Experimental Cell Research

Title

Induction of pluripotency by defined factors

Keisuke Okita\textsuperscript{1*} & Shinya Yamanaka\textsuperscript{1\textasciitilde4}

\textsuperscript{1} Center for iPS Cell Research and Application (CiRA), Institute for Integrated Cell-Material Sciences, Kyoto University, Kyoto 606-8507, Japan

\textsuperscript{2} Department of Stem Cell Biology, Institute for Frontier Medical Sciences, Kyoto University, Kyoto 606-8507, Japan

\textsuperscript{3} Yamanaka iPS Cell Special Project, Japan Science and Technology Agency, Kawaguchi 332-0012, Japan.

\textsuperscript{4} Gladstone Institute of Cardiovascular Disease, San Francisco, CA 94158, USA

* To whom correspondence should be addressed:
Keisuke Okita, Center for iPS Cell Research and Application (CiRA), Institute for Integrated Cell-Material Sciences, Kyoto University, 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan. Phone: +81-75-366-7012, Fax: +81-75-751-4632, e-mail: okita@cira.kyoto-u.ac.jp
Abstract

Somatic cells can be reprogrammed into pluripotent stem cells by introducing a combination of several transcription factors. The induced pluripotent stem (iPS) cells from a patient’s somatic cells could be useful source of cells for drug discovery and cell transplantation therapies. However, most human iPS cells are made by viral vectors, such as retrovirus and lentivirus, which integrate the reprogramming factors into host genomes and may increase the risk of tumor formation. Studies of the mechanisms underlying the reprogramming and establishment of non-integration methods contribute evidence to resolve the safety concerns associated with iPS cells. On the other hand, patient-specific iPS cells have already been established and used for recapitulating disease pathology.

Key words

Reprogramming, iPS cell, pathology, cell transplantation therapy
Introduction

Cell differentiation is a process of limiting their differentiation potential. The first segregation of cell lineage after fertilization occurs at the morula stage of mouse embryos, where outer cells become extraembryonic tissue, trophectoderm. The inner cells make small cell clump called inner cell mass, which contributes to the embryo proper in subsequent development. ES cells are pluripotent stem cells established from the inner cell mass. ES cells transplanted into the morula can populate the embryo proper but not the trophoblast lineage suggesting a limitation cell fate. Terminally differentiated cells like fibroblasts or lymphocytes were believed to lose the potential of producing other cell types. However, successful cloning experiments in amphibians[1] and mammals[2] showed their states are reversible. Fusion with an enucleated oocyte gives the somatic cells pluripotency to produce an adult animal. Reprogramming has been extensively investigated based on those findings.

Reprogramming of mouse somatic cells with defined factors was reported in 2006[3]. The iPS cells can be generated by the addition of several combinations of transcription factors (Oct3/4, Sox2, Klf4 and c-Myc) and chemical compounds in mouse, rat, pig, monkey, and human. Mouse iPS cells can differentiate into all three germ cell layers and contribute to chimeric mice after they are injected into blastocysts which
indicate their pluripotency. Human iPS cells can differentiate into neurons and cardiomyocytes *in vitro*[4]. An important feature of iPS cells is their unlimited proliferation *in vitro*, while maintaining their pluripotency. These characteristics could allow the iPS cells to supply patient-specific pluripotent stem cells. The iPS cells have raised interest in the fields of the disease pathogenesis, drug discovery, oncology, and cell transplantation therapy.

