DOI 10.1002/bies.200800204

Genomic analysis of induced pluripotent stem (iPS) cells: routes to reprogramming

Ashlin Kanawaty and Jeffrey Henderson *

Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, M5S 3M2, Canada

The phenomenal proliferation of scientific studies into the nature ofi nduced pluripotent stem (iPS) cells following publication of the findings of Takahashi and Yamanaka little more than 2 years ago, have significantly expanded our understanding of cellular mechanisms relating to cell lineage, differentiation, and proliferation. While the full potential ofi PS cell lineages for both scientific tool and therapeutic applications is as yet unclear, findings from several lines ofi nvestigation suggests that multipotential and terminally differentiated cells from an array of cell types are competent to undergo epigenetic reprogramming to a pluripotential state. The nature of this pluripotential state appears to be similar to, but not identical with that previously described for embryonic stem (ES) cells. Understanding the nature of this induced reprogrammed state will be critical to determining the full potential ofi PS cells. Recently, this issue has been examined through an integrated analysis of the genome in fully and partially reprogrammed iPS cell lineages. These results provide a window onto the temporal components of reprogramming and suggest mechanisms by which the efficacy of reprogramming can be enhanced.

Introduction

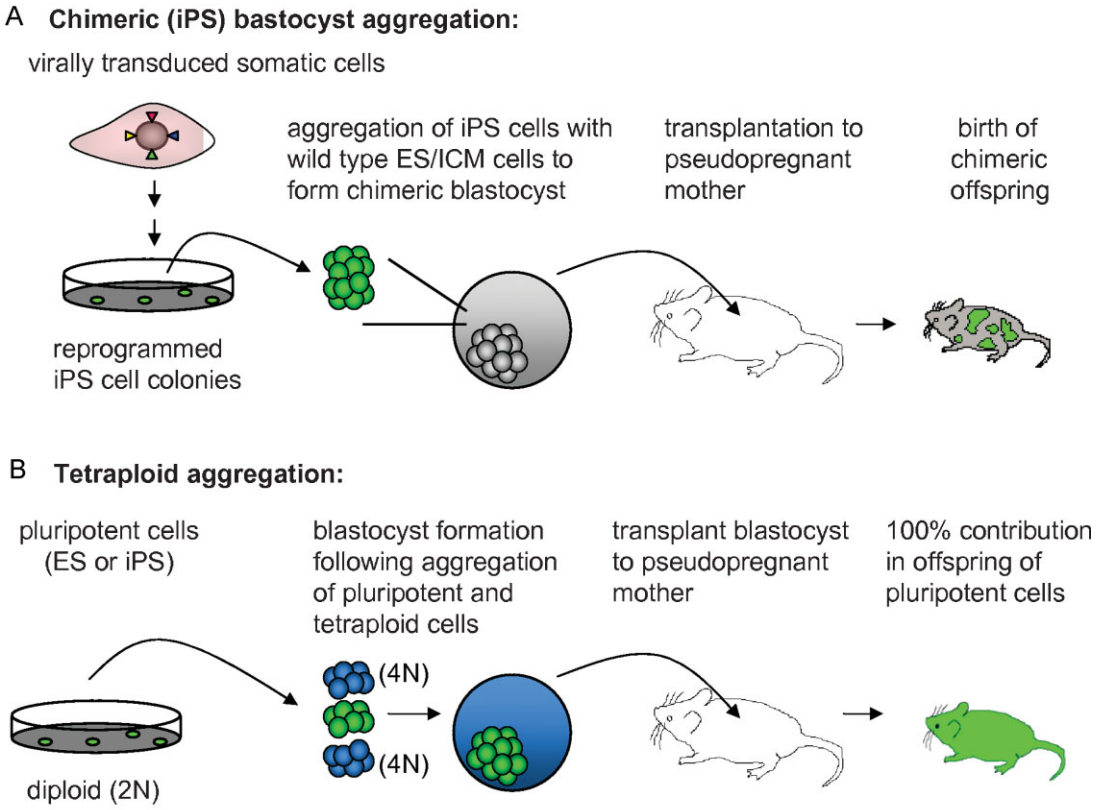
Following the initial description by Takahashi and Yamanaka [1] that the enforced retroviral-mediated expression of four transcription factors (Oct3/4, Sox2, c-Myc, and Klf4) in murine fibroblasts resulted in the appearance of cells exhibiting pluripotential character, a number ofl aboratories over the past 2 years have described similar findings for an array of somatic cell types. While the chimeric blastocysts (Fig. 1A) of Takahashi and Yamanaka constructed using induced pluripotent stem (iPS) cells failed to complete embryonic development in July 2007, subsequent studies performed by Yamanaka and coworkers [2] and a team headed by Rudolph Jaenisch [3] successfully demonstrated live-born chimeras using mixed iPS cell blastocysts. Consistent with this, chimeras were subsequently shown to be capable of germ-line transmission ofi PS-specific characteristics. These findings demonstrate that iPS cells were, at a minimum, capable of contributing to a diverse set of somatic and germ cell lineages, providing an important early test of the potential functional similarities between iPS cells and pluripotent embryonic stem (ES) cells. Because mixed population chimeric blastocysts (Fig. 1 Fig. 1A) are capable of masking subtle functional deficiencies present within a given subpopulation through functional compensation, a more stringent testing of pluripotent character is the construction of tetraploid aggregation chimeras (Fig. 1 Fig. 1B). In such chimeras, blastocysts are constructed such that the embryo proper develops solely from the introduced pluripotential population. The demonstration of viable late-term embryos using this technique from derived iPS cells was a further powerful demonstration that such cell lines possessed sufficient development potential to correctly generate all the somatic lineages necessary to construct the embryo. [3]

