A peer-reviewed version of this preprint was published in PeerJ on 11 May 2018.

View the peer-reviewed version (peerj.com/articles/4370), which is the preferred citable publication unless you specifically need to cite this preprint.

Omole AE, Fakoya AOJ. 2018. Ten years of progress and promise of induced pluripotent stem cells: historical origins, characteristics, mechanisms, limitations, and potential applications. PeerJ 6:e4370
https://doi.org/10.7717/peerj.4370
Ten years of progress and promise of induced pluripotent stem cells: historical origins, characteristics, mechanisms, limitations, and potential applications

Adekunle Ebenezer Omole Correspond. 1, Adegbenro Omotuyi John Fakoya Correspond. 2

1 Anatomical Sciences, Atlantic University School of Medicine, St. John's, Antigua
2 Anatomical Sciences, All Saints University, School of Medicine, Roseau, Dominica

Corresponding Authors: Adekunle Ebenezer Omole, Adegbenro Omotuyi John Fakoya
Email address: kunlesty@yahoo.com, gbenrofakoya@gmail.com

The discovery of induced pluripotent stem cells (iPSCs) by Shinya Yamanaka in 2006 was heralded as a major breakthrough of the decade in stem cell research. The ability to reprogrammed human somatic cells to a pluripotent embryonic stem cell-like state through the ectopic expression of a combination of embryonic transcription factors was greeted with great excitement by scientists and bioethicists. The reprogramming technology offers the opportunity to generate patient-specific stem cells for modeling human diseases, drug development and screening, and individualized regenerative cell therapy. However, fundamental questions have been raised regarding the molecular mechanism of iPSCs generation, a process still poorly understood by scientists. The efficiency of reprogramming of iPSCs remains low due to the effect of various barriers of reprogramming. There is also the risk of chromosomal instability and oncogenic transformation associated with the use of viral vectors, such as retrovirus and lentivirus, which deliver the reprogramming transcription factors by integration in the host cell genome. These challenges can hinder the therapeutic prospects and promise of iPSCs and their clinical applications. Consequently, extensive studies have been done to elucidate the molecular mechanism of reprogramming and novel strategies have been identified which help to improve the efficiency of reprogramming methods and overcome the safety concerns linked with iPSCs generation. Distinct barriers and enhancers of reprogramming have been elucidated and non-integrating reprogramming methods have been reported. Here, we summarize the progress and the recent advances that have been made over the last 10 years in the iPSCs field, with emphasis on the molecular mechanism of reprogramming, strategies to improve the efficiency of reprogramming, characteristics and limitations of iPSCs, and the progress made in the applications of iPSCs in the field of disease modelling, drug discovery and regenerative medicine. Additionally, this study appraised the role of genomic editing technology in the generation of healthy iPSCs.
Ten years of progress and promise of induced pluripotent stem cells: historical origins, characteristics, mechanisms, limitations, and potential applications.

Adekunle Ebenezer Omole¹ and Adegbenro Omotuyi John Fakoya²

¹Department of Anatomical Sciences, Atlantic University School of Medicine, St. John’s, Antigua.
²Department of Anatomical Sciences, All Saints University School of Medicine, Roseau, Dominica.

Abstract

The discovery of induced pluripotent stem cells (iPSCs) by Shinya Yamanaka in 2006 was heralded as a major breakthrough of the decade in stem cell research. The ability to reprogrammed human somatic cells to a pluripotent embryonic stem cell-like state through the ectopic expression of a combination of embryonic transcription factors was greeted with great excitement by scientists and bioethicists. The reprogramming technology offers the opportunity to generate patient-specific stem cells for modelling human diseases, drug development and screening, and individualized regenerative cell therapy. However, fundamental questions have been raised regarding the molecular mechanism of iPSCs generation, a process still poorly understood by scientists. The efficiency of reprogramming of iPSCs remains low due to the effect of various barriers of reprogramming. There is also the risk of chromosomal instability and oncogenic transformation associated with the use of viral vectors, such as retrovirus and lentivirus, which deliver the reprogramming transcription factors by integration in the host cell genome. These challenges can hinder the therapeutic prospects and
promise of iPSCs and their clinical applications. Consequently, extensive studies have been done to elucidate the molecular mechanism of reprogramming and novel strategies have been identified which help to improve the efficiency of reprogramming methods and overcome the safety concerns linked with iPSCs generation. Distinct barriers and enhancers of reprogramming have been elucidated and non-integrating reprogramming methods have been reported. Here, we summarize the progress and the recent advances that have been made over the last 10 years in the iPSCs field, with emphasis on the molecular mechanism of reprogramming, strategies to improve the efficiency of reprogramming, characteristics and limitations of iPSCs, and the progress made in the applications of iPSCs in the field of disease modelling, drug discovery and regenerative medicine. Additionally, this study appraised the role of genomic editing technology in the generation of healthy iPSCs.

Keywords: Induced pluripotent stem cells, reprogramming, reprogramming factors, embryonic stem cells, gene editing technology.

1. Introduction

The birth of iPSCs in 2006 by Shinya Yamanaka was a remarkable breakthrough that was made possible on the basis of many research findings by past and current scientists in related fields. In 1962, Sir John Gurdon achieved the first example of cellular reprogramming by reporting the generation of tadpoles from enucleated unfertilized frog egg cells that had been transplanted with the nucleus from intestinal epithelial somatic cells of tadpoles. This remarkable method of reprogramming somatic cells to the pluripotent embryonic state with the same genetic makeup was termed somatic cell nuclear transfer (SCNT). This discovery led to the birth of cloning. Thirty-five years later, Sir Ian Wilmut and his team used the same SCNT strategy of cellular reprogramming in the cloning of Dolly the sheep, the first mammalian to be generated by somatic cloning.

These two scientific breakthroughs in somatic cloning proved that the nuclei of differentiated somatic cells contain all the necessary genetic information to generate a whole organism, and that the egg cell contains the necessary factors to bring about such reprogramming. In 2001, Tada et al. further lent credence to the somatic cloning hypothesis through another novel strategy of reprogramming termed cell fusion. The cell fusion of somatic cells with ESCs to generate cells capable of expressing pluripotency-related genes showed that ESCs do contain some factors that can reprogram somatic cells. There are two other important landmarks- the generation of mouse ESCs cell lines in 1981 by Sir Martin Evans, Matthew Kaufman and Gail R. Martin and
the subsequent generation of human ESCs in 1998 by James Thomson\(^4,5,6\). The ESCs are developed from
pre-implantation embryo and are capable of generating any cell type in the body, an inherent characteristics
termed as pluripotency. Their discoveries shed light on the appropriate culture conditions and transcription
factors that will be necessary for the maintenance of pluripotency. The merging of all these essential historical
landmarks led to the discovery of iPSCs (Figure 1).

But why the need for iPSCs since they are pluripotent just like ESCs? Firstly, the use of ESCs was fraught
with strong ethical concerns related to embryo destruction and this has hindered its clinical application.
Secondly, there were the safety concerns related to immune rejection of the ESCs. Finally, due to its source
from the embryo, ESCs will be limited in supply and this will limit a broader therapeutic application. Hence,
there was urgent need for another substitute for ESCs that bypass these important drawbacks. Indeed, the
iPSCs serves as an alternative source of pluripotent stem cells with the same differentiation potential as
embryonic stem cells (ESCs) while avoiding the ethical issue associated with the latter.

Shinya Yamanaka and Kazutoshi Takahashi developed the mouse iPSCs in 2006 through a different
method of reprogramming: the use of a retrovirus to deliver into a somatic cell (mouse fibroblast), a
combination of 4 reprogramming transcription factors, including Oct 3/4 (Octamer binding transcription
factor-3/4), Sox2 (Sex determining region Y)-box 2, Klf4 (Kruppel Like Factor-4), and c-Myc nicknamed
the “OSKM factors”\(^7\). A year later in 2007, Yamanaka and his team applied the same reprogramming method
for adult human fibroblast to generate human iPSCs (hiPSCs) and James Thomson’s group reported the
generation of the same hiPSCs though using a different delivery system, the lentivirus and a different set of 4
factors: Oct 3/4, Sox2, Nanog and Lin 28\(^8,9\). For their remarkable revolutionary discovery, Shinya Yamanaka
and John B. Gurdon were awarded the 2012 Nobel prize for Physiology or Medicine\(^10\). Like ESCs, the iPSCs
have a self-renewal capability in culture and can differentiate into cell types from all three germ cell layers
(ectoderm, mesoderm and endoderm). The iPSC technology holds great promise for personalized cell-based
therapy, human disease modelling and drug development and screening. However, this technology is by no
means free of its own challenges. The reprogramming efficiency is low and tedious and there is associated
risk of chromosomal instability and tumorigenesis from insertional mutagenesis due to the viral vectors
delivery method\(^7,8,9\). These drawbacks will have a significant impact on the clinical application of iPSCs.
Much progress has since been made to improve the efficiency of reprogramming and to reduce the risk associated with the technology. Novel strategies already employed to improve reprogramming includes the inhibition of barriers of reprogramming, use of non-integrative delivery methods, overexpression of enhancing genes and the use of certain small molecules which enhanced reprogramming. Factors that influences the reprogramming process have been studied, namely, the choice of the somatic cell source, reprogramming transcription factors, delivery methods and culture conditions. Extensive research on the molecular mechanisms of reprogramming have improved the efficiency of reprogramming.

In this review, we provide an overview of the progress made in iPSC technology in the last decade. First, we briefly define iPSCs by providing a summary of Yamanaka’s key findings and the characterization of iPSCs, and then provide a summary of the current knowledge on the molecular mechanism of reprogramming, it’s limitations and the various strategies employed to address the drawbacks of this technology. We will then briefly discuss the potential application of iPSCs in the field of disease modelling, drug development and regenerative medicine.

2. Methods
The data for this review were obtained from Medline on OvidSP, which includes PubMed, Embase by the US National Library of Medicine as well as a search through the University of Bristol Library services.

2.1. Search strategy
A thorough search was carried out by signing into Ovid, Wolters, and Kluwer portal and “All Resources” was selected. Three separate keywords were used for the search. The first search with the keyword “induced pluripotent stem cells” yielded a total number of 5,975 publications. The second search with the keyword “cellular reprogramming” gave a total number of 3,002 publications. The third search with the keyword “transcription factors” gave a total number of 299,870 publications.

A combination of the search for “induced pluripotent stem cells” using the Boolean operator “AND” with “cellular reprogramming” and “transcription factor” yielded a total number of 200 publications. We now hand screened these 200 publications to see which one fit into the inclusion criteria for the study and we arrived at a total of 114 publications.

Furthermore, other data were included in this review and these were obtained from the University of Bristol Library services using the search phrase “induced pluripotent stem cells”, “cellular reprogramming” and “transcription factors”. The publications generated were hand screened to fit the inclusion criteria and 61
publications were selected. Also included were relevant references from previously selected publications as well as many other recommended publications. A total of 228 articles were reviewed.

2.2. Inclusion criteria

The publications selected were thoroughly analyzed to see if they focused on the study which was on the molecular mechanism of cellular reprogramming of somatic cells into induced pluripotent stem cells using transcription factors and other small molecules. We included studies that focused on the barriers and enhancers of cellular reprogramming and those that emphasized the various novel strategies for enhancing the kinetics and efficiency of the process. Also considered are articles on the limitations and potential of induced pluripotent stem cells and the progress made to address such limitations. Publications that included the role of genomic editing technology in the generation of iPSCs were also considered.

3. Generation of iPSCs: A brief overview

Briefly, iPSCs can be defined as ‘embryonic stem cell-like’ cells derived from the reprogramming of adult somatic cells by the introduction of specific pluripotent-associated genes. Prior to the discovery of iPSCs, ESCs which are derived from the inner cell mass (ICM) of a blastocyst of pre-implantation stage embryo, are the most well known pluripotent stem cells. Just like ESCs, iPSCs have the ability to proliferate extensively in culture and can give rise to the three germ cell layers, namely, endoderm, mesoderm and ectoderm.

Takahashi and Yamanaka set out to identify the genes that help in the maintenance of pluripotency in mouse ES cells. The search started in year 2000 leading to a list of 24 candidate reprogramming factors chosen for their links to ES-cell pluripotency. A screening method was developed to test a pool of 24 pluripotency-associated candidate factors for the ability to induced pluripotency. These genes were transduced into mouse embryonic fibroblast (MEFs) using a retroviral delivery system. The mouse fibroblast was generated by the fusion of the mouse F-box only protein 15 (Fbxo15) gene locus with a β-galactosidase (β-geo) cassette. The expression of β-geo is to be used as a reporter of Fbxo15 expression and activity, as cells expressing β-geo would be resistant to the selection marker geneticin (G418). The ESC-specific Fbxo-15 locus is not expressed in somatic cells and this cells are not resistant to G418 treatment. The Fbxo15-β-geo MEFs was used to screen the pool of 24 transcription factors by transducing different combinations of the candidate genes and assessing the capability of the MEFs to survive in G418 treatment (Figure 2). Consecutive rounds of elimination of each individual factors then led to the identification of a minimal core set of four genes, comprising Oct3/4, Sox2, Klf4 and c-Myc (OSKM cocktail/factors)⁷. These factors were already shown to be important in early
embryonic development and vital for ES cell identity\textsuperscript{11-14}. The reprogrammed cell colonies, which were named as iPSCs, demonstrated ES cell-like morphology, express major ES cell marker genes like SSEA-1 and Nanog and form teratomas upon injection into immunocompromised mice\textsuperscript{7} (Table 1).

Takahashi and Yamanaka demonstrated that ectopic expression of defined transcription factors is able to reprogram mouse fibroblast back to a pluripotent state thus circumventing the ethical concerns surrounding the use of ESCs. However, these “first generation” iPSCs demonstrated a lower levels of key ES pluripotency genes and failed to generate adult chimeras or contribute to the germline\textsuperscript{7}. These latter characteristics suggest that the iPSCs were only partially reprogrammed. In 2007, Yamanaka and other laboratories modified the induction protocols to generate fully reprogrammed iPSCs that are competent for adult chimera and germline transmission\textsuperscript{15-17}. The technology has also been successfully translated to human fibroblast\textsuperscript{8-9,18} and then to other somatic cell types, such as pancreatic β cells\textsuperscript{19}, neural stem cells\textsuperscript{20,21}, stomach and liver cells\textsuperscript{22}, mature B lymphocytes\textsuperscript{23}, melanocytes\textsuperscript{24}, adipose stem cells\textsuperscript{25} and keratinocytes\textsuperscript{26}, demonstrating the universality of cellular reprogramming. The advantages of iPSCs technology is its reproducibility and simplicity, thus encouraging many laboratories to modify and improve upon the reprogramming technique. Consequently, remarkable progress has been made in the last decade in the field of iPSCs technology.

4. Technical advances and progress in iPSC generation.

If iPSCs are to fulfil their promise (that they are viable and possibly superior substitutes for ESCs in disease modelling, drug discovery and regenerative medicine), limitations and obstacles on the road to their clinical application need to be cleared. The initial reports of iPSCs generation were inefficient (\~0.001-1\%)\textsuperscript{7,8,16,27,28}, that is, on average only 1 out of 10,000 somatic cells formed iPSCs. The overexpression of oncogenes such
as c-Myc and Klf4 during the generation of iPSCs raises safety concerns. Indeed, in the original report of
germline-competent iPSCs, ~20% of the offspring developed tumor attributable to the reactivation of c-Myc transgene\(^\text{16}\). Furthermore, there is the risk of insertional mutagenesis due to virus based delivery methods\(^\text{7,8,9}\).

