PROSPECT
Journal of Cellular Biochemistry 105:940–948 (2008)

Journal of Cellular Biochemistry

Retroviral Vector Silencing During iPS Cell Induction: An Epigenetic Beacon that Signals Distinct Pluripotent States

Akitsu Hotta[1] and James Ellis[1,2*]

[1]*Developmental and Stem Cell Biology Program, SickKids Hospital, Toronto, Ontario, Canada*

[2]*Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada*

ABSTRACT

Retroviral vectors are transcriptionally silent in pluripotent stem cells. This feature has been potently applied in studies that reprogram somatic cells into induced pluripotent stem (iPS) cells. By delivering the four Yamanaka factors in retroviral vectors, high expression is obtained in fibroblasts to induce the pluripotent state. Partial reprogramming generates Class I iPS cells that express the viral transgenes and endogenous pluripotency genes. Full-reprogramming in Class II iPS cells silences the vectors as the endogenous genes maintain the pluripotent state. Thus, retroviral vector silencing serves as a beacon marking the fully reprogrammed pluripotent state. Here we review known silencer elements, and the histone modifying and DNA methylation pathways, that silence retroviral and lentiviral vectors in pluripotent stem cells. Both retroviral and lentiviral vectors are influenced by position effects and often exhibit variegated expression. The best vector designs facilitate full-reprogramming and subsequent retroviral silencing, which is required for directed-differentiation. Current retroviral reprogramming methods can be immediately applied to create patient-specific iPS cell models of human disease, however, future clinical applications will require novel chemical or other reprogramming methods that reduce or eliminate the integrated vector copy number load. Nevertheless, retroviral vectors will continue to play an important role in genetically correcting patient iPS cell models. We anticipate that novel pluripotent-specific reporter vectors will select for isolation of high quality human iPS cell lines, and select against undifferentiated pluripotent cells during regenerative medicine to prevent teratoma formation after transplantation. J. Cell. Biochem. 105: 940–948, 2008.

KEY WORDS: INDUCED PLURIPOTENT STEM CELLS; RETROVIRUS SILENCING; HISTONE MODIFICATIONS; POSITION EFFECT VARIEGATION

Retroviral vectors are powerful tools to introduce foreign genes into host chromosomes in stem cells. However, retroviral transcription is subject to silencing or attenuation in pluripotent stem cells, including embryonic carcinoma (EC) cells, embryonic stem (ES) cells and early preimplantation embryos [Pannell and Ellis, 2001]. It is now evident that retroviral expression is silenced after direct reprogramming of somatic cells into induced pluripotent stem (iPS) cells [Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007]. This exciting breakthrough for regenerative medicine has many important applications. One unexpected outcome is that it not only supports the previous observations of retroviral silencing in ES cells, but also provides a novel iPS cell experimental platform to further interrogate the complex epigenetic mechanisms of vector silencing.

In the pioneering iPS cell report from Yamanaka's group [Takahashi and Yamanaka, 2006], an elegant screen was performed to identify a set of critical genes for inducing pluripotency. This screen employed a retroviral vector system for high efficiency gene delivery to introduce combinations of 24 candidate genes into mouse fibroblasts. Despite the challenge of integrating multiple genes into a single cell, they successfully identified a gene cocktail (Oct-4, Sox2, Klf4, and c-Myc) that directly reprograms mouse somatic cells into a pluripotent state. The authors correctly speculated that retroviral silencing must occur in their iPS cells because they were able to differentiate them into other cell types, which would be prevented if the pluripotency genes continued to express. Here, we review the use of retroviral vectors, and more particularly the epigenetic

Grant sponsor: Canadian Institutes for Health Research (CIHR); Grant number: FRN 81129; Grant sponsor: Stem Cell Network (SCN).

*Correspondence to: Dr. James Ellis, Senior Scientist, Developmental and Stem Cell Biology Program, SickKids Hospital, Toronto Medical Discovery Tower, MaRS Centre, 101 College St., Room 13-310, Toronto, Ontario, Canada M5G 1L7. E-mail: jellis@sickkids.ca

Received 1 August 2008; Accepted 6 August 2008 • DOI 10.1002/jcb21912 • 2008 Wiley-Liss, Inc.

Published online 4 September 2008 in Wiley InterScience (www.interscience.wiley.com).

mechanisms of retroviral silencing, during direct reprogramming into iPS cells.

RETROVIRAL VECTOR INDUCTION OF THE PLURIPOTENT STEM CELL STATE

iPS cells are artificially induced into different states of pluripotency. Pluripotency is defined as the ability to give rise to all cell types except for extra-embryonic tissues. Experimentally, pluripotency is determined either by teratoma formation in mice to form organized tissues belonging to the three germ layers, or by the generation of chimeric mice to test tissue contribution in vivo [Jaenisch and Young, 2008]. However, the first indication of pluripotency during iPS cell induction is often the activation of a pluripotent-specific reporter gene accompanied by colony morphology changes reminiscent of ES cells. In the mouse system, the first reporter used was the β-geo selectable marker knocked into the Fbx15 locus [Takahashi and Yamanaka, 2006], but better Oct-4-neo knockins [Wernig et al., 2007] or Nanog-puro BAC transgenes [Okita et al., 2007] subsequently proved to be more accurate indicators of reprogramming.

When Yamanaka's group used Fbx15 expression for selection of mouse iPS cells, the resulting Fbx15-iPS cells were pluripotent by teratoma formation. However, global gene expression and DNA methylation patterns were distinct from ES cells, and these Fbx15-iPS cells did not generate adult chimeric mice [Takahashi and Yamanaka, 2006]. On the other hand, selection for Nanog or Oct-4 expression by several groups has shown that these mouse iPS cells successfully contribute to adult chimeras and transmit through the germ line [Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007]. In addition, their global gene expression and DNA methylation patterns were much more similar to ES cells [Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007]. These observations suggest that Fbx15 selection tends to isolate "partially reprogrammed" iPS cells, and Nanog (or Oct-4) selection enriches for "fully reprogrammed" iPS cells [Jaenisch and Young, 2008].

