEF expression ratios for selected groups of chromatin-modifying enzymes or chromatin-modifying complexes. The red line indicates overall median expression change from (a). (ii-vi) Expression changes for indicated individual complex subunits or specific enzymes between MEFs, pre-iPS cells and iPS cells, normalized to the MEF value. Pre-iPS cells represent embryonic-stem-cell-like colonies that arise during the reprogramming process but do not express pluripotency genes and can be clonally expanded. Expression changes for Taf7 (green), Taf71 (light green), Taf5 (orange), Dpy30 (maroon), Wdr5 (purple), Smarcc1 (BAF155, red) and Smarcc2 (BAF170, blue) are highlighted and discussed in the text. ex., example; Dnmt, DNA methyltransferase; FDR, false discovery rate; TFIIID, transcription factor IID; MLL, mixed-lineage leukemia.

In addition to the changes in specific gene programs, reprogramming fundamentally alters the cell in several important ways. For instance, mouse ES/iPS cells have an altered cell cycle with a shortened G<sub>1</sub> phase [49]. Thus, reprogrammed cells have a reduced doubling time, and a greater fraction of these cells reside in the later phases of the cell cycle [49]. In order to protect genomic integrity during early development, ES/iPS cells have an enhanced capacity for DNA repair [50,51]. Pluripotent cells also have an increased nuclear to cytoplasmic ratio when compared with differentiated cells, as shown by electron microscopy [52].

In accordance with the reduction in membrane surface area and secretory function relative to MEFs, iPS cells generally express genes whose products function outside of the nucleus at comparatively lower levels. Significantly enriched gene ontology (GO) terms within the list of genes whose expression is reduced at least twofold from MEFs to iPS cells include: Golgi apparatus, endoplasmic reticulum and extracellular matrix (Figure 2a). Conversely, genes whose expression is up at least twofold in iPS cells relative to MEFs act primarily within the nucleus and are enriched for GO terms such as nuclear lumen, chromosome and chromatin (Figure 2a).

One important class of nuclear proteins whose gene expression is increased dramatically in ES/iPS cells relative to MEFs is chromatin-modifying complexes (Figure 2b) [53]. These molecular machines modulate gene expression partly by covalent and non-covalent modification of nucleosomes. The expression levels of physically associated subunits within these complexes are largely coordinately regulated during reprogramming. For example, transcripts encoding the components of the PRC2 polycomb complex, responsible for H3K27me<sub>3</sub>, are highly upregulated as cells progress to the pluripotent state (Figure 2b). The DNA methyltransferases, which are not stably associated, also experience similar increases in their expression as reprogramming proceeds (Figure 2b). On the other hand, the transcription factor IID (TFIID) and mixed-lineage leukemia (MLL)/Set complexes are more moderately upregulated as a whole, yet they contain highly upregulated individual subunits, which play important roles in pluripotency and reprogramming (Figure 2b; Taf7, Taf7l and Taf5 of TFIID; Dpy30 and Wdr5 of MLL/Set) [54-56]. Expression switches within chromatin-modifying complexes may affect the induction of pluripotency. In agreement with this notion, Smarcc1 (BAF155) replaces Smarcc2 (BAF170) in the specific form of the BAF complex expressed in pluripotent cells and is critical for their self-renewal (Figure 2b) [57].

The presence of increased levels of chromatin-modifying complexes in ES/iPS cells may serve one of two purposes. First, these proteins may contribute to the maintenance of the self-renewing, undifferentiated state. Examples of this class, where loss-of-function disrupts self-renewal, include Smarca4 (Brg1), Chd1 and Wdr5 [54,57,58]. Second, while a given protein may not be required for normal growth of ES/iPS cells, its presence may be required for the proper execution of subsequent developmental events. Thus, a loss-of-function phenotype will only be detected upon differentiation, as is seen for PRC2, G9a and TAF3, and the DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b [59-63].

### Chromatin changes during reprogramming

Epigenetic changes during reprogramming, most frequently seen in the posttranslational modification status of histone tails, are likely to be both cause and consequence of the previously mentioned changes in gene expression. Differences in H3K4me<sub>2</sub> and H3K27me<sub>3</sub> are detected rapidly upon reprogramming factor induction and often at times precede transcriptional upregulation of the underlying loci [39]. Shifts in the balance of active and inactive chromatin marks at proximal gene regulatory elements are highly correlated with transcriptional changes during reprogramming. ChIP experiments in MEFs and iPS cells demonstrate that the promoter regions of many genes

with the greatest expression increase in the transition from MEFs to iPS cells lose H3K27me<sub>3</sub> and gain H3K4me<sub>3</sub> [10,12]. The low efficiency of reprogramming makes it difficult to study the chromatin state of reprogramming intermediates with population studies such as ChIP, particularly towards the end of the process where the majority of cells have not progressed down the reprogramming path. Pre-iPS cells, which are a clonal population of cells expanded from Nanog-negative colonies with an ES-cell-like morphology, are thought to represent a relatively homogeneous late reprogramming state amenable to ChIP [11,12,22,33]. Similar to what has been observed regarding changes in gene expression, the resetting of chromatin marks does not appear to occur all at once because pre-iPS cells display an intermediate pattern of a subset of chromatin modifications that lies between the MEF and iPS states, both globally and near transcription start sites [12,64].

High-throughput sequencing coupled with ChIP has allowed for the identification of putative distal regulatory elements based on combinations of chromatin marks. These 'enhancer' regions have been mainly defined by the presence of H3K4me<sub>1</sub> and H3K4me<sub>2</sub> at sites that lie at a distance from transcription start sites, which are frequently marked by H3K4me<sub>3</sub> [39,65,66]. Chromatin at these distal sites is reset to an ES-cell-like state over the course of reprogramming [39,65]. In addition to promoting the proper expression of pluripotency-related genes, these sites may contribute to the developmental potential of pluripotent cells by maintaining a poised state that allows for the upregulation of lineage-specific genes in response to the appropriate signals [65,66]. Future studies that analyze more histone marks and incorporate machine learning techniques will help to better characterize these regions as well as other important chromatin states in cells at different stages of reprogramming, which will require the isolation or at least enrichment of cells that will undergo faithful reprogramming.