Much progress have been made in the past decade to address these limitations and to improve the reprogramming technique. New methods for induced reprogramming have been developed. The following sections presents an overview of the advancement made to improve the reprogramming technique, with emphasis on the reprogramming factors and the delivery systems for iPSCs generation.

4.1. Reprogramming factors

The conventional OSKM cocktail by Yamanaka’s group has been used extensively by researchers on a wide range of human somatic cells and delivery systems\(^\text{29}\). Thomson’s group provided an alternative combination of four factors: Oct 3/4, Sox2, Nanog and Lin 28 (OSNL)\(^\text{9}\). The generation of iPSCs by Yamanaka and Thomson’s group using different cocktails of transcription factors may suggest that different transcription factors activate the same reprogramming pathway by reinforcing each other’s synthesis. The OSKM and OSNL reprogramming cocktails have proved efficient on a wide range of delivery systems, albeit at a variably low efficiency rate\(^\text{29,30}\). Consequently, researchers have sought to discover new molecules that will enhance the reprogramming technique and improve its efficiency (Table 2). We will refer to these molecules as reprogramming ‘enhancers’. Some other molecules discovered are ‘barriers’ of reprogramming technique. So the strategy employed to increase the efficiency of reprogramming includes the inhibition of such barriers and the overexpression and administration of the enhancers.

Pluripotency-associated transcription factors. Many of the transcription factors used for reprogramming somatic cells are part of a core pluripotency circuitry. These factors are pluripotency-associated genes expressed early during embryonic development and are involved in the maintenance of pluripotency and self-renewal. The expression of other pluripotency-associated genes along with the minimal pluripotency factors (OSKM) can enhance the reprogramming efficiency or even replace some of the reprogramming factors. For example, the expression of UTF1 or SALL4 with OSKM/OSK, improved the reprogramming efficiency\(^\text{31,32}\). Non-coding RNA’s like LincRoR and Let7 are involved in the regulation of expression of core transcriptional factors. LincRoR is a reprogramming enhancer while Let7 act as a barrier by blocking the activation of
pluripotency factors c-Myc, Lin 28 and SALL4. Thus, Let7 inhibition and the expression of LincRoR both enhance reprogramming efficiency. Nanog and Lin 28 can replace Klf4 and c-Myc respectively and ESRRβ can replace Klf4. A recent single-cell gene expression study for partially reprogrammed cells showed that SALL4, ESRRβ, Nanog and Lin 28 (rather than OSKM) was enough for reprogramming fibroblast into iPSCs, albeit with low efficiency. These observations proved, that most of these enhancer genes are possibly part of the reprogramming circuitry network activated by OSKM. Consequently, a detailed analysis of the downstream targets of OSKM may help us to understand the molecular mechanism of reprogramming, thus opening the way on how to increase its efficiency.

Cell cycle-regulating genes. As they move towards pluripotency, somatic cells also gain the ability to proliferate indefinitely. Not surprisingly, two of the minimal pluripotency factors, c-Myc and Klf4, are oncogenes that enhance cellular proliferation. Apparently, there will be other regulators in this cell cycle pathway. The p53 tumor suppressor protein promotes senescence and inhibit growth, thus has an inhibitory effect on iPSCs generation. Thus, many studies have shown that p53 inhibition can greatly enhance reprogramming efficiency. Cell cycle-dependent kinase inhibitors like INK4A and ARF (which are linked to the p53-p21 pathway) can block iPSCs reprogramming. Conversely, overexpression of p53 inhibitor proteins (such as SV40 large T antigen, REM2, and MDM2), increased the efficiency of reprogramming (up to 23-fold increase compared to OSKM alone). So researchers have used the strategy of overexpressing reprogramming enhancers to eliminate the barriers on the road toward pluripotency.

Epigenetic modifiers. The reprogramming of somatic cells into iPSCs is characterized by epigenetic changes, from DNA methylation to histone modifications. Chromatin remodeling is a rate-limiting step in the reprogramming process, and thus researchers have studied chemical compounds that modified the epigenetic process. For example, DNA methyltransferase inhibitor 5-azacytidine and histone deacetylase (HDAC) inhibitors (like suberoylanilide hydroxamic acid (SAHA), trichostatin A (TSA) and valproic acid (VPA)) enhanced reprogramming efficiency in MEFs. VPA also promote somatic cell reprogramming with Oct4 and Sox2 alone. The combination of CHIR99021 (a GSK3 inhibitor) with Parnate (a lysine-specific demethylase 1 inhibitor) causes the reprogramming of human keratinocyte with only Oct4 and Klf4. Similarly, G9a histone methyltransferase promote epigenetic repression of Oct4 during embryonic development, which is why a G9a inhibitor (BIX-01294) enhances MEFs reprogramming with only Oct4 and Klf4. DOT1L, MBD3, RCOR2, Sirt6, and miR766 (a Sirt6 inhibitor) are all involved in
chromatin remodeling, thus affecting the efficiency of reprogramming when inhibited or overexpressed. Vitamin C improves cellular reprogramming efficiency, in part by promoting the activity of histone demethylases JHDM1A (KDM2A) and JHDM1B (KDM2B)\textsuperscript{56}, alleviating cell senescence\textsuperscript{57} and inducing DNA demethylation\textsuperscript{58}. In conclusion, microRNA (miRNA) have been used to increase reprogramming efficiency. The miRNA’s mostly work by inhibiting the TGF\(\beta\) signalling pathway, thereby inhibiting the epithelial to mesenchymal transition (EMT). The combination of miR-291-3p, miR-294 and miR-295 with OSK cocktail promotes iPSC generation\textsuperscript{59}. More recently, miR302, miR367, miR369, miR372 and miR200c have been used either alone or in combinations to enhanced the reprogramming process in humans by replacing the OSKM traditional nuclear factors\textsuperscript{60-64}. The miRNA can specifically target multiple pathways thus reducing the need and the amount of transcription factors for reprogramming\textsuperscript{64}. In the nearest future, miRNA-based reprogramming may provide an effective way of cellular reprogramming than traditional nuclear factor (OSKM) method.

Table 2. Reprogramming factors capable of reprogramming human cells. Adapted from (82).

4.2. Delivery methods

A number of different delivery methods have been used to introduce reprogramming factors into somatic cells (Figure 3). The reprogramming methods can be grouped into 2 categories -**Integrative systems** (involving the integration of exogenous genetic material into the host genome) and **Non-Integrative systems** (involving no integration of genetic material into the host genome). The integrative delivery methods include the use of viral vectors (retrovirus, lentivirus and inducible lentivirus) and non-viral vectors (linear/plasmid DNA fragment and transposons). Similarly, the non-integrative delivery methods include the use of viral vectors (adenovirus and Sendai virus) and non-viral vectors (episomal DNA vectors, mRNA and proteins). This session is focused on the reprogramming methods currently available.

4.2.1. Integrative delivery systems

4.2.1.1. Viral integrative vectors. **Retroviruses** was used for the delivery of transcription factors in the original studies on iPSCs generation\textsuperscript{7,15-17}. Retroviruses are efficient and relatively easy form of delivery
They can integrate randomly into the host genome leading to an increased risk of insertional mutagenesis. They require an actively dividing somatic cell to integrate well in the genome. iPSC is considered to be fully reprogrammed only after the upregulation of endogenous pluripotency genes and the downregulation or silencing of the integrated transgenes expression. Though retroviral vectors are usually silenced in ESCs\textsuperscript{65,66} and iPSCs\textsuperscript{18,67}, the silencing is not always efficient and the silenced transgenes may be reactivated later on. Certainly, in the original report of germline-competent iPS\textsuperscript{16}Cs, \textasciitilde20\% of the offspring developed tumor attributable to the reactivation of c-Myc transgene\textsuperscript{16}.

\textbf{Lentivirus} have also been successfully used for the introduction of transgenes during cellular reprogramming\textsuperscript{9,68}. Compared with retroviral vectors, lentivirus integrate into the host genome with risk of insertional mutagenesis, and inefficient silencing and transgene reactivation are possible. Unlike retroviruses, they can integrate into both dividing and non-dividing cells, thus iPSCs can be generated from most somatic cell type. The original studies on iPSCs generation by Yamanaka involve the use of different types of retroviruses, each delivering only one type of transcription factor\textsuperscript{8}. This can create many uncontrollable integration sites with increased risks of transgene reactivation, inefficient transgene silencing and diminished efficiency of reprogramming. The creation of \textit{polycistronic} viral vectors (for retrovirus\textsuperscript{69} and lentivirus\textsuperscript{70,71}) allow for the expression of all reprogramming factors driven by a single promoter, with the genes separated by self-deleting peptide sequences. This method remarkably reduces the amount of genomic insertions thus improving the safety and efficiency of the reprogramming process. Moreover, the introduction of both the excisable vector (\textit{Cre/loxP} system)\textsuperscript{72,73} and inducible (\textit{tetracycline/doxycycline inducible system})\textsuperscript{26,74-76} vector systems has allowed for a better control of transgene expression thus reducing the effects of inefficient silencing and transgene reactivation.

\subsection*{4.2.1.2. Non-viral integrative vectors.} An alternative to viral vectors is the transfection of DNA (plasmid/linear) into cells using liposomes or electroporation. Using this method, the transduction efficiency is much lower with only a few cells capturing the full set of reprogramming factors. However, the use of polycistronic vectors to express all cDNAs from a single promoter has helped to improve the reprogramming efficiency. Kaji \textit{et al.} successfully generate iPSCs from mouse fibroblasts with a non-viral polycistronic vector combined with an excisable \textit{Cre/loxP} system for deleting the reprogramming construct\textsuperscript{77}.

\textbf{Transposon}. Kaji \textit{et al.} and Woltjen \textit{et al.} applied the non-viral single vector system for the generation of human iPSCs using a \textit{piggybac} (PB) transposon-based delivery system\textsuperscript{77,78}. The PB is a mobile genetic element which includes an enzyme PB transposase (that mediate gene transfer by insertion and excision), and
a donor plasmid (transposon) co-transfected with a helper plasmid (expressing the transposase enzyme) that mediates the integration\textsuperscript{29}. Once the reprogramming is achieved, the enzyme can precisely delete the transgenes without any genetic scars thus avoiding the risk of insertional mutagenesis. Drawbacks to the use of PB systems include the risks of integrating back into the genome, and the knowledge that human genome contain endogenous PB transposon elements which may be acted upon by the transposase enzyme meant for the transgene excision\textsuperscript{70-82}. The recent introduction of another transposon, the Sleeping Beauty (SB), has help to overcome these limitations of the PB transposon\textsuperscript{83,84}. SB integrates less compare to the PB and there are no SB-like elements in the human genome. However, the reprogramming efficiency of transposons are low compared to viral vectors and their use involves multiple rounds of excision, thus increasing the risk of re-integration.

Overall, integrative delivery system comes with a risk of integration into the genome leading to insertional mutagenesis. This lack of safety may limit their therapeutic application. Non-integrative delivery system will later address this major limitation.

\subsection{4.2.2. Non-Integrative delivery systems}

\subsubsection{4.2.2.1. Non-integrative viral vectors.} Stadtfeld \textit{et al.} reported the generation of the first integration-free iPSCs from adult mouse hepatocytes using nonintegrating adenovirus\textsuperscript{85}. Transgene-free iPSCs will later be generated from human fibroblasts by Zhou \textit{et al.} using similar adenoviral vectors\textsuperscript{86}. However, the reprogramming process requires multiple viral infection, and the production of adenovirus is very labour-intensive. Most importantly, the reprogramming efficiency using adenoviruses are several orders of magnitude lower compare to lenti- or retroviruses.

Another non-integrating viral vector that has been successfully used for iPSCs generation is the Sendai virus (SeV)\textsuperscript{87-93}. They are very efficient in transferring genes (in the form of negative –strand single stranded RNA) into a wide range of somatic cells\textsuperscript{94-97}. Although they are very effective, the viral vector’s RNA replicase is very sensitive to the transgene sequence content. Additionally, the vectors may be difficult to eliminate from the somatic cells because they replicate constitutively\textsuperscript{87}. A new improved Sendai virus (SeV dp) has since been developed\textsuperscript{98-99}.

\subsubsection{4.2.2.2. Non-integrative non-viral delivery.} Episomal vectors provides an alternative to the integrative-defective viruses. Episomes are extrachromosomal DNA capable of replicating within a cell independently of the chromosomal DNA. The reprogramming factors can be directly and transiently transfected into the
somatic cells through the episomal vectors as **plasmids**\textsuperscript{100-107} or as **minicircle** DNA\textsuperscript{108,109}. Unlike retro- and lentiviruses, this technique is relatively simple and easy to use and does not integrate into the host genome. However, since their expression is only transient, they require multiple transfections. In general, their reprogramming efficiency is low although when compared to plasmids, the minicircle has a higher transfection efficiency (probably due to it’s smaller size) and a longer ectopic expression of the transgenes (due to a lowered silencing mechanisms)\textsuperscript{110,111}.

**RNA delivery.** iPSCs have been generated by direct delivery of synthetic mRNA into somatic cells\textsuperscript{112,113}. This method has the highest reprogramming efficiency when compared with other non-integrative delivery systems. RNA have short half lives, thus repeated transfection is required to sustain the reprogramming process. RNA-based methods are also highly immunogenic.

**Protein delivery.** Reprogramming factors can be directly delivered as recombinant proteins into somatic cells for iPSCs generation\textsuperscript{114,115}. The reprogramming efficiency is low and repeated transfection is also required to maintain the intracellular protein level for reprogramming. Overall, integrative delivery methods have a higher reprogramming efficiency than non-integrating methods, but they are less safe due to the risk of insertional mutagenesis. Therefore, the use of non-integrating methods will appeal more for iPSCs generation and usage in the clinical settings.

Figure 3. Schematic representation of various delivery methods of iPSC induction.

5. **Molecular mechanism of induced pluripotency**

The reprogramming of somatic cells into iPSCs is a long and complex process involving the activation of ES-cell-specific transcription network, combinatorial overexpression of multiple transcription factors and epigenetic modifications. Understanding the molecular mechanisms of cellular reprogramming is critical for the generation of safe and quality iPSCs for therapeutic application. This section reviews the molecular mechanism leading to induced pluripotency.