With respect to retroviral silencing, partially reprogrammed iPS cells still express high levels of the exogenous pluripotent transgenes present in the retrovirus, whereas fully reprogrammed iPS cells consistently exhibit retroviral silencing or attenuation [Okita et al., 2007; Jaenisch and Young, 2008] (Fig. 1). Retroviral expression assessed by fluorescence never overlaps with endogenous Nanog [Nakagawa et al., 2008] or Sox2 [Stadtfeld et al., 2008] in reactivated iPS cells. This implies that expression of exogenous Yamanaka factors is required only for the initial activation of self-renewal genes. Thus, fully reprogrammed iPS cells maintain their pluripotency independent of transgene expression. This conclusion is supported by experiments that employ doxycycline-inducible systems to limit the temporal expression of the four transgenes to the first 10–14 days of reprogramming [Maherali et al., 2007; Brambrink et al., 2008; Stadtfeld et al., 2008]. The specific order of reprogramming events appears to be activation of the alkaline phosphatase gene and silencing of fibroblast specific expression, followed by expression of SSEA-1, and finally progressive silencing of the retroviral vectors as the endogenous pluripotent Oct-4 and

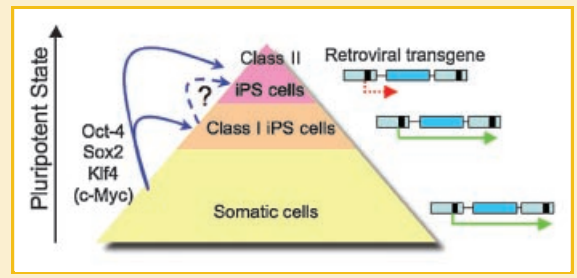

Fig. 1. Somatic cells are reprogrammed into iPS cells by retroviral gene delivery of the four Yamanaka factors (Oct-4, Sox2, Klf4, c-Myc). As the epigenetic modifications are reset, the partially reprogrammed Class I iPS cells continue to express the retroviral transgenes. Upon full reprogramming, Class II iPS cells resemble ES cells in their epigenetic landscapes and transcriptional profiles and also silence retroviral transgenes.

Nanog genes are induced by day 16 [Brambrink et al., 2008; Stadtfeld et al., 2008; Wernig et al., 2008a].

In the human system, there is no effective and easily introduced pluripotent stem cell reporter, and therefore potential human iPS cells are isolated by colony morphology or live-cell staining for endogenous surface markers prior to colony picking. Using these approaches, it was shown that the same four Yamanaka factors induce human iPS cell formation [Takahashi et al., 2007]. As with mouse iPS cells, the C-MYC gene can be omitted but leads to slower and less efficient reprogramming by the remaining three factors [Nakagawa et al., 2008]. In addition, a four factor combination of OCT-4, SOX2, NANOG, and LIN28 can reprogram human somatic cells [Yu et al., 2007], and additional genes such as TERT and SV40 large T may facilitate the process [Mali et al., 2008; Park et al., 2008]. While human iPS cells take longer to be induced and are grown in different culture conditions than mouse iPS cells, they share the phenomena of partially reprogrammed colony formation and retroviral silencing [Takahashi et al., 2007; Lowry et al., 2008; Masaki et al., 2008; Park et al., 2008].

Here, we relate the reprogramming state of iPS cells to their retrovirus expression: Class I are partially reprogrammed iPS cells that are permissive for retroviral expression (i.e., mouse Fbx15-iPS cells), and Class II are fully reprogrammed iPS cells that are repressive for retroviral expression (i.e., mouse Nanog-iPS cells). The chromatin state of Class I and Class II iPS cells are more similar to each other than to fibroblasts, but Class I and Class II cells differ in being permissive and repressive environments for retroviral expression respectively (Fig. 1). These two closely related classes of iPS cells should prove useful for answering fascinating unresolved questions of retroviral silencing. In particular, how do pluripotent stem cells silence retroviral transgenes? Which factors are involved in initiating and maintaining this process?

RETROVIRAL VECTOR SILENCER ELEMENTS

Yamanaka's group and Jaenisch's group used Moloney murine leukemia virus (MoMLV) derived vectors (pMXs and pLIB,

respectively) for iPS cell derivation. In these vectors, transgene transcription is initiated at the MoMLV long terminal repeat (LTR) promoter. Historically, MoMLV has been extensively studied for retroviral silencing since Teich et al. [1977] found that MoMLV did not replicate in EC cells some 30 years ago. This finding lead to many studies attempting to identify negative regulatory elements called silencers in MoMLV sequences and their cognate binding factors. To date, the known silencers in MoMLV have primarily been mapped to LTR regions including the negative control region (NCR), direct repeat (DR) enhancer, CpG rich promoter and primer binding site (PBS). Detailed molecular characterizations have identified that YY-1 binds to the NCR, ELP (also known as Nr5a1, nuclear receptor subfamily five group A member one) binds between the NCR and DR, and Factor A (recently identified as a large complex including a Trim28 subunit [Wolf and Goff, 2007]) binds to the PBS (Fig. 2). These silencers play an important role in transcriptional shut down of MoMLV LTR expression in ES cells. Other retroviral vectors designed to overcome stem cell silencing have deleted or mutated some of the silencers. One example already used in iPS cell experiments is the murine stem cell virus (MSCV) derived vector (pMIG). In this case, MSCV transduced pluripotency transgenes are still silenced in human iPS cells [Park et al., 2008], which is an important feature of fully reprogramed Class II iPS cells that facilitates their subsequent differentiation into embryoid bodies or specific cell types.

