Cell Stem Cell
Perspective

Genetic and Epigenetic Variations in iPSCs: Potential Causes and Implications for Application

Gaoyang Liang[1,2,3] and Yi Zhang[1,2,3,4,5,*]
[1]Howard Hughes Medical Institute
[2]Program in Cellular and Molecular Medicine
[3]Division of Hematology/Oncology, Department of Pediatrics
Boston Children's Hospital, 300 Longwood Avenue, Boston, MA 02115, USA
[4]Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA
[5]Harvard Stem Cell Institute, WAB-149G, 200 Longwood Avenue, Boston, MA 02115
*Correspondence: yzhang@genetics.med.harvard.edu
http://dx.doi.org/10.1016/j.stem.2013.07.001

The ability to reprogram somatic cells to induced pluripotent stem cells (iPSCs) has revolutionized the field of regenerative medicine. However, recent studies on the genetic and epigenetic variations in iPSCs have raised concerns that these variations may compromise the utility of iPSCs. In this Perspective, we review the current understanding of genetic and epigenetic variations in iPSCs, trace their causes, discuss the implications of these variations for iPSC applications, and propose approaches to cope with these variations.

Introduction

iPSCs, derived from transcription-factor-mediated reprogramming, are pluripotent stem cells (PSCs) with molecular and functional properties similar to embryonic stem cells (ESCs) (Stadtfeld and Hochedlinger, 2010). The iPSC technology holds tremendous promise for regenerative medicine. iPSCs offer autologous cell sources for replacement therapy, and patient-specific iPSCs can serve as in vitro models for disease mechanism studies and drug screening (Robinton and Daley, 2012). Yet, this promise is obscured by recent findings of genetic and epigenetic variations in iPSCs. These variations exist between iPSC lines, between iPSC and ESC lines, between different passages of the same iPSC line, and even between different populations at a specific passage of the same iPSC line. Such variations potentially affect the properties of iPSCs and undermine their accountability in downstream applications. In this Perspective, we discuss the genetic and epigenetic variations in iPSCs and their causes, the implications of these variations in iPSC applications, and potential approaches to cope with these variations.

Genetic Variations in iPSCs

An iPSC genome may harbor a wide range of variations, including aneuploidy, subchromosomal copy number variation (CNV), and single nucleotide variations (SNVs). These variations can be introduced into the iPSCs from different sources during iPSC generation and maintenance (Figure 1). First, genetic variations in iPSCs may originate from the heterogeneous genetic makeup of source cell population. Due to the low efficiency and clonal nature of iPSC derivation, individual iPSC lines are capable of capturing genetic variations from individual starting cells, even if the variations only occur at low frequencies among the source cells (Figure 1AI). Moreover, if certain genetic variations in source cells facilitate the derivation of iPSCs, those variations will be preferentially propagated in the derived iPSC lines (Figure 1AII). Second, the reprogramming process may be mutagenic, which potentially introduces de novo variations (Figure 1B). Third, like ESCs, prolonged culturing of iPSCs may introduce or select for genetic alterations that facilitate cell propagation (Figure 1C). In addition to these causes, certain variations may arise from innate genetic instability of the in vitro pluripotent state. In the following sections, we will discuss each type of genetic variation and look into its potential causes.

Aneuploidy

Recurrent aneuploidy. Aneuploidy, an abnormality in chromosome number, is frequently reported in in vitro cultured PSCs, including iPSCs and ESCs. One comprehensive study by the International Stem Cell Initiative revealed that approximately one in three analyzed human ESC (hESC) or iPSC (hiPSC) lines have karyotype abnormalities in at least one passage (Amps et al., 2011), while a second study estimated that ~13% of hESC and hiPSC cultures bear aberrant karyotypes (Taapken et al., 2011). Recurrent gains of specific chromosomes account for more than half of the total karyotype abnormalities, with trisomy 12 being the most common in both hESCs and hiPSCs. Other less frequent whole-chromosome gains include trisomy of chromosome 8 and chromosome X (Amps et al., 2011; Mayshar et al., 2010; Taapken et al., 2011). For unknown reasons, trisomy 17, which occurs frequently in hESCs, is rarely detected in hiPSCs (Mayshar et al., 2010; Taapken et al., 2011). In mouse ESC (mESC) and iPSC (miPSC) lines, whole-chromosome gain occurs frequently for chromosomes 8 and 11, and the latter shares significant syntenic regions with human chromosome 17 (Ben-David and Benvenisty, 2012).

The recurrent aneuploidy patterns in PSCs have long been thought to reflect the adaptation of these cells to their in vitro culture conditions (Baker et al., 2007). The occurrence frequency generally increases through continuous passaging, although the abnormalities can be detected at early passages, and normal karyotypes can be found at late passages (Amps et al., 2011; Taapken et al., 2011). In addition, recurrent aneuploidy can be detected in a particular subpopulation of hESC or hiPSC culture. The fact that these subpopulations expand along passaging suggests that the abnormalities are positively selected during culturing (Amps et al., 2011; Mayshar et al., 2010; Taapken et al., 2011). Gaining an extra copy of certain chromosomes

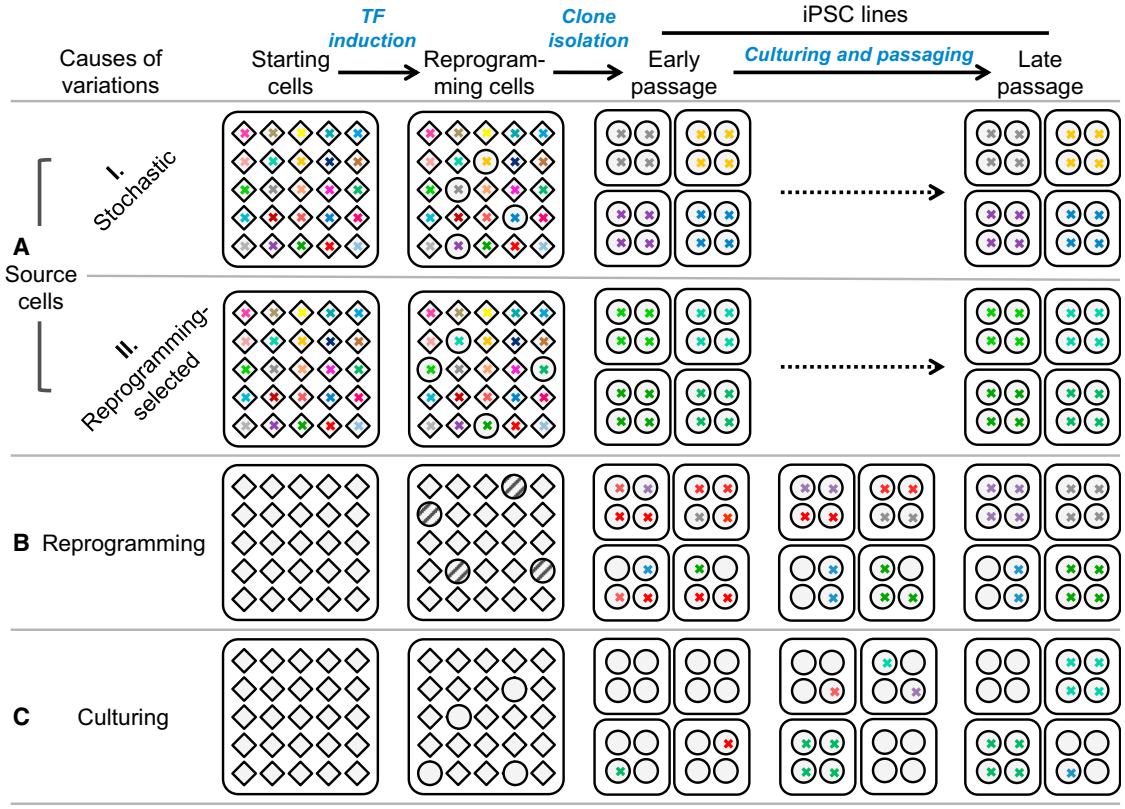

Figure 1. Sources of Genetic Variations in iPSC Lines

Genetic variations of iPSC lines may have different sources.