5.1. **The Fantastic Four (OSKM)**

Takahashi and Yamanaka showed that four exogenous reprogramming factors, Oct4, Sox2, Klf4 and c-Myc, all have key roles in iPSCs generation\textsuperscript{7}. They discover that Oct3/4 and Sox2 are essential for iPSCs generation,
c-Myc and Klf4 were essential factors and Nanog was dispensable. Though exogenous Nanog (not part of
the “fantastic four”) is not an essential factor and is not required to initiate the reprogramming process, it’s
possible that exogenous Oct 4, Sox2 and other reprogramming factors induce expression of endogenous
Nanog to levels adequate enough to achieve full reprogramming. Genetic studies have shown that Oct4, Sox2 and Nanog (OSN) are key regulators of embryonic development and they are critical for pluripotency maintenance. These factors are expressed both in pluripotent ESCs and in the ICM of blastocysts. Oct 3/4, Sox2 and Nanog knockout embryos die at the blastocyst stage and when cultured in vitro, their ESCs lose its pluripotency and differentiate. Klf4 play key roles in cellular processes, like development, proliferation, differentiation and apoptosis. It is expressed in ESCs and can interact with Oct4-Sox2 complex to activate certain ESCs genes. Klf4 can revert epiblast derive stem cells to ESC state. Its interaction with Oct4-Sox2 complex and its tumor suppressor activity is thought to be important in iPSCs generation. c-Myc is a potent oncogene associated with apoptosis, cell proliferation and cell cycle regulation. Though iPSCs can be generated without Klf4 and c-Myc, the marked reduction in the efficiency of the process greatly emphasize their importance in cellular reprogramming.

5.2. Autoregulatory loop driving pluripotency

Experimental studies using chromatin immunoprecipitation and genome-wide location analysis in human and murine ESCs to identify genes occupied by Oct4, Sox2 and Nanog has provided much understanding on how these transcription factors contribute to pluripotency. The studies reveal that Oct4, Sox2, and Nanog bind together to activate the promoters of both their own genes and the genes of each other, hence forming an autoregulatory loop (Figure 4). The three factors function cooperatively to maintain their own expression, thus enhancing the stability of pluripotency gene expression. Since the initial hypothesis, several other studies have provided strong verifiable evidence for the existence of the autoregulatory circuitry.

5.3. Transcriptional regulatory network

The experimental studies also demonstrated that Oct4, Sox2 and Nanog target several hundreds of other ESCs genes, collectively co-occupying these genes cooperatively to maintain a transcriptional regulatory network required for pluripotency. This may explain why efficient iPSC generation seems to require the combinatorial overexpression of multiple transcription factors. The cascades of genes targeted were found to be both transcriptionally active and inactive genes (Figure 5). The actively transcribed genes all have a key
role in the maintenance of ESCs pluripotency and self-renewal. They include various ESC transcription factors, chromatin modifying enzymes and ESC-signal transduction genes. Conversely, the inactive genes are developmental transcription factors that are silent in ESCs, whose expression is associated with cellular differentiation and lineage commitment\textsuperscript{130,131}. Altogether, Oct4, Sox2 and Nanog appear to be master regulators of induced pluripotency by enhancing transcription of pluripotency genes, while at the same time silencing genes related to development and differentiation. Therefore, to achieve pluripotency, the autoregulatory loop and the transcriptional regulatory network will need to be resuscitated in reprogrammed somatic cells.

**Figure 4.** The autoregulatory loop. Oct4, Sox2 and Nanog form an interconnected autoregulatory circuit, by binding together to activates the promoters of both their own genes and the genes of each other. The 3 master regulators are able to maintain their own expression, thus maintaining pluripotency. Proteins are depicted as brown colors and gene promoters as dark slate gray. Adapted from (117).

**Figure 5.** The Oct4, Sox2 and Nanog trio contributes to ES cell pluripotency by repressing genes linked to lineage commitment, and activating genes involved in pluripotency. Proteins are depicted as brown colors and gene promoters as dark slate gray. Adapted from (117).

### 5.4. Epigenetic changes during iPSC reprogramming
iPSCs have a unique epigenetic signature that distinguish them from differentiated somatic cells (**Figure 6**). PSCs have open, active chromatin conformations, with activating histone H3 lysine-4 trimethylation marks (H3K4me3), histone acetylation and hypomethylated DNA around their pluripotency genes. In contrast, lineage-commitment leads to the silencing of these pluripotency genes, with repressive H3K27me3 and H3K9me3 histone marks, hypermethylated DNA and a closed heterochromatin conformation. During the reprogramming process, epigenetic signature of the somatic cell must be erased in order to adopt a stem cell-like epigenome. These epigenetic changes include chromatin remodeling, DNA demethylation of promoter regions of pluripotency genes, reactivation of the somatically silenced X chromosome and histone post-translational modifications\textsuperscript{8,15,17, 136-138}. 
Figure 6. Model of sequential steps in the reprogramming of somatic cells. (A) Sequential changes of phenotypes and activation of Oct4, Sox2 and Nanog. Following transduction with OSKM factors, the infected fibroblasts assumed a transformed phenotype. The endogenous Oct4 or Nanog genes become transcribed at a low level that is sufficient enough for drug resistance in cells carrying the neo gene but not sufficient to produce GFP expression in Oct4-GFP or Nanog-GFP cells. After 2-3 weeks endogenous Oct4 and Nanog genes become fully activated as shown by the appearance of GFP+ iPSCs in Oct4-GFP or Nanog-GFP fibroblasts\textsuperscript{15-17}. (B) During the reprogramming process, repressive H3K9me3 histone marks are gradually replaced by the transcriptionally active H3K4me3 histone marks while the DNA are gradually demethylated (open lollipops). (C) Molecular circuitry during reprogramming. During reprogramming process, the de novo methyltransferases Dnmt3a and Dnmt3b become activated and in turn de novo methylate and silence the virally transduced factors. The pluripotent state is now maintained by the autoregulatory loop of expression of the three master regulatory factors, Oct4, Sox2 and Nanog. Adapted from (116).

DNA methylation in iPSC reprogramming. DNA methylation is an epigenetic barrier of iPSCs generation\textsuperscript{139-141}. The methylation occurs at C5 position of cytosine on the target gene promoters in mammalian somatic cells\textsuperscript{138}. Promoter DNA methylation is inversely associated with gene expression\textsuperscript{142}. The epigenome of PSCs are transcriptionally active and are characterized by demethylation at the promoter regions of key pluripotency genes, like Oct4, Sox2 and Nanog (Figure 6). These genes are silenced by \textit{de novo} DNA methylation during lineage commitment and differentiation. The methylation is established by \textit{de novo} methyltransferases Dnmt3a and Dnmt3b and preserved by the maintenance methyltransferase Dnmt1\textsuperscript{143}. During reprogramming, the methylation marks are removed from these endogenous pluripotency genes in order to allow for their transcription, and tissue-specific genes are hypermethylated\textsuperscript{144,145}. Indeed, manipulation of the DNA and chromatin modifications by certain small molecules can significantly improve iPSCs formation\textsuperscript{37-50}. (See Reprogramming factors-epigenetic modifiers). Likewise, mice lacking DNA methyltransferases remain non-viable or die within weeks\textsuperscript{146-147}. These observations show that epigenetic modifications are key to cellular differentiation, and it is reasonable to conclude that these same events have to be reversed during induced reprogramming.

Histone modifications in iPSC reprogramming. Histone modification patterns differ between PSCs and differentiated somatic cells. The silencing of developmental genes in PSCs is controlled in a remarkable way. The differentiation-related genes carry “bivalent” domains (i.e. repressive histone H3 lysine-27 trimethylation marks (H3K27me3) and activating histone H3 lysine-4 trimethylation marks (H3K4me3)) in their genome loci\textsuperscript{148}. The H3K4me3 marks of the bivalent domains allows for transcription initiation on the developmental genes. Transcription are repressed on these genes by the action of Polycomb group, a family of proteins that regulate developmental gene expression through gene silencing by binding to repressive H3K27me3 marks.
Thus, lineage-commitment genes with bivalent domains can have their expression quickly turned on or switched off via erasure of H3K27me3 or H3K4me3, respectively. The bivalent domains are almost only found in PSCs and their restoration represent a vital step in the reprogramming process. During reprogramming, repressive H3K9me3 marks present on the endogenous pluripotency genes (Oct4, Sox2 and Nanog) are gradually replaced by the transcriptionally active H3K4me3\textsuperscript{144} (Figure 6 & 7). The loss of the H3K9me3 marks allow an access of OSKM transgenes on their target regions thus activating the autoregulatory loop.

5.5. Role of microRNAs in iPSC reprogramming.

miRNA are small non-coding RNA molecules that binds to protein coding messenger RNA (mRNA) to regulate their degradation or translation. They regulate gene expression by post-transcriptional gene silencing\textsuperscript{149}. Some miRNA promotes iPSC reprogramming (See Reprogramming factors-epigenetic modifiers), while others are barriers to iPSC reprogramming. Let-7 miRNAs are expressed in somatic cells and upregulated in ES cell differentiation\textsuperscript{150}. Lin 28 (one of the factors used by Thomson et al. to substitute for c-Myc and Klf4)\textsuperscript{9}, promotes reprogramming by inhibiting let-7 miRNAs\textsuperscript{151}.

5.6. The role of Reprogramming factors in iPSC reprogramming.

Following the introduction of exogenous OSKM factors into the somatic cells, exogenous Oct4 and Sox2 may directly induce the expression of endogenous Oct4, Sox2 and Nanog via the autoregulatory circuitry, through which they continue to maintain their own expression. Thereafter, these factors activate the pluripotent transcriptional network. Hence, the autoregulatory loop and the transcriptional network that are repressed in somatic cells, are now ‘resuscitated’ during the reprogramming process (Figure 8).

c-Myc is a vital component of active chromatin and associate with histone acetyltransferase (HAT) complexes. Thus, it facilitates an open chromatin conformation through global histone acetylation, thereby allowing Oct4 and Sox2 to target their genomic loci\textsuperscript{21,117,152}. As a well-known oncogene, it facilitates cancer-like transformation of somatic cells, conferring immortality and rapid proliferative potential on the PSCs\textsuperscript{153}. Thus, cellular division driven by c-Myc may provide somatic cells an opportunity to reset their epigenome,
thereby enhancing their reprogramming\textsuperscript{116}. As was mentioned in \textit{Reprogramming factors-cell cycle regulating genes}, p53 tumor suppressor proteins have inhibitory effect on iPSCs generation by promoting senescence, apoptosis and cell cycle inhibition\textsuperscript{39-43}. Hyperexpression of c-Myc can lead to increase in p53 levels and Klf4 can block the resulting apoptotic effect of c-Myc by suppressing p53 levels\textsuperscript{154}. Furthermore, Klf4 can suppress proliferation by activating p21 (a cyclin-dependent kinase inhibitor), and c-Myc can inhibit this anti-proliferative effect of Klf4 by suppressing p21\textsuperscript{155,156}. Thus, we can conclude that c-Myc and Klf4 are mutually complementary and a balance between their expressions is necessary for successful reprogramming\textsuperscript{117,153}. The overall summary of the roles of reprogramming factors is shown in \textbf{Figure 9}\textsuperscript{157}.

\textbf{Figure 8.} Exogenous Oct4 and Sox2 resuscitate the interconnected autoregulatory loop during reprogramming. In infected fibroblasts, endogenous Oct4, Sox and Nanog are reactivated by ectopic expression of Oct4, Sox2 and other factors. The endogenous genes (in dark slate gray) continue to maintain their own expression while the transgene expression is gradually silenced by de novo DNA methylation. This indicate that exogenous factors are required only for the induction of pluripotency. Adapted from (117).

\textbf{Figure 9.} The roles of OSKM factors in the induction of iPSCs. Pluripotent stem cells are immortal with open and active chromatin structure. It is probable that c-Myc induce these two properties by binding to several sites on the genome and by the recruitment of multiple histone acetylase complexes. However, c-Myc also induces apoptosis and senescence and this effect may be antagonized by Klf4. Oct3/4 probably changes the cell fate from tumor cells to ES-like cells while Sox2 helps to drive pluripotency. Adapted from (157).

5.7. \textbf{Two-phase Model of Induced Reprogramming: A gradual, stochastic process.}

Several studies have shown how exactly the ectopic expression of OSKM in somatic cells induces the transition to a pluripotent state\textsuperscript{157-162}. Based on these studies, we now know the order of events of the reprogramming process, and we can posit that the reprogramming process consists of two broad phases: An initial, stochastic \textbf{early} phase (phase 1) and a more deterministic and hierarchical \textbf{late} phase (phase 2) (\textbf{Figure 10}).

\textbf{Phase 1}

The earliest event in phase 1 is the \textbf{downregulation of lineage-specific genes}. This may be due to the direct repression effect of OSKM on these developmental genes or indirectly through the restoration of bivalent
histone marks on the same genes\textsuperscript{117}. The next event is the upregulation of a subset of ESC-specific genes, such as alkaline phosphatase (AP), Fbx15 and SSEA1. These two events may produce a partially reprogrammed iPSC with ESC-like morphology, but can quickly revert back to the differentiated state once the transgene expression is terminated. The next step is the \textbf{global chromatin remodeling of the full array of pluripotency genes}. This event involves the gradual unfolding of condensed heterochromatin to form an open euchromatin conformation and the removal of repressive H3K9me3 histone marks. The latter event is brought on by the effect of c-Myc, Klf4, histones modification enzymes (acetyltransferases and demethylases) and other small molecules. The removal of the repressive histone marks requires multiple rounds of cell division, and that is why reactivation of endogenous Oct4, Sox2 and Nanog occurs late in the reprogramming process.

\textbf{Phase 2}

After the completion of the global chromatin remodeling, exogenous Oct4 and Sox2 are now able to target and activate the loci of endogenous Oct4, Sox2 and Nanog genes leading to the \textbf{resuscitation of the autoregulatory loop}. The completion of chromatin remodeling at other pluripotency genes further leads to the gradual \textbf{resuscitation of the full ESC transcription network}. This lead to the establishment of full-blown pluripotency, characterized by reactivations of telomerase, inactivated X chromosome and ESC signalling cascades. As the autoregulatory loop continue to self-maintained the expression of the endogenous pluripotency genes, the \textbf{transgene silencing} previously initiated in phase 1 comes to a completion. The pluripotent state is now completely dependent on the endogenous autoregulatory circuitry.

\textbf{Figure 10.} Two-phase model of induced reprogramming. Adapted from (117).

\textbf{5.8. iPSC Reprogramming- An inefficient process}

As was mentioned in \textbf{Technical advances and progress in iPSC generation}, low reprogramming efficiency is one of the limitations of induced reprogramming\textsuperscript{7,8,16,27,28}. The \textit{Elite}, \textit{Stochastic} and \textit{Deterministic} models have been posited to explain the reason why only a small part of the transduced cells become pluripotent.

\textit{Elite model}. This model postulates that only a few, rare, ‘elites’ somatic cells (with stem cells characteristics)
present within the somatic cell population, can be induced towards pluripotency.\textsuperscript{163,164} In contrast to these ‘special’ cells, differentiated cells within the population are resistant to OSKM-mediated induction (Figure \textsuperscript{11a}). Although, somatic cell population are heterogeneous and contains stem cells\textsuperscript{165}, we now know that fully differentiated cells can be reprogrammed, thus disproving the elite model\textsuperscript{19,22,23}. Most of the somatic cells initiate the reprogramming process but majority never complete it.

**Stochastic and Deterministic models.** Assuming all somatic cells are transduced by the OSKM, the next path to pluripotency could occur by two mechanisms: a “stochastic” manner in which iPSCs appear at different, random, unpredictable periods; or a “deterministic” manner in which iPSCs appear at a fixed, predictable period (Figures \textsuperscript{11b} and \textsuperscript{11c}). Both types of mechanism might be involved in the reprogramming process.