Lentiviral vectors were originally reported to escape silencing in ES cells and transgenic mice [Lois et al., 2002; Pfeifer et al., 2002], and therefore may be less suitable for reprogramming experiments. In fact, lentiviral vectors are often completely silent in ES cells [Yao et al., 2004; Ellis, 2005]. Consistent with this observation, lentiviral vectors have been used for successful induction of mouse and human iPS cells, although it was not shown that they were silenced [Blelloch et al., 2007; Yu et al., 2007]. When retroviral and lentiviral vectors do express in ES cells, they are always attenuated to significantly lower levels than in fibroblasts, suggesting reduced levels of certain activating factors or more potent repressor factors. Finally, retroviral or lentiviral vector infected ES cells can also be variegated in which subsets of genetically identical cells either express the vector or are silent [Swindle et al., 2004; Yao et al., 2004]. Together, these results suggest that position effects influence vector expression. These features of vector silencing need to be incorporated into mechanistic models of retroviral silencing and their implications for iPS cell induction considered.

Many studies have shown that multiple epigenetic pathways are involved in the silencing of retroviruses and lentiviruses [Ellis, 2005]. Specifically, histone modifications and DNA methylation have been evaluated by several investigators, and play a central role in marking the silenced state (Fig. 3). Here, we first discuss several epigenetic players that participate in retroviral silencing before addressing the role of position effects.

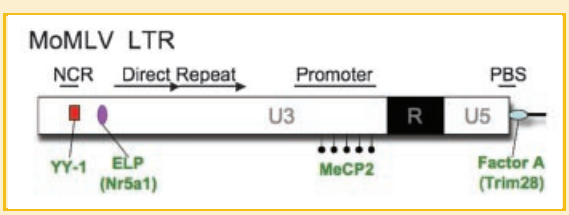

Fig. 2. Retroviral silencers identified in the MoMLV LTR promoter region. Trans-acting factors (indicated in Green) directly recognize specific DNA sequences or DNA methylation (lollipops). NCR: Negative control region, PBS: Primer binding site.

HISTONE DEACETYLATION

Histone acetylation is the best characterized mark associated with transcriptional activation. Acetylation of histones leads to open euchromatin formation and active transcription, whereas deacetylation corresponds to the silent heterochromatin state. Indeed, acetylation marks are enriched on actively expressing retrovirus while silent retrovirus is deacetylated. Deacetylation of histone H3 is more pronounced in comparison to histone H4 on silent retroviral vectors in ES cells [Yao et al., 2004]. A similar decrease in histone H3 acetylation is found at silent lentivirus in P19 EC cells [He et al., 2005]. To date, there is no histone deacetylase (HDAC) that specifically catalyzes deacetylation of histone H3, making it difficult to speculate which HDAC participates in retroviral silencing in pluripotent stem cells. Possibly, a co-factor of HDACs may determine the specificity (i.e., HDAC1/2 can interact with Nanog in ES cells [Wang et al., 2006; Liang et al., 2008]). Alternatively, reduced levels of histone acetyltransferases (HAT) may contribute to the loss of acetylation. It is notable that treatment with HDAC inhibitors such as trichostatin A (TSA) appears to be most effective at reactivating retrovirus expression early after infection of ES cells [Lorincz et al., 2000], suggesting that HDACs may play a functional role during establishment that may not be required for maintenance of silencing.

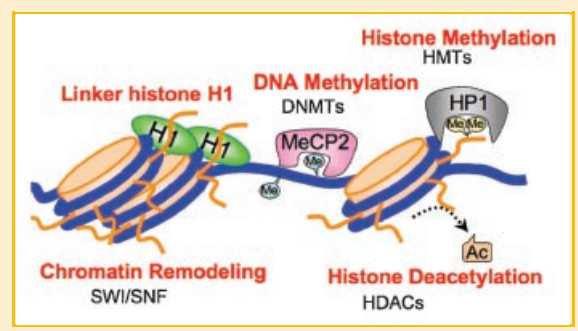

Fig. 3. Multiple epigenetic pathways are involved in retroviral silencing through structural or positional alterations to nucleosomes (linker histone H1 and chromatin remodeling factors), DNA methylation and histone modifications (methylation and deacetylation). Many of the pathways reinforce each other through mutual recruitment. DNMTs: DNA methyltransferases, HMTs: Histone methyltransferases, HDACs: Histone deacetylases.

HISTONE METHYLATION

Methylation of histone tails specifically at histone 3 lysine 9 (H3K9) or 27 (H3K27) is recognized as being repressive, but histone 3 lysine 4 (H3K4) methylation is an active transcription mark. Histone methylation marks have recently been identified on silent retrovirus. Wolf and Goff [2007] identified enriched H3K9 di-methylation and HP1γ recruited by Trim28 near the PBS silencer element (Fig. 3). In addition, on silent lentivirus in P19 EC cells there is loss of H3K4 di-methylation and increase of H3K9 di-methylation [He et al., 2005]. Currently, several histone methyltransferases (HMTs) have been identified that write H3K9 methylation marks. For example, Ehmt1 (=GLP) and Ehmt2 (=G9a) are responsible for mono- and di-methylation, Setdb1 (=ESET) catalyses di- and tri-methylation, and Suv39h1 and Suv39h2 add the tri-methylation mark [Kouzarides, 2007]. These marks are read by the HP1 family of proteins that may also contribute to the recruitment of DNA methyltransferases. Finally, the YY-1 factor recruits polycomb group complexes to methylate H3K27 and deacetylate histones [Srinivasan and Atchison, 2004]. It will be of great interest to determine the function of these factors on retroviral silencing in pluripotent stem cells.