(A) Individual starting somatic cells (diamond) within a culture (rounded rectangle) bear subtle genetic variations (colored crosses), which can be captured and manifested in the iPSC (circle) lines for the clonal nature of the transcription factor (TF)-mediated iPSC derivation process. (AI) Given that reprogramming occurs stochastically among the starting cell population, the genetic variations captured in iPSC lines may have random patterns. (AII) If reprogramming preferentially takes place in cells bearing genetic variations conferring selective advantage (green crosses), the iPSC-manifested variations may show functional enrichment.

(B) The reprogramming process per se may introduce variations. The cells that undergo reprogramming may have enhanced genomic instability (striped circles), resulting in de novo mutations in iPSCs. Early-passage iPSCs may display mosaicism of de novo mutations, which are subjected to selection along passaging. Mutations conferring advantage in self-renewal or proliferation (green crosses) eventually prevail the culture; those deleterious for cell survival (red crosses) are selected against in culture; while other neutral mutations (crosses with other colors) undergo genetic drift.

(C) Mutations that arise during prolonged culturing are subjected to similar selection patterns described in (B).

can confer growth advantage by increasing the dosage of genes beneficial for self-renewal or proliferation. For example, human chromosome 12 harbors pluripotency genes *NANOG* and *GDF3*, which may explain the frequent trisomy of this chromosome in cultured hESC and hiPSC lines (Draper et al., 2004; Mayshar et al., 2010). Despite a general correlation between recurrent aneuploidy and prolonged culturing, it is not clear whether whole-chromosome stability is generally compromised in PSCs.

Aneuploidy from the source cells in iPSC generation. While both ESCs and iPSCs are subject to the risk of acquiring aneuploidy due to long-term culturing and/or their innate properties, the chromosomal states of iPSCs can also be influenced by source cells and the reprogramming process. In fact, transcription-factor-mediated reprogramming has been shown to be compatible with aneuploid starting cells. Fibroblasts with trisomy 21 from Down syndrome patients can be successfully reprogrammed into iPSCs (Park et al., 2008). One study on trisomy 21 iPSCs containing a selectable transgene on one copy of chromosome 21 showed that the rate of spontaneous loss of this chromosome is approximately 10^{-4} (Li et al., 2012a); however, it is unknown whether euploid or other aneuploid iPSCs bear a similar rate of chromosome loss. Furthermore, iPSCs can also be derived from mouse embryonic fibroblasts (MEFs) in aneuploid-susceptible genetic backgrounds with efficiency comparable to that of normal MEFs (Hamada et al., 2012). Interestingly, high rates of aneuploidy in source cells do not always translate into similarly high rates in the derived iPSC lines. Deficiency in either BubR1, a core component of spindle assembly complex, or RanBP2, a regulator of chromosome decatenation, results in similarly high rates of aneuploidy in MEFs. However, the aneuploidy rate is substantially different between the iPSC lines from these two genetic backgrounds. While nearly all iPSC lines with BubR1 deficiency show higher rates of aneuploidy relative to the starting MEFs, iPSCs derived from RanBP2-deficient background are largely devoid of aneuploidy and even have a lower aneuploidy rate compared to wild-type iPSCs (Hamada et al., 2012). It seems likely that aneuploidy stemming from different causes may be subjected to different selective pressures during reprogramming, resulting in different

levels of tolerance in the reprogramming cells. Therefore, aneuploidy in iPSC lines can be inherited from source cells, but whether reprogramming selects for or against aneuploidy depends on how aneuploidy is generated.

Subchromosomal Variations

CNV. Similar to recurrent aneuploidy, hiPSCs and hESCs may share Mb-scale CNVs, which can be detected by karyotyping and gene expression meta-analysis. Some of these CNVs occur at specific chromosomal locations, for example, around pluripotency gene *NANOG* on human chromosome 12 and *DNMT3B* on human chromosome 20 (Laurent et al., 2011; Martins-Taylor et al., 2011; Mayshar et al., 2010). Such location preferences indicate that these CNVs may be selected during in vitro propagation.

In addition, CNVs can be introduced into iPSCs from source cells and/or during the reprogramming process. Comparative genomic analyses based on single nucleotide polymorphism (SNP) array detected high CNV frequency in hiPSCs when compared to hESCs, source cells, or human non-PSCs (Hussein et al., 2011; Laurent et al., 2011). The CNVs specifically detected in iPSCs have been suspected to be generated from the reprogramming process. One study showed that the iPSC-specific CNVs constitute genetic mosaicism in iPSC lines at early passages. The mosaicism is gradually lost through passaging, consistent with the finding that most of these CNVs are deletions, which is likely disadvantageous to cell survival (Hussein et al., 2011). Further investigation of the iPSC-specific CNVs indicates they are enriched in common fragile sites where the replication fork is inclined to stall and collapse (Hussein et al., 2011). This leads to the hypothesis that the reprogramming process de novo generates CNVs for the elevated replication stress encountered by the reprogramming cells (Hussein et al., 2011, 2013). A second study reported the detection of high-frequency deletions in early-passage iPSCs. Interestingly, many of the deletions occur near tumor suppressor genes, indicating potential roles of these CNVs in reprogramming (Laurent et al., 2011). However, neither the findings of early-passage mosaicism nor CNV enrichment around tumor suppressor genes in iPSCs could be confirmed in a later study (Abyzov et al., 2012). One caveat for the array-based studies is that low-frequency variations in source cells might be beyond the detection limit of this approach. This limitation might explain the discrepancy between results from the array-based studies (Hussein et al., 2011; Laurent et al., 2011) and those from studies using next generation sequencing, which is generally a more sensitive method (Abyzov et al., 2012; Cheng et al., 2012; Gore et al., 2011; Quinlan et al., 2011; Young et al., 2012).

Overall, sequencing-based studies have revealed much fewer or no detectable de novo CNVs in iPSCs (Abyzov et al., 2012; Cheng et al., 2012; Gore et al., 2011; Quinlan et al., 2011; Young et al., 2012), arguing against the notion that the reprogramming process is susceptible to de novo CNVs. In particular, one study, which represents perhaps the most comprehensive study so far of CNVs using a deep sequencing approach, showed that on average each hiPSC line manifests only two CNVs that are undetectable in the bulk population of starting fibroblasts (Abyzov et al., 2012). However, at least half of the hiPSC-manifested CNVs are present in rare populations (estimated to be <15%) of parental fibroblasts when more sensitive PCR-based detection is applied (Abyzov et al., 2012). This finding suggests that iPSC generation does not necessarily lead to de novo generation of CNVs. Instead, the low-grade genetic mosaicism of CNVs in source cells is clonally captured by the iPSC derivation process. Although it is possible that some of the CNVs might arise during the reprogramming process, low-grade genetic mosaicism of the somatic cells appears to be the major source of CNVs in iPSCs.

SNV. SNVs have been studied in both hiPSC and miPSC lines by high-throughput sequencing of whole genome or exome (protein-coding regions). Genome-wide sequencing has revealed more than a thousand SNVs for each examined hiPSC line and hundreds of SNVs in the mouse counterpart (Cheng et al., 2012; Young et al., 2012). The average numbers of SNVs in protein-coding genes are estimated to be no more than a dozen in an individual hiPSC or miPSC line (Cheng et al., 2012; Gore et al., 2011; Ji et al., 2012; Ruiz et al., 2013; Young et al., 2012). The number of SNVs in each iPSC line appears to be independent of the delivery methods of reprogramming factors (viral vector, episomal vector, or mRNA transduction) or source cell types (Gore et al., 2011; Ruiz et al., 2013). Half or more than half of the exome SNVs in iPSCs can be traced back to the source cells (Gore et al., 2011; Young et al., 2012). It is therefore important to note that to date only a single study has argued that the majority of SNVs come from the reprogramming process (Ji et al., 2012). However, the limited sequencing depth of this study may not have been able to uncover rare SNVs in the source cells, leading to an overestimation of the SNVs generated during reprogramming. In addition, mutation rate in the source cells was not experimentally assessed in this study, which may also contribute to the discrepancy. Therefore, although reprogramming-induced point mutations cannot be excluded, source cell contribution appears to be the major source of SNVs manifested in iPSCs.