Over the course of reprogramming, cells experience dramatic global increases in a variety of active histone acetylation and methylation marks, while H3K27me<sub>3</sub> levels remain unchanged [64]. The majority of these changes occur during the late stages of reprogramming – between the pre-iPS and fully reprogrammed states [64]. Additionally, the number of heterochromatin foci per cell, as marked by HP1 $\alpha$  (heterochromatin protein 1 $\alpha$ ), is reduced in iPS cells when compared with MEFs [64]. In accordance with this observation, electron spectroscopic imaging demonstrates that lineage-committed cells have compacted blocks of chromatin near the nuclear envelope that are not seen in the pluripotent state [67,68]. The specific increase in active chromatin is somewhat surprising given that the expression levels of chromatin-modifying complexes associated with both the deposition

of active and inactive marks increase as reprogramming proceeds. Overall, changes in chromatin structure and histone marks coupled with increased transcription of repeat regions indicate that the pluripotent state may possess a unique, open chromatin architecture [53].

Another epigenetic modification, DNA methylation, plays an important role in silencing key pluripotency genes, including *Oct4* and *Nanog*, as cells undergo differentiation [69]. The promoter regions of pluripotency genes are demethylated in ES cells but strongly methylated in fibroblasts [11]. The lack of DNA methylation within these promoters in faithfully reprogrammed iPS cells strongly suggests that during reprogramming, this repressive mark must be erased in order to allow for the establishment of induced pluripotency [5,9-11]. Bisulfite sequencing suggests that removal of DNA methylation from pluripotency loci is a late event that can be placed between the pre-iPS and iPS cell states in the reprogramming continuum [11]. Furthermore, reprogramming efficiency is increased in response to the DNA methyltransferase inhibitor 5-azacytidine [11]. This enhancement is greatest when it is added in a brief window towards the end of the reprogramming process, thus reinforcing the importance of the late-stage removal of DNA methylation [11].

Several other components of the chromatin-modifying machinery have also been shown to affect reprogramming efficiency. Knockdown of *LSD1*, as well as chemical inhibition of histone deacetylases, leads to enhanced reprogramming [70]. Also, overproduction of the histone demethylases Jhdmla and Jhdmlb/Kdm2b, and the SWI/SNF complex components Brg1 and Baf155, increases the efficiency of iPS cell generation [71,72]. In contrast, knockdown of Chd1 and Wdr5 inhibits reprogramming in a cell-proliferation-independent manner [54,58]. Knockdown of candidate chromatin-modifying proteins during human reprogramming identified the histone methyltransferases DOT1L and SUV39H1, and members of the PRC1 and PRC2 polycomb complexes as modulators of reprogramming activity [73]. Reducing the levels of DOT1L and SUV39H1 led to enhanced reprogramming, while reductions in Polycomb complex subunits (BMI1, RING1, SUZ12, EZH2 and EED) resulted in decreased reprogramming efficiency [73]. Recently, Utx/Kdm6a was also shown to be critical for several types of reprogramming, including iPS cell generation from MEFs [74]. The action of this protein is important to remove H3K27me3 from repressed genes in MEFs and prevent the acquisition of H3K27me3 by pluripotency genes as reprogramming proceeds [74]. Finally, Parp1 and Tet2, which both contribute to chromatin modification of the silenced *Nanog* locus early in reprogramming, are each required for iPS cell formation [75].

Through the results mentioned above, several general themes have emerged. First, heterochromatin-associated marks, namely histone deacetylation, H3K9me3 and DNA methylation, represent a barrier whose removal leads to increased reprogramming efficiency. Second, proteins that contribute to an active chromatin environment by writing or reading the H3K4me3 mark are important for achieving pluripotency. Finally, removal of marks associated with transcriptional elongation (H3K36me2/3 and H3K79me2) surprisingly enhances reprogramming. Mechanistically, removal of H3K36me2/3 by Jhdmlb, which is stimulated by ascorbic acid, has been shown to overcome cell senescence by repressing the *Ink4/Arf* locus [76]. Inhibition of DOT1L leads to reduced H3K79me2 at mesenchymal genes, thereby facilitating their downregulation [73].

### **Molecular mechanisms of reprogramming factor activity**

From comparing their binding profiles between pre-iPS cells and iPS cells [12], it is thought that O, S and K vary considerably in their DNA-binding patterns over the course of reprogramming. Eventually, however, they adopt an ES-cell-like binding configuration upon reaching the iPS cell state [12]. Genes that exhibit the largest expression changes during reprogramming are frequently bound by all three reprogramming factors in ES and iPS cells [12]. Increased factor binding at gene promoters in iPS cells is associated with higher levels of transcription, indicating that O, S and K work together to regulate genes primarily as transcriptional activators as described for ES cells [11,12,44,45].

Reprogramming factors must navigate a dynamic chromatin landscape at the various stages of iPS cell generation. While it is plausible that DNA binding differences may be due in part to changes in local chromatin accessibility, O, S and K do not appear to be blocked by the presence of the repressive mark H3K27me3, as promoters enriched for this chromatin mark also can be bound by O, S and K [12,45,77]. In contrast, binding of overproduced OCT4 to the enhancers of silenced genes is associated with nucleosome depletion and the absence of DNA methylation, suggesting that nucleosomes and DNA methylation may comprise a physical barrier that inhibits factor binding [78,79]. Future work may identify additional chromatin signatures that enable or inhibit reprogramming factor binding. Mapping of O, S and K binding in the early stages of reprogramming should reveal chromatin states and nucleosome positions that allow the factors to access target genes.