CHROMATIN REMODELING, LINKER HISTONES, AND RNA MEDIATED SILENCING

Other chromatin factors also influence retroviral silencing. Brm, a catalytic subunit of the SWI/SNF complex that controls nucleosome positioning, is required for consistent viral expression in human tumor cell lines [Mizutani et al., 2002]. In addition, the linker histone H1 facilitates the folding of chromatin into a condensed 30 nm fiber. Histone H1 is generally associated with transcriptional repression, and is detected on silent retrovirus and lentivirus in ES cells [Yao et al., 2004]. Both SWI/SNF and H1 may have effects on DNA methylation [Fan et al., 2005] and the DNA methylation binding factor MeCP2 [Harikrishnan et al., 2005], demonstrating the complexity of potential epigenetic pathways involved in retroviral silencing. Finally, small RNAs (rasiRNA, repeat associated siRNA) have been shown to repress endogenous retrotransposons in mouse oocytes [Watanabe et al., 2006] and could have effects on retroviral silencing in stem cells.

DNA METHYLATION

Cytosine methylation of DNA has long been implicated in retroviral silencing. Silent retrovirus is heavily methylated at CpG sites [Cherry et al., 2000; Minoguchi and Iba, 2008], which are bound by MeCP2 [Lorincz et al., 2001; Yao et al., 2004]. Silent lentiviral vectors are also heavily methylated in P19 EC cells [He et al., 2005]. The question of the functional importance of DNA methylation in retroviral silencing has been studied using ES cells that are null for the DNA methyltranferases. These experiments show that the de novo methyltransferases Dnmt3a and Dnmt3b are not required to establish retroviral silencing over the first 10 days [Pannell et al., 2000]. In addition, DNA methylation is not detected immediately after infection and inhibitors of DNA methylation such as 5-azaC require up to 5 days before significant levels of retroviral expression is reactivated [Yao et al., 2004; Ramunas et al., 2007]. We interpret this data as indicating that DNA methylation is not required to establish silencing, but when present plays a functional role in the later maintenance of silencing. Notably, histone H1 has been shown to affect DNA methylation of imprinted genes in knockout ES cells [Fan et al., 2005], suggesting a potential role for histone H1 mediated recruitment of DNA methylation.

In terms of silencing in iPS cells, there is some evidence connecting DNA methylation and retroviral silencing. For example, retroviral transcription is inversely correlated with the Dnmt3a2 (Dnmt3a, isoform 2) expression levels [Okita et al., 2007]. Dnmt3a levels peak at day 3 of reprogramming while Dnmt3b and Trim28 levels gradually increase with their highest expression at day 13 of iPS cell induction [Stadtfeld et al., 2008]. Finally, silent retrovirus in Class II iPS cells can be reactivated by knock-down of Dnmt1, which is responsible for the maintenance of DNA methylation [Wernig et al., 2007]. Interestingly, Class I iPS cells have an abnormal DNA methylation pattern at ES-specific gene promoters that is intermediate between fibroblasts and ES cells, whereas Class II iPS cells are more similar to ES cells. In addition, 5-azaC treatment of Class I iPS cells induced full-reprogramming that eventually led to retroviral silencing [Mikkelsen et al., 2008]. These data indicate that epigenetic remodeling in Class I cells is incomplete, and that the DNA methylation state must be reset to complete the transition to fully-reprogrammed Class II iPS cells. Once the pluripotent cell epigenetic and transcriptional networks are established, retroviral silencing also occurs.

RETROVIRAL SILENCING IN PLURIPOTENT STEM CELLS

It is now abundantly evident that several epigenetic pathways are involved in retroviral silencing that may have a redundant temporal order, but it is not clear what initiates silencing in pluripotent stem cells. It is puzzling that the known factors that recognize MoMLV silencers are expressed ubiquitously. For example, the Factor A repressor is enriched in stem cells but its recently identified Trim28 component is detected in partially reprogrammed iPS cells [Stadtfeld et al., 2008] and other cell types that do not silence retrovirus [Wolf and Goff, 2007]. Potentially a co-factor of Trim28 that is part of the Factor A complex, such as one of the large family of KRAB-ZNF DNA binding proteins that interact with Trim28, may contribute to the specificity of retroviral silencing [Ellis et al., 2007]. Alternatively, it could be that there is no genuinely stem cell specific factor for recognition of the vector and establishment of retroviral silencing. Since the relative levels of reprogramming factors such as Oct-4 can dramatically affect the fate of the cell [Niwa et al., 2000], it is conceivable that retroviral silencing is a consequence of similar minor differences of factor concentrations in stem cells. For example, higher levels of Trim28 and a specific KRAB-ZNF protein might predispose stem cells to form the Factor A repressor complex and silence retroviral vectors via the PBS element.

To address the initiator of retroviral silencing in stem cells, it will be important to move beyond the current focus on screening the function or binding of candidate epigenetic factors. This approach can investigate which epigenetic modifications correlate with the silent state, and can test which modifying enzymes are critical for writing and reading the marks for functional silencing. However, to identify the factors that directly recognize the vector and initiate silencing in stem cells, it will be important to employ broader screens. Recently, candidate siRNA screens identified genes such as HDAC1 and HP1γ that reactivate an avian sarcoma virus (ASV) vector in a TSA-treatment sensitive clone of HeLa cells [Poleshko et al., 2008]. Similar full genome screens may be required to identify the factors involved in establishment of silencing in ES cells and iPS cells.