While such findings raise the promise ofi PS cells for the investigation of both scientific and therapeutic applications, several critical issues remain. Principle among these are the sequence and nature of epigenetic reprogramming events, the tumorigenic potential of differentiated iPS cell derivatives, and definitive proof that derived iPS cell clones do not arise directly or indirectly as a result of pre-existing multipotential somatic cells. While several recent studies have examined this latter issue (the results consistent with a true ''de-differentiation'' event arising in somatic cells), [4,5] the recent work of Mikkelsen et al.[6] seeks to address this first issue concerning the nature and mechanics of epigenetic reprogramming. As such, it provides significant new information on the sequential nature of these cellular events, adding to data previously provided by both Yamanaka and coworkers [2,7] and Hochedlinger and coworkers. [8]

Abbreviations: iPS cells, induced pluripotent stem cells; ES cells, embryonic stem cells; MEFs, mouse embryonic fibroblasts; GFP, green fluorescent protein; AZA, 5-aza-cytidine.

*Correspondence to: Dr. J. Henderson, Department of Pharmaceutical Sciences, Leslie Dan Faculty of Pharmacy, University of Toronto, 144 College Street, Room 903, Toronto, ON, Canada M5S 3M2.
E-mail: jeff.henderson@utoronto.ca

BioEssays 31:134–138,



C. **Identified transcription factor compositions used for reprogramming:**

Oct-3/4, Sox-2, Klf4, n/c-Myc 1	Oct-3/4, Sox-2, Klf4 2
Oct-3/4, Sox-2, Nanog, Lin28 3	Oct-4, Klf4 4

*5 SV-40 large T

Figure 1. Analyses of developmental potential and transcription factors used in iPS cells. (A) Chimeric blastocysts: Chimeric blastocysts are typically generated following isolation of pluripotent stem cells away from supporting mitomycin-inhibited fibroblast feeders. Potential pluripotent cells are re-plated onto feeder-free plates (1–2 days) and cell clusters 8–15 cells in size are injected into host blastocyst or aggregated with host morula to form the chimeric embryo. These pre-implantation embryos are then placed in to a pseudopregnant recipient. Truly pluripotent cells such as some ES cell lineages will ultimately form live-born chimeric pups. (B) Tetraploid blastocysts: Pluripotent cells of normal diploid (2n) karyotype are derived as above and sandwiched between clusters of cells previously made tetraploid (4n), often through electrofusion at the two-cell stage. These tetraploid cells are subsequently cultured (eight-cell stage) prior to aggregation with pluripotent cells. Because tetraploid cells cannot contribute to the embryo proper (instead forming the extra-embryonic tissues), resulting embryos if developmentally successful will be derived solely from the diploid cells provided. (C) Transcription factor combinations utilized to reprogram somatic cells. Transcription factor which has been successfully reported to reprogram various somatic cells are indicated. 1, Original transcription factor combination utilized by Takahashi and Yamanaka[1] and Blelloch *et al.*;[16] 2, reprogramming induced in the absence of c-Myc, as initially reported by Nakagawa *et al.*;[14] 3, reprogramming by an alternative set of four factors, as initially reported by Yu *et al.*;[17] 4, Oct4, Klf4, and Oct4, c-Myc based reprogramming of neural progenitors as reported by Kim *et al.*[13] In these cells, elevated levels of Sox-2 and c-Myc appear to obviate the need for exogenous addition of these factors; 5, addition of SV40 large T to transcription factor mixes 1 and 3 have been reported to increase rates of iPS cell colony generation by Mali *et al.*[18] Most current methods utilize retro- or lentiviruses to maximize cellular transfection, however adenoviral methods have also been reported by Stadtfeld *et al.*[10]