While there is considerable overlap between the ChIP profiles of all three factors in ES and iPS cells, Oct4 and Sox2 are found together most frequently, whereas Klf4

binds to approximately twice as many sites genome-wide as either of the other factors [12,44,45]. Oct4 and Sox2 can bind cooperatively to composite sox-oct motifs that are frequently found within the regulatory elements of important pluripotency genes [80-82]. These genes include those that encode Oct4 and Sox2 themselves, indicating that these two factors act within autoregulatory positive feedback loops that help to reinforce the pluripotent state [80,81].

Each reprogramming factor contains a highly conserved domain that functions primarily to bind DNA in a sequence-specific manner (Figure 3a). The DNA-binding domains of O, S and K each have distinct evolutionary origins with differing modes of interacting with the double helix. Klf4 binds DNA through three tandem C<sub>2</sub>H<sub>2</sub> zinc fingers that wrap around the major groove [83]. Arginine and histidine side chains that project into the major groove and make contacts with the electronegative surface presented by guanine dictate the GC-rich DNA-binding motif of Klf4 (Figure 3b) [83]. Sox2 binds an AT-rich motif (Figure 3b) through its high-mobility group (HMG) box, which forms an L-shaped binding surface that exclusively contacts the minor groove [84]. This unique shape, along with amino acid side chains that intercalate between the DNA base pair stacks, create a substantial bend in the DNA that is important for its ability to activate transcription [84,85]. Oct4 interacts with DNA through two separate domains containing helix-turn-helix (POU) motifs that each contact half sites within its DNA-binding motif (Figure 3b) in a cooperative manner [86].

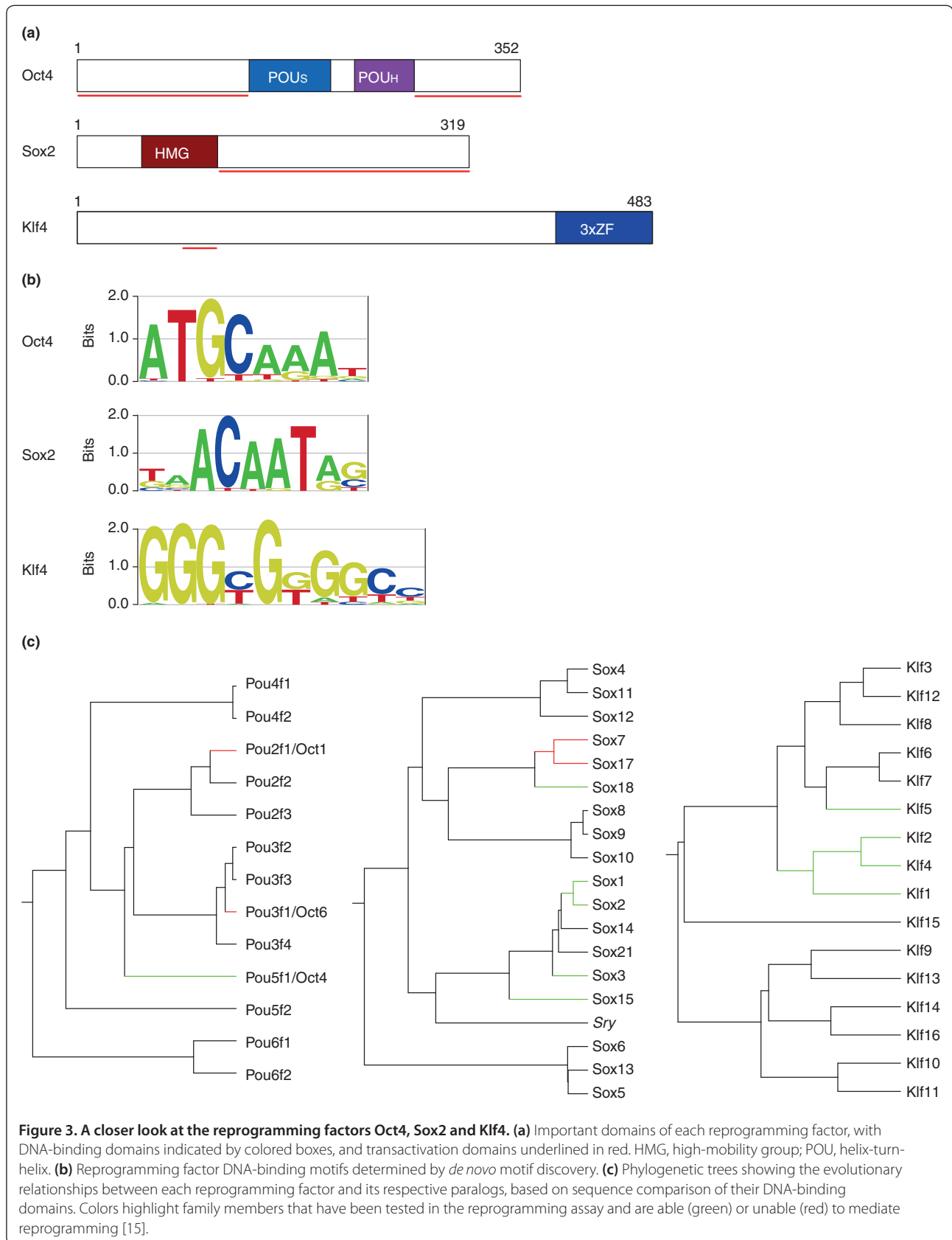
Reprogramming factors can sometimes be functionally replaced by paralogs within their respective families (Figure 3c). Comparison of O, S and K with their paralogs grouped in terms of functional redundancy may provide insight into their mechanisms of action during reprogramming. The binding pattern in ES cells and DNA-binding specificity *in vitro* measured for Klf4 overlaps substantially with Klf2 and Klf5 [87]. Only triple knockdown of all three of these proteins together is sufficient to induce the loss of pluripotency [87]. However, each of these factors may also play more nuanced roles in maintaining self-renewal of pluripotent cells [88]. During reprogramming, Klf2, Klf5 and another close family member, Klf1, have been reported to replace Klf4 with varying degrees of efficiency (Figure 3c) [15]. Sox2, on the other hand, can be replaced by several diverse family members from across its phylogenetic tree, but not others (Figure 3c) [15]. Interestingly, reprogramming activity can be activated in Sox17, a reprogramming-incompetent paralog, by point mutation of a single glutamate within helix 3 of its HMG domain to the corresponding lysine residue present in Sox2 [89]. This change enables cooperative binding with Oct4 at the