**iPS induction methods**

The original method of iPS induction used a retrovirus vector for transgene expression. MMLV (Moloney murine leukemia virus)-based vectors with the infection efficiency over $5 \times 10^6$ TU/ml were employed[5]. The vector can robustly infect a variety of cell types and introduce their coding genes into the host genome by reverse transcriptase which thereby enables constant transgene expression during reprogramming. The inactivation of the retroviral promoter is observed in ES cells as well as in iPS cells maybe due to epigenetic modifications, such as histone methylation[6]. Therefore the expression of retroviral transgenes continues until the cells become iPS cells. This guided reprogramming and automatic silencing mechanism is thought to provide effective iPS induction in somatic cells. Most patient-specific iPS cells have been
established with retrovirus vectors. However, the retrovirally derived iPS cells have numerous transgene integrations in the genome, and the integrations may result in leaky expression which could disturb endogenous transcription factor network and lead to failure of differentiation. Another important problem of transgene integration is tumorigenic risk after transplantation. In particular, c-Myc, one of the reprogramming factors, is a well-known oncogene, and its reactivation could give rise to transgene derived-tumor formation in chimeraic mice[7]. There have been several reports of improvements of the transduction method for making safe iPS cells. Removal of the c-Myc transgene from reprogramming cocktail is one important approach. Human and mouse iPS cells can be established from fibroblasts with only Oct3/4, Sox2, and Klf4, although the efficiency is significantly reduced[8]. The chimeraic mice produced with c-Myc-free iPS cells did not show enhanced of tumor formation during the observation period (6 months) in comparison to control mice. However, the overexpression of Oct3/4 and Klf4 can cause tumor formation, and various human tumors express OCT3/4, SOX2 and KLF4. In addition, the retroviral insertion to the genome itself may disturb endogenous gene structure and increase the risk of tumors[9]. Another approach is the reduction of integration sites by putting the reprogramming factors into a single vector with IRES or 2A self-cleavage peptide. This reprogramming cassette was used with a
lentivirus system containing a loxP sequence in the LTR and produced iPS cells with only single insertions[10]. The expression of Cre recombinase successfully cut out the cassette. Although it left an incomplete LTR in the iPS genome, this method minimizes the genomic alteration. A transposon system encoding a reprogramming cassette has also been used for iPS induction[11, 12]. The transduction of a plasmid-based transposon vector can integrate into the host genome with the help of transposase, and induces iPS colony formation. The re-expression of the transposase after the establishment of iPS cells recognizes the terminal repeat of the integrated transposon vector, and excises it from the genome. The excision of the transposon does not leave a footprint in most cases, so it maintains the original endogenous sequences. Several other methods accomplished iPS induction by the transient expression of reprogramming factors. These methods include viral vectors (adenovirus and sendaivirus), DNA vectors (plasmid, episomal plasmid, and minicircle vector), and direct protein delivery. Their efficiencies of iPS cell induction are lower than that with retrovirus vectors, possibly due to low transduction efficiency, and unstable expression. However they could potentially become standard methods in the future.

The mixture of specific reprogramming factors has been evaluated. The standard mixture contains Oct3/4, Sox2, Klf4, and c-Myc, and that mixture can induce
reprogramming in mouse, human, rat, pig, and dog. Yu et al. reported human iPS induction with a slightly different set of reprogramming factors, including Oct3/4, Sox2, Nanog, and LIN28[13]. Inclusion of Oct3/4 and Sox2 in both sets indicates their importance for reprogramming. The reprogramming efficiency are enhanced by the addition of extra factors, such as ESRRB, UTF1, Sall4, Tbx3, miRNAs (miR-291-3p, miR-294 and miR-295), and shRNAs for p53 or p21. The improvement of reprogramming efficiency seemed to be accomplished by a direct or indirect effect. Reprogramming events would include stochastic steps like epigenetic change and the microenvironment where the cell is cultured. Therefore increments of cell number are indirectly associated with high iPS colony formation. Hanna and colleagues found some reprogramming factors, such as Lin28 and shRNA for p53, mainly regulate the reprogramming efficiency through the control of cell proliferation[14]. In contrast, Nanog is seemed to enhance the efficiency of reprogramming through affecting the process itself. Tbx3 would also affect the process because it improves the germ line transmission efficiency of mouse iPS cells[15].