Several studies have also probed into the potential relationship between SNVs and iPSC generation. Most studies revealed no shared SNV between any of the iPSC lines examined (Cheng et al., 2012; Gore et al., 2011; Ji et al., 2012; Ruiz et al., 2013), suggesting the stochastic nature of iPSC generation and/or reprogramming-related mutagensis. However, one study showed that, in one of the reprogramming experiments, all selected iPSC clones shared a set of variations from somatic cells (Young et al., 2012). This result raises the possibility that certain genetic compositions of the source cells may favor reprogramming and therefore be preferentially selected for during iPSC generation (Figure 1AII). Nonetheless, it is currently unknown how this set of variation confers a reprogramming advantage. While most of the studies indicate no specific functional enrichment among the genes displaying SNVs in iPSCs (Cheng et al., 2012; Ji et al., 2012; Ruiz et al., 2013; Young et al., 2012), one report claims that exome SNVs in hiPSCs are enriched in genes mutated in certain cancers (Gore et al., 2011). However, these SNVs are not detected at the mutated spots found in cancers, nor are they shared by multiple iPSC lines, which would argue against the idea that they are selected for their oncogenic potential. Moreover, a follow-up study examining the functionality of iPSC-manifested SNVs indicates that, in general, genes with these SNVs do not facilitate iPSC generation (Ruiz et al., 2013). Thus, it seems that most iPSC-manifested SNVs are randomly

distributed in the genome and functionally irrelevant to iPSC generation.

p53 and DNA damage response in reprogramming and iPSC genome integrity. Subchromosomal genome alteration, which may contribute to CNVs or SNVs in iPSCs, mainly results from DNA damage and unsuccessful damage repair. The tumor suppressor p53 is a key protein that mediates the DNA damage response (DDR) and guards genomic integrity. Recent studies have demonstrated a role for p53 and DDR in iPSC generation. It has been shown that γH2AX, a marker for DNA double-strand breaks (DSBs), is elevated in at least a portion of reprogramming cells, indicating that DSBs are triggered when reprogramming is initiated (Gonzalez et al., 2013; Kawamura et al., 2009). This might be due to the oncogenic properties of the reprogramming factors as well as the genome editing activities of viral transduction used in reprogramming. It is believed that the highly proliferative program initiated by the reprogramming factors may cause replication stress, which leads to genotoxicity and DDR (Hussein et al., 2013). However, cells with γH2AX foci only account for a minor population (1%–5%) of the cells undergoing reprogramming (Gonzalez et al., 2013; Marión et al., 2009). Due to the presence of p53, the cells with significant genomic lesions are unlikely to proceed to the full iPSC state. Consistently, somatic cells that are vulnerable to DNA damage, including those with a deficiency in Atm (Kinoshita et al., 2011; Marión et al., 2009) or Brca1/Brca2 (Gonzalez et al., 2013), display reduced reprogramming capacity.

Concurrent with the DDR response, the p53 pathway is activated upon the induction of reprogramming factors (Banito et al., 2009; Kawamura et al., 2009). As a result, most of the reprogramming cells undergo cell cycle arrest or apoptosis (Smith et al., 2010). Given that p53 senses and integrates diverse stress signals, activation of p53-related pathways can be due to DDR and/or other stresses initiated by the reprogramming factors. For example, oxidative stress is associated with an increase in p53 levels in reprogramming cells (Utikal et al., 2009; Yoshida et al., 2009). In contrast, ascorbate (or vitamin C), a scavenging agent for reactive oxidative species, is capable of lowering p53 levels and enhancing reprogramming efficiency (Esteban et al., 2010). Therefore, p53 presumably prevents stressed cells from proceeding to pluripotency. Consistent with this notion, when p53-deficient cells are subjected to reprogramming, cell proliferation is not checked under reprogramming-induced stresses, leading to widespread DNA lesions (Marión et al., 2009). Interestingly, despite these lesions, reprogramming is able to proceed due to the unchecked cell cycle in the p53-deficient cells, suggesting that for the establishment of pluripotency, compromised genome integrity may be tolerated to a certain extent. Consequently, increased reprogramming efficiency is observed when p53-deficient MEFs are used (Banito et al., 2009; Hanna et al., 2009; Kawamura et al., 2009; Marión et al., 2009; Utikal et al., 2009). Of note, although p53 deficiency appears to enhance reprogramming efficiency, so far, to our knowledge, there is no report for de novo p53 mutation found in iPSC lines, nor do p53-deficient iPSCs bear increased mutation loads (Gore et al., 2011). Thus, currently there is no definitive evidence for the contribution of p53 to the genetic variations detected in iPSCs.

The DNA damage signaling and repair pathways utilized in iPSCs are similar to those of ESCs, which are thought to have high efficiency and fidelity (Momčilović et al., 2011). In ESCs, the mutation rate is low, presumably because genes involved in damage signaling and repair are expressed at high levels compared to differentiated cells (Maynard et al., 2008; Momčilović et al., 2011). Furthermore, homologous recombination, a repair pathway more precise than nonhomologous end joining, is preferred for damage repair in ESCs (Serrano et al., 2011; Tichy et al., 2010). It has also been shown that, in ESCs and iPSCs, cells with DNA damage can be effectively excluded from the self-renewing pools by the induction of differentiation (Li et al., 2012b; Lin et al., 2005) or apoptosis (Aladjem et al., 1998). Therefore, it seems that after pluripotency is established, genomic integrity is effectively safeguarded in the iPSC lines.

In summary, subchromosomal variations present in iPSCs are mainly derived from source cells. These somatic genetic variations, even in low abundance, can be captured in the iPSC generation process and can be amplified in the established iPSC lines. The possible contribution of the iPSC generation process to the variations remains to be shown. Isolating iPSC-destined cells and monitoring their genomic integrity throughout the reprogramming process should shed light on whether the reprogramming process is mutagenic. Although DDR and the p53 pathway are activated by the reprogramming factors, their relevancy to the genetic variations of iPSCs remains to be shown.

Epigenetic Variations in iPSCs

The generation of iPSCs involves resetting epigenetic landscapes (Liang and Zhang, 2013). However, due to incomplete reprogramming, epigenetic variations may exist between ESCs and iPSCs and between different iPSC lines. In addition, the epigenetic status of the cells may change during prolonged culturing, which can also contribute to epigenetic variations observed in iPSCs. In this section, we will discuss these epigenetic variations.

Variations in X Chromosome Inactivation

In mice and humans, female somatic cells achieve dosage compensation through X chromosome inactivation (XCI), in which one of the two X chromosomes is inactivated. During reprogramming of mouse female somatic cells, the inactive X chromosome (Xi) is reactivated, resulting in two active X chromosomes (XaXa) in miPSCs. The reactivated X chromosome is capable of undergoing XCI, when miPSCs are induced to differentiate (Maherali et al., 2007). These observations in miPSCs are consistent with the epigenetic state of the X chromosomes in mESCs. Unlike miPSCs/mESCs, hiPSCs/hESCs are highly variable in terms of the epigenetic state of the X chromosomes (Wutz, 2012). Under conventional derivation and culturing conditions, the XaXi state from somatic cells is largely retained during the reprogramming process (Anguera et al., 2012; Cheung et al., 2011; Mekhoubad et al., 2012; Pomp et al., 2011; Tchieu et al., 2010), although XaXa have been reported in some hiPSC lines (Marchetto et al., 2010) or subpopulations of other lines (Anguera et al., 2012). Upon differentiation, XaXa hiPSCs undergo XCI, while XaXi cells keep their X chromosome status in differentiated cells. Similar to hiPSCs, the XaXi status is also predominantly detected in hESCs (Silva et al., 2008), suggesting that the XCI status may reflect the innate properties of these human PSCs or the culturing conditions that are commonly applied to them.

The variability of X chromosome inactivation status is further shown by the tendency for cells to lose the inactive status of Xi during prolonged culturing. The transcriptional repression of Xi in XaXi hiPSCs is prone to defect in late passages, as shown by loss of Xist expression and repressive chromatin modifications such as H3K27 methylation and DNA methylation (Anguera et al., 2012; Mekhoubad et al., 2012; Nazor et al., 2012; Silva et al., 2008; Tchieu et al., 2010). This results in hiPSCs with one Xa and one X chromosome with "eroded" inactivation (Xe). While some reports suggest that Xe is still transcriptionally inactive at certain examined loci (Anguera et al., 2012; Tchieu et al., 2010), others show that erosion of XCI is associated with increased gene expression (Mekhoubad et al., 2012; Nazor et al., 2012). Importantly, XaXe cells appear to have a growth advantage and gradually take over the hiPSC population. This advantage is likely due to the enhanced expression of oncogenes in XaXe hiPSC lines (Anguera et al., 2012; Mekhoubad et al., 2012; Tchieu et al., 2010). Consistent with the potential cancer-like properties, XaXe hiPSCs show inefficient differentiation when subjected to differentiation cues. In addition, the eroded state of Xe is also passed onto differentiated cells and Xe never undergoes XCI, indicating different properties between Xe and Xa (Anguera et al., 2012; Mekhoubad et al., 2012). Finally, the erosion of Xi is also observed in hESC cultures (Silva et al., 2008), again suggesting that the intrinsic properties of human PSCs and/or their shared culturing conditions might be the cause of the epigenetic variability on X chromosomes.