canonical subset of sox-oct motifs [89]. Thus, the physical association between Sox2 and Oct4 when bound to DNA is likely to be critical for the induction of pluripotency. Oct4 cannot be replaced by Oct1 or Oct6 in reprogramming, suggesting that it may possess divergent activity not seen in other family members (Figure 3c) [15]. This difference in reprogramming activity among the different Oct factors may not be simply due to differences in DNA-binding preference. Oct1 and Oct4 both bind cooperatively to sox-oct elements in the *Egf4* enhancer, but only Oct4 promotes transcriptional activation of the gene due to its ability to form an active ternary complex with Sox2 [82,90].

Residues that lie outside of the highly conserved DNA-binding domains in O, S and K are also important for their ability to activate transcription and mediate reprogramming (Figure 3a). Klf4 possesses an acidic transactivation domain (TAD) that interacts non-covalently with SUMO-1 [91]. Oct4 contains TADs both amino-terminal and carboxy-terminal of its DNA-binding domains, while Sox2 contains several regions with transactivation activity carboxy-terminal of its HMG box (Figure 3a) [92]. Since these regions were characterized using assays from different developmental contexts, future work is needed to determine which of these TADs function in reprogramming and to identify the co-activators that act through these domains.

Reprogramming efficiency can be enhanced by fusing TADs from other proteins to the reprogramming factors. Addition of a TAD from VP16 to Oct4 or Sox2 increases reprogramming efficiency [93,94]. Fusion of the MyoD TAD to either terminus of Oct4 accelerates and enhances the induction of pluripotency [95]. This enhancement activity is highly specific, since a variety of other known TADs were unable to accomplish the same feat [95]. Additionally, the MyoD TAD was unable to replace the transactivation regions within the Oct4 protein, indicating that these TADs are functionally distinct [95]. Collectively, these results imply that the Oct4 TADs make contact with reprogramming-specific cofactors that cannot be recruited by other well-studied TADs. However, the presence of these TADs fused to the full-length protein likely brings in additional co-activators that enhance the induction of pluripotency. Further investigation is needed to elucidate the exact mechanisms through which these TADs cooperate with the reprogramming factors to enhance reprogramming.

The reprogramming factors are likely to effect changes in transcription through interactions between their TADs and protein cofactors that recruit the RNA polymerase machinery or modify the local chromatin structure. Several of these cofactors have been identified thus far. For instance, Sox2 and Oct4 have been reported to bind





to a complex of XPC, RAD23B and CENT2 to mediate the transactivation of *Nanog* [96]. Loss-of-function experiments demonstrated that these proteins are important for ES cell pluripotency and somatic cell reprogramming [96]. Additionally, several proteomic studies have identified a multitude of candidate O,S,K-interacting proteins that warrant further study [97-100].

Reprogramming factor activity can also be modulated by posttranslational modifications (PTMs). Oct4 phosphorylation at S229 within the POU homeodomain reduces its transactivation activity, possibly by impairing DNA binding as a result of the disruption of a hydrogen bond with the DNA backbone [84,101]. Reprogramming activity is completely abolished in a phosphomimetic mutant (S229D) protein [102]. Additionally, Oct4 can be O-GlcNAcylated at T228 [102]. Mutation of this residue to alanine substantially reduces reprogramming activity, indicating that this PTM may be important for the induction of pluripotency [102]. Given these results, it will be important to examine the effects of other known PTMs within O, S and K during reprogramming.

## Conclusion

Incredibly, somatic cells can revert to the pluripotent state through the forced expression of defined reprogramming factors. The identification and study of these factors has helped to provide insight into the mechanism of induced pluripotency. Conversely, the reprogramming process serves as a robust functional assay that allows us to advance our understanding of Oct4, Sox2, Klf4 and other essential regulators. Much remains to be learned regarding the logic of where these factors bind in the genome and the transcriptional changes that they then induce at these sites. This is not a trivial task given the heterogeneity and inefficiency of the reprogramming process. In a broad sense, knowledge gained through the study of somatic cell reprogramming may be applicable to other gene regulatory events that transform the epigenome and drive embryonic development.

## Abbreviations

ChIP, chromatin immunoprecipitation; ES, embryonic stem; GO, gene ontology; HMG, high-mobility group; iPS, induced pluripotent stem; K, Klf4; MET, mesenchymal-to-epithelial transition; MLL, mixed-lineage leukemia; MEF, mouse embryonic fibroblast; O, Oct4; POU, helix-turn-helix; PTM, posttranslational modifications; S, Sox2; TGF, transforming growth factor; TAD, transactivation domain; TFIIID, transcription factor IID.

## Competing interests

The authors declare that they have no competing interests.

## Acknowledgements

KP is supported by grants from the NIH and CIRM, and the Broad Stem Cell Center at UCLA. RS is supported by the Medical Scientist Training Program training grant from the NIH.

Published: 22 October 2012

## References

We apologize to authors whose work could not be cited owing to space constraints.