**Molecular mechanisms of reprogramming**

ES cells are maintained with strict regulation of the transcription factors network, which
includes Oct3/4, Sox2, Nanog, and Klf family. Some of these factors have direct protein-protein interactions, and they also bind promoter regions of a lot of common genes and control them. Oct3/4, Sox2, and Nanog induce the expression of stemness genes, such as STAT3 and ZIC3, with RNA polymerase II in human ES cells[16]. On the other hand, these factors are thought to have the opposite function, to repress differentiation related genes like PAX6 and ATBF1 with SUZ12[17]. Therefore forced expression of some members of the network enables both induction and suppression of genes for pluripotency. The inactivation of a differentiation related gene, Thy-1, and the activation of SSEA-1, a stemness marker gene, occur at early time point of reprogramming during iPS induction from mouse embryonic fibroblasts (MEF)[18]. More specific pluripotent markers like endogenous Oct3/4 and Sox2, and activation of telomerase subsequently became detectable. Genome wide analysis of transcription factor binding sites reveals the importance of activation of genes co-occupied with Oct4, Sox2, and Klf4 for full reprogramming[19]. Live cell imaging allows a detailed analysis of the reprogramming steps. Retroviral transduction of reprogramming factors stimulates MEF to divide several times in a morphologically symmetric manner, while maintaining a fibroblastic shape[20]. Most descendant cells fail to reprogram and undergo cell death. Only a small portion of cells are gradually transformed into an
ES-like shape and became iPS cells. Chan et al. observed human iPS cell induction in detail and found that there are three types of human iPS cells based on their expression profiles of cell surface markers and retroviral silencing[21]. The three types differ in the methylation status of the promoter region in Nanog and Oct3/4 loci and their differentiation potential. The most well reprogrammed type is positive for the pluripotency markers, SSEA-4 and TRA-1-60, and negative for the fibroblast marker, CD13, and show inactivation of the retroviral promoter. Only this cell type can make teratomas containing tissues of all three germ layers. Discrimination of the high pluripotency cells from iPS cell induction cultures is necessary because the other two types of partial reprogrammed cells are morphologically similar to the correctly reprogrammed iPS cells. Reprogramming seems to continue even after establishment of iPS cells. Chin [22] et al. reported that the continual cultivation of iPS cells yields a gene expression profile more similar to human ES cells than that of early passages.

Many scientists want to know the mechanisms underlying reprogramming process. First of all, expression of reprogramming factors cannot induce iPS generation in all cells. In fact, the retroviral vector can infect over 90% of MEF. However only a small number of iPS colonies emerged from 1x10^6 cells (around 0.001%). The low efficiency suggests that the origin of iPS cells is some type of tissue stem cells that
represent a small population of the primary culture. However, a lineage tracing analysis revealed that albumin-positive cells in mouse adult liver can become iPS cells[23]. Mouse iPS cells were also established from pancreatic β-cells[24]. Terminally differentiated mature B lymphocytes can be reprogrammed with the addition of C/EBPα[25]. In addition, improvement of gene delivery methods increases the reprogramming efficiency up to 5%, which is higher than the estimated stem cell population in primary cultures. On the other hand, genetic manipulation can make reprogrammable mouse iPS cells which have drug-inducible reprogramming factors in its genome[26]. The secondary system enables to uniform expression of reprogramming factor for iPS induction. Eminli [27] and colleagues produced a chimeric mouse with the iPS cells, obtained from hematopoietic cells at different stages of differentiation, and examined their reprogramming potential by drug treatment. Their data clearly indicate that the reprogramming efficiency varies with the stage of differentiation stage of the original cells. These results suggest that reprogramming would ideally occur not only in stem cells but also in all cell types with divergent efficiencies.

The expression level and balance of reprogramming factor is an important feature of iPS cell induction. Over expression of some reprogramming factors, such as
Nanog, c-Myc and Klf4, can maintain a pluripotent state in mouse ES cells. On the other hand, the forced overexpression of Oct3/4 or Sox2 results in the differentiation of mouse ES cells. A small degree of imbalance can be compensated in the mutual regulation networks because these reprogramming factors constitute a transcriptional circuit and maintain their expression level. Reprogramming seems to be highly dependent on Oct3/4 transgenes. The increment of Oct3/4 transgene in the reprogramming mixture enhances iPS cell induction whereas increments of other three factors, Sox2, Klf4, and c-Myc, reduce the efficiency[28]. However, the excess expression of Oct3/4 leads to the suppression of the reprogramming. Therefore there should be an optimum proportion of transgene expression. The duration of the expression of reprogramming factors affects iPS induction. MEF need around 1 week of continual transgene expression, and human fibroblasts require around 2 weeks. The optimal time period depends on the cell source. For example, human keratinocytes can be reprogrammed within a week, earlier than fibroblasts[29]. Once iPS cells become pluripotent, endogenous expression of Oct3/4, Sox2, and Klf4 maintain their undifferentiated state without the need of exogenous factors.