Temporal analysis of reprogramming in iPS clones

To analyze reprogramming, it is necessary to minimize potential variability in cellular response resulting from positional or numeric differences in lentiviral integration within the genome. The authors derived fibroblasts from murine lines carrying integrated versions of each of the four transcription factors (Oct4, Sox2, Klf4, c-Myc) whose expression was inducible by doxycycline. The attainment of a pluripotent state was made visually detectable by incorporating – in addition – a *Nanog*-green fluorescent protein (GFP) reporter gene: *Nanog* has been shown to be reliably and selectively expressed in pluripotent cells.[2] Using cellular fluorescence from the *Nanog*-GFP reporter as an initial measure of cellular reprogramming, the authors analyzed the number of resulting GFP-positive clones as a consequence of variable periods of doxycycline induction. Similar to previous studies, they noted that, while few cells expressed stem cell-related marker SSEA1 upon less than 8 days of induction, more than 20% of cells expressed this marker following 16 days of doxycycline induction. However, of this population, little more than 1% actually expressed the *Nanog*-GFP reporter, suggesting both the presence of heterogeneity in the reprogrammed population, and that activation of the *Nanog* locus required more extensive reprogramming than that required to express SSEA1. By expressing the four transcription factors for variable periods using their doxycycline-inducible system, the authors were also able to generate a series of partially reprogrammed cell lines, which were then analyzed by chromatin and expression profiling. By analyzing the nature of the genetic changes observed following partial reprogramming and variable periods of induction, the authors were able to correlate changes in DNA methylation pattern, gene expression, and chromatin status with steps along the road toward the ''completely'' reprogrammed state. Analysis of iPS cell lineages that failed to achieve the fully reprogrammed state following prolonged (16 days) induction were also used to gain insight into dominant mechanisms that serve to inhibit cellular reprogramming, thus reducing reprogramming efficiency. With respect to gene modification and expression, the trends observed in the current study for the most fully reprogrammed iPS clones *versus* ES cells are consistent with those previously reported from other groups.[2,3,5] These findings demonstrate an early down-regulation of lineage-specific markers, a trend toward activation of bivalent gene loci (seen by genome-wide enrichment in the trimethylation status of histone H3 at lysine 4, H3K4) and repression of cell type-specific transcription factors and cyclin-dependent kinase inhibitors (as examined by H3K27 methylation status and gene profiling).

A key finding of the present study arose from the analysis of chromatin and DNA methylation data in partially reprogrammed iPS cell clones. Such partially reprogrammed clones were observed to exhibit hypermethylation at one or more pluripotency-related genes. The authors hypothesized that specific (Dnmt1 inhibition) or global [5-aza-cytidine (AZA)-treated] measures to induce hypomethylation would relieve this inhibitory block, releasing ''trapped'' partially reprogrammed iPS cell clones, thus enhancing rates of full cellular reprogramming. As shown by the authors, both measures (AZA $>$ Dmnt1 inhibition) substantially increased the number of *Nanog*-GFP expressing clones at induction times greater than 8 days. The authors suggest that relief of this methylation block is a later event in the process of reprogramming, as demonstrated by the inability of global hypomethylation to enhance *Nanog*-GFP expression in cells that have been induced for periods of less than 8 days. The inability of AZA treatment to enhance *Nanog*-GFP expression from chimeric mouse embryonic fibroblasts (MEFs) induced with doxycycline for periods shorter than 8 days may be related to the expression of higher levels of lineage-specific markers in such cells, since induction of hypomethylation is known to induce apoptosis and/or necrosis in a number of differentiated cell types.