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. Gurdon JB, Elsdale TR, Fischberg M: **Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei.** *Nature* 1958, **182**:64-65.
3. Evans MJ, Kaufman MH: **Establishment in culture of pluripotential cells from mouse embryos.** *Nature* 1981, **292**:154-156.
4. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM: **Embryonic stem cell lines derived from human blastocysts.** *Science* 1998, **282**:1145-1147.
5. Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K, Bernstein BE, Jaenisch R: **In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state.** *Nature* 2007, **448**:318-324.
6. Zhao XY, Li W, Lv Z, Liu L, Tong M, Hai T, Hao J, Guo Cl, Ma QW, Wang L, Zeng F, Zhou Q: **iPS cells produce viable mice through tetraploid complementation.** *Nature* 2009, **461**:86-90.
7. 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.
8. Boland MJ, Hazen JL, Nazor KL, Rodriguez AR, Gifford W, Martin G, Kupriyanov S, Baldwin KK: **Adult mice generated from induced pluripotent stem cells.** *Nature* 2009, **461**:91-94.
9. Okita K, Ichisaka T, Yamanaka S: **Generation of germline-competent induced pluripotent stem cells.** *Nature* 2007, **448**:313-317.
10. Maherali N, Sridharan R, Xie W, Utikal J, Eminli S, Arnold K, Stadtfeld M, Yachechko R, Tchieu J, Jaenisch R, Plath K, Hochedlinger K: **Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution.** *Cell Stem Cell* 2007, **1**:55-70.
11. Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, Bernstein BE, Jaenisch R, Lander ES, Meissner A: **Dissecting direct reprogramming through integrative genomic analysis.** *Nature* 2008, **454**:49-55.
12. Sridharan R, Tchieu J, Mason MJ, Yachechko R, Kuoy E, Horvath S, Zhou Q, Plath K: **Role of the murine reprogramming factors in the induction of pluripotency.** *Cell* 2008, **136**:364-377.
13. González F, Boué S, Izpisua Belmonte JC: **Methods for making induced pluripotent stem cells: reprogramming à la carte.** *Nat Rev Genet* 2011, **12**:231-242.
14. Grskovic M, Javaherian A, Strulovici B, Daley GQ: **Induced pluripotent stem cells – opportunities for disease modelling and drug discovery.** *Nat Rev Drug Discov* 2011, **10**:915-929.
15. Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochizuki Y, Takizawa N, Yamanaka S: **Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts.** *Nat Biotechnol* 2008, **26**:101-106.
16. 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.
17. Rahl PB, Lin CY, Seila AC, Flynn RA, Mccuine S, Burge CB, Sharp PA, Young RA: **c-Myc regulates transcriptional pause release.** *Cell* 2010, **141**:432-445.
18. Nakagawa M, Takizawa N, Narita M, Ichisaka T, Yamanaka S: **Promotion of direct reprogramming by transformation-deficient Myc.** *Proc Natl Acad Sci U S A* 2010, **107**:14152-14157.
19. Tsubooka N, Ichisaka T, Okita K, Takahashi K, Nakagawa M, Yamanaka S: **Roles of *Sall4* in the generation of pluripotent stem cells from blastocysts and fibroblasts.** *Genes Cells* 2009, **14**:683-694.
20. Maekawa M, Yamaguchi K, Nakamura T, Shibukawa R, Kodanaka I, Ichisaka T, Kawamura Y, Mochizuki H, Goshima N, Yamanaka S: **Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1.** *Nature* 2011, **474**:225-229.
21. Hanna J, Saha K, Pando B, Zon Jv, Lengner CJ, Creighton MP, Oudenaarden Av, Jaenisch R: **Direct cell reprogramming is a stochastic process amenable to acceleration.** *Nature* 2009, **462**:595-601.
22. Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, Barrandon O, Wray J, Yamanaka S, Chambers I, Smith A: **Nanog is the gateway to the pluripotent ground state.** *Cell* 2009, **138**:722-737.
23. Theunissen TW, van Oosten AL, Castelo-Branco G, Hall J, Smith A, Silva JC: **Nanog overcomes reprogramming barriers and induces pluripotency in minimal conditions.** *Curr Biol*, **21**:65-71.