Variations in Local Epigenetic Status

Apart from the whole-chromosome epigenetic variability on the X chromosome, iPSCs also bear local epigenetic variations in other parts of the genome. Several studies have identified differences in epigenetic profiles between iPSC and ESC lines and between iPSC lines. While histone modifications generally show little difference (Chin et al., 2010; Guenther et al., 2010), variations in DNA methylation have been reported in multiple comparative studies (Bock et al., 2011; Lister et al., 2011; Nishino et al., 2011; Ruiz et al., 2012). The variations in DNA methylation in iPSCs can be attributed to either source cell memory or aberrant methylation generated during reprogramming. Continued passaging generally reduces the variations of DNA methylation in iPSCs, although they may persist or, in some cases, even increase with passaging (Nazor et al., 2012; Nishino et al., 2011).

Source cell memory. Due to incomplete reprogramming, hiPSC or miPSC lines may retain some of the epigenetic signature from source cells. One type of source cell memory is insufficient silencing of lineage-specific genes from the source cells. At these loci, DNA hypomethylation and/or transcription-permissive histone modifications are maintained in iPSCs similar to the source cells (Bar-Nur et al., 2011; Kim et al., 2010, 2011; Lister et al., 2011; Ohi et al., 2011; Polo et al., 2010; Ruiz et al., 2012). Genome-wide DNA methylation analyses showed that insufficient DNA methylation accounts for most memory-related differentially methylated regions (DMRs) between hiPSCs and hESCs (Lister et al., 2011; Ohi et al., 2011; Ruiz et al., 2012). In addition, source cell memory also includes chromatin constraint at gene loci that specify lineages other than source cells. These loci are hypermethylated in iPSCs compared to ESCs, rendering them incapable of being activated upon differentiation (Kim et al., 2010, 2011). Consequently, iPSCs with source cell memory have skewed differentiation potentials in favor of the source cell lineage (Bar-Nur et al., 2011; Kim et al., 2010, 2011; Polo et al., 2010).

These remnant somatic epigenetic modifications can be erased by continued passaging, cross-lineage differentiation followed by serial reprogramming, or chemical inhibition of DNA methylation and histone deacetylation (Kim et al., 2010; Polo et al., 2010). The observation that iPSCs, at least in early passages, retain portions of the source cell DNA methylation pattern is consistent with the finding that the resetting of DNA methylation patterns takes place late during the reprogramming process and may remain incomplete after the establishment of iPSC lines (Polo et al., 2012). Thus, incomplete reprogramming of epigenetic profiles contributes to the differences between iPSCs and ESCs and between iPSCs derived from different source cells.

Aberrant DNA methylation. During reprogramming, DNA methylation status may be erroneously altered, leading to iPSC-specific DNA methylation patterns distinct from those in source cells and ESCs. Depending on the loci, aberrant methylation can be specific for individual iPSC lines, or it can be common to multiple iPSC lines. The iPSC-specific differentially methylated loci include certain imprinted loci as well as other genomic regions.

Genomic imprinted loci are expressed in an allele-specific manner and are subject to regulation by allele-specific DNA methylation. The radical change in the epigenetic landscape during reprogramming raises concerns that such changes may potentially interfere with the DNA methylation status of imprinted loci. Multiple studies in hiPSCs indicate that some imprinted loci are vulnerable to epigenetic alteration during reprogramming or prolonged culturing (Chamberlain et al., 2010; Nazor et al., 2012; Nishino et al., 2011; Pick et al., 2009). Depending on specific loci, reprogramming can induce either hypermethylation or hypomethylation. Extensive passaging of hiPSCs/hESCs is also associated with aberrant methylation status at some imprinted loci, including *H19* (Nazor et al., 2012; Nishino et al., 2011). The DNA methylation changes, at least for some imprinted loci, have been correlated with loss of allele-specific expression, implicating a potential functional effect for these changes (Nazor et al., 2012; Pick et al., 2009). In addition, studies in miPSCs show that the imprinted locus *Dlk1-Dio3* frequently undergoes aberrant repression during the reprogramming process (Liu et al., 2010; Stadtfeld et al., 2010). Decreased expression at this locus has been shown to be the sole detectable difference in gene expression between miPSCs and mESCs when effects of genetic background are removed (Stadtfeld et al., 2010). Aberrant silencing of this locus in iPSCs has been functionally associated with failure to generate all-iPSC mice through tetraploid complementation (Liu et al., 2010; Stadtfeld et al., 2010). Interestingly, altering the stoichiometry of reprogramming factors (Carey et al., 2011) or including ascorbate in the reprogramming cocktail (Stadtfeld et al., 2012) can avoid aberrant silencing of the *Dlk1-Dio3* locus, suggesting that optimization of the reprogramming protocol and/or culturing conditions can prevent aberrant DNA methylation.

In addition to imprinted loci, reprogramming-induced aberrations in DNA methylation have been found in other genomic regions. Methylome profiling with whole-genome bisulphite

sequencing showed that DMRs between hESCs and hiPSCs can be detected in both CG dinucleotides and non-CG sites. For CG DMRs, hypomethylation in hiPSCs are prevalent for both memory-related and reprogramming-induced DMRs; however, all hypermethylated CG DMRs in hiPSCs are recognized as reprogramming-induced aberrancies (Lister et al., 2011). Interestingly, in another study using reduced representative bisulphite sequencing (RRBS), hypermethylation was reported as the predominant form of reprogramming-induced methylation abnormality (Ruiz et al., 2012). The discrepancy between the two studies may be from the selection of genomic regions in the latter studies and/or line-specific differences of iPSCs. With regard to non-CG DMRs, they are mainly found in Mb scale at the regions proximal to centromeres and telomeres. Methylation at these regions appears to be depleted during reprogramming because the regions are partially methylated in somatic cells and highly methylated in ESCs, but hypomenthylated in iPSCs (Lister et al., 2011). Interestingly, genomic sites enriched for H3K9me3 has been found in these large iPSC-specific hypomethylated regions, indicating that the heterochromatic state may interfere with the reprogramming events taking place in these regions (Lister et al., 2011). A recent study suggests that large-scale regional enrichment of H3K9me3 might be caused by in vitro culturing (Zhu et al., 2013). Hence, culturing stress might be partly responsible for the non-CG DNA methylation aberrancy in iPSCs. Of note, reprogramming-induced epigenetic abnormalities and their potential effect on transcription can be transmitted through differentiation and can potentially alter the properties of differentiated cells (Lister et al., 2011; Ruiz et al., 2012).

To date, it remains controversial whether iPSCs can be distinguished from ESCs in terms of their epigenetic and/or transcriptional profiles. In particular, some studies have identified hotspots or sets of genes in hiPSCs whose DNA methylation and transcription statuses are clearly different from that of hESCs. The expression statuses of these ''signature'' genes, including *FAM19A5*, *FZD10*, *TCERG1L,* and *TEME132D*, have been reported by several groups to be different in hiPSCs and hESCs (Lister et al., 2011; Nishino et al., 2011; Ruiz et al., 2012). However, another comprehensive study concluded that the variations of DNA methylation between hESCs and hiPSC lines are not greater than those between different hESC lines (Bock et al., 2011). The discrepancy between these studies is probably due to the differences in sample size, iPSC derivation methods, and methods of DNA methylation analysis. Nevertheless, for each iPSC line, cell-line-specific DNA methylation patterns have been widely reported and have to be characterized and considered when iPSCs are used for downstream functional studies and therapeutic applications.

Application Concerns

The development of the iPSC technology has made the generation of patient-specific PSCs feasible and readily accessible (Robinton and Daley, 2012). Patient-specific iPSCs can be used not only for disease modeling but also for drug screening. To realize the potential of the iPSC technology, concerns about the genetic and epigenetic variations of iPSCs have to be addressed.