Comparative analysis of gene expression in iPS versus ES cell clones

A key issue for a number of both scientific and therapeutic applications of iPS cells concerns their developmental potential in comparison with ES cells. Analysis of the most fully reprogrammed iPS cell clones obtained by the authors using their doxycycline-inducible/lentiviral system supports earlier findings that iPS cell clones are substantially similar to, but not identical with, ES cells based upon comparative analyses of gene expression, DNA methylation, and chromatin status.

In particular, the gene-profiling data demonstrate that the expression of a small but significant number of genes differ between iPS and ES cells. The consistency with which individual gene expression patterns differ between ES *versus* iPS cell lines will no doubt be important in understanding the true character of genetic reprogramming in iPS cells. In the event that such differences are not consistent among ES or iPS cell lines, or that the character of this variability is similar to that seen between bona fide ES cell lines, such changes may be more closely tied to differences in culture conditions, cell handling, and treatment conditions involved in the genesis of iPS clones.

Careful comparative analysis of gene-profiling data for different iPS cell lines and genesis methods will need to be examined and indexed as a function of the initial somatic population, to tease out the relative importance of each of these effects. As an example, the varied results seen in the

development of chimeric iPS embryos of Takahashi and Yamanaka[1] and Wernig *et al.*[3] did not ultimately define intrinsic limits to iPS cell reprogramming, but likely reflect technical differences in iPS cell genesis, similar to effects seen initially using ES cells.[9]

iPS cells: expansion of rare multipotential progenitors?

An ongoing issue in the mechanics of iPS cell induction has been the low efficiency of "complete" reprogramming seen within primary populations of somatic cells using any current formulation of transcription factors. Certainly a number of straightforward technical reasons explain why full reprogramming might be expected to be relatively rare. For example, primary somatic cells must be successfully transfected and actively transcribe each of the requisite transcription factors, perhaps expressing these factors at specific ratios, and/or over defined periods of time for complete reprogramming to occur. In addition, factors such as cell culture conditions or promoter utilization for each of the transcription factors may not yet be optimal for cellular reprogramming. However, alternative possibilities exist. iPS cells might arise as a result of additional genetic modifications occurring either as a result of induced genetic instability, or as a consequence of positional integration of retro/lentiviral vectors within the genome. Alternatively, iPS colonies might arise from the expansion or partial de-differentiation of pre-existing rare multipotential progenitors.

The study of Mikkelsen *et al.*[6] utilized a genetically homogeneous starting population of primary cells in which the positional integration of each of the doxycycline-inducible vectors is similarly fixed for all cells. Rates of primary iPS cell isolation observed using this system were consistent with that previously described.[1,3] However, the authors also observed that chemically induced hypomethylation, specific inhibition of Dnmt1, or inhibition of transcription factors such as Pax-3 or Pax-7 resulted in a significant increase in the numbers of *Nanog*-GFP-positive clones derived from partially reprogrammed iPS cells.[6] This is interesting given recent reports from several laboratories using similar drug-inducible systems in genetically fixed vector lineages of somatic keratinocytes and fibroblasts.[4,5] In these studies, iPS cells were allowed to differentiate, after which reprogramming was re-induced with determinations of the number of iPS cells that could be derived. These results demonstrated a substantial increase in the number of secondary iPS clones. Taken together, these data argue against the derivation of iPS clones either directly, or as a result of partial reprogramming from rare pre-existing multipotential progenitors.

With respect to the issue of potential additional (unknown integration-induced) repressive events being required for the generation of fully reprogrammed iPS clones, the data obtained by Mikkelsen *et al.*[6] and others[4,5] tend to argue against this, given that the position of any epigenetic insertion is similarly fixed for the entire population. Each of the studies demonstrates similar levels of iPS cell generation, despite logical differences in the precise position of their epigenetic insertions. In addition similar efficiencies of iPS cell generation have recently been demonstrated using non-integrating adenoviral vectors.[10] Taken together, these data suggest that while much remains to be learned regarding the precise mechanism of epigenetic reprogramming in iPS cells, the potential effects of vector integration have likely not played a major role in contributing to the functional consequences of iPS cell reprogramming.