24. Hong H, Takahashi K, Ichisaka T, Aoi T, Kanagawa O, Nakagawa M, Okita K, Yamanaka S: **Suppression of induced pluripotent stem cell generation by the p53-p21 pathway.** *Nature* 2009, **460**:1132-1135.
25. Li H, Collado M, Villasante A, Strati K, Ortega S, Cañamero M, Blasco MA, Serrano M: **The Ink4/Arf locus is a barrier for iPSC cell reprogramming.** *Nature* 2009, **460**:1136-1139.
26. Utikal J, Polo JM, Stadtfeld M, Maherali N, Kulalert W, Walsh RM, Khalil A, Rheinwald JG, Hochedlinger K: **Immortalization eliminates a roadblock during cellular reprogramming into iPSC cells.** *Nature* 2009, **460**:1145-1148.
27. Banito A, Rashid ST, Acosta JC, Li S, Pereira CF, Geti I, Pinho S, Silva JC, Azuara V, Walsh M, Vallier L, Gil J: **Senescence impairs successful reprogramming to pluripotent stem cells.** *Genes Dev* 2009, **23**:2134-2139.
28. Anokye-Danso F, Trivedi CM, Jühr D, Gupta M, Cui Z, Tian Y, Zhang Y, Yang W, Gruber PJ, Epstein JA, Morrissey EE: **Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency.** *Cell Stem Cell* 2011, **8**:376-388.
29. Loewer S, Cabili MN, Guttman M, Loh Y-H, Thomas K, Park IH, Garber M, Curran M, Onder T, Agarwal S, Manos PD, Datta S, Lander ES, Schlaeger TM, Daley GQ, Rinn JL: **Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells.** *Nat Genet* 2010, **42**:1113-1117.
30. Feng B, Jiang J, Kraus P, Ng J-H, Heng J-CD, Chan Y-S, Yaw L-P, Zhang W, Loh Y-H, Han J, Vega VB, Cacheux-Rataboul V, Lim B, Lufkin T, Ng HH: **Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb.** *Nat Cell Biol* 2009, **11**:197-203.
31. Kawamura T, Suzuki J, Wang YV, Menendez S, Morera LB, Raya A, Wahl GM, Izpisua Belmonte JC: **Linking the p53 tumour suppressor pathway to somatic cell reprogramming.** *Nature* 2009, **460**:1140-1144.
32. Lyssiotis CA, Foreman RK, Staerk J, Garcia M, Mathur D, Mankoulaki S, Hanna J, Lairson LL, Charette BD, Bouchez LC, Bollong M, Kunick C, Brinker A, Cho CY, Schultz PG, Jaenisch R: **Reprogramming of murine fibroblasts to induced pluripotent stem cells with chemical complementation of Klf4.** *Proc Natl Acad Sci U S A* 2009, **106**:8912-8917.
33. Ichida JK, Blanchard J, Lam K, Son EY, Chung JE, Egli D, Loh KM, Carter AC, Di Giorgio FP, Koszka K, Huangfu D, Akutsu H, Liu DR, Rubin LL, Eggan K: **A small-molecule inhibitor of TGF-beta signaling replaces Sox2 in reprogramming by inducing nanog.** *Cell Stem Cell* 2009, **5**:491-503.
34. Maherali N, Hochedlinger K: **Tgfbeta signal inhibition cooperates in the induction of iPSCs and replaces Sox2 and cMyc.** *Curr Biol* 2009, **19**:1718-1723.
35. Shi Y, Despons C, Do JT, Hahn HS, Schöler HR, Ding S: **Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds.** *Cell Stem Cell* 2008, **3**:568-574.
36. Heng JC, Feng B, Han J, Jiang J, Kraus P, Ng J-H, Orlov YL, Huss M, Yang L, Lufkin T, Lim B, Ng HH: **The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells.** *Cell Stem Cell* 2010, **6**:167-174.
37. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA: **Induced pluripotent stem cell lines derived from human somatic cells.** *Science* 2007, **318**:1917-1920.
38. Samavarchi-Tehrani P, Golipour A, David L, Sung H-k, Beyer TA, Datti A, Woltjen K, Nagy A, Wrana JL: **Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming.** *Cell Stem Cell* 2010, **7**:64-77.
39. Koche RP, Smith ZD, Adli M, Gu H, Ku M, Gnirke A, Bernstein BE, Meissner A: **Reprogramming factor expression initiates widespread targeted chromatin remodeling.** *Cell Stem Cell* 2011, **8**:96-105.
40. Li R, Liang J, Ni S, Zhou T, Qing X, Li H, He W, Chen J, Li F, Zhuang Q, Qin B, Xu J, Li W, Yang J, Gan Y, Qin D, Feng S, Song H, Yang D, Zhang B, Zeng L, Lai L, Esteban MA, Pei D: **A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts.** *Cell Stem Cell* 2010, **7**:51-63.
41. Smith ZD, Nachman I, Regev A, Meissner A: **Dynamic single-cell imaging of direct reprogramming reveals an early specifying event.** *Nat Biotechnol* 2010, **28**:521-526.
42. Stadtfeld M, Maherali N, Breault DT, Hochedlinger K: **Defining molecular cornerstones during fibroblast to iPSC cell reprogramming in mouse.** *Cell Stem Cell* 2008, **2**:230-240.
43. Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, Suh H, Jaenisch R: **Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells.** *Cell Stem Cell* 2008, **2**:151-159.
44. 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.
45. Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH: **Integration of external signaling pathways with the core transcriptional network in embryonic stem cells.** *Cell* 2008, **133**:1106-1117.
46. Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, DeCoste C, Schafer X, Lun Y, Lemischka IR: **Dissecting self-renewal in stem cells with RNA interference.** *Nature* 2006, **442**:533-538.
47. Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A: **Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells.** *Cell* 2003, **113**:643-655.
48. Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S: **The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells.** *Cell* 2003, **113**:631-642.
49. White J, Dalton S: **Cell cycle control of embryonic stem cells.** *Stem Cell Rev* 2005, **1**:131-138.
50. Saretzki G, Armstrong L, Leake A, Lako M, von Zglinicki T: **Stress defense in murine embryonic stem cells is superior to that of various differentiated murine cells.** *Stem Cells* 2004, **22**:962-971.
51. Hong Y, Cervantes RB, Tichy E, Tischfield JA, Stambrook PJ: **Protecting genomic integrity in somatic cells and embryonic stem cells.** *Mutat Res* 2007, **614**:48-55.
52. Sampath P, Pritchard DK, Pabon L, Reinecke H, Schwartz SM, Morris DR, Murry CE: **A hierarchical network controls protein translation during murine embryonic stem cell self-renewal and differentiation.** *Cell Stem Cell* 2008, **2**:448-460.
53. Efroni S, Duttgupta R, Cheng J, Dehghani H, Hoepfner DJ, Dash C, Bazett-Jones DP, Le Grice S, McKay RD, Buetow KH, et al: **Global transcription in pluripotent embryonic stem cells.** *Cell Stem Cell* 2008, **2**:437-447.
54. Ang Y-S, Tsai S-Y, Lee D-F, Monk J, Su J, Ratnakumar K, Ding J, Ge Y, Darr H, Chang B, Gingeras TR, Misteli T, Meshorer E: **Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network.** *Cell* 2011, **145**:183-197.
55. Geronne A, Tai X, Zhang J, Wu G, Zhu J, Yoshimoto A, Hanson J, Cultraro C, Chen Q-R, Guinter T, Yang Z, Hathcock K, Singer A, Rodriguez-Canales J, Tessarollo L, Mackem S, Meerzaman D, Buetow K, Singer DS: **The general transcription factor TAF7 is essential for embryonic development but not essential for the survival or differentiation of mature T cells.** *Mol Cell Biol* 2012, **32**:1984-1997.
56. Jiang H, Shukla A, Wang X, Chen WY, Bernstein BE, Roeder RG: **Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains.** *Cell* 2011, **144**:513-525.
57. Ho L, Ronan JL, Wu J, Staahl BT, Chen L, Kuo A, Lessard J, Nesvizhskii AI, Ranish J, Crabtree GR: **An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency.** *Proc Natl Acad Sci U S A* 2009, **106**:5181-5186.
58. Gaspar-Maia A, Alajem A, Polesso F, Sridharan R, Mason MJ, Heidersbach A, Ramalho-Santos J, McManus MT, Plath K, Meshorer E, Ramalho-Santos M: **Chd1 regulates open chromatin and pluripotency of embryonic stem cells.** *Nature* 2009, **460**:863-868.
59. Chamberlain SJ, Yee D, Magnuson T: **Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency.** *Stem Cells* 2008, **26**:1496-1505.
60. Lei H, Oh SP, Okano M, Juttermann R, Goss KA, Jaenisch R, Li E: **De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells.** *Development* 1996, **122**:3195-3205.
61. Okano M, Bell DW, Haber DA, Li E: **DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.** *Cell* 1999, **99**:247-257.
62. Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H, Shinkai Y: **G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis.** *Genes Dev* 2002, **16**:1779-1791.
63. Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S, Sakaue M, Matsuoka C, Shimotohno K, Ishikawa F, Li E, Ueda HR, Nakayama J, Okano M: **Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b.** *Genes Cells* 2006, **11**:805-814.