The gene expression is regulated not only by transcription factors but also by epigenetic modifications, such as DNA methylation and methylation/acetylation of
histones. iPS cells have epigenetic modification similar to those of ES cells in terms of DNA methylation and histone modifications. The promoter regions of Oct3/4 and Nanog in fibroblasts are highly DNA methylated and inactive, while these regions are demethylated and active in iPS cells. There are several studies that the efficiency of iPS cell induction was increased by the treatment with epigenetic modification drugs, such as DNA methyltransferase inhibitor (5’-azacytidine and RG108), histone deacetylase inhibitors (valproic acid, suberoylanilide hydroxamic acid, and trichostatin A), or G9a histone methyltransferase inhibitor (BIX-01294). Therefore, the alteration of epigenetic modifications is also important for iPS cell induction. Bhutani et al. showed AID to be involved in the active DNA de-methylation of somatic cell nuclei in the reprogramming induced by cell fusion and its loss attenuated endogenous Oct3/4 and Nanog reactivation[30]. The AID may also participate in iPS generation. Global DNA methylation pattern of iPS cells resemble that of ES cells, but a recent study revealed their differences. Doi et al. examined approximately 4.6 million CpG sites containing almost all CpG islands in the human genome, and found that 71 differential methylation regions (DMRs) between ES cells and iPS cells[31]. The DMRs showed significant accumulation in genes associated with developmental processes that were hypermethylated in iPS cells in comparison to ES cells, which could leads
differentiation failure of iPS cells. A difference in the gene expression profiles was also reported. These data would show iPS cells have some “memory” of their somatic origin, and are not identical to ES cells. There is still no sufficient evidence to determine whether the memory of human iPS cells is fatal for cell therapy. However, Miura and colleagues reported that mouse iPS cells established from fetal and adult fibroblasts vary in their potential to differentiate into a neuronal lineage[32]. Although both iPS cells can contribute to chimeric mice when transplanted into an early embryo, they show clear differences in their in vitro differentiation procedure. Almost all iPS cells derived from MEF became neuronal cells and only small portion of cells remained in an undifferentiated state like ES cells. On the other hand, iPS cells established from adult tail fibroblasts tend to maintain their undifferentiated state. These undifferentiated cells could form tumors after transplantation into the mouse brain. The results may indicate the memory of iPS cells influence on their safety. Human iPS cells are able to make functional neuronal cells, blood cells, hepatocytes, and retinal cells. However there are some reports that human iPS cells show attenuated differentiation potential into neuronal or hematopoietic lineages in comparison to ES cells[33, 34]. These results indicate limited application of human iPS cells and suggest the need for improvement of reprogramming quality.
Disease pathogenesis and drug discovery

iPS cells can be established from a patient’s own somatic cells, and can be used for in vitro study of numerous medical applications, such as elucidation of disease pathogenesis and discovery of new drugs. Retrovirus derived-iPS cells are sufficient for such applications. There are already many reports of iPS cells establishment from patients of deaminase deficiency-related severe combined immunodeficiency, Shwachman-Bodian-Diamond syndrome, spinal muscular atrophy, Duchenne muscular dystrophy, Parkinson disease, Huntington disease, and dyskeratosis congenita[35]. Lee et al. used lentivirus encoding Oct3/4, Sox2, Klf4, and c-Myc and established iPS cells from patients with familial dysautonomia (FD), which is a fatal autosomal recessive disease affecting the sensory and autonomic nervous system[36]. Most patients have a mutation in the intron of the I-κ-B kinase complex-associated protein (IKBKAP) gene resulting in abnormal mRNA splicing and lack of the exon 20. However, the mechanism underlying the disease development is still elusive because of the absence of a disease model system. Lee et al. performed the directed differentiation of the patient’s iPS cells into peripheral neurons. They found that the iPS cells recapitulated the mutant splicing of IKBKAP mRNA and showed impaired differentiation. They used these cells for the
screening of known candidate drugs. Kinetin is a plant hormone, which reduces the levels of the mis-splicing in FD-derived lymphoblast cell lines. Epigallocatechin gallate and tocotrienol also affect splicing and the absolute levels of IKBKAP. Their screening clearly showed that kinetin but neither epigallocatechin gallate nor tocotrienol reduced the level of abnormal splicing and improved neural differentiation. The study provided proof of principle for application of iPS cells in medical research. However, most diseases do not develop from a simple cause. Diseases are derived from a summary combination of genetic/epigenetic issues, exposure of chemical materials, environment, and aging etc, in a complicated relationship between several cell types in the body. It is therefore necessary to establish a way to recapitulate late onset disease and the environmental effects either in vitro or in an animal model.