Figure 11. Mechanistic insights into transcription factor-mediated reprogramming. (a) The Elite model, (b) The Deterministic model, and (c) The Stochastic model. Adapted from (164).

The generation of iPSCs require a precise, limited-range expression levels of the transduced factors and the process involves tightly regulated levels of pluripotency genes. Specific stoichiometry balance of the OSKM factors is fundamental for a successful reprogramming\textsuperscript{166,167}. Thus, maintaining this delicate balance appropriately can be a difficult, even a rare event. Additionally, somatic cells have to overcome many barriers on the road to pluripotency (See Two-phase Model of Induced Reprogramming: A gradual, stochastic process). Furthermore, random transgene integration can create a heterogeneous transgene expression that is achieved by very few cells. The lower chance of completing these stochastic reprogramming events and the need to overcome reprogramming barriers altogether contribute to the low efficiency of reprogramming.

There are other variables that can affect the efficiency of induced reprogramming such as, reprogramming factors, delivery methods, donor cell types and culture conditions\textsuperscript{29,82}. We have already considered the effects of Reprogramming factors and Delivery methods earlier in this review. Under the same culture conditions,
keratinocytes reprogramme 100 times more efficiently and two times faster than fibroblasts\textsuperscript{168}. Haematopoetic stem cells generate iPSC colonies 300 times more than B and T cells, suggesting that the differentiation status of the donor cell type is important\textsuperscript{169}. Hypoxic culture conditions (5% O\textsubscript{2}) greatly enhance reprogramming efficiency in mouse and human cells\textsuperscript{170}. Taken together, donor cell types and culture conditions can modulate reprogramming efficiencies.

6. iPSCs versus ESCs

Are iPSCs different from ESCs? Some recent comprehensive studies reveal only a few differences in global gene expression and DNA methylation patterns, which were more obvious in early passages of iPSCs\textsuperscript{171-173}. However, comparison studies with relatively smaller cell clones of iPSCs and ESCs revealed more significant differences in either gene expression or DNA methylation patterns\textsuperscript{174-176}. Some of the differences were attributed to differential activation of promoters by pluripotency factors and variables such as the exogenous factors combination, culture conditions and delivery methods. Altogether, these studies have conflicting conclusions, thus the answer to the question raised above is not straightforward. A study reveals a similarity in DNA methylation patterns between the iPSCs and the donor somatic cells, suggesting that iPSCs have a residual epigenetic ‘memory’ marks\textsuperscript{177,178}. Even among ESCs population, there exist epigenetic heterogeneity and variable differentiation potential\textsuperscript{179-180}. Thus, the current consensus is that iPSCs and ESCs are neither identical or distinct, but are overlapping cell populations with genetic and epigenetic differences that reflect their origins. Further experiments are essential to ascertain if these noticeable differences have any impact on the potential therapeutic utility of iPSCs.

Though iPSCs offer many advantages as compared with ESCs, there are some limitations associated with iPSCs as well. The Table 3 below shows the advantages and limitations of the iPSCs technology as compared with ESCs.

Table 3. Advantages and limitations of iPSCs technology.

7. Potential applications of iPSCs.

The iPSCs technology offers the opportunity to generate disease-specific and patient-specific iPSCs for modelling human diseases, drug development and screening, and individualized regenerative cell therapy. These three concepts are illustrated in Figure 12 and are discussed in this section.
7.1. Disease modelling
Genetically matched iPSC lines can be generated in unlimited quantities from patients afflicted with diseases of known or unknown causes. These cells can be differentiated in vitro into the affected cell types, thus recapitulating the ‘disease in a Petri dish’ models. The differentiated specialized cells or disease models, offers the opportunity to gain mechanistic insights into the disease and to use the cells to identify novel disease-specific drugs to treat the disorder; for example, drugs to prevent the death of medium spiny neurons in patients suffering from Huntington’s disease (Figure 12). The ability of iPSCs to proliferate extensively in culture and differentiate into all types of cells in the human body ensured that they can be use as disease models to study those diseases. Certainly, many studies have demonstrated the generation of iPSC lines from patients with various genetically inherited and sporadic diseases (Table 4). These in vitro studies give the first proof of principle that disease modelling using iPSCs technology is a viable option. However, the aim of disease modelling is to understand the molecular mechanism of diseases, with the ultimate goal of developing drugs for their treatment.

Figure 12. A schematic showing the potential applications of human iPSC technology for disease modelling, drug discovery and cell therapy using Huntington’s disease (HD) as an example. In HD patients, there is progressive loss of striatal GABAergic medium spiny neurons (MSNs). HD-specific iPSCs generated by cellular reprogramming can be differentiated into striatal MSNs in order to establish an in vitro model of the disease, and potential drugs can be screened leading to discovery of novel drugs that will prevent the degenerative process. Alternatively, if known, the disease-causing mutation (i.e. mutant HTT gene) could be repaired in iPSCs by gene targeting prior to their differentiation into healthy MSNs, followed by transplantation into the patient’s brain.

7.2. Drug development and cytotoxicity studies
Lee et al. utilized iPSCs to show disease modelling and drug screening for familial dysautonomia, a rare genetic disorder of the peripheral nervous system (Table 4). The generated familial dysautonomia-iPSCs were screened with multiple compounds and the authors revealed that a plant hormone, kinetin, can partly normalize the disease phenotype. Loss of neurons following in vitro differentiation of spinal muscular atrophy-iPSCs was ameliorated by exposure to experimental drugs. These studies and many others (see Table 4) show that iPSCs can facilitate drugs screening and discovery. Indeed, several clinical drug candidates
have been derived from iPSC studies currently in clinical trials\textsuperscript{193-196}. iPSCs are also used for testing for the toxic and non-toxic effect of therapeutic drugs. Itzhaki and colleagues use long QT syndrome cardiomyocytes-iPSCs to test the potency and efficacy of existing and new pharmacological drugs, and to assess the cardiotoxic effects and safe dose levels of drugs\textsuperscript{197}. As a powerful tool for disease models, drug discovery and cytotoxicity studies, iPSCs offers more advantage over animal models and clinical testing. Animal models does not perfectly mirror the true human disease phenotype, and iPSCs toxicity models are less expensive and saves time when compared with conventional testing systems. Additionally, different response to drug toxicity in animals due to species differences could prevent the recapitulation of full human disease phenotype.

Table 4. Summary of published human iPSC disease models. Adapted from (190). ND- not determined.

7.3. Regenerative medicine.

The iPSC technology offers an exciting opportunity of generating patient-specific stem cells for autologous transplantation. In regenerative medicine, the stem cells are used to promote endogenous regenerative repair or to replace injured tissues after cellular transplantation. The clinical translation of iPSC-based cell therapy is no longer futuristic, as the dream has now been realized. Two ground-breaking preclinical studies provided a proof-of-concept that led to the realization of this dream. In 2007, Jaenisch and colleagues used homologous recombination (gene targeting method) to repair the disease-causing mutations in iPSCs generated from humanized mouse model of sickle cell anemia (SCA)\textsuperscript{198}. The repaired SCA-iPSCs were differentiated into hematopoietic progenitor cells and subsequently transplanted into the affected transgenic mice. This resulted in the rescue and correction of the disease phenotype. The following year, Wernig and colleagues (from Jaenisch research group) reported an improvement in the dopaminergic function and behavioral symptoms in a rat model of Parkinson’s disease, after the transplantation of iPSC-derived dopaminergic neurons\textsuperscript{199}. These two successful iPSC-based cell therapies spurred the stem cell research community into exploring iPSCs therapy in humans. The first clinical trial using human iPSC was initiated in 2014 by transplanting human iPSC-derived retinal pigment epithelial (RPE) cells to treat macular degeneration\textsuperscript{200}. The progression of the
macular degeneration was halted in the first patient, with improved vision\textsuperscript{201}. However, the trial was placed on hold due to discovery of mutations in the iPSCs of the second patient\textsuperscript{200}. The researchers at RIKEN institute are hoping to resume the study using HLA-matched allogeneic iPSCs\textsuperscript{202,203}.

The recent advances in genome editing technology now allows for the introduction of genetic changes into iPSCs in a site-specific manner. We can now repair disease-causing gene mutations in patient-derived iPSCs, thus generating genetically healthy human iPSCs lines for iPSC-based cell therapy (See Figure 12). Similarly, we can also introduce specific mutations into non-diseased iPSCs, and generate genetically-matched isogenic iPSC lines that mimic the true pathology of the disease of interest, to be used for human iPSC-based disease models. Gene editing technologies like \textit{zinc-finger nuclease}s\textsuperscript{204,205}, \textit{transcription activator-like effector nuclease}s (TALENs)\textsuperscript{206-208}, and CRISPR-Cas\textsuperscript{9,209-212} technology has greatly improved the efficiency of gene editing in both human ESCs and iPSCs via DNA double-stranded breaks at the site of gene alteration. The combination of human iPSC platform with gene editing technologies can make iPSC-based cell therapy a more powerful and viable stem cell therapy option. The following section present an in-depth information regarding gene editing technology in iPSCs generation.

8. Genome editing technology in iPSCs generation

iPSCs have been indisputably proven to be a discovery that will transform medicine with respect to understanding the genetic etiology of diseases while equally providing the so needed genetic therapies. Its current combination with genome editing has further enhanced the diagnostic and therapeutic power of the iPSCs\textsuperscript{213}. Several methods have been used in the past to genetically target pluripotent stem cells. The process of gene targeting basically means modifying a specific genomic locus on a host DNA and the locus is replaced with an exogenous sequence by supplementation with a targeting vector. The technique of gene targeting has availed scientists the ability to control cellular genome\textsuperscript{213}. Gene targeting has however been shown to be way more challenging in human pluripotent stem cells than in mouse ES cells\textsuperscript{213} and this has been attributed to differences in developmental stages rather than species-related differences\textsuperscript{214}. Conventional gene targeting has recorded only a limited amount of success\textsuperscript{215} hence the drive towards developing better methods of gene targeting.

Gene editing technologies have remarkably improved over the years with the recent technologies enabled to introduce genetic changes in a site specific manner to the iPSCs\textsuperscript{216}. The more recent technologies induce double-stranded DNA breaks at the region of gene modification\textsuperscript{216}. These programmable site-specific nuclease have evolved from Zinc-finger nucleases (ZFN)\textsuperscript{204,205} to transcription activator-like effector nucleases (TALENs)\textsuperscript{207,208} and the RNA guided engineered nucleases (RGEN) gotten from the bacterial clustered regularly interspaced short palindromic repeat (CRISPR)-Cas (CRISPR-associated) 9 system\textsuperscript{210,211}. These technologies can easily correct pathology-causing genetic mutations derived from diseased patients and similarly can be used to induce specific mutations in disease-free wild-type iPSCs\textsuperscript{216}. Thus with this
approach, genetically matched, isogenic iPSCs can be generated, while ensuring that true pathologies are reliably identified and not confused with genetic back ground variations or epiphenomena associated with line-to-line disparities\textsuperscript{216}. In as much as the three nucleases possess similar mechanism of action which is the cleavage of chromosomal DNA in a location-specific manner, each of the nucleases still has its own unique characteristics\textsuperscript{217}. The well documented study done by Kim et al.\textsuperscript{217} on the nucleases has been briefly summarized in Table 5. Of the three nucleases, the CRISPR-Cas9 system has however gained wide acceptance and usage in the editing of human iPSC because it is simple to design and use\textsuperscript{216}, thus necessitating a little more review below.

Cas9 is a large multifunctional protein having two putative nuclease domains, the HNH and RuvC-like\textsuperscript{218}. The HNH and the RuvC-like domains cleave the complementary 20-nucleotide sequence of the crRNA and the DNA strand opposite the complementary strand respectively\textsuperscript{218}. Several variants of the CRISPR-Cas9 system exists and hence the subtle diversity to their modes of action: (1) The original CRISPR-Cas9 system functions by inducing DNA double-stranded breaks which is triggered by the wild-type Cas9 nuclease directed by a single RNA\textsuperscript{216}. However, its major challenge is the possibility of off-target effects\textsuperscript{216}, (2) The nickase variant of Cas9(D10A mutant) which is generated by the mutation of either the Cas9 HNH or the RuvC-like domain\textsuperscript{219,220} directed by paired guide RNAs, (3) Engineered nuclease variant of Cas9 with enhanced specificity (eSpCas9)\textsuperscript{221,222}. The nickase (D10A mutant) and the eSpCas9 variants have both been shown to substantially reduce off-targets effects while still maintaining their meticulous on-target cleavage\textsuperscript{221,222}, (4) Catalytically dead Cas9 (dCas9) variant is generated by mutating both domains (HNH and RuvC-like)\textsuperscript{219-220}. dCas9, when merged with a transcriptional suppressor or activator can be used to modify transcription of endogenous genes (CRISPRa or CRISPRi) or when fused with fluorescent protein can be used to image genomic loci\textsuperscript{221-223}, (5) A modified CRISPR-Cas9 variant has been used to efficiently introduce DNA sequences in an exact monoallelic or biallelic manner\textsuperscript{224}, and (6) CRISPR-Cas9 fused with cytidine deaminase, results in a variant which induces the direct conversion of cytidine to uridine, hence circumventing the DNA double-stranded break\textsuperscript{225}.

Hotta and Yamanaka have extensively reviewed how these nucleases have been used to mediate gene editing in pluripotent stem cells\textsuperscript{213}. Thus it is anticipated that the combination of these two technologies (gene editing and iPSCs) might be the dawn of a new phase of gene therapy.
9. Future perspective

The promise that iPSCs are viable and possibly superior substitutes for ESCs in disease modelling, drug discovery and regenerative medicine have not yet been fulfilled. Despite great successes in animal models, there are still many obstacles on the road to the clinical application of iPSCs. A major limitation is the heterogeneity nature of the cell population and differentiation potential of iPSCs. Hopefully, the CRISPR-Cas9 system can be used to address this limitation since the technology can improve the disease phenotype of differentiated cells\textsuperscript{213,226}. Another major limitation is the lack of robust lineage-specific differentiation protocols to generate large quantities of purified and matured iPSC-differentiated cells. More basic research on reprogramming technology are critical for the development of novel protocols for the generation of standardized human iPSC. A more current biotechnology, the microRNA switch\textsuperscript{227}, is expected to facilitate the maturation and purification of iPSC-differentiated cells and to reduce clonal variation.

While we wait for these limitations to be addressed, it will be wise to bank iPSCs from patients with specific diseases. Doing so will allow us the time to guarantee the quality of these cells thus saving time and cost when they are made available for transplantation.

10. Conclusion

The discovery of iPSCs by Takahashi and Yamanaka is truly a major breakthrough of the decade in stem cell science. The year 2016 marks the 10\textsuperscript{th} anniversary of this landmark discovery. The last decade has witnessed remarkable advancement in our understanding of the molecular mechanisms of induced pluripotency and we move from the ‘bench to the bedside’ in 2014. The more recent long-term study involving the application of human iPSC-derived dopaminergic neurons in primate Parkinson’s disease (PD) models at the Center for iPS Cell Research and Application, Kyoto University, Japan, reveals that human iPSCs are clinically applicable
for the treatment of patients with PD\textsuperscript{228}. The iPSC-based cell therapy is still at its infancy stage. The remaining barriers blocking the path to successful translation of this technology into clinical therapy has to be overcome. I believe many of these challenges are only technical in nature and with time ‘\textit{this too shall pass away}’. The combination of the human iPSC technology with genome-editing technologies may trigger a new era of gene therapy utilizing iPSCs.