One of the major concerns is whether the genetic and epigenetic variations in iPSCs change their differentiation potential. Indeed, different lines of iPSCs have been shown to have varied differentiation efficiency and developmental capacity (Feng et al., 2010; Hu et al., 2010; Kim et al., 2010; Miura et al., 2009; Polo et al., 2010; Stadtfeld et al., 2010). One of the causes for the varied differentiation capacity is source cell memory, which biases iPSC differentiation into the source cell lineage (Kim et al., 2010, 2011; Polo et al., 2010). Aberrant epigenetic statuses on certain loci (for example, the imprinted *Dlk1-Dio3* locus) can also contribute to differential developmental potentials (Liu et al., 2010; Stadtfeld et al., 2010). Furthermore, culture-adapted hiPSCs, often harboring recurrent aneuploidy, recurrent CNVs, or eroded XCI, often differentiate poorly when subjected to differentiation conditions (Ben-David and Benvenisty, 2011). In addition, genetic variations inherited from source cells or introduced during reprogramming might also affect iPSC differentiation capacity, particularly when the variations occur at developmentally important loci.

A second concern for disease modeling and cell replacement therapy is that the genetic and epigenetic variations detected in iPSCs may potentially cause unexpected phenotypic changes after differentiation of iPSCs into target cells. For disease modeling, this may lead to acquisition of disease-unrelated phenotypes or disappearance of disease-related phenotypes. A recent study utilizing female hiPSCs to model the X-linked Lesch-Nyhan syndrome (LNS) showed that, when extensively passaged patient-specific hiPSCs are used for neural differentiation, the neural disease phenotype caused by a mutated *HPRT* gene on the active X chromosome can be rescued by expression of the wild-type gene on the eroded inactive X chromosome (Mekhoubad et al., 2012). This finding calls attention to caveats associated with using high-passage female hiPSCs to model X-linked disease. For cell replacement therapy, any phenotypic abnormality is undesirable and limits the use of iPSC-derived somatic cells. It has been reported that hemangioblasts derived from hiPSCs exhibit limited proliferative capacity compared to hESC-derived hemangioblasts (Feng et al., 2010), which diminishes the clinical potential of these hiPSCs. Such a functional difference may result from the genetic or epigenetic variations in hiPSCs.

When iPSCs are derived for cell replacement therapy, it is important to evaluate the tumorigenic potential of these iPSCs and their derivatives before clinical application. First, increased tumorigenic potential has been found in culture-adapted iPSCs, which should be avoided. Second, genetic variations inherited from source cells or introduced during reprogramming may also facilitate oncogenesis if variations affect oncogene or tumor suppressor gene functions. Extensive genomic and epigenomic profiling should be performed to exclude the iPSC lines with potential oncogenic risk. Finally, an earlier study demonstrated that, for the oncogenic potential of reprogramming factors, especially that of c-Myc, leaky silencing or reactivation of transgenes would highly increase the oncogenic potential of iPSCs (Okita et al., 2007); however, this risk can be eliminated by using nonintegrating transduction methods, which now are robust enough for routine iPSC derivation.

Although iPSCs are presumed to be autologous to the donor animals, the possibility that iPSCs or iPSC-derived somatic cells might be immunogenic was recently investigated and discussed (Araki et al., 2013; Guha et al., 2013; Okita et al., 2011; Zhao

et al., 2011). So far, there is no evidence supporting the immunogenicity from iPSC-derived terminally differentiated cells, which should be the most commonly used cells for replacement therapy (Araki et al., 2013; Guha et al., 2013). However, it remains possible that immune response can be elicited by aberrant gene expression caused by genetic or epigenetic variations in iPSCs. In addition, although less studied and not the focus of this Perspective, differentiation from iPSCs to somatic cells may also introduce aberrations in the genome or epigenome. In fact, it has been shown that the in vitro differentiation process is associated with accumulation of genetic alterations (Laurent et al., 2011). The potential differentiation-induced abnormalities can cause functional consequences similar to those caused by genetic or epigenetic variations in iPSCs. Further studies are required to clarify the genomic and epigenomic stability during the in vitro differentiation process.

Coping with Genetic and Epigenetic Variations in iPSCs

At least two strategies can be used to cope with the genetic and epigenetic variations in iPSCs. One is to minimize genetic and epigenetic variations between iPSC lines; the other is to comprehensively characterize the iPSC lines and avoid using iPSC lines with potentially problematic variations.

Reduction of Variations

A logical way to minimize genetic and epigenetic variations between iPSC lines is to aim at reducing the causes of variations in source cells, reprogramming protocols, and culture conditions. With regards to source cells, ideally cells with the least accumulated genetic mutations should be used. Therefore, when selecting source cells, cells from embryonic or juvenile tissues are preferred over cells from adult or aged ones. In addition, selection of somatic cell types should also be considered. A recent study indicated that cell types with higher reprogramming efficiency correlate with fewer DNA methylation abnormalities during reprogramming (Ruiz et al., 2012), suggesting that optimizing source cell type may reduce epigenetic variability of hiPSC lines. Finally, investigations of somatic cell genetic mosaicism, which are relatively sparse, may help to identify ideal source cell types for iPSC generation.

Another factor affecting iPSC variability and safety is the reprogramming protocol utilized. Nonintegrating methods for introduction of reprogramming factors can eliminate the risk of transgene reactivation and genome editing-associated mutagenesis; hence, they should be routinely applied to derivation of iPSCs that are designated for clinical application. A deeper understanding of reprogramming mechanisms would help us develop an iPSC generation protocol that maximally protects genomic integrity and maintains epigenetic fidelity during reprogramming. Lessons can be learned from somatic cell nuclear transfer (SCNT), which has also been successfully used in human cells recently (Tachibana et al., 2013), as SCNT appears to be more efficient and presumably introduces less stress to somatic cell nucleus. Recently, oocyte factor Zsan4 has been shown to facilitate iPSC generation and reduce the amplitude of DDR in at least certain settings of reprogramming, indicating that Zsan4 may reduce the risk of genome instability during reprogramming (Hirata et al., 2012; Jiang et al., 2013). For epigenetic variations, manipulating the stoichiometry of reprogramming factors and inclusion of certain chemicals during reprogramming has been shown to reduce epigenetic aberrancy in iPSCs (Carey et al., 2011; Stadtfeld et al., 2012). Therefore, mechanistic studies in iPSC generation can provide hints for increasing the genetic and epigenetic fidelity of iPSCs.

Propagation of iPSCs after reprogramming is necessary for obtaining sufficient cell numbers for downstream studies and applications. The genetic and epigenetic variability of iPSCs can be reduced by cell passaging because this reduces certain mosaic CNVs and eliminates somatic cell memory. However, certain genetic and epigenetic alterations favoring iPSC propagation can dominate the culture after extensive passaging. Therefore, it is important to balance these two factors when considering the optimal passage number of iPSCs for downstream applications. As far as we know, few culturing conditions or manipulations have been correlated to the genetic or epigenetic variability detected in iPSCs. One exception is that the XaXa epigenetic status of X chromosomes in hiPSCs is linked to coculture of LIF-secreting feeder cells (Tomoda et al., 2012). Nevertheless, stress reduction in culture is a common theme for maintaining homeostasis of in vitro cultured PSCs, and usage of chemically defined medium should help eliminate culture variations that may influence the status of iPSCs. Finally, lab-specific variability between iPSC lines (Newman and Cooper, 2010) can be dismissed if systematic evaluation with multiple iPSC lines is carried out simultaneously within the same settings.

Although it is feasible to reduce genetic and epigenetic variations by limiting their sources, having a certain level of genetic and epigenetic variability is a unanimous feature of all biological systems, including PSCs. Therefore, rather than trying to derive iPSC lines completely free of alterations, it is more practical to obtain iPSC lines that can be used for application purposes but may contain trivial variations that are functionally negligible. Alternatively, other approaches that do not involve reprogramming to the pluripotent stage—for example, direct lineage conversion—should be explored and compared to the iPSC-based therapeutic strategy.