64. Mattout A, Biran A, Meshorer E: **Global epigenetic changes during somatic cell reprogramming to iPSCs.** *J Mol Cell Biol* 2011, **3**:341-350.
65. Creighton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R: **Histone H3K27ac separates active from poised enhancers and predicts developmental state.** *Proc Natl Acad Sci U S A* 2010, **107**:21931-21936.
66. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J: **A unique chromatin signature uncovers early developmental enhancers in humans.** *Nature* 2011, **470**:279-283.
67. 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* 2010, **5**:e10531.
68. Hiratani I, Ryba T, Itoh M, Rathjen J, Kulik M, Papp B, Fussner E, Bazett-Jones DP, Plath K, Dalton S, Rathjen PD, Gilbert DM: **Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis.** *Genome Res* 2010, **20**:155-169.
69. Li J-Y, Pu M-T, Hirasawa R, Li B-Z, Huang Y-N, Zeng R, Jing N-H, Chen T, Li E, Sasaki H, Xu G-L: **Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog.** *Mol Cell Biol* 2007, **27**:8748-8759.
70. Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, Chen AE, Melton DA: **Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds.** *Nat Biotechnol* 2008, **26**:795-797.
71. Liang G, He J, Zhang Y: **Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming.** *Nat Cell Biol* 2012, **14**:457-466.
72. Singhal N, Graumann J, Wu G, Araúz-Bravo MJ, Han DW, Greber B, Gentile L, Mann M, Schöler HR: **Chromatin-remodeling components of the BAF complex facilitate reprogramming.** *Cell* 2010, **141**:943-955.
73. Onder TT, Kara N, Cherry A, Sinha AU, Zhu N, Bernt KM, Cahan P, Mancarci OB, Unternaehrer J, Gupta PB, Lander ES, Armstrong SA, Daley GQ: **Chromatin-modifying enzymes as modulators of reprogramming.** *Nature* 2012, **483**:598-602.
74. Mansour AA, Gafni O, Weinberger L, Zviran A, Ayyash M, Rais Y, Krupalnik V, Zerbib M, Amann-Zalcenstein D, Maza I, Geula S, Viukov S, Holtzman L, Pribluda A, Canaani E, Horn-Saban S, Amit I, Noverstern N, Hanna JH: **The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming.** *Nature* 2012, **488**:409-413.
75. Doege CA, Inoue K, Yamashita T, Rhee DB, Travis S, Fujita R, Guarnieri P, Bhagat G, Vanti WB, Shih A, Levine RL, Nik S, Chen EI, Abeliovich A: **Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2.** *Nature* 2012, **488**:652-655.
76. Wang T, Chen K, Zeng X, Yang J, Wu Y, Shi X, Qin B, Zeng L, Esteban MA, Pan G, Pei D: **The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner.** *Cell Stem Cell* 2011, **9**:575-587.
77. Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, Kosaki H, Fuchikami T, Abe K, Murray HL, Zucker JP, Yuan B, Bell GW, Herbolzheimer E, Hannett NM, Sun K, Odum DT, Otte AP, Volkert TL, Bartel DP, Melton DA, Gifford DK, Jaenisch R, Young RA: **Control of developmental regulators by Polycomb in human embryonic stem cells.** *Cell* 2006, **125**:301-313.
78. Taberlay PC, Kelly TK, Liu C-C, You JS, De Carvalho DD, Miranda TB, Zhou XJ, Liang G, Jones PA: **Polycomb-repressed genes have permissive enhancers that initiate reprogramming.** *Cell* 2011, **147**:1283-1294.
79. You JS, Kelly TK, De Carvalho DD, Taberlay PC, Liang G, Jones PA: **OCT4 establishes and maintains nucleosome-depleted regions that provide additional layers of epigenetic regulation of its target genes.** *Proc Natl Acad Sci U S A* 2011, **108**:14497-14502.
80. Chew JL, Loh YH, Zhang W, Chen X, Tam WL, Yeap LS, Li P, Ang YS, Lim B, Robson P, Ng HH: **Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells.** *Mol Cell Biol* 2005, **25**:6031-6046.
81. Masui S, Nakatake Y, Toyooka Y, Shimamoto D, Yagi R, Takahashi K, Okochi H, Okuda A, Matoba R, Sharov AA, Ko MS, Niwa H: **Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells.** *Nat Cell Biol* 2007, **9**:625-635.
82. Ambrosetti DC, Basilico C, Dailey L: **Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites.** *Mol Cell Biol* 1997, **17**:6321-6329.
83. Schuetz A, Nana D, Rose C, Zocher G, Milanovic M, Koenigsmann J, Blasig R, Heinemann U, Carstanjen D: **The structure of the Klf4 DNA-binding domain links to self-renewal and macrophage differentiation.** *Cell Mol Life Sci* 2011, **68**:3121-3131.