\textbf{Authors’ contribution:}

AE wrote the manuscript; AO wrote the manuscript. AE and AO reviewed and approved the manuscript for publication.

\textbf{Acknowledgement}

The author wants to acknowledge Professor James Uney and Dr. Liang-Fong Wong for their guidance. Thanks to Kingsley Nnawuba for his assistance with some of the figures and tables.

\textbf{Conflict of Interest:}

The authors declare that there is no conflict of interest financial or otherwise.

\textbf{REFERENCES}

1. Gurdon, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. \textit{J. Embryol. Exp. Morphol.} \textbf{10}, 622-640.
2. Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H. (1997). Viable offspring derived from fetal and adult mammalian cells. \textit{Nature} \textbf{385}, 810-813.
3. Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. \textit{Curr. Biol.} \textbf{11}, 1553-1558.
4. Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. \textit{Nature} \textbf{292}, 154-156.
5. Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium
conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci. USA* **78**, 7634-7638.

6. Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145-1147.

7. Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676.

8. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872.

9. Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., Slukvin, I.I., Thomson, J.A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917-1920.

10. The 2012 Nobel Prize in Physiology or Medicine-Press Release. Nobel Media AB. 8 October, 2012. (Online). (Accessed 31 March 2017). Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/2012/press.html.

11. Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N, Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **17**, 126-140.

12. Cartwright, P., McLean, C., Sheppard, A., Rivett, D., Jones, K, Dalton, S. (2005). LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* **132**, 885-896.

13. Li, Y., McClintick, J., Zhong, L., Edenberg, H.J., Yoder, M.C., Chan, R.J. (2005). Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. *Blood* **105**, 635-637.

14. Niwa, H., Miyazaki, J., Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self renewal of ES cells. *Nature Genetics* **24**, 372-376.

15. Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., et al. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES cell-like state. *Nature* **448**, 318-324.

16. Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313-317.

17. Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., et al. (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1**, 55-70.

18. Park, I.H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., et al. (2008). Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141-146.
19. Stadtfeld, M., Brenmand, K., and Hochedlinger, K. (2008). Reprogramming of pancreatic β cells into induced pluripotent stem cells. *Curr. Biol.* **18**, 890-894.

20. Eminli, S., Utikal, J., Arnold, K., Jaenisch, R., and Hochedlinger, K. (2008). Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. *Stem Cells* **26**, 2467-2474.

21. Kim, J.B., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V. et al. (2008). Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* **454**, 646-650.

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

23. Hanna, J., Markoula, S., Schorderet, P., Carey, B.W., Beard, C., Wernig, M. et al. (2008). Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* **133**, 250-264.

24. Utikal, J., Maherali, N., Kulalurt, W., and Hochedlinger, K. (2009). Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells. *J. Cell Sci.* **122**, 3502-3510.

25. Sun, N., Panetta, N.J., Gupta, D.M., Wilson, K.D., Lee, A., Jia, F. et al. (2009). Feeder-free derivation of induced pluripotent stem cells from adult human adipose stem cells. *Proc. Natl Acad. Sci. USA* **106**, 15720-15725.

26. Maherali, N., Ahfeldt, T., Rigamonti, A., Utikal, J., Cowan, C., and Hochedlinger, K. (2008). A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* **3**, 340-345.

27. Lowry, W.E., Richter, L., Yacheck, R., Pyle, A.D., Tchiew, J., Sridharan, R., Clark, A.T., and Plath, K. (2008). Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl Acad. Sci. USA* **105**, 2883-2888.

28. Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D.A. (2008). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat.Biotechnol* **26**, 1269-1275.

29. Gonzalez, F., Boue, S., and Izpisua Belmonte, J.C. (2011). Methods for making induced pluripotent stem cells: reprogramming a la carte. *Nature Reviews Genetics* **12**, 231-242.

30. Yakubov, E., Rechavi, G., Rozenblatt, S., and Givol, D. (2010). Reprogramming of human fibroblasts to pluripotent stem cells using mRNA of four transcription factors. *Biochem. Biophys. Res Commun* **394**,
31. Zhao, Y., Yin, X., Qin, H., Zhu, F., Liu, H., Yang, W., et al. (2008). Two supporting factors greatly improve the efficiency of human iPSC generation. *Cell Stem Cell* 3, 475-479.

32. Tsubooka, N., Ichisaka, T., Okita, K., Takahashi, K., Nakagawa, M., and Yamanaka, S. (2009). Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. *Genes Cells* 14, 683-694.

33. Loewer, S., Cabili, M.N., Guttman, M., Loh, Y.H., Thomas, K., Park, I.H., et al. (2010). Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nature Genetics* 42, 1113-1117.

34. Wang, Y., Xu, Z., Jiang, J., Xu, C., Kang, J., Xiao, L., et al. (2013). Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Developmental Cell* 25, 69-80.

35. Melton, C., Judson, R.L., Blclloch, R. (2010). Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature* 463, 621-626.

36. Worninger, K.A., Rand, T.A., Hayashi, Y., Sami, S., Takahashi, K., Tanabe, K., Narita, M., Srivastava, D., and Yamanaka, S. (2014). The Let-7/LIN-41 pathway regulates reprogramming to human induced pluripotent stem cells by controlling expression of prodifferentiation genes. *Cell Stem Cell* 14, 40-52.

37. Feng, B., Jiang, J., Kraus, P., Ng, J.H., Heng, J.C., Chan, Y.S., et al. (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat. Cell Biol.* 11, 197-203.

38. Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovitch, E., Markoulaki, S., Ganz, K., et al. (2012). Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 150, 1209-1222.

39. Kawamura, T., Suzuki, J., Wang, Y.V., Menendez, S., Morera, L.B., Raya, A., Wahl, G.M., and Izpisua Belmonte, J.C. (2009). Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature* 460, 1140-1144.

40. Marion, R.M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., et al. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* 460, 1149-1153.

41. Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., et al. (2009). Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. *Nature* 460, 1145-1148.

42. Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka,
905 S. (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. *Nature* 460, 1132-1135.

907 43. Banito, A., Rashid, S.T., Acosta, J.C., Li, S., Pereira, C.F., Geti, I., et al. (2009). Senescence impairs successful reprogramming to pluripotent stem cells. *Genes Dev.* 23, 2134-2139.

909 44. Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Canamero, M., Blasco, M.A., and Serrano, M. (2009). The Ink4/Arf locus is a barrier for iPSC cell reprogramming. *Nature* 460, 1136-1139.

911 45. Mali, P., Ye, Z., Hommond, H.H., Yu, X., Lin, J., Chen, G., Zou, J., and Cheng, L. (2008). Improved efficiency and pace of generating induced pluripotent stem cells from human adult and fetal fibroblasts. *Stem Cells* 26, 1998-2005.

914 46. Edel, M.J., Menchon, C., Menendez, S., Consiglio, A., Raya, A., and Izpisua Belmonte, J.C. (2010). Rm2 GTPase maintains survival of human embryonic stem cells as well as enhancing reprogramming by regulating p53 and cyclin D1. *Genes Dev.* 24, 561-573.

917 47. Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nature Biotechnology* 26, 795-797.

920 48. Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D.A. (2008). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nature Biotechnology* 26, 1269-1275.

923 49. Li, W., Zhou, H., Abujarour, R., Zhu, S., Young Joo, J., Lin, T., Hao, E., Scholer, H.R., Hayek, A., and Ding, S. (2009). Generation of human-induced pluripotent stem cells in the absence of exogenous Sox2. *Stem Cells* 27, 2992-3000.

926 50. Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct3/4 during early embryogenesis. *Nature Cell Biol.* 8, 188-194.

929 51. Shi, Y., Desponts, C., Do, J.T., Hahn, H.S., Scholer, H.R., and Ding, S. (2008). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* 3, 568-574.

932 52. Onder, T.T., Kara, N., Cherry, A., Sinha, A.U., Zhu, N., Bernt, K.M., et al. (2012). Chromatin-modifying enzymes as modulators of reprogramming. *Nature* 483, 598-602.

934 53. Rais, Y., Zviran, A., Geula, S., Gafni, O., Chomsky, E., Viukov, S., et al. (2013). Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* 502, 65-70.
54. Yang, P., Wang, Y., Chen, J., Li, H., Kang, L., Zhang, Y., Chen, S., Zhu, B., and Gao, S. (2011). RCOR2 is a subunit of the LSD1 complex that regulates ESC property and substitutes for Sox2 in reprogramming somatic cells to pluripotency. *Stem Cells* **29**, 791-801.

55. Sharma, A., Diecke, S., Zhang, W.Y., Lan, F., He, C., Mordwinkin, N.M., Chua, K.F., and Wu, J.C. (2013). The role of SIRT6 in aging and reprogramming of human induced pluripotent stem cells. *The Journal of Biological Chemistry* **288**, 18439-18447.

56. Wang, T., Chen, K., Zeng, X., Yang, J., Wu, Y., Shi, X., Qin, B., Zeng, L., Esteban, M.A., Pan, G., and Pei, D. (2011). The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin C-dependent manner. *Cell Stem Cell* **9**, 575-587.

57. Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., et al. (2010). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* **6**, 71-79.

58. Chung, T.L., Brena, R.M., Kolle, G., Grimmond, S.M., Berman, B.P., Laird, P.W., Pera, M.F., and Wolvetang, E.J. (2010). Vitamin C promotes widespread yet specific DNA demethylation of the epigenome in human embryonic stem cells. *Stem Cells* **28**, 1848-1855.

59. Judson, R.L., Babiarz, J.E., Venere, M., and Blelloch, R. (2009). Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nature Biotechnology* **27**, 459-461.

60. Anokye-Danso, F., Trivedi, C.M., Juhr, D., Gupta, M., Cui, Z., Tian, Y., et al. (2011). Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* **8**, 376-388.

61. Lin, S.L., Chang, D.C., Lin, C.H., Ying, S.Y., Leu, D., and Wu, D.T. (2011). Regulation of somatic cell reprogramming through inducible mir-302 expression. *Nucleic Acids Research* **39**, 1054-1065.

62. Lin, S.L., Chang, D.C., Ying, S.Y., Leu, D., and Wu, D.T. (2010). MicroRNA miR-302 inhibits the tumorigenicity of human pluripotent stem cells by coordinate suppression of the CDK2 and CDK4/6 cell cycle pathways. *Cancer Research* **70**, 9473-9482.

63. Miyoshi, N., Ishii, H., Nagano, H., Haraguchi, N., Dewi, D.L., Kano, Y., et al. (2011). Reprogramming of mouse and human stem cells to pluripotency using mature microRNAs. *Cell Stem Cell* **8**, 633-638.

64. Subramanyam, D., Lamouille, S., Judson, R.L., Liu, J.Y., Bucay, N., Derynck.R., and Blelloch, R. (2011). Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nature Biotechnology* **29**, 443-448.

65. Jahner, D., Stuhlmann, H., Stewart, C.L., Harbers, K., Lohler, J., Simon, I., and Jaenisch, R. (1982). De novo methylation an expression of retroviral genomes during mouse embryogenesis. *Nature* **298**, 623-
66. Stewart, C.L., Stuhlmann, H., Jahner, D., and Jaenisch, R. (1982). De novo methylation, expression, and infectivity of retroviral genomes introduced into embryonal carcinoma cells. Proc. Natl Acad. Sci. USA 79, 4098-4102.

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

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

69. Rodriguez-Piza, I., Richaud-Patin, Y., Vassena, R., Gonzalez, F., Barrero, M.J., Veiga, A., Raya, A., and Izpisua-Belmonte, J.C. (2010). Reprogramming of human fibroblasts to induced pluripotent stem cells under xeno-free conditions. Stem Cells 28, 36-44.

70. Carey, B.W., Markoulaki, S., Hanna, J., Saha, K., Gao, Q., Mitalipova, M., and Jaenisch, R. (2009). Reprogramming of murine and human somatic cells using a single polycistronic vector. Proc. Natl Acad. Sci. USA 106, 157-162.

71. Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochdlinger, K., Kotton, D.N., and Mostoslavsky, G. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cells 27, 543-549.

72. Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G.W., Cook, E.G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M., Isacson, O., and Jaenisch, R. (2009). Parkinson’s disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. Cell 136, 964-977.

73. Chang, C.W., Lai, Y.S., Pawlik, K.M., Liu, K., Sun, C.W., Li, C., Schoeb, T.R., and Townes, T.M. (2009). Polycistronic lentiviral vector for “hit and run” reprogramming of adult skin fibroblasts to induced pluripotent stem cells. Stem Cells 27, 1042-1049.

74. Hockemeyer, D., Soldner, F., Cook, E.G., Gao, Q., Mitalipova, M., and Jaenisch, R. (2008). A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. Cell Stem Cell 3, 346-353.

75. Wernig, M., Lengner, C.J., Hannah, J., Lodato, M.A., Steine, E., Foreman, R., Staerk, J., Markoulaki, S., and Jaenisch, R. (2008). A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. Nature Biotechnology 26, 916-924.

76. Staerk, J., Dawlaty, M.M., Gao, Q., Maetzel, D., Hanna, J., Sommer, C.A., Mostoslavsky, G., and
Jaenisch, R. (2010). Reprogramming of peripheral blood cells into induced pluripotent stem cells. *Cell Stem Cell* 7, 20-24.

Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P., and Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* 458, 771-775.

Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hamalainen, R., et al. (2009). PiggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* 458, 766-770.

Newman, J.C., Bailey, A.D., Fan, H.Y., Pavelitz, T., and Weiner, A.M. (2008). An abundant evolutionarily conserved CSB-PiggyBac fusion protein expressed in Cockayne syndrome. *PLoS Genetics* 4, e100031.

Feschotte, C. (2006). The piggyBac transposon holds promise for human gene therapy. *Proc. Natl Acad. Sci. USA* 103, 14981-14982.

Grabundzija, I., Irgang, M., Mates, L., Belay, E., Matrai, J., Golgo-Doring, A., et al. (2010). Comparative analysis of transposable element vector systems in human cells. *Mol Ther.* 18, 1200-1209.

Brouwer, M., Zhou, H., and Kasri, N.N. (2016). Choices for induction of pluripotency: Recent developments in human induced pluripotency stem cell reprogramming strategies. *Stem Cell Rev.* 12, 54-72.

Grabundzija, I., Wang, J., Sebe, A., Erdei, Z., Kajdi, R., Devaraj, A., et al. (2013). Sleeping beauty transposon-based system for cellular reprogramming and targeted gene insertion in induced pluripotent stem cells. *Nuclei Acids Research* 41, 1829-1847.

Davis, R.P., Nemes, C., Varga, E., Freund, C., Kosmidis, G., Gkatzis, K., et al. (2013). Generation of induced pluripotent stem cells from human foetal fibroblasts using the Sleeping Beauty transposon gene delivery system. *Differentiation: Research in Biological Diversity* 86, 30-37.

Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G., and Hochedlinger, K. (2008). Induced pluripotent stem cells generated without viral integration. *Science* 322, 945-949.

Zhou, W., and Freed, C.R. (2009). Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. *Stem Cells* 27, 2667-2674.

Fusaki, N., Ban, H., Nishiyama, A., Saeki, K., and Hasegawa, M. (2009). Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc. Jpn Acad. Ser. B Phys. Biol. Sci.* 85, 348-362.

Seki, T., Yuasa, S., Oda, M., Egashira, T., Yae, K., Kusumoto, D., et al. (2010). Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. *Cell Stem Cell* 7, 11-14.
89. Ban, H., Nishishita, N., Fusaki, N., Tabata, T., Saeki, K., Shikamura, M., et al. (2011). Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors. Proc. Natl Acad. Sci. USA 108, 14234-14239.

90. Nishishita, N., Shikamura, M., Takenaka, C., Takada, N., Fusaki, N., and Kawamata, S. (2012). Generation of virus-free induced pluripotent stem cell clones on a synthetic matrix via a single cell subcloning in the naive state. PLoS One 7, e38389.

91. Ono, M., Hamada, Y., Horiuchi, Y., Matsuo-Takasaki, M., Imoto, Y., Satomi, K., et al. (2012). Generation of induced pluripotent stem cells from nasal epithelial cells using a Sendai virus vector. PLoS One 7, e42855.

92. Seki, T., Yuasa, S., and Fukuda, K. (2012). Generation of induced pluripotent stem cells from a small amount of human peripheral blood using a combination of activated T cells and Sendai virus. Nat. Protoc. 7, 718-728.

93. Macarthur, C.C., Fontes, A., Ravinder, N., Kuninger, D., Kaur, J., Bailey, M., et al. (2012). Generation of human pluripotent stem cells by a non-integrating RNA Sendai virus vector in feeder-free or xeno-free conditions. Stem Cells Int. 2012, 564612.

94. Li, H.O., Zhu, Y.F., Asakawa, M., Kuma, H., Hirata, T., Ueda, Y., et al. (2000). A cytoplasmic RNA vector derived from nontransmissible Sendai virus with efficient gene transfer and expression. J. Virol. 74, 6564-6569.

95. Tokusumi, T., Iida, A., Hirata, T., Kato, A., Nagai, Y., and Hasegawa, M. (2002). Recombinant Sendai viruses expressing different levels of a foreign reporter gene. Virus Res. 86, 33-38.

96. Inoue, M., Tokusumi, Y., Ban, H., Kanaya, T., Tokusumi, T., Nagai, Y., Iida, A., and Hasegawa, M. (2003). Nontransmissible virus-like particle formation by F-deficient Sendai virus is temperature sensitive and reduced by mutations in M and HN proteins. J. Virol. 77, 3238-3246.

97. Nakanishi, M., and Otsu, M. (2012). Development of Sendai virus vectors and their potential applications in gene therapy and regenerative medicine. Curr. Gene Ther. 12, 410-416.

98. Nishimura, K., Sano, M., Ohtaka, M., Furuta, B., Umemura, Y., Nakajima, Y., et al. (2011). Development of defective and persistent Sendai virus vector: a unique gene delivery/expression system ideal for cell reprogramming. J. Biol. Chem. 286, 4760-4771.

99. Kawagoe, S., Higuchi, T., Otaka, M., Shimada, Y., Kobayashi, H., Ida, H., et al. (2013). Morphological features of iPS cells generated from Fabry disease skin fibroblasts using Sendai virus vector (SeV dp). Molecular Genetics and Metabolism 109, 386-389.
100. Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., and Yamanaka, S. (2008). Generation of mouse induced pluripotent stem cells without viral vectors. *Science* **322**, 949-953.

101. Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, I.I., Thomson, J.A. (2009). Human induced pluripotent stem cells free of vector and transgene sequences. *Science* **324**, 797-801.

102. Gonzalez, F., Barragan Monasterio, M., Tiscornia, G., Montserrat Pulido, N., Vassena, R., Battle Morera, L., Rodriguez Piza., I., and Izpisua Belmonte, J.C. (2009). Generation of mouse-induced pluripotent stem cells by transient expression of a single nonviral polycistronic vector. *Proc. Natl Acad. Sci USA* **106**, 8918-8922.

103. Okita, K., Hong, H., Takahashi, K., and Yamanaka, S. (2010). Generation of mouse-induced pluripotent stem cells with plasmid vectors. *Nature Protoc.* **5**, 418-428.

104. Cheng, L., Hansen, N.F., Zhao, L., Du, Y., Zou, C., Donovan, F., et al. (2012). Low incidence of DNA sequence variation in human induced pluripotent stem cells generated by nonintegrating plasmid expression. *Cell Stem Cell* **10**, 337-344.

105. Montserrat, N., Garreta, E., Gonzalez, F., Gutierrez, J., Eguizabal, C., Ramos, V., Borros, S., and Izpisua Belmonte, J.C. (2011). Simple generation of human induced pluripotent stem cells using poly-beta-amino esters as the non-viral gene delivery system. *The Journal of Biological Chemistry* **286**, 12417-12428.

106. Si-Tayeb, K., Noto, F.K., Sepac, A., Sedlic, F., Bosnjak, Z.J., Lough, J.W., and Duncan, S.A. (2010). Generation of human induced pluripotent stem cells by simple transient transfection of plasmid DNA encoding reprogramming factors. *BMC Dev Biol.* **10**, 81.

107. Okita, K., Matsumura, Y., Sato, Y., Okada, A., Morizane, A., Okamoto, S., Hong, H., Nakagawa, M., Tanabe, K., Tezuka, K., Shibata, T., Kunisada, T., Takahashi, M., Takahashi, J., Saji, H., and Yamanaka, S. (2011). A more efficient method to generate integration-free human iPS cells. *Nat. Methods* **8**, 409-412.

108. Jia, F., Wilson, K.D., Sun, N., Gupta, D.M., Huang, M., Li, Z., et al. (2010). A nonviral minicircle vector for deriving human iPS cells. *Nature Methods* **7**, 197-199.

109. Narsinh, K.H., Jia, F., Robbins, R.C., Kay, M.A., Longaker, M.T., and Wu, J.C. (2011). Generation of adult human induced pluripotent stem cells using nonviral minicircle DNA vectors. *Nature Protocols* **6**, 78-88.

110. Chen, Z.Y., He, C.Y., Ehrhardt, A., and Kay, M.A. (2003). Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. *Mol. Ther.* **8**, 495-500.

111. Chen, Z.Y., He, C.Y., and Kay, M.A. (2005). Improved production and purification of minicircle DNA
vector free of plasmid bacterial sequences and capable of persistent transgene expression in vivo. *Hum. Gene Ther.* **16**, 126-131.

112. Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.H., Li, H., Lau, F., et al. (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* **7**, 618-630.

113. Warren, L., Ni, Y., Wang, J., and Guo, X. (2012). Feeder-free derivation of human induced pluripotent stem cells with messenger RNA. *Scientific Reports* **2**, 657.

114. Kim, D., Kim, C.H., Moon, J.I., Chung, Y.G., Chang, M.Y., Han, B.S., et al. (2009). Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* **4**, 472-476.

115. Zhou, H., Wu, S., Joo, J.Y., Zhu, S., Han, D.W., Lin, T., et al. (2009). Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* **4**, 381-384.

116. Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* **132**, 567-582.

117. Scheper, W., and Copray, S. (2009). The molecular mechanism of induced pluripotency: A two-stage switch. *Stem Cell Rev. and Rep.* **5**, 204-223.

118. Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., et al. (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat. Cell Biol.* **9**, 625-635.

119. Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-655.

120. Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on Sox2 function. *Genes and Development* **17**, 126-140.

121. Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-391.

122. Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-642.

123. 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.

Dang, D.T., Pevsner, J., and Yang, V.W. (2000). The biology of the mammalian Kruppel-like family of transcription factors. *Int. J. Biochem. Cell Biol.* **32**, 1103-1121.

Nakatake, Y., Fukui, N., Iwamatsu, Y., Masui, S., Takahashi, K., Yagi, K., et al. (2006). Klf4 cooperates with with Oct3/4 and Sox2 to activate the Lefty1 core promoter in embryonic stem cells. *Mol Cell Biol* **26**, 7772-7782.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **136**, 1063-1069.

Dang, C. V., O’Donnell, K.A., Zeller, K.I., Nguyen, T., Osthus, R.C., and Li, F. (2006). The c-Myc target gene network. *Seminars in Cancer Biology* **16**, 253-264.

Lebofsky, R., and Walter, J.C. (2007). New Myc-anism for DNA replication and tumorigenesis? *Cancer cell* **12**, 102-103.

Patel, J.H., Loboda, A.P., Showe, M.K., Showe, L.C., and McMahon, S.B. (2004). Analysis of genome targets reveals complex functions of MYC. *Nature Reviews Cancer* **4**, 562-568.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-956.

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**, 431-440.

Chew, J.L., Loh, Y.H., Zhang, W., Chen, X., Tam, W.L., Yeap, L.S., et al. (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol. Cell. Biol.* **25**, 6031-6046.

Kuroda, T., Tada, M., Kubota, H., Kimura, H., Hatano, S.Y., Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol. Cell. Biol.* **25**, 2475-2485.

Okumura-Nakanishi, S., Saito, M., Niwa, H., and Ishikawa, F. (2005). Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J. Biol. Chem.* **280**, 5307-5317.

Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H, Wang, B., Ng, H.H., and Robson, P. (2005). Transcriptional regulation of Nanog by OCT4 and SOX2. *J. Biol. Chem.* **280**, 24731-24737.

Fussner, E., Djuric, U., Strauss, M., Hotta, A., Perez-Iratxeta, C., Lanner, F., Dilworth, F.J., Ellis, J., and Bazett-Jones, D.P. (2011). Constitutive heterochromatin reorganization during somatic cell reprogramming. *EMBO J.* **30**, 1178-1189.
Buganim, Y., Faddah, D.A., and Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. *Nat. Rev. Genet.* 14, 427-439.

Gonzalez, F., and Huangfu, D. (2016). Mechanisms underlying the formation of induced pluripotent stem cells. *Wiley Interdiscip. Rev. Biol.* 5, 39-65.

Nishino, K., Toyoda, M., Yamazaki-Inoue, M., Fukawatase, Y., Chikazawa, E., Sakaguchi, H., et al. (2011). DNA methylation dynamics in human induced pluripotent stem cells over time. *PLOS Genet.* 7, 5-8.

Doege, C.A., Inoue, K., Yamashita, T., Rhee, D.B., Travis, S., Fujita, R., et al. (2012). Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* 488, 625-655.

Gao, Y., Chen, J., Li, K., Wu, T., Huang, B., Liu, W., et al. (2013). Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell* 12, 453-469.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6-21.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* 14, 204-220.

Gladych, M., Andrzejewska, A., Oleksiewicz, U., and Estecio, M.R. (2015). Epigenetic mechanisms of induced pluripotency. *Contemp. Oncol. (Pozn)* 19, A30-A38.

Berdasco, M., and Estellar, M. (2011). DNA methylation in stem cell renewal and multipotency. *Stem Cell Res. Ther.* 2, 42.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915-926.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247-257.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-326.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism and function. *Cell* 116, 281-297.

Roush, S., and Slack, F.J. (2008). The let-7 family of microRNAs. *Trends Cell Biol.* 18, 505-516.

Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. *Science* 320, 97-100.

Knoepfler, P.S., Zhang, X.Y., and Cheng, P.F. (2006). Myc influences global chromatin structure. *EMBO Journal* 25, 2723-2734.
Yamanaka, S. (2007). Strategies and new developments in the generation of patient-specific pluripotent stem cells. *Cell Stem Cell* 1, 39-49.

Rowland, B.D., Bernards, R., and Peeper, D.S. (2005). The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene. *Nat. Cell Biol.* 7, 1074-1082.

Zhang, W., Geiman, D.E., Shields, J.M., Dang, D.T., Mahatan, C.S., Kaestner, K.H., et al. (2000). The gut-enriched Kruppel-like factor (Kruppel-like factor 4) mediates the transactivating effect of p53 on the *p21*WAF1/Cip1 promoter. *J Biol Chem* 275, 18391-18398.

Seoane, J., Le, H.V., and Massague, J. (2002). Myc suppression of the p21(Cip1) Cdk inhibitor influences the outcome of the p53 response to DNA damage. *Nature* 419, 729-734.

Yamanaka, S. (2009). Elite and stochastic models for induced pluripotent stem cell generation. *Nature* 460, 49-52.

Takahashi, K., and Yamanaka, S. (2016). A decade of transcription factor-mediated reprogramming to pluripotency. *Nat. Rev. Mol. Cell Biol.* 17, 183-193.

Goodell, M.A., Nguyen, H., and Shroyer, N. (2015). Somatic stem cell heterogeneity: diversity in the blood, skin and intestinal stem cell compartments. *Nat. Rev. Mol. Cell Biol.* 16, 299-309.
Optimal reprogramming factor stoichiometry increases colony numbers and affects molecular characteristics of murine induced pluripotent stem cells. *Cytometry A.* **79**, 426-435.

Yamaguchi, S., Hirano, K., Nagata, S., and Tada, T. (2011). Sox2 expression effects on direct reprogramming efficiency as determined by alternative somatic cell fate. *Stem Cell Res.* **6**, 177-186.

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., et al. (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nature Biotechnol.* **26**, 1276-1284.

Eminli, S., Foudi, A., Stadtfeld, M., Maherali, N., Ahfeldt, T., Mostoslavsky, G., Hock, H., and Hochedlinger, K. (2009). Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nature Genet.* **41**, 968-976.

Yoshida, Y., Takahashi, K., Okita, K., Ichisaka, T., and Yamanaka, S. (2009). Hypoxia enhances the generation of induced pluripotent stem cells. *Cell Stem Cell* **5**, 237-241.

Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., et al. (2011). Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* **144**, 439-452.

Guenther, M.G., Frampton, G.M., Soldner, F., Hockemeyer, D., Mitalipova, M., Jaenisch, R., and Young, R.A. (2010). Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell* **7**, 249-257.

Newman, A.M., and Cooper, J.B. (2010). Lab-specific gene expression signatures in pluripotent stem cells. *Cell Stem Cell* **7**, 258-262.

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., et al. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* **5**, 111-123.

Marchetto, M.C., Yeo, G.W., Kainohana, O., Marsala, M., Gage, F.H., and Muotri, A.R. (2009). Transcriptional signature and memory retention of human-induced pluripotent stem cells. *PLoS ONE* **4**, e7076.

Lister, R., Pelizzola, M., Kida, Y.S., Hawkins, R.D., Nery, J.R., Hon, G., et al. (2011). Hot-spots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68-73.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., et al. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285-290.

Kim, K., Zhao, R., Doi, A., Ng, K., Unvermaehr, J., Cahan, P., et al. (2010). Donor cell type can influence
the epigenome and differential potential of human induced pluripotent stem cells. *Nat. Biotechnol.* **29,** 1117-1119.