Detection and Monitoring of Variations

Detecting adverse genetic and epigenetic variations in iPSCs and monitoring iPSCs throughout passages enables us to avoid using potentially problematic cell lines for downstream applications. It also helps to maintain homeostasis of individual iPSC lines during passaging. Depending on the requirements of specific downstream applications, a wide range of analyses can be performed. Basic characterizations, such as karyotyping and expression analysis of pluripotent markers, should be performed routinely for iPSCs during their passaging. Functional analyses, including embryoid body formation or teratoma analysis, can be used to assess the differentiation and oncogenic potential of iPSCs and should be performed if these properties are critical for downstream applications. The iPSC lines destined for therapeutic application should be characterized more extensively. Genome-wide sequencing, expression analysis, and DNA and histone modification analysis have been used for comprehensive genetic and epigenetic profiling of iPSC lines. More cost-effectively, locus-specific analysis may be performed at developmentally important genes, cancer-related genes, and altered hotspots (e.g., aberrant methylation hotspots or X chromosomes). In the long run, to satisfy the need for therapeutic replacement of different tissue and cell types, it would be ideal

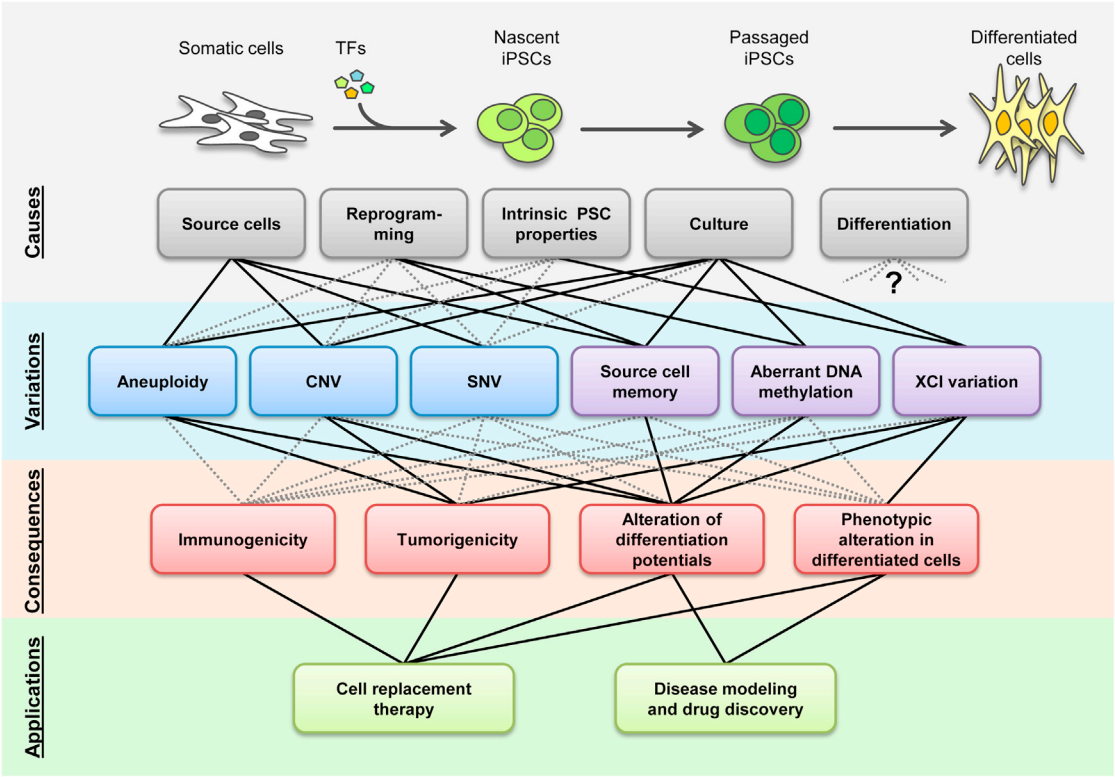

Figure 2. Genetic and Epigenetic Variations and Their Causes, Functional Consequences, and Impacts on Applications
iPSCs derived from transcription factor (TF)-mediated reprogramming may bear different types of genetic (blue boxes) or epigenetic (purple boxes) variations that can be introduced from varied sources (gray boxes) during the derivation and manipulations of iPSCs. These variations may lead to different functional consequences (red boxes) that need to be considered when iPSCs or their derivatives are used for applications (green boxes). Solid lines, reported or definite connections; dotted lines, potential connections.

to set up a bank of hiPSC and hESC lines from various MHC types, and each line should contain detailed genetic and epigenetic profiles and differentiation potential "score cards" (Bock et al., 2011; Taylor et al., 2012).

Conclusion

Genetic and epigenetic variations in iPSCs come from different sources. Some of the variations may be inherited from donor somatic cells, induced or selected by the reprogramming process, or accumulated during culturing; others may simply reflect the innate genetic and epigenetic stability of the pluripotent state of iPSCs (Figure 2). Although each variation is not relevant to the functionality of iPSCs, certain variations may change the properties of iPSCs and their derivatives. For example, the variations may alter the differentiation potential of iPSCs, cause phenotypic changes in iPSC-derived somatic cells, or increase the tumorigenicity or immunogenicity of iPSCs and their derivatives. These adverse changes directly affect the utility of iPSCs (Figure 2). Optimizing the reprogramming strategy and culture conditions helps reduce the occurrence of variations. Comprehensive characterization and rigorous monitoring of genome and epigenome integrity can ensure the quality of iPSCs designated for downstream applications. Additional studies on the mechanism of iPSC generation will further reveal the ontology of genetic and epigenetic variations and will provide better solutions for overcoming the limitations caused by these variations.

ACKNOWLEDGMENTS

We would like to thank Diana Cai for critical reading of this manuscript. We apologize to the people whose work cannot be cited due to space limitations. Stem-cell-related work in our lab is supported by NIH (U01DK089565) and HHMI. Y.Z. is an Investigator of the HHMI.

REFERENCES

Abyzov, A., Mariani, J., Palejev, D., Zhang, Y., Haney, M.S., Tomasini, L., Ferrandino, A.F., Rosenberg Belmaker, L.A., Szekely, A., Wilson, M., et al. (2012). Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. Nature *492*, 438–442.

Aladjem, M.I., Spike, B.T., Rodewald, L.W., Hope, T.J., Klemm, M., Jaenisch, R., and Wahl, G.M. (1998). ES cells do not activate p53-dependent stress responses and undergo p53-independent apoptosis in response to DNA damage. Curr. Biol. *8*, 145–155.

Amps, K., Andrews, P.W., Anyfantis, G., Armstrong, L., Avery, S., Baharvand, H., Baker, J., Baker, D., Munoz, M.B., Beil, S., et al.; International Stem Cell Initiative. (2011). Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. Nat. Biotechnol. *29*, 1132–1144.

Anguera, M.C., Sadreyev, R., Zhang, Z., Szanto, A., Payer, B., Sheridan, S.D., Kwok, S., Haggarty, S.J., Sur, M., Alvarez, J., et al. (2012). Molecular signatures of human induced pluripotent stem cells highlight sex differences and cancer genes. Cell Stem Cell *11*, 75–90.

Araki, R., Uda, M., Hoki, Y., Sunayama, M., Nakamura, M., Ando, S., Sugiura, M., Ideno, H., Shimada, A., Nifuji, A., and Abe, M. (2013). Negligible immunogenicity of terminally differentiated cells derived from induced pluripotent or embryonic stem cells. Nature *494*, 100–104.

Baker, D.E., Harrison, N.J., Maltby, E., Smith, K., Moore, H.D., Shaw, P.J., Heath, P.R., Holden, H., and Andrews, P.W. (2007). Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. Nat. Biotechnol. *25*, 207–215.

Banito, A., Rashid, S.T., Acosta, J.C., Li, S., Pereira, C.F., Geti, I., Pinho, S., Silva, J.C., Azuara, V., Walsh, M., et al. (2009). Senescence impairs successful reprogramming to pluripotent stem cells. Genes Dev. *23*, 2134–2139.

Bar-Nur, O., Russ, H.A., Efrat, S., and Benvenisty, N. (2011). Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. Cell Stem Cell *9*, 17–23.

Ben-David, U., and Benvenisty, N. (2011). The tumorigenicity of human embryonic and induced pluripotent stem cells. Nat. Rev. Cancer *11*, 268–277.

Ben-David, U., and Benvenisty, N. (2012). High prevalence of evolutionarily conserved and species-specific genomic aberrations in mouse pluripotent stem cells. Stem Cells *30*, 612–622.

Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., Ziller, M., Croft, G.F., Amoroso, M.W., Oakley, D.H., et al. (2011). Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell *144*, 439–452.

Carey, B.W., Markoulaki, S., Hanna, J.H., Faddah, D.A., Buganim, Y., Kim, J., Ganz, K., Steine, E.J., Cassady, J.P., Creyghton, M.P., et al. (2011). Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. Cell Stem Cell *9*, 588–598.

Chamberlain, S.J., Chen, P.F., Ng, K.Y., Bourgois-Rocha, F., Lemtiri-Chlieh, F., Levine, E.S., and Lalande, M. (2010). Induced pluripotent stem cell models of the genomic imprinting disorders Angelman and Prader-Willi syndromes. Proc. Natl. Acad. Sci. USA *107*, 17668–17673.

Cheng, L., Hansen, N.F., Zhao, L., Du, Y., Zou, C., Donovan, F.X., Chou, B.K., Zhou, G., Li, S., Dowey, S.N., et al.; NISC Comparative Sequencing Program. (2012). Low incidence of DNA sequence variation in human induced pluripotent stem cells generated by nonintegrating plasmid expression. Cell Stem Cell *10*, 337–344.

Cheung, A.Y., Horvath, L.M., Grafodatskaya, D., Pasceri, P., Weksberg, R., Hotta, A., Carrel, L., and Ellis, J. (2011). Isolation of MECP2-null Rett Syndrome patient hiPS cells and isogenic controls through X-chromosome inactivation. Hum. Mol. Genet. *20*, 2103–2115.

Chin, M.H., Pellegrini, M., Plath, K., and Lowry, W.E. (2010). Molecular analyses of human induced pluripotent stem cells and embryonic stem cells. Cell Stem Cell *7*, 263–269.

Draper, J.S., Smith, K., Gokhale, P., Moore, H.D., Maltby, E., Johnson, J., Meisner, L., Zwaka, T.P., Thomson, J.A., and Andrews, P.W. (2004). Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. Nat. Biotechnol. *22*, 53–54.

Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S., et al. (2010). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. Cell Stem Cell *6*, 71–79.

Feng, Q., Lu, S.J., Klimanskaya, I., Gomes, I., Kim, D., Chung, Y., Honig, G.R., Kim, K.S., and Lanza, R. (2010). Hemangioblastic derivatives from human induced pluripotent stem cells exhibit limited expansion and early senescence. Stem Cells *28*, 704–712.

Gonzalez, F., Georgieva, D., Vanoli, F., Shi, Z.D., Stadtfeld, M., Ludwig, T., Jasin, M., and Huangfu, D. (2013). Homologous Recombination DNA Repair Genes Play a Critical Role in Reprogramming to a Pluripotent State. Cell reports.

Gore, A., Li, Z., Fung, H.L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E., et al. (2011). Somatic coding mutations in human induced pluripotent stem cells. Nature *471*, 63–67.

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.

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.

Hamada, M., Malureanu, L.A., Wijshake, T., Zhou, W., and van Deursen, J.M. (2012). Reprogramming to pluripotency can conceal somatic cell chromosomal instability. PLoS Genet. *8*, e1002913.

Hanna, J., Saha, K., Pando, B., van Zon, J., Lengner, C.J., Creyghton, M.P., van Oudenaarden, A., and Jaenisch, R. (2009). Direct cell reprogramming is a stochastic process amenable to acceleration. Nature *462*, 595–601.

Hirata, T., Amano, T., Nakatake, Y., Amano, M., Piao, Y., Hoang, H.G., and Ko, M.S. (2012). Zscan4 transiently reactivates early embryonic genes during the generation of induced pluripotent stem cells. Scientific Reports *2*, 208.

Hu, B.Y., Weick, J.P., Yu, J., Ma, L.X., Zhang, X.Q., Thomson, J.A., and Zhang, S.C. (2010). Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. Proc. Natl. Acad. Sci. USA *107*, 4335–4340.

Hussein, S.M., Batada, N.N., Vuoristo, S., Ching, R.W., Autio, R., Närvä, E., Ng, S., Sourour, M., Hämäläinen, R., Olsson, C., et al. (2011). Copy number variation and selection during reprogramming to pluripotency. Nature *471*, 58–62.

Hussein, S.M., Elbaz, J., and Nagy, A.A. (2013). Genome damage in induced pluripotent stem cells: assessing the mechanisms and their consequences. BioEssays *35*, 152–162.

Ji, J., Ng, S.H., Sharma, V., Neculai, D., Hussein, S., Sam, M., Trinh, Q., Church, G.M., McPherson, J.D., Nagy, A., and Batada, N.N. (2012). Elevated coding mutation rate during the reprogramming of human somatic cells into induced pluripotent stem cells. Stem Cells *30*, 435–440.

Jiang, J., Lv, W., Ye, X., Wang, L., Zhang, M., Yang, H., Okuka, M., Zhou, C., Zhang, X., Liu, L., and Li, J. (2013). Zscan4 promotes genomic stability during reprogramming and dramatically improves the quality of iPS cells as demonstrated by tetraploid complementation. Cell Res. *23*, 92–106.

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

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I., et al. (2010). Epigenetic memory in induced pluripotent stem cells. Nature *467*, 285–290.

Kim, K., Zhao, R., Doi, A., Ng, K., Unternaehrer, J., Cahan, P., Huo, H., Loh, Y.H., Aryee, M.J., Lensch, M.W., et al. (2011). Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. Nat. Biotechnol. *29*, 1117–1119.

Kinoshita, T., Nagamatsu, G., Kosaka, T., Takubo, K., Hotta, A., Ellis, J., and Suda, T. (2011). Ataxia-telangiectasia mutated (ATM) deficiency decreases reprogramming efficiency and leads to genomic instability in iPS cells. Biochem. Biophys. Res. Commun. *407*, 321–326.

Laurent, L.C., Ulitsky, I., Slavin, I., Tran, H., Schork, A., Morey, R., Lynch, C., Harness, J.V., Lee, S., Barrero, M.J., et al. (2011). Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. Cell Stem Cell *8*, 106–118.

Li, L.B., Chang, K.H., Wang, P.R., Hirata, R.K., Papayannopoulou, T., and Russell, D.W. (2012a). Trisomy correction in Down syndrome induced pluripotent stem cells. Cell Stem Cell *11*, 615–619.

Li, M., He, Y., Dubois, W., Wu, X., Shi, J., and Huang, J. (2012b). Distinct regulatory mechanisms and functions for p53-activated and p53-repressed DNA damage response genes in embryonic stem cells. Mol. Cell *46*, 30–42.

Liang, G., and Zhang, Y. (2013). Embryonic stem cell and induced pluripotent stem cell: an epigenetic perspective. Cell Res. *23*, 49–69.

Lin, T., Chao, C., Saito, S., Mazur, S.J., Murphy, M.E., Appella, E., and Xu, Y. (2005). p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. Nat. Cell Biol. *7*, 165–171.

Lister, R., Pelizzola, M., Kida, Y.S., Hawkins, R.D., Nery, J.R., Hon, G., Antosiewicz-Bourget, J., O'Malley, R., Castanon, R., Klugman, S., et al. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature *471*, 68–73.

Liu, L., Luo, G.Z., Yang, W., Zhao, X., Zheng, Q., Lv, Z., Li, W., Wu, H.J., Wang, L., Wang, X.J., and Zhou, Q. (2010). Activation of the imprinted Dlk1-Dio3

region correlates with pluripotency levels of mouse stem cells. J. Biol. Chem. *285*, 19483–19490.

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

Marchetto, M.C., Carromeu, C., Acab, A., Yu, D., Yeo, G.W., Mu, Y., Chen, G., Gage, F.H., and Muotri, A.R. (2010). A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. Cell *143*, 527–539.

Marión, R.M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., Fernandez-Capetillo, O., Serrano, M., and Blasco, M.A. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. Nature *460*, 1149–1153.

Martins-Taylor, K., Nisler, B.S., Taapken, S.M., Compton, T., Crandall, L., Montgomery, K.D., Lalande, M., and Xu, R.H. (2011). Recurrent copy number variations in human induced pluripotent stem cells. Nat. Biotechnol. *29*, 488–491.

