Cell Death and Differentiation (2011) 18, 745–753

www.nature.com/cdd

Genomic instability in induced stem cells

CE Pasi[1,8], A Dereli-Öz[2,8], S Negrini[2,8], M Friedli[3], G Fragola[1,4], A Lombardo[5], G Van Houwe[2], L Naldini[5], S Casola[4], G Testa[1], D Trono[3], PG Pelicci*[,1,6] and TD Halazonetis*[,2,7]

The ability to reprogram adult cells into stem cells has raised hopes for novel therapies for many human diseases. Typical stem cell reprogramming protocols involve expression of a small number of genes in differentiated somatic cells with the *c-Myc* and *Klf4* proto-oncogenes typically included in this mix. We have previously shown that expression of oncogenes leads to DNA replication stress and genomic instability, explaining the high frequency of *p53* mutations in human cancers. Consequently, we wondered whether stem cell reprogramming also leads to genomic instability. To test this hypothesis, we examined stem cells induced by a variety of protocols. The first protocol, developed specifically for this study, reprogrammed primary mouse mammary cells into mammary stem cells by expressing *c-Myc*. Two other previously established protocols reprogrammed mouse embryo fibroblasts into induced pluripotent stem cells by expressing either three genes, *Oct4*, *Sox2* and *Klf4*, or four genes, OSK plus *c-Myc*. Comparative genomic hybridization analysis of stem cells derived by these protocols revealed the presence of genomic deletions and amplifications, whose signature was suggestive of oncogene-induced DNA replication stress. The genomic aberrations were to a significant degree dependent on *c-Myc* expression and their presence could explain why *p53* inactivation facilitates stem cell reprogramming.
Cell Death and Differentiation (2011) **18,** 745–753; doi:10.1038/cdd.2011.9; published online 11 February 2011

Recent studies have demonstrated that it is possible to reprogram somatic cells into pluripotent stem cells by expressing a specific combination of transcription factors.[1–5] The typical mix of transcription factors used for this purpose includes the oncogenes *c-Myc* and *Klf4*. Interestingly, oncogenes have the potential to induce genomic instability,[6] which raises the possibility that induced pluripotent stem (iPS) cells have aberrant genomes.

Given the excitement in the stem cell field by the ability to transform adult differentiated cells into pluripotent stem cells, it is not surprising that research has focused mainly on methods that enhance reprogramming efficiency, with less attention been paid to the genomic status of the generated iPS cells.[7–9] However, the presence of genomic aberrations in iPS cells could be one of the reasons why the efficiency with which these cells produce live mice in tetraploid complementation assays is very low.[10]

The possibility that aberrant genomes are prevalent in iPS cells became more likely after the demonstration that *p53* inactivation facilitates reprogramming.[11–14] One study suggested that the absence of *p53* enhances reprogramming, because it enhances cell proliferation.[15] However, *p53* is a DNA damage response gene,[16] raising the possibility that reprogramming is accompanied by DNA damage and genomic instability. Consistent with this interpretation, iPS cells generated from mouse embryo fibroblasts (MEFs) lacking *p53* function, form malignant tumors when injected in donor mice.[17]

In this study, we examined directly the genomes of induced stem cells by array-based comparative genomic hybridization (cGH) analysis. Using three different experimental systems, we report genomic aberrations in induced stem cells. These aberrations were associated with oncogene-induced DNA replication stress.

Results

Genomic instability in induced mammary stem cells. As a first step in exploring whether stem cell reprogramming is

[1]Department of Experimental Oncology at the IFOM-IEO Campus, Istituto Europeo di Oncologia, Milan, Italy; [2]Department of Molecular Biology, University of Geneva, Geneva, Switzerland; [3]School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland; [4]IFOM, FIRC Institute of Molecular Oncology Foundation, Milan, Italy; [5]San Raffaele Telethon Institute for Gene Therapy, Vita Salute San Raffaele University, Milan, Italy; [6]Dipartimento di Medicina, Chirurgia ed Odontoiatria, Universita degli Studi di Milano, Milan, Italy and [7]Department of Biochemistry, University of Geneva, Geneva, Switzerland
*Corresponding authors: PG Pelicci, Istituto Europeo di Oncologia (IEO), Department of Experimental Oncology at the IFOM-IEO Campus, Via Adamello 16, 20139 Milan, Italy. Tel: + 39 02 5748 9831; Fax: + 39 02 5748 9851; E-mail: piergiuseppe.pelicci@ifom-ieo-campus.it
or TD Halazonetis, Department of Molecular Biology, University of Geneva, quai Ernest-Ansermet 30, 1205 Geneva, Switzerland. Tel: + 41 22 379 6112; Fax: + 41 22 379 68 68; E-mail: thanos.halazonetis@unige.ch
[8]These authors contributed equally to this work.
Keywords: induced pluripotent stem cells; genomic instability; c-Myc
Abbreviations: iPS, induced pluripotent stem; OSK, *Oct4*, *Sox2*, *Klf4*; OSKC, *Oct4*, *Sox2*, *Klf4*, *c-Myc*; TAM, 4-hydroxytamoxifen; MEF, mouse embryo fibroblast; CFS, common fragile site; cGH, comparative genomic hybridization; CNC, copy number change; *Rora*, retinoic acid receptor-related orphan receptor A; *Pde4D*, phosphodiesterase 4D; *Ptprg*, protein tyrosine phosphatase receptor type G; *Fhit*, fragile histidine triad; *Jarid2*, jumonji AT rich interactive domain 2; *Dtnbp1*, dystrobrevin-binding protein 1; *Odz3*, odd Oz/ten-m homolog 3; *Cdh13*, cadherin 13; *Fto*, fat mass and obesity associated; *Dock3*, dedicator of cytokinesis 3; *Cadm2*, cell adhesion molecule 2; *Lsamp*, limbic system-associated membrane protein; *Naaladl2*, *N*-acetylated alpha-linked acidic dipeptidase-like 2; *Erc2*, ELKS–RAB6-interacting/CAST family member 2
Received 10.1.11; revised 13.1.11; accepted 13.1.11; Edited by G Melino; published online 11.2.11

accompanied by genomic instability, we developed a protocol for inducing mammary stem cells. The protocol involves expressing *c-Myc*, one of the four original stem cell reprogramming genes, in mouse mammary cells and examining whether these cells acquire stem cell properties, such as the ability to form mammospheres *in vitro* and to repopulate cleared mouse fat pads.[18]

We first attempted to reprogram nearly homogeneous populations of mammary progenitor cells. These cells, chosen because they lack stem cell properties, were obtained using a previously described PKH26-based label-retaining protocol.[19] Briefly, primary mammary cells were pulse-labeled with the lipophilic, fluorescent-dye PKH26 and then cultured as mammospheres for two passages. At this time, PKH26-high cells (about 0.3% of all cells, representing stem cells) and PKH26-negative cells (about 30% of all cells, representing progenitor cells) were isolated by flow sorting.[19] The PKH26-negative cells were then infected with a control lentivirus or a lentivirus-expressing MycER, a c-Myc protein containing a modified estrogen receptor hormone-binding domain at its C-terminus.[20] Five thousand infected cells were cultured under non-adherent conditions to generate mammospheres, which were then passaged on a weekly basis. The control-infected progenitor cells formed mammospheres with very low efficiency even at the first passage and both the mammosphere number and