Martinez, Y., Bena, F., Gimelli, S., Tirefort, D., Dubois-Dauphin, M., Krause, K.H., Preynat-Seauve, O. (2012). Cellular diversity within embryonic stem cells: pluripotent clonal sublines show distinct differentiation potential. *J. Cell Mol. Med.* **16,** 456-467.

Osafune, K., Caron, L., Borowiak, M., Martinez, R.J., Fitz-Gerald, C.S., Sato, Y., et al. (2008). Marked difference in differentiation propensity among human embryonic stem cell lines. *Nature Biotechnol.* **26,** 313-315.

Guha, P., Morgan, J.W., Mostoslavsky, G., Rodrigues, N.P., and Boyd, A.S. (2013). Lack of immune response to differentiated cells derived from syngeneic induced pluripotent stem cells. *Cell Stem Cell* **12,** 407-412.

Chun, Y.S., Byun, K., and Lee, B. (2011). Induced pluripotent stem cells and personalized medicine: current progress and future perspectives. *Anat. Cell Biol.* **44,** 245-255.

Wobus, A.M., and Loser, P. (2011). Present state and future perspectives of using pluripotent stem cells in toxicology research. *Arch. Toxicol.* **85,** 79-117.

Choi, M.S., Kim, Y., Shim, J.S., Park, J.T., Wang, R., Leach, S.D. et al. (2013). Efficient drug screening and gene correction for treating liver disease using patient-specific stem cells. *Hepatology* **57,** 2458-2468.

Hochedlinger, K., Yamada, Y., Beard, C., and Jaenisch, R. (2005). Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell* **121,** 465-477.

Park, E.T., Gum, J.R., Kakar, S., Kwon, S.W., Deng, G., and Kim, Y.S. (2008). Aberrant expression of SOX2 upregulates MUC5AC gastric foveolar mucin in mucinous cancers of the colorectum and related lesions. *Int. J. Cancer* **122,** 1253-1260.

Ghaleb, A.M., Nandan, M.O., Chanchevalap, S., Dalton, W.B., Hisamuddin, I.M., and Yang, V.W. (2005). Kruppel-like factors 4 and 5: the yin and yang regulators of cellular proliferation. *Cell Res.* **15,** 92-96.

Kuttler, F., and Mai, S. (2006). c-Myc, genomic instability and disease. *Genome Dyn.* **1,** 171-190.

Pappas, J.J., and Yang, P.C. (2008). Human ESC vs iPSC- Pros and Cons. *J. Cardiovasc. Trans. Res.* **1,** 96-99.

Wu, S.M., and Hochedlinger, K. (2011). Harnessing the potential of induced pluripotent stem cells for regenerative medicine. *Nat. Cell Biol.* **13,** 497-505.

Lee, G., Papapetrou, E.P., Kim, H., Chambers, S.M., Tomishima, M.J., Fasano, C.A., et al. (2009).
Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature* **461**, 402-406.

192 Ebert, A.D., Yu, J., Rose, F.F. Jr., Mattis, V.B., Lorson, C.L., Thomson, J.A., and Svendsen, C.N. (2009). Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature* **457**, 277-280.

193 Bright, J., Hussain, S., Dang, V., Wright, S., Cooper, B., Byun, T., et al. (2015). Human secreted tau increases amyloid-beta production. *Neurobiol. Aging* **36**, 693-709.

194 Naryshkin, N.A., Weetall, M., Dakka, A., Narasimhan, J., Zhao, X., Feng, Z., et al. (2014). SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* **345**, 688-693.

195 Mullard, A. (2015). Stem-cell discovery platforms yield first clinical candidates. *Nat. Rev. Drug Discov.* **14**, 589-591.

196 McNeish, J., Gardner, J.P., Wainger, B.J., Woolf, C.J., and Eggnan, K. (2015). From dish to bedside: lessons learned while translating findings from a stem cell model of disease to a clinical trial. *Cell Stem Cell* **17**, 8-10.

197 Itzhaki, I., Maizels, L., Huber, I., Zwi-Dantsis, L., Caspi, O., Winterstern, A., et al. (2011). Modelling the long QT syndrome with induced pluripotent stem cells. *Nature* **471**, 225-229.

198 Hanna, J., Wernig, M., Markoulaki, S., Sun, C.W., Meissner, A., Cassady, J.P., et al. (2007). Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* **318**, 1920-1923.

199 Wernig, M., et al. (2008). 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.

200 Kimbrel, E.A. and Lanza, R. (2015). Current status of pluripotent stem cells: moving the first therapies to the clinic. *Nat. Rev. Drug Discov.* **14**, 681-692.

201 Scudellari, M. (2016). How iPS cells changed the world. *Nature* **534**, 310-312.

202 Trounson, A., and DeWitt, N.D. (2016). Pluripotent stem cells progressing to the clinic. *Nat. Rev. Mol. Cell Biol.* **17**, 194-200.

203 Cell Stem Cell Editorial Team. (2016). 10 questions: clinical outlook of iPSCs. *Cell Stem Cell* **18**, 170-173.

204 Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., et al. (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat.
Zou, J., Maeder, M.L., Mali, P., Pruett-Miller, S.M., Thibodeau-Beganny, S., Chou, B.K., et al. (2009). Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. Cell Stem Cell 5, 97-110.

Christian, M, Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., and Voytas, D.F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186, 757-761.

Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. Nat. Biotechnol. 29, 731-734.

Sanjana, N.E., Cong, L., Zhou, Y, Cunniff, M.M., Feng, G., and Zhang, F. (2012). A transcription activator-like effector toolbox for genome engineering. Nat. Protoc. 7, 171-192.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genomic engineering using CRISPR-Cas systems. Science 339, 819-823.

Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat. Methods 10, 973-976.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelson, T., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343, 84-87.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821.

Hotta, A., and Yamanaka, S. (2015). From genomics to gene therapy: Induced pluripotent stem cells meet genome editing. Annu. Rev. Genet. 49, 47-70.

Shi, Y., Inoue, H., Wu, J.C., and Yamanaka, S. (2017). Induced pluripotency stem cell technology: a decade of progress. Nature Reviews 16, 115-130.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell 4, 487-492.

Urbach, A., Schuldiner, M., and Benvenisty, N. (2004). Modeling for Lesch-Nyhan disease by gene targeting in human embryonic stem cells. Stem Cells 22, 635-641.

Kim, H., and Kim, J. (2014). A guide to genome engineering with programmable nucleases. Nature Reviews 15, 321-334.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 346 (6213), 1258096.
1339 219. Li, T., Huang, S., Jiang, W.Z., Wright, D., Spalding, M.H., Weeks, D.P., and Yang, B. (2011). TAL
1340 nuclease (TALNs): hybrid proteins composed of TAL effectors and Fok1 DNA-cleavage domain.
1341 *Nucleic Acids Res* **39**, 359-372.
1342 220. Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., and
1343 Voytas, D.F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757-761.
1344 221. Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z, Yang, Z., et al. (2013). Chromosomal deletions and
1345 inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.* **41**, e141.
1346 222. Gupta, A., Hall, V.L., Kok, F.O., Shin, M., McNulty, J.C., Lawson, N.D., and Wolfe, S.A. (2013).
1347 Targeted chromosomal deletions and inversions in zebrafish. *Genome Res.* **23**, 1008-1017.
1348 223. Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J. (2014). Analysis of off-target
1349 effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* **24**, 132-141.
1350 224. Gaj, T., Gersbach, C.A., and Barbas, C.F. 3rd. (2013). ZFN, TALEN, and CRISPR/Cas-based methods
1351 for genome engineering. *Trends Biotechnol.* **31**, 397-405.
1352 225. Segal, D.J., and Meckler, J.F. (2013). Genome engineering at the dawn of the golden age. *Annu. Rev.
1353 Genom. Hum. Genet.* **14**, 135-158.
1354 226. Deleidi, M., and Yu, C. (2016). Genome editing in pluripotent stem cells: research and therapeutic
1355 applications. *Biochem. Biophys. Res. Commun.* **473**, 665-674.
1356 227. Miki, K., Endo, K., Takahashi, S., Funakoshi, S., Takei, I., Katayama, S., et al. (2015). Efficient detection
1357 and purification of cell populations using synthetic microRNA switches. *Cell Stem Cell* **16**, 699-711.
1358 228. Kikuchi, T., Morizane, A., Doi, D., Magotani, H., Onoe, H., Hayashi, T., et al (2017). Human iPS cell-
1359 derived dopaminergic neurons function in a primate Parkinson’s disease model. *Nature* **548**, 592-596.
Figure 1 (on next page)

Historical timeline showing events that led to the development of iPSCs.
1962 1981 1997 1998 2001 2006 2007 2012

- John Gurdon: Cloning in frog using SCNT
- Ian Wilmut: Cloning in sheep using SCNT
- Martin Evans & Mathew Kaufman, Gail Martin: Generation of mouse ESCs
- James Thomson: Generation of human ESCs
- Shinya Yamanaka: Generation of human iPS Cs with OSKM
- Masako Tada: Reactivation of pluripotency-related genes in mouse somatic cells fusion with mouse ESCs
- James Thomson: Generation of human iPS Cs with OSNL
- Shinya Yamanaka & Kazutoshi Takahashi: Generation of mouse iPS Cs with OSKM
- Shinaya Yamanaka & John Gurdon: The award of 2012 Nobel prize for Physiology and Medicine
Figure 2 (on next page)

Generation of iPSCs from MEF cultures via 24 factors by Yamanaka.
G418 sensitive

Cellular reprogramming

G418 resistant

MEF

Fbxo15 promoter

β-geo

Retroviral infection with 24 candidate factors

ES like cells

Fbxo15 promoter

β-geo
Figure 3 (on next page)

Schematic representation of various delivery methods of iPSC induction.
The autoregulatory loop. Oct4, Sox2 and Nanog form an interconnected autoregulatory circuit, by binding together to activates the promoters of both their own genes and the genes of each other. The 3 master regulators are able to maintain their own expression, thus maintaining pluripotency. Proteins are depicted as brown colors and gene promoters as dark slate gray. Adapted from (117)
The Oct4, Sox2 and Nanog trio contributes to ES cell pluripotency by repressing genes linked to lineage commitment, and activating genes involved in pluripotency. Proteins are depicted as brown colors and gene promoters as dark slate gray. Adapted from (117).
Model of sequential steps in the reprogramming of somatic cells. (A) Sequential changes of phenotypes and activation of Oct4, Sox2 and Nanog. Following transduction with OSKM factors, the infected fibroblasts assumed a transformed phenotype. The endogenous Oct4 or Nanog genes become transcribed at a low level that is sufficient enough for drug resistance in cells carrying the neo gene but not sufficient to produce GFP expression in Oct4-GFP or Nanog-GFP cells. After 2-3 weeks endogenous Oct4 and Nanog genes become fully activated as shown by the appearance of GFP+ iPSCs in Oct4-GFP or Nanog-GFP fibroblasts\textsuperscript{15-17}. (B) During the reprogramming progress, repressive H3K9me3 histone marks are gradually replaced by the transcriptionally active H3K4me3 histone marks while the DNA are gradually demethylated (open lollipops). (C) Molecular circuitry during reprogramming. During reprogramming process, the de novo methyltransferases Dnmt3a and Dnmt3b become activated and in turn de novo methylate and silence the virally transduced factors. The pluripotent state is now maintained by the autoregulatory loop of expression of the three master regulatory factors, Oct4, Sox2 and Nanog. Adapted from (116).
Schematic representation of the chromatin rearrangement occurring during somatic cell reprogramming and differentiation of pluripotent stem cells. Adapted from (144).
Exogenous Oct4 and Sox2 resuscitate the interconnected autoregulatory loop during reprogramming. In infected fibroblasts, endogenous Oct4, Sox and Nanog are reactivated by ectopic expression of Oct4, Sox2 and other factors. The endogenous genes (in dark slate gray) continue to maintain their own expression while the transgene expression is gradually silenced by de novo DNA methylation. This indicate that exogenous factors are required only for the induction of pluripotency. Adapted from (117).
The roles of OSKM factors in the induction of iPSCs. Pluripotent stem cells are immortal with open and active chromatin structure. It is probable that c-Myc induce these two properties by binding to several sites on the genome and by the recruitment of multiple histone acetylase complexes. However, c-Myc also induces apoptosis and senescence and this effect may be antagonized by Klf4. Oct3/4 probably changes the cell fate from tumor cells to ES-like cells while Sox2 helps to drive pluripotency. Adapted from (157).
Somatic Cells → c-Myc → Apoptosis, senescence

Immortalization, open chromatin → Oct-3/4 → Nullipotent ES-like Cells

Nullipotent ES-like Cells → Sox2 → Pluripotent iPS Cells

Klf4
Figure 10 (on next page)

Two-phase model of induced reprogramming. Adapted from (117).
Phase I

Downregulation of lineage genes
- Direct repression
- Restoration of bivalent domains

Activation of specific ES cell genes
- E.g. SSEA1, Fbx15, AP

Chromatin remodelling at pluripotency genes
- Unfolding of condensed chromatin
- Removal of repressive chromatin marks

Phase II

Resuscitation of autoregulatory loop

Full reactivation of ES cell transcriptional network
- Reactivation of telomerase, ES cell signal cascades, etc.

Completion of transgene silencing
Mechanistic insights into transcription factor-mediated reprogramming.