POSITION EFFECTS AND SILENCING IN ES CELLS

Because the same retroviral construct can be silent, variegated or attenuated in ES cells, position effects emanating from the genomic integration site must play a role in defining the chromatin and expression state. Strictly speaking, position effects could be used to argue that silent virus has integrated into heterochromatin, and expressing virus is in euchromatin. In this scenario, variegated virus would be expected to have integrated near heterochromatin that spreads dynamically over time and through multiple cell generations. As would be expected by this model, retroviral integration sites in mouse ES cells can be dynamically variegated as observed using EGFP real time imaging of progeny cells over several generations [Katz et al., 2007; Ramunas et al., 2007]. Most cells either maintain silencing or expression over this time frame, but some are caught in the act of switching on or off. Interestingly, the cells that do express are subject to transcriptional noise that reveals itself as small but distinct fluctuations of expression. These small transcriptional effects can lead to dramatically different expression levels in genetically identical cells over time. In this respect, retroviral vector expression in stem cells may mimic the variegated expression pattern of the endogenous Nanog gene [Chambers et al., 2007].

Position effect variegation is generally explained by the spread of nearby heterochromatin into the transgene. We have argued previously that retroviral variegation is a consequence of the balance between activator and repressor factors binding to the virus combined with the influence of the surrounding genomic position effects. However, recent genome wide chromatin evaluations of mouse ES cells indicates that pluripotent stem cells have more open chromatin than mature differentiated cells [Azuara et al., 2006], and this chromatin is hyperdynamic dependent on the kinetics of HP1 and histone H1 [Meshorer et al., 2006]. Moreover, ES cells have extensive regions of bivalent chromatin that are marked by both silent (H3K27me) and active (H3K4me) histone modifications [Bernstein et al., 2006]. These reports suggest that most retroviral integration sites in ES cells are in open or bivalent chromatin, consistent with their known preference for integrating into promoter elements. Even lentivirus vectors prefer to integrate into the body of expressing genes marked by H3Ac, H4Ac, and H3K4me1, but disfavored to H3K27me3 and DNA methylation sites [Wang et al., 2007]. Given this discovery and the absence of global integration site data to support retroviral variegation being near genomic heterochromatin, it may be worthwhile to consider alternative mechanisms. Conceivably these mechanisms could include accumulated transcriptional noise effects due to nuclear fluctuations in trans-factor concentrations. Since genes that are transcriptionally active tend to be located near transcription factories in more central neighborhoods of the nucleus, it is also possible that variegated retroviral vectors are directed to peripheral nuclear locations when silent and more central locations when active [Wiblin et al., 2005]. Finally, it has been reported that lentiviral LTRs loop to interact with each other when transcribed for proper RNA processing [Perkins et al., 2008]. It is possible that variegated virus temporarily loops to interact with nearby endogenous regulatory elements that repress or activate the virus. To fully understand retroviral variegation in pluripotent stem cells, these issues must be examined carefully.

POSITION EFFECTS AND RETROVIRAL VARIEGATION IN iPS CELLS

What impact might position effects and variegation have on iPS cell induction? When using retroviral vectors to reprogram somatic cells, it is likely that the initial integration sites in fibroblasts are in euchromatin. Positive position effects from surrounding enhancers or other elements may influence the expression level of each retroviral integration. Some cells may obtain an advantageous stoichiometry or relative levels of the reprogramming factors from this process or through differences in vector copy numbers. However, once the progeny of that cell is induced into a fully reprogrammed iPS cell colony many days later, these integration sites may have altered chromatin structure [Mikkelsen et al., 2008]. Only housekeeping genes would stay open, while tissue specific genes would likely become closed via repressive histone marks and/or DNA methylation. Some key developmental genes would obtain bivalent marks and be poised to express at very low levels. These chromatin structure reorganizations might reduce the viral expression levels as is seen in some Class II iPS cell lines. If negative position effects are now present from heterochromatin surrounding the virus after reprogramming, this may complete the transcriptional shut off or facilitate variegation [Wernig et al., 2008a]. However, we have already discussed that pluripotent stem cells generally have more open chromatin than fibroblasts. Therefore we propose that complete silencing is dependent on active recognition of the virus by factors that are enriched in stem cells, or by complexes that form specifically in stem cells. Comparison of the differences between Class I and II iPS cells may be exceptionally valuable to make progress in this area.

One aspect of position effects is the possibility that preferred integration sites contribute to iPS cell formation. If the stoichiometry of the four factors is crucial to establish iPS cells, then there may be a subset of integration sites that favor the optimal factor ratios. Alternatively, the retrovirus may integrate near to and activate an additional gene required for pluripotency. Analysis of retroviral integration sites in four different iPS cell lines suggests

that there is no single common integration site among the lines [Aoi et al., 2008], although contributions from many independent genes cannot be formally excluded [Hawley, 2008]. Importantly, once the fibroblast has reprogrammed into a Class II iPS cell the virus might be expected to be silenced. We therefore consider it unlikely that surrounding genes would continue to be activated in iPS cells.

IMPORTANCE OF VIRAL SILENCING IN THE INDUCTION OF iPS CELLS

What is the significance of retroviral silencing in iPS cell biology? It is reported that Oct-4 expression levels are precisely regulated in ES cells to maintain their pluripotency. For example, just two-fold overexpression of Oct-4 causes differentiation into primitive endoderm and mesoderm, and down regulation of Oct-4 causes differentiation into trophectoderm [Niwa et al., 2000]. This means that the total amount of Oct-4 must be consistent during the intermediate stages of reprogramming (Fig. 4). If viral transgenes get silenced earlier than the establishment of the endogenous pluripotency transcriptional network, then partially reprogrammed iPS cells could easily revert back to the somatic state, or progress down the trophectoderm lineage. At the same time, if viral transgenes do not get silenced after the establishment of pluripotency, then the cells could be forced to differentiate due to excess amounts of Oct-4 or may become stable Class I iPS cell lines. In fact, extended expression of the factors by a doxycycline-inducible promoter does stabilize them into Class I iPS cells [Mikkelsen et al., 2008]. Therefore, accurate switching between exogenous and endogenous gene expression while maintaining factor levels consistent with pluripotency could be a rare event that contributes to the low efficiency of iPS cell induction. Of course the use of doxycycline-inducible vectors allows greater control of the timing of the switch without relying on retroviral silencing mechanisms. Nevertheless, doxycycline-inducible systems generally use a transgenic mouse expressing the Tet transactivator protein. It may be more challenging to apply this system to human cells because of the need to introduce the reprogramming factors as well as the Tet transactivator, and these inducible systems are frequently leaky preventing complete transgene shutdown.