Future challenges

Answers to a number of scientific and technical issues must be resolved before the full potential of iPS stem cells can be realized. In the near term, a key issue for such findings will be to validate the true developmental potential of partially reprogrammed iPS cell clones that have undergone subsequent chemically induced hypomethylation to induce *Nanog*-GFP positive sub-clones. The derivation of live-born chimeras following tetraploid aggregation of iPS clones such as MCV8.1 would be a powerful functional demonstration of the degree of epigenetic reprogramming that occurs following chemical reactivation.

Given the number of independent studies demonstrating isolation of iPS-like cells from somatic tissues of neural ectodermal origin,[1,3,5,11–13] it can now be reasonably assumed that epigenetic reprogramming induced *via* the enforced expression of Oct4, Sox2, Klf4, and c-Myc represents a fairly generalizable phenomenon in mammalian cells. A number of modifications to this initial set of transcription factors have already been described (Fig. 1 Fig. 1C), including the dispensability of c-Myc for iPS induction, and recent two component reprogramming mixtures (Oct4, Klf4 or Oct4, c-Myc) in neural stem cells.[11,13,14] In addition, numerous groups are now working on small molecule and other non-viral methods to alter key transcription factor levels sufficiently to allow iPS cell derivation. In time, such methods may have an important influence upon the potential therapeutic utilization of iPS cells, since non-viral methods will likely be required for any future human utilization of iPS generated cells. However, it is still too early to tell if Thompson's vision of ES cells as a potential "historical anomaly" will come to pass.[15]

For the sake of being concise, we apologize to those we have not acknowledged whose research has contributed to these discoveries.

References

1. **Takahashi, K. and Yamanaka, S.,** Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006. **126**(4), 663–676.
2. **Okita, K., Ichisaka, T. and Yamanaka, S.,** Generation of germline-competent induced pluripotent stem cells. *Nature* 2007. **448**(7151), 313–317.
3. **Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M.,** *et al.* In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 2007. **448**(7151), 318–324.
4. **Hockemeyer, D., Soldner, F., Cook, E. G., Gao, Q., Mitalipova M. and Jaenisch, R.,** A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell* 2008. **3**(3), 346–353.
5. **Maherali, N., Ahfeldt, T., Rigamonti, A., Utikal, J., Cowan C. and Hochedlinger, K.,** A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* 2008. **3**(3), 340–345.
6. **Mikkelsen, T. S., Hanna, J., Zhang, X., Ku, M., Wernig, M.,** *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* 2008. **454**(7200), 49–55.
7. **Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T.,** *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 2007. **131**(5), 861–872.
8. **Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S.,** *et al.* Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* 2007. **1**(1), 55–70.
9. **Abeyta, M. J., Clark, A. T., Rodriguez, R. T., Bodnar, M. S., Pera, R. A. and Firpo, M. T.,** Unique gene expression signatures of independently derived human embryonic stem cell lines. *Hum Mol Genet* 2004. **13**(6), 601–608.
10. **Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. and Hochedlinger, K.,** Induced pluripotent stem cells generated without viral integration. *Science* 2008.
11. **Eminli, S., Utikal, J. S., Arnold, K., Jaenisch, R. and Hochedlinger, K.,** Reprogramming of neural progenitor cells into iPS cells in the absence of exogenous Sox2 expression. *Stem Cells* 2008. **26**(10), 2467–2474.
12. **Henderson, J. T.,** Lazarus's gate: challenges and potential of epigenetic reprogramming of somatic cells. *Clin Pharmacol Ther* 2008. **83**(6), 889–893.
13. **Kim, J. B., Zaehres, H., Wu, G., Gentile, L., Ko, K.,** *et al.* Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* 2008. **454**(7204), 646–650.
14. **Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T.,** *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol* 2008. **26**(1), 101–106.
15. **Cyranoski, D.,** Stem cells: 5 things to know before jumping on the iPS bandwagon. *Nature* 2008. **452**(7186), 406–408.
16. **Blelloch, R., Venere, M., Yen, J. and Ramalho-Santos, M.,** Generation of induced pluripotent stem cells in the absence of drug selection. *Cell Stem Cell* 2007. **1**(3), 245–247.
17. **Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L.,** *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* 2007. **318**(5858), 1917–1920.
18. **Mali, P., Ye, Z., Hommond, H. H., Yu, X., Lin, J.,** *et al.* Improved efficiency and pace of generating induced pluripotent stem cells from human adult and fetal fibroblasts. *Stem Cells* 2008. **26**(8), 1998–2005.