84. Remenyi A, Lins K, Nissen LJ, Reinbold R, Scholer HR, Wilmanns M: **Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers.** *Genes Dev* 2003, **17**:2048-2059.
85. Scaffidi P, Bianchi ME: **Spatially precise DNA bending is an essential activity of the sox2 transcription factor.** *J Biol Chem* 2001, **276**:47296-47302.
86. Schöler HR, Ruppert S, Suzuki N, Chowdhury K, Gruss P: **New type of POU domain in germ line-specific protein Oct-4.** *Nature* 1990, **344**:435-439.
87. Jiang J, Chan Y-S, Loh Y-H, Cai J, Tong G-Q, Lim C-A, Robson P, Zhong S, Ng H-H: **A core Klf circuitry regulates self-renewal of embryonic stem cells.** *Nat Cell Biol* 2008, **10**:353-360.
88. Hall J, Guo G, Wray J, Eyres I, Nichols J, Grotewold L, Morfopoulou S, Humphreys P, Mansfield W, Walker R, Tomlinson S, Smith A: **Oct4 and LIF/Stat3 additively induce Kruppel factors to sustain embryonic stem cell self-renewal.** *Cell Stem Cell* 2009, **5**:597-609.
89. Jauch R, Aksoy I, Hutchins AP, Ng CKL, Tian XF, Chen J, Palasingam P, Robson P, Stanton LW, Kolatkar PR: **Conversion of Sox17 into a pluripotency reprogramming factor by reengineering its association with Oct4 on DNA.** *Stem Cells* 2011, **29**:940-951.
90. Yuan H, Corbi N, Basilico C, Dailey L: **Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3.** *Genes Dev* 1995, **9**:2635-2645.
91. Du JX, McConnell BB, Yang VW: **A small ubiquitin-related modifier-interacting motif functions as the transcriptional activation domain of Kruppel-like factor 4.** *J Biol Chem* 2010, **285**:28298-28308.
92. Ambrosetti DC, Scholer HR, Dailey L, Basilico C: **Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer.** *J Biol Chem* 2000, **275**:23387-23397.
93. Wang Y, Chen J, Hu J-L, Wei X-X, Qin D, Gao J, Zhang L, Jiang J, Li J-S, Liu J, Lai KY, Kuang X, Zhang J, Pei D, Xu GL: **Reprogramming of mouse and human somatic cells by high-performance engineered factors.** *EMBO Rep* 2011, **12**:373-378.
94. Hirai H, Katoku-Kikyo N, Karian P, Firpo M, Kikyo N: **Efficient iPSC cell production with the MyoD transactivation domain in serum-free culture.** *PLoS ONE* 2012, **7**:e34149.
95. Hirai H, Tani T, Katoku-Kikyo N, Kellner S, Karian P, Firpo M, Kikyo N: **Radical acceleration of nuclear reprogramming by chromatin remodeling with the transactivation domain of MyoD.** *Stem Cells* 2011, **29**:1349-1361.
96. Fong YW, Inouye C, Yamaguchi T, Cattoglio C, Grubisic I, Tjian R: **A DNA repair complex functions as an Oct4/Sox2 coactivator in embryonic stem cells.** *Cell* 2011, **147**:120-131.
97. Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH: **A protein interaction network for pluripotency of embryonic stem cells.** *Nature* 2006, **444**:364-368.
98. van den Berg DL, Snoek T, Mullin NP, Yates A, Bezstarosti K, Demmers J, Chambers I, Poot RA: **An Oct4-centered protein interaction network in embryonic stem cells.** *Cell Stem Cell* 2010, **6**:369-381.
99. Pardo M, Lang B, Yu L, Prosser H, Bradley A, Babu MM, Choudhary J: **An expanded Oct4 interaction network: implications for stem cell biology, development, and disease.** *Cell Stem Cell* 2010, **6**:382-395.
100. Mallanna SK, Ormsbee BD, Iacovino M, Gilmore JM, Cox JL, Kyba M, Washburn MP, Rizzino A: **Proteomic analysis of Sox2-associated proteins during early stages of mouse embryonic stem cell differentiation identifies Sox21 as a novel regulator of stem cell fate.** *Stem Cells* 2010, **28**:1715-1727.
101. Saxe JP, Tomilin A, Schöler HR, Plath K, Huang J: **Post-translational regulation of Oct4 transcriptional activity.** *PLoS ONE* 2009, **4**:e4467.
102. Jang H, Kim TW, Yoon S, Choi SY, Kang TW, Kim SY, Kwon YW, Cho EJ, Youn HD: **O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network.** *Cell Stem Cell* 2012, **11**:62-74.

doi:10.1186/gb-2012-13-10-251

Cite this article as: Schmidt R, Plath K: The roles of the reprogramming factors Oct4, Sox2 and Klf4 in resetting the somatic cell epigenome during induced pluripotent stem cell generation. *Genome Biology* 2012, **13**:251.