(a) The Elite model, (b) The Deterministic model, and (c) The Stochastic model. Adapted from (164).
A schematic showing the potential applications of human iPSC technology for disease modelling, drug discovery and cell therapy using Huntington’s disease (HD) as an example. In HD patients, there is progressive loss of striatal GABAergic medium spiny neurons (MSNs). HD-specific iPSCs generated by cellular reprogramming can be differentiated into striatal MSNs in order to establish an *in vitro* model of the disease, and potential drugs can be screened leading to discovery of novel drugs that will prevent the degenerative process. Alternatively, if known, the disease-causing mutation (i.e. mutant HTT gene) could be repaired in iPSCs by gene targeting prior to their differentiation into healthy MSNs, followed by transplantation into the patient’s brain.
Treatment of HD patients with drug

Healthy medium spiny neurons

GABAergic medium spiny neurons are dying

Transplantation of healthy medium spiny neurons

HD patient

Skin biopsy

HD-iPSCs

Add reprogramming factors

Klf4
Oct4
Sox2
c-Myc

Genetically healthy medium spiny neurons

In vitro differentiation

Repaired HD-iPSCs

Use of gene targeting to repair mutant HTT gene thus silencing the mutant HTT expression

Unhealthy HD medium spiny neurons

Screening of drugs that prevent medium spiny neurons death

Disease modelling

In vitro differentiation

Mutant HTT gene

(CAG)n
The characterization of iPSCs. Adapted from (82).
| **Morphology** | Flat, cobblestone-like cells, ES like morphology |
|               | Tightly packed colonies with sharp edges |
| **Pluripotency markers** | Alkaline phosphatase assay (as a live marker) |
|               | Increase levels of pluripotency proteins such as Oct4, Nanog, SSEA3/4, TRA-1-60 and TRA-1-81. |
| **Differentiation potential** | Teratoma formation- can form ectoderm, mesoderm and endoderm, the three germ layers. |
|               | Embryoid body formation-can form ectoderm, mesoderm and endoderm, the three germ layers. |
| **Genetic Analyses** | Diploid karyotype. |
|               | Transgene silencing after reprogramming. |
| **Epigenetic Analyses** | DNA methylation of lineage-committed genes |
|               | DNA demethylation of key pluripotency genes like Oct4, Sox2, Nanog |
Table 2 (on next page)

Reprogramming factors capable of reprogramming human cells. Adapted from (82).
| Reprogramming factors | Function                                      | Affected pathway                  | Effect on pluripotency | References      |
|-----------------------|-----------------------------------------------|-----------------------------------|------------------------|-----------------|
| Oct4                  | Maintenance of pluripotency and self-renewal. | Core transcriptional circuitry    | +                      | 8               |
| Sox2                  | Maintenance of pluripotency and self-renewal  | Core transcriptional circuitry    | +                      | 8               |
| Klf4                  | Maintenance of pluripotency and self-renewal  | Core transcriptional circuitry    | +                      | 124,125,126     |
| c-Myc                 | Maintenance of pluripotency and self-renewal  | Core transcriptional circuitry    | +                      | 8               |
| Lin28                 | Maintenance of pluripotency, translational enhancer, inhibits let7 | Core transcriptional circuitry    | +                      | 9,38            |
| Nanog                 | Maintenance of pluripotency and self-renewal  | Core transcriptional circuitry    | +                      | 9,38            |
| Sall4                 | Maintenance of pluripotency and self-renewal  | Core transcriptional circuitry    | +                      | 32,38           |
| Utf1                  | Maintenance of pluripotency                    | Core transcriptional circuitry    | +                      | 31,38           |
| p53                   | Induces senescence, tumor suppressor           | Apoptosis/ cell cycle             | -                      | 39-43           |
| Reprogramming factors | Function | Affected pathway | Effect on pluripotency | References |
|-----------------------|----------|-----------------|-----------------------|------------|
| Oct4                  | Maintenance of pluripotency and self-renewal | Core transcriptional circuitry | + | 8 |
| Sox2                  | Maintenance of pluripotency and self-renewal | Core transcriptional circuitry | + | 8 |
| Klf4                  | Maintenance of pluripotency and self-renewal | Core transcriptional circuitry | + | 124,125, 126 |
| c-Myc                 | Maintenance of pluripotency and self-renewal | Core transcriptional circuitry | + | 8 |
| Lin28                 | Maintenance of pluripotency, translational enhancer, inhibits let7 | Core transcriptional circuitry | + | 9,38 |
| Nanog                 | Maintenance of pluripotency and self-renewal | Core transcriptional circuitry | + | 9,38 |
| Salt4                 | Maintenance of pluripotency and self-renewal | Core transcriptional circuitry | + | 32,38 |
| Utf1                  | Maintenance of pluripotency | Core transcriptional circuitry | + | 31,38 |
| p53                   | Induces senescence, tumor suppressor | Apoptosis/ cell cycle | - | 39-43 |

### Non-coding RNA

| Reprogramming factors | Function | Affected pathway | Effect on pluripotency | References |
|-----------------------|----------|-----------------|-----------------------|------------|
| miR367                | Inhibits EMT | TGFβ | + | 60 |
| LincRNA-ROR           | Regulates expression of core transciational factors | Core transcriptional circuitry | + | 33,34,35,36 |
| miR302                | Inhibits EMT/stimulates oct4 expression | TGFβ; Core transcriptional circuitry; apoptosis | + | 60-62, 64 |
| miR766                | Inhibits Sirt6 | Chromatin remodeling | - | 55 |
| miR200c               | Inhibits EMT/TGFβ pathway | TGFβ | + | 63 |
| miR369                | Inhibits EMT/TGFβ pathway | TGFβ | + | 63 |
| miR372                | Inhibits EMT/TGFβ pathway | TGFβ | + | 64 |
| Let7                  | Regulates expression of core transciational factors and prodifferentiation genes | Core Transcriptional circuitry/ TGFβ | - | 33,34,35,36 |
| Reprogramming factors | Function                                | Affected pathway             | Effect on pluripotency | Reference |
|-----------------------|-----------------------------------------|------------------------------|------------------------|-----------|
| Vitamin C             | Alleviates cell senescence/antioxidant  | Hypoxia response             | +                      | 56-58     |
| Valproic acid         | Inhibits histone deacetylases           | Chromatin remodeling         | +                      | 47        |
| CHIR99021             | GSK 3-inhibitor                        | PI3k; Wnt/β-catenin          | +                      | 49        |
| Parnate               | Lysine-specific demethylase 1 inhibitor | Chromatin remodeling         | +                      | 49        |
| BIX-01294             | Methyltransferase G9a inhibitor         | Chromatin remodeling         | +                      | 50,51     |
| 5-azacytidine         | DNA methyltransferase inhibitor         | Chromatin remodeling         | +                      | 47        |
| Trichostatin A        | Inhibits histone deacetylases           | Chromatin remodeling         | +                      | 47        |
Table 3 (on next page)

Advantages and limitations of iPSCs technology
| Advantages                                                                 | Limitations                                                                 |
|---------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| Eliminates ethical issues and religious concerns associated with ESCs use | Efficiency of reprogramming is generally low<sup>7,8,27,28</sup>             |
| Risk of immune rejection is reduced<sup>181</sup>                         | Tumorigenesis<sup>16</sup>                                                  |
| Donor cell is easily and non-invasively obtained, no embryo destruction   | Risk of insertional mutagenesis from virus based delivery methods<sup>7,8,9,16</sup> |
| Accessible to large number of patients, unlike ESCs limited by ethical concerns | Increased chances of development of diseases due to factors used<sup>185-188</sup> |
| Personalization of treatment with patient-specific stem cells and drugs<sup>182</sup> | Very early days in this field, more basic research are needed                |
| Use for disease modelling-they carry the same disease-causing factor as the patient | Complex and polygenic diseases are difficult to be modeled.                  |
| High-throughput screening for drugs and toxicity prediction<sup>183,184</sup> | High costs associated with production and characterization of each cell line |
| Allows for gene targeting and gene editing technology to correct mutations<sup>184</sup> | Suboptimal standardization<sup>189</sup>. Stringent protocols are still needed. |
Table 4 (on next page)

Summary of published human iPSC disease models. Adapted from (190). ND- not determined.
| Disease type | Disease name | Genetic cause | Number of lines | Cell type | Control line | Phenotype | Drug test | PMID    |
|--------------|--------------|---------------|----------------|-----------|--------------|-----------|-----------|---------|
| Neurological | Parkinson’s disease | Polygenic (with LRRK2 mutation) | 23 | Dopaminergic neurons | hiPSC | No obvious defect | Yes | 19269371 |
|              | Amyotrophic lateral sclerosis | Polygenic | 3 | Motor neurons | hESC | ND | ND | 18669821 |
|              | Spinal muscular atrophy | Monogenic | 2 | Motor neurons | hiPSC | Loss of neuron formation, loss of SMN gene expression | Yes | 19098894 |
|              | Familial dysautonomia | Monogenic | 2 | Neural crest cells | hiPSC, hESC | Loss of neural crest cells | Yes | 19693009 |
|              | RETT syndrome | Monogenic | 4 | Neurons | hiPSC | Loss of synapses, reduced spine density, smaller soma size | Yes | 21074045 |
|              | Huntington’s disease | Monogenic | 2 | ND | hiPSC, hESC | ND | ND | 18691744 |
| Blood        | Friedreich ataxia | Monogenic | 6+ | ND | hESC | Changes GAA-TTC repeat | ND | 21040903 |
|              | Fanconi anaemia | Monogenic | 19 | Blood cells | hiPSC, hESC | Corrected loss of FANCA function | ND | 15483674 |
|              | Fragile X syndrome | Monogenic | 11 | ND | hiPSC, hESC | Loss of FMR1 expression | ND | 20452313 |
| Disease type                  | Disease name                          | Genetic cause | Number of lines | Cell type                                      | Control line | Phenotype                                                                 | Drug test | PMID       |
|------------------------------|---------------------------------------|---------------|-----------------|------------------------------------------------|--------------|---------------------------------------------------------------------------|-----------|------------|
| Cardiac and Vascular         | Long QT 1 syndrome                    | Monogenic     | 6               | Cardiomyocytes                                   | hPSC         | Increased cardiomyocyte depolarization                                    | Yes       | 20666394   |
|                              | Long QT 2 syndrome                    | Monogenic     | Not reported    | Cardiomyocytes                                   | hPSC         | Increased cardiomyocyte depolarization                                    | Yes       | 21240260   |
|                              | LEOPARD syndrome                      | Monogenic     | 6               | Cardiomyocytes                                   | hPSC, hESC   | Increased cardiomyocyte size, decreased MAPK signalling                   | ND        | 20535210   |
|                              | Timothy syndrome                      | Monogenic     | 16              | Cardiomyocytes                                   | hPSC         | Increased cardiomyocyte depolarization                                    | Yes       | 21307850   |
|                              | Hutchinson Gilford Progeria           | Monogenic     | 4               | Smooth muscle cells, mesenchymal stem cells      | hPSC, hESC   | Smooth muscle and mesenchymal cells-apoptosis                              | ND        | 21185252   |
|                              | Monogenic                              | 6               | Smooth muscle cells                                   | hPSC         | Smooth muscle cell nuclear morphology and ageing phenotype                 | ND        | 21346760   |
|                              | Duchenne muscular dystrophy           | Monogenic     | 2               | ND                                               | hPSC, hESC   | ND                                                                        | ND        | 18691744   |
| Pancreatic                   | Type 1 diabetes                       | Polygenic     | 4               | Insulin- and glucagon-producing cells            | hESC         | ND                                                                        | ND        | 19720998   |
|                              | Hepatic                               | Monogenic     | 19              | Hepatocytes                                       | hPSC         | Loss of A1-antitrypsin expression                                          | Yes       | 20739751   |
|                              | Others                                | Monogenic     | 4               | Neurons                                           | hPSC, hESC   | Imprint disorder                                                          | ND        | 20956530   |
|                              | Angelman and Prader-Willi syndrome    | Monogenic     | 13              | Neurons                                           | hPSC, hESC   | Loss of paternal UBE3A expression                                          | ND        | 20876107   |
|                              | Down syndrome                         | Monogenic     | 2               | ND                                               | hPSC, hESC   | ND                                                                        | ND        | 18691744   |
Table 5 (on next page)

Summary of the nucleases used in genome editing for iPSCs generation.

a) ZFN b) TALENS c) RGEN
| Nuclease | Composition | Availability | Targetable sites | Pitfalls |
|----------|-------------|--------------|------------------|----------|
| ZFN      | ZFN is composed of a modular structure which has two domains: a DNA-binding Zinc-finger protein (ZFP) domain and a nuclease domain gotten from the *FokI* restriction enzyme. The *FokI* nuclease domain has to dimerize in order to cleave DNA. ZFPs determines the ZFNs sequence specificity, which comprise of C2H2 zinc-fingers tandem arrays—the DNA-binding motif that is most common in higher eukaryotes. | By modular assembly of pre-characterized zinc-fingers, it is quite convenient to construct new ZFPs with desired specificities. Available resources for programmable nucleases have been extensively elucidated by Kim et al.217 | Sites that can be successfully targeted are often rich in guanines and consists of 5'-GNN-3' (where N stands for nucleotide) repeat sequences. | The ZFNs created through the convenient method of zinc-fingers pre-characterization are often devoid of DNA targeting activity or are often cytotoxic owing to off-target effects. Constructing ZFNs with high activity and low cytotoxicity still remains a challenge with the use of publicly available resources. The use of ZFNs are hampered by poor targeting densities Presently no available open-source collection of 64 zinc-fingers that can cover all the likely combinations of triplet sites. Chromosomal DNA cannot be cleaved efficiently by all newly assembled ZFNs, especially those having 3 zinc-fingers. |
### Table 5b

| Nuclease | Composition | Availability | Targetable Sites | Pitfalls |
|----------|-------------|--------------|------------------|----------|
| TALENs   | Although the TALENs use a different category of DNA-binding domains named transcription activator-like effectors (TALEs), they however, still contain the FokI nuclease domain at their carboxyl termini. The TALEs are made up of 33-35 amino acid repeats. Repeat variable diresidues (RVDs) determines the nucleotide specificity of each repeat domain. The 4 different RVDs include: Asn-Ile, His-Asp, Asn-Asn, Asn-Gly-these are most widely used to recognize adenine, cytosine, guanine and thymine respectively. New TALENs with desired sequence specificities can be easily designed because of the one-to-one correspondence between the 4 bases and the 4 RVD modules. Available resources for programmable nucleases have been extensively elucidated by Kim et al.\(^\text{217}\) | The crucial advantage of TALENs over the other nucleases is that it can be designed to target almost any desired DNA sequence. Although conventional TALENs do not cleave target DNA containing methylated cytosine, interestingly, a methylated cytosine is identical to thymine in the major groove. Therefore, Asn-Gly RVD repeat (which recognizes thymines) can be used to replace His-Asp RVD repeat (which recognizes cytosines) and thus generate TALENs that cleave methylated DNA. | The fact that TALENs frequently consists of about 20 RVDs and that highly homologous sequences can fuse with one another in cells, make the construction of DNA segments that encode TALE arrays challenging and time-consuming. The need for a thymine to be at the 5ʹ of the target sequence for recognition by two amino-terminal cryptic repeat folds appear to be the only limitation to the construction of the TALENs. |
| Nuclease | Composition | Availability | Targetable Sites | Pitfalls |
|----------|-------------|--------------|-----------------|----------|
| RGEN     | The organisms bacteria and archaea capture small fragment of the DNA (~20bp) form the DNA of invading plasmids and phages and fuses these sequences (named protospacers) with their own genome thus forming a CRISPR. For type II CRISPR, the CRISPR sites are first transcribed as pre-CRISPR RNA (pre-crRNA) and further processed to form target-specific CRISPR RNA (crRNA). Also contributing to the processing of the pre-crRNA is the invariable target-independent trans-activating crRNA (tracrRNA), which is also transcribed from the locus. New RGEN formation does not require complicated protein engineering because Cas9 stays the same. Available resources for programmable nucleases have been extensively elucidated by Kim et al.²¹⁷ RGENs cleave methylated DNA as opposed to TALENs and ZFNs. | 20-bp guide DNA sequences can be cloned into vectors that encode either crRNA or sgRNA and this easily generates new RGEN plasmids. A 23 –bp target DNA sequence is cleaved by the formed DNA endonuclease, this target DNA sequence is made up of the 20-bp guide sequence in the crRNA (which is the protospacer) and the 5′-NGG-3′, also 5′-NAG-3′ (but to a lesser degree) a sequence regarded as the protospacer adjacent motif (PAM), recognizable by Cas9 itself. RGENs in cells do not efficiently cleave all sequences that contain the PAM sequence. | The need for a PAM sequence is a limitation for the RGEN target sites. The need for guanine to be at the 5′ end is also another limitation for the targetable sites as RNA polymerase III transcribes guide RNAs under the guidance of the U6 promoter in cells. | The need for guanine to be at the 5′ end is also another limitation for the targetable sites as RNA polymerase III transcribes guide RNAs under the guidance of the U6 promoter in cells. |
tracrRNA, this simplifies the RGEN components.