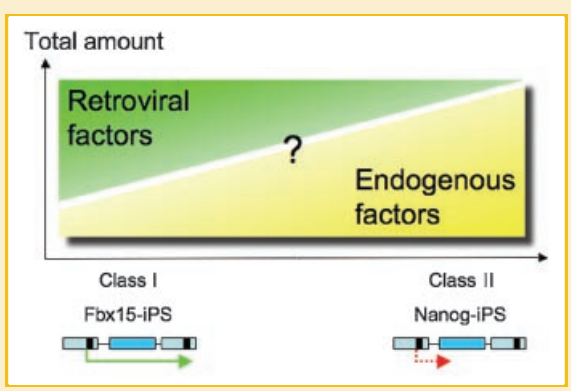

Fig. 4. Class I iPS cells identified through Fbx15 selection continue to express the retroviral transgenes. To complete reprogramming into Class II iPS cells as identified under Nanog selection, the total expression of the Yamanaka factors must be conserved to maintain the pluripotent state as the retroviral transgenes are silenced. If the viral transgene shuts off before full reprogramming, or still expresses after activation of the endogenous gene, the cells may not be able to maintain their pluripotency.

iPS CELLS TO MODEL HUMAN DISEASE

The current retroviral vector based reprogramming methods are being used to make models of mouse and human disease. For example, by generating iPS cells from patients with neurological or cardiac disease it will be possible to differentiate these cells into neurons or beating cardiomyocytes to study the disease process. To be successful, retroviral silencing of the reprogramming factors is extremely important to allow differentiation to proceed using protocols established for ES cells. If the Yamanaka factors continue to express, they could interfere with endogenous cell signaling pathways affecting the normal differentiation of pluripotent stem cells [Brambrink et al., 2008; Takahashi and Yamanaka, 2006].

The choice of vector design continues to be important for patient derived iPS cell models. First generation MoMLV based vectors using the LTR as a promoter are a better choice than MSCV or self-inactivating lentiviral vectors that employ a ubiquitous internal promoter for introducing the reprogramming factors. Simple improvements in design could include the use of loxP sites to remove the reprogramming vectors after iPS cell induction [Hanna et al., 2007], or the use of doxycycline-inducible transgenes on vectors that co-express the Tet transactivator. To reduce the copy number of integrated vectors, some transient chemical methods have been developed that use the HDAC inhibitor Valproic Acid (VPA) [Huangfu et al., 2008] or the G9a histone methyltransferase inhibitor (BIX-01294) [Shi et al., 2008] to enhance the frequency of mouse iPS cell reprogramming and replace individual factors such as c-Myc and Oct-4. These approaches once again emphasize the importance of epigenetics in the induction of iPS cell lines. Finally, reprogramming of mouse neural stem cells does not require the Sox2 factor, and Oct-4 and Klf4 are sufficient [Kim et al., 2008], but even the Oct-4 vector can be replaced by BIX-01294 action [Shi et al., 2008]. As C-MYC, TERT, and SV40 Large T are all oncogenic, it will be preferable to use fewer factor infections without them to generate patient specific iPS cells and avoid secondary transformation effects. By choosing the optimal cell type and reprogramming method, it is likely that patient specific human iPS cell lines can also be made with reduced oncogenic potential and lower vector copy number loads through the combined use of retroviral vectors and/or appropriate chemicals.

Regardless of the reprogramming method used, independent iPS cell lines derived from the same patient will continue to be heterogeneous in their differentiation potential and each one will have slightly different expression and epigenetic profiles. To characterize patient-specific iPS cell lines to study disease, it will be important to quickly identify the most appropriate lines. Generation of chimeric mice or tetraploid complementation to test germ line transmission is the ultimate test for fully reprogrammed Class II mouse iPS cell lines. However, these experiments are not only

technically demanding and time consuming, they cannot be applied to human iPS cells due to ethical issues and legal regulations. Currently, teratoma formation is the standard way to show in vivo functional pluripotency, however, it is difficult to assess the quality of iPS cells (Class I vs. Class II) because it is not a quantitative assay. One way to prioritize newly isolated human iPS cell lines for analysis after using the retroviral vector system would be to show that all the retroviral transgenes are silent by qRT-PCR, together with quantitative expression of other definitive endogenous pluripotent markers for Class II iPS cells. This early evaluation can be quickly performed to narrow down candidate lines for testing by teratoma formation and lineage-directed differentiation in vitro.

It is not clear whether the iPS cell lines must be fully pluripotent to be useful in modeling human disease. For example, a line that preferentially differentiates into a relatively pure population of beating cardiomyocytes but fails to efficiently make cells derived from other germ layers may actually be well suited to study cardiac disease. It is important to recognize that differentiation procedures may need to be subtly optimized for each iPS cell line. In the future, somatic cells may be directly reprogrammed into lineage restricted progenitors and fully reprogrammed iPS cell lines may not be required for all disease models.