**Possible medical application**

The first cell transplantation model of iPS cells used a humanized mouse model of sickle cell anemia[37]. It is a blood disorder which makes abnormal, sickle shaped red blood cells. A mutation in the \( \beta \) -globin gene causes the disease. Homozygous model mice for \( \beta \) -globin genes shows characteristic symptoms including severe anemia due to erythrocyte sickling, splenic infarcts, urine concentration defects and poor health.
Hanna et al. obtained tail fibroblasts from the mouse and established iPS cells by retroviral transduction of reprogramming factors. They used a c-Myc transgene flanked with two loxP sites and removed the c-Myc by Cre recombinase expression in the iPS cells to reduce the oncogenic potential of iPS cells. They also corrected the mutation of \( \beta \)-globin by homologous gene targeting. Hematopoietic progenitors were differentiated from the corrected iPS cells and transplanted into sickle mouse. All three treated mice demonstrated stable engraftment of the cells and the symptoms were relieved. The study showed the promise of cell transplantation therapy using iPS cells. However, some other studies reported some disadvantages of transplantation therapy using stem cells. Possible contamination of pluripotent stem cells or incorrectly differentiated cells lose control after transplantation. A small population of Nanog- or Oct3/4- positive undifferentiated cells are remained after weeks of in vitro neural differentiation from ES cells and iPS cells[32]. Wernig et al. depleted these cells using fluorescence-activated cell sorting of SSEA-1 positive cells to minimized the risk of tumor formation[38]. Better systems for reprogramming, differentiation, and purification of required cells is required to make better therapeutic cells. In addition, new methods are also needed to guarantee the quality and the safety of such cell transplantation. Drug inducible suicide genes can therefore be employed to achieve this aim.
**Conclusion**

iPS cells are artificial pluripotent stem cells. However they can produce chimeric animals in mouse and rat. A tetraploid complementation experiment demonstrated that mouse iPS cells have the ability to autonomously generate full-term mice[39, 40]. These results clarified the existence of pure pluripotent stem cells in iPS cells. On the other hand, hepatocyte-derived mouse iPS cells show a high peri-natal death rate which would indicate abnormal differentiation of iPS cells[23]. Methods for the generation of iPS cells remain at a developmental stage. A reliable method to evaluate the iPS cells must be established. Although many problems still remained to be resolved, iPS cells may be applicable for medical treatment in the future. Studies of disease pathogenesis and drug discovery have already been launched, and the results thereof could provide important relief to countless people throughout the world.


[15] J. Han, P. Yuan, H. Yang, J. Zhang, B.S. Soh, P. Li, S.L. Lim, S. Cao, J. Tay, Y.L. Orlov, T. Luftin, H.H. Ng, W.L. Tam, B. Lim, Tbx3 improves the germ-line competency of induced pluripotent stem cells, Nature.


[34] Q. Feng, S.J. Lu, I. Klimanskaya, I. Gomes, D. Kim, Y. Chung, G.R. Honig, K.S. Kim, R. Lanza, Hemangioblastic Derivatives from Human Induced Pluripotent Stem
Cells Exhibit Limited Expansion and Early Senescence, Stem Cells.