Maynard, S., Swistowska, A.M., Lee, J.W., Liu, Y., Liu, S.T., Da Cruz, A.B., Rao, M., de Souza-Pinto, N.C., Zeng, X., and Bohr, V.A. (2008). Human embryonic stem cells have enhanced repair of multiple forms of DNA damage. Stem Cells *26*, 2266–2274.

Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J.C., Yakir, B., Clark, A.T., Plath, K., Lowry, W.E., and Benvenisty, N. (2010). Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. Cell Stem Cell *7*, 521–531.

Mekhoubad, S., Bock, C., de Boer, A.S., Kiskinis, E., Meissner, A., and Eggan, K. (2012). Erosion of dosage compensation impacts human iPSC disease modeling. Cell Stem Cell *10*, 595–609.

Miura, K., Okada, Y., Aoi, T., Okada, A., Takahashi, K., Okita, K., Nakagawa, M., Koyanagi, M., Tanabe, K., Ohnuki, M., et al. (2009). Variation in the safety of induced pluripotent stem cell lines. Nat. Biotechnol. *27*, 743–745.

Momčilović, O., Navara, C., and Schatten, G. (2011). Cell cycle adaptations and maintenance of genomic integrity in embryonic stem cells and induced pluripotent stem cells. Results Probl. Cell Differ. *53*, 415–458.

Nazor, K.L., Altun, G., Lynch, C., Tran, H., Harness, J.V., Slavin, I., Garitaonandia, I., Müller, F.J., Wang, Y.C., Boscolo, F.S., et al. (2012). Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. Cell Stem Cell *10*, 620–634.

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

Nishino, K., Toyoda, M., Yamazaki-Inoue, M., Fukawatase, Y., Chikazawa, E., Sakaguchi, H., Akutsu, H., and Umezawa, A. (2011). DNA methylation dynamics in human induced pluripotent stem cells over time. PLoS Genet. *7*, e1002085.

Ohi, Y., Qin, H., Hong, C., Blouin, L., Polo, J.M., Guo, T., Qi, Z., Downey, S.L., Manos, P.D., Rossi, D.J., et al. (2011). Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. Nat. Cell Biol. *13*, 541–549.

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

Okita, K., Nagata, N., and Yamanaka, S. (2011). Immunogenicity of induced pluripotent stem cells. Circ. Res. *109*, 720–721.

Park, I.H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M.W., Cowan, C., Hochedlinger, K., and Daley, G.Q. (2008). Disease-specific induced pluripotent stem cells. Cell *134*, 877–886.

Pick, M., Stelzer, Y., Bar-Nur, O., Mayshar, Y., Eden, A., and Benvenisty, N. (2009). Clone- and gene-specific aberrations of parental imprinting in human induced pluripotent stem cells. Stem Cells *27*, 2686–2690.

Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., Apostolou, E., Stadtfeld, M., Li, Y., Shioda, T., et al. (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. Nat. Biotechnol. *28*, 848–855.

Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J., et al. (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. Cell *151*, 1617–1632.

Pomp, O., Dreesen, O., Leong, D.F., Meller-Pomp, O., Tan, T.T., Zhou, F., and Colman, A. (2011). Unexpected X chromosome skewing during culture and reprogramming of human somatic cells can be alleviated by exogenous telomerase. Cell Stem Cell *9*, 156–165.

Quinlan, A.R., Boland, M.J., Leibowitz, M.L., Shumilina, S., Pehrson, S.M., Baldwin, K.K., and Hall, I.M. (2011). Genome sequencing of mouse induced pluripotent stem cells reveals retroelement stability and infrequent DNA rearrangement during reprogramming. Cell Stem Cell *9*, 366–373.

Robinton, D.A., and Daley, G.Q. (2012). The promise of induced pluripotent stem cells in research and therapy. Nature *481*, 295–305.

Ruiz, S., Diep, D., Gore, A., Panopoulos, A.D., Montserrat, N., Plongthongkum, N., Kumar, S., Fung, H.L., Giorgetti, A., Bilic, J., et al. (2012). Identification of a specific reprogramming-associated epigenetic signature in human induced pluripotent stem cells. Proc. Natl. Acad. Sci. USA *109*, 16196–16201.

Ruiz, S., Gore, A., Li, Z., Panopoulos, A.D., Montserrat, N., Fung, H.L., Giorgetti, A., Bilic, J., Batchelder, E.M., Zaehres, H., et al. (2013). Analysis of protein-coding mutations in hiPSCs and their possible role during somatic cell reprogramming. Nat. Comm. *4*, 1382.

Serrano, L., Liang, L., Chang, Y., Deng, L., Maulion, C., Nguyen, S., and Tischfield, J.A. (2011). Homologous recombination conserves DNA sequence integrity throughout the cell cycle in embryonic stem cells. Stem Cells Dev. *20*, 363–374.

Silva, S.S., Rowntree, R.K., Mekhoubad, S., and Lee, J.T. (2008). X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. Proc. Natl. Acad. Sci. USA *105*, 4820–4825.

Smith, Z.D., Nachman, I., Regev, A., and Meissner, A. (2010). Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. Nat. Biotechnol. *28*, 521–526.

Stadtfeld, M., and Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. Genes Dev. *24*, 2239–2263.

Stadtfeld, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., Kono, T., Shioda, T., and Hochedlinger, K. (2010). Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. Nature *465*, 175–181.

Stadtfeld, M., Apostolou, E., Ferrari, F., Choi, J., Walsh, R.M., Chen, T., Ooi, S.S., Kim, S.Y., Bestor, T.H., Shioda, T., et al. (2012). Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. Nat. Genet. *44*, 398–405, S391–392.

Taapken, S.M., Nisler, B.S., Newton, M.A., Sampsell-Barron, T.L., Leonhard, K.A., McIntire, E.M., and Montgomery, K.D. (2011). Karotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. Nat. Biotechnol. *29*, 313–314.

Tachibana, M., Amato, P., Sparman, M., Gutierrez, N.M., Tippner-Hedges, R., Ma, H., Kang, E., Fulati, A., Lee, H.S., Sritanaudomchai, H., et al. (2013). Human embryonic stem cells derived by somatic cell nuclear transfer. Cell *153*, 1228–1238.

Taylor, C.J., Peacock, S., Chaudhry, A.N., Bradley, J.A., and Bolton, E.M. (2012). Generating an iPSC bank for HLA-matched tissue transplantation based on known donor and recipient HLA types. Cell Stem Cell *11*, 147–152.

Tchieu, J., Kuoy, E., Chin, M.H., Trinh, H., Patterson, M., Sherman, S.P., Aimiuwu, O., Lindgren, A., Hakimian, S., Zack, J.A., et al. (2010). Female human iPSCs retain an inactive X chromosome. Cell Stem Cell *7*, 329–342.

Tichy, E.D., Pillai, R., Deng, L., Liang, L., Tischfield, J., Schwemberger, S.J., Babcock, G.F., and Stambrook, P.J. (2010). Mouse embryonic stem cells, but not somatic cells, predominantly use homologous recombination to repair double-strand DNA breaks. Stem Cells Dev. *19*, 1699–1711.

Tomoda, K., Takahashi, K., Leung, K., Okada, A., Narita, M., Yamada, N.A., Eilertson, K.E., Tsang, P., Baba, S., White, M.P., et al. (2012). Derivation

conditions impact X-inactivation status in female human induced pluripotent stem cells. Cell Stem Cell *11*, 91–99.

Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., Khalil, A., Rheinwald, J.G., and Hochedlinger, K. (2009). Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. Nature *460*, 1145–1148.

Wutz, A. (2012). Epigenetic alterations in human pluripotent stem cells: a tale of two cultures. Cell Stem Cell *11*, 9–15.

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.

Young, M.A., Larson, D.E., Sun, C.W., George, D.R., Ding, L., Miller, C.A., Lin, L., Pawlik, K.M., Chen, K., Fan, X., et al. (2012). Background mutations in parental cells account for most of the genetic heterogeneity of induced pluripotent stem cells. Cell Stem Cell *10*, 570–582.

Zhao, T., Zhang, Z.N., Rong, Z., and Xu, Y. (2011). Immunogenicity of induced pluripotent stem cells. Nature *474*, 212–215.

Zhu, J., Adli, M., Zou, J.Y., Verstappen, G., Coyne, M., Zhang, X., Durham, T., Miri, M., Deshpande, V., De Jager, P.L., et al. (2013). Genome-wide chromatin state transitions associated with developmental and environmental cues. Cell *152*, 642–654.