iPS CELLS FOR THERAPEUTIC APPLICATIONS

iPS cell technology has the potential for future stem cell therapy applications that avoid immune rejection through the use of the patient's own cells. In this case the most realistic candidate diseases are those that are treatable by transplantation of well characterized adult stem cells or specific cell types rather than complex whole organs such as kidney or lung. It has already been demonstrated that iPS cells derived from mice with sickle cell anemia can be genetically corrected by homologous recombination at the β-globin locus prior to generation of hematopoietic stem cells and transplantation back into recipient mice [Hanna et al., 2007]. Parkinson's disease in a rat model can also be ameliorated by transplantation of iPS cell derived neural stem cells [Wernig et al., 2008b]. While these experiments are proof of principle that iPS cell therapy can succeed, the real concern of teratoma formation by residual undifferentiated iPS cells [Wernig et al., 2008b] and the possibility of oncogenesis caused by the retroviral vectors remains [Okita et al., 2007]. While these risks may be overstated, there are no current clinical trials using autologous human ES cells, which may make it difficult to translate the current iPS cell technology into clinical applications. Ultimately, clinical trials will require improved reprogramming methods that preferably employ transient protocols such as chemicals, non-integrating vectors, or factor administration through protein transduction to reduce or eliminate the vector copy number. Genetic correction of disease mutations in patient iPS cells by homologous recombination [Hanna et al., 2007] may be extremely challenging given the difficulties encountered with gene targeting in human ES cells. Therefore, gene therapy approaches may be more successful, and could easily involve gene transfer via retroviral or lentiviral vectors.

Finally, to enhance the throughput of human iPS cell induction experiments, the development of small pluripotent reporter genes that can be easily transferred into primary patient derived tissues would have great utility. Through a better understanding of the regulatory elements that control the pluripotent transcription network, it may soon be possible to create an effective reporter gene that is activated as reprogramming progresses equivalent to the endogenous Nanog or Oct-4 loci. By delivering such a reporter at high efficiency in a self-inactivating retroviral or lentiviral vector, it should be possible to positively select for high quality iPS cell lines. In addition, such a selectable marker would be a valuable tool for screens of novel compounds and methods of reprogramming human somatic cells into iPS cells. Another potential use for a pluripotent-specific retroviral vector is to deliver a HSV1 thymidine kinase (TK) or other suicide gene into iPS cell lines. Negative selection against TK using gancyclovir could be used to eliminate any undifferentiated iPS cells and prevent teratoma formation in vivo.

In summary, it will be important to understand how retroviral silencing is initiated and maintained in pluripotent stem cells. Factors that play a central role in the recognition and initiation of retroviral silencing may also be functionally important for maintaining pluripotency, or may help to convert partially reprogrammed iPS cells into the fully reprogrammed state.

ACKNOWLEDGMENTS

We thank Aaron Cheung and Arash Khairandish for thoughtful comments on iPS cells and vector variegation respectively, and lab members for useful discussion. We acknowledge funding from the Canadian Institutes for Health Research (CIHR Grant FRN 81129) and the Stem Cell Network (SCN) to JE. AH is supported by a Senior Fellowship Award from SickKids Hospital Restracomp.

REFERENCES

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

Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, Fisher AG. 2006. Chromatin signatures of pluripotent cell lines. Nat Cell Biol 8:532–538.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125:315–326.

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

Brambrink T, Foreman R, Welstead GG, Lengner CJ, Werning M, Suh H, Jaenisch R. 2008. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell 2:151–159.

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

Cherry SR, Biniszkiewicz D, van Parijs L, Baltimore D, Jaenisch R. 2000. Retroviral expression in embryonic stem cells and hematopoietic stem cells. Mol Cell Biol 20:7419–7426.

Ellis J. 2005. Silencing and variegation of gammaretrovirus and lentivirus vectors. Hum Gene Ther 16:1241–1246.

Ellis J, Hotta A, Rastegar M. 2007. Retrovirus silencing by an epigenetic TRIM. Cell 131:13–14.

Fan Y, Nikitina T, Zhao J, Fleury TJ, Bhattacharyya R, Bouhassira EE, Stein A, Woodcock CL, Skoultchi AI. 2005. Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. Cell 123:1199–1212.

Hanna J, Wernig M, Markoulaki S, Sun CW, Meissner A, Cassady JP, Beard C, Brambrink T, Wu LC, Townes TM, Jaenisch R. 2007. Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. Science 318:1920–1923.

Harikrishnan KN, Chow MZ, Baker EK, Pal S, Bassal S, Brasacchio D, Wang L, Craig JM, Jones PL, Sif S, El-Osta A. 2005. Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. Nat Genet 37:254–264.

Hawley RG. 2008. Does retroviral insertional mutagenesis play a role in the generation of induced pluripotent stem cells? Mol Ther 16:1354–1355.

He J, Yang Q, Chang LJ. 2005. Dynamic DNA methylation and histone modifications contribute to lentiviral transgene silencing in murine embryonic carcinoma cells. J Virol 79:13497–13508.

Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, Chen AE, Melton DA. 2008. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol 26:795–797.

Jaenisch R, Young R. 2008. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell 132:567–582.

Katz RA, Jack-Scott E, Narezkina A, Palagin I, Boimel P, Kulkosky J, Nicolas E, Greger JG, Skalka AM. 2007. High-frequency epigenetic repression and silencing of retroviruses can be antagonized by histone deacetylase inhibitors and transcriptional activators, but uniform reactivation in cell clones is restricted by additional mechanisms. J Virol 81:2592–2604.

Kim JB, Zaehres H, Wu G, Gentile L, Ko K, Sebastiano V, Arauzo-Bravo MJ, Ruau D, Han DW, Zenke M, Scholer HR. 2008. Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. Nature 454:646–650.

Kouzarides T. 2007. Chromatin modifications and their function. Cell 128:693–705.

Liang J, Wan M, Zhang Y, Gu P, Xin H, Jung SY, Qin J, Wong J, Cooney AJ, Liu D, Songyang Z. 2008. Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat Cell Biol 10:731–739.

Lois C, Hong EJ, Pease S, Brown EJ, Baltimore D. 2002. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. Science 295:868–872.

Lorincz MC, Schubeler D, Goeke SC, Walters M, Groudine M, Martin DI. 2000. Dynamic analysis of proviral induction and De Novo methylation: Implications for a histone deacetylase-independent, methylation density-dependent mechanism of transcriptional repression. Mol Cell Biol 20:842–850.

Lorincz MC, Schubeler D, Groudine M. 2001. Methylation-mediated proviral silencing is associated with MeCP2 recruitment and localized histone H3 deacetylation. Mol Cell Biol 21:7913–7922.

Lowry WE, Richter L, Yachechko R, Pyle AD, Tchieu J, Sridharan R, Clark AT, Plath K. 2008. Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc Natl Acad Sci USA 105:2883–2888.

Maherali N, Sridhara R, Xie W, Utikal J, Eminli S, Arnold K, Stadtfeld M, Yachechko R, Tchieu J, Jaenisch R, Plath K, Hochedlinger K. 2007. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell 1:55–70.

Mali P, Ye Z, Hommond HH, Yu X, Lin J, Chen G, Zou J, Cheng L. 2008. Improved efficiency and pace of generating induced pluripotent stem cells from human adult and fetal fibroblasts. Stem Cells 26:1998–2005.

Masaki H, Ishikawa T, Takahashi S, Okumura M, Sakai N, Haga M, Kominami K, Migita H, McDonald F, Shimada F, Sakurada K. 2008. Heterogeneity of pluripotent marker gene expression in colonies generated in human iPS cell induction culture. Stem Cell Res 1:1–2.

Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. 2006. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev Cell 10:105–116.

Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, Bernstein BE, Jaenisch R, Lander ES, Meissner A. 2008. Dissecting direct reprogramming through integrative genomic analysis. Nature 454: 49–55.

Minoguchi S, Iba H. 2008. Instability of retroviral DNA methylation in embryonic stem cells. Stem Cells 26:1166–1173.

Mizutani T, Ito T, Nishina M, Yamamichi N, Watanable A, Iba H. 2002. Maintenance of integrated proviral gene expression requires Brm, a catalytic subunit of SWI/SNF complex. J Biol Chem 277:15859–15864.

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

Niwa H, Miyazaki J, Smith AG. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet 24:372–376.

Okita K, Ichisaka T, Yamanaka S. 2007. Generation of germline-competent induced pluripotent stem cells. Nature 448:313–317.

Pannell D, Ellis J. 2001. Silencing of gene expression: Implications for design of retrovirus vectors. Rev Med Virol 11:205–217.

Pannell D, Osborne CS, Yao S, Sukonnik T, Pasceri P, Karaiskakis A, Okano M, Li E, Lipshitz HD, Ellis J. 2000. Retrovirus vector silencing is de novo methylase independent and marked by a repressive histone code. EMBO J 19:5884–5894.

Park IH, Zhao R, West JA, Yabuuchi A, Huo H, Ince TA, Lerou PH, Lensch MW, Daley GQ. 2008. Reprogramming of human somatic cells to pluripotency with defined factors. Nature 451:141–146.

Perkins KJ, Lusic M, Mitar I, Giacca M, Proudfoot NJ. 2008. Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. Mol Cell 29:56–68.

Pfeifer A, Ikawa M, Dayn Y, Verma IM. 2002. Transgenesis by lentiviral vectors: Lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos. Proc Natl Acad Sci USA 99:2140–2145.

Poleshko A, Palagin I, Zhang R, Boimel P, Castagna C, Adams PD, Skalka AM, Katz RA. 2008. Identification of cellular proteins that maintain retroviral epigenetic silencing: Evidence for an antiviral response. J Virol 82:2313–2323.

Ramunas J, Montgomery HJ, Kelly L, Sukonnik T, Ellis J, Jervis EJ. 2007. Real-time fluorescence tracking of dynamic transgene variegation in stem cells. Mol Ther 15:810–817.

Shi Y, Do JT, Desponts C, Hahm HS, Schöler HR, Ding S. 2008. A combined chemical and genetic approach for the generation of induced pluripotent stem cells. Cell Stem Cell 2:525–528.

Srinivasan L, Atchison ML. 2004. YY1 DNA binding and PcG recruitment requires CtBP. Genes Dev 18:2596–2601.

Stadtfeld M, Maherali N, Breault DT, Hochedlinger K. 2008. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell 2:230–240.

Swindle CS, Kim HG, Klug CA. 2004. Mutation of CpGs in the murine stem cell virus retroviral vector long terminal repeat represses silencing in embryonic stem cells. J Biol Chem 279:34–41.

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

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

Teich NM, Weiss RA, Martin GR, Lowy DR. 1977. Virus infection of murine teratocarcinoma stem cell lines. Cell 12:973–982.

Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH. 2006. A protein interaction network for pluripotency of embryonic stem cells. Nature 444:364–368.

Wang GP, Ciuffi A, Leipzig J, Berry CC, Bushman FD. 2007. HIV integration site selection: Analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. Genome Res 17:1186–1194.

Watanabe T, Takeda A, Tsukiyama T, Mise K, Okuno T, Sasaki H, Minami N, Imai H. 2006. Identification and characterization of two novel classes of small RNAs in the mouse germline: Retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev 20:1732–1743.

Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K, Bernstein BE, Jaenisch R. 2007. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature 448:318–324.

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

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

Wiblin AE, Cui W, Clark AJ, Bickmore WA. 2005. Distinctive nuclear organisation of centromeres and regions involved in pluripotency in human embryonic stem cells. J Cell Sci 118:3861–3868.

Wolf D, Goff SP. 2007. TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. Cell 131:46–57.

Yao S, Sukonnik T, Kean T, Bharadwaj RR, Pasceri P, Ellis J. 2004. Retrovirus silencing, variegation, extinction, and memory are controlled by a dynamic interplay of multiple epigenetic modifications. Mol Ther 10:27–36.

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. 2007. Induced pluripotent stem cell lines derived from human somatic cells. Science 318:1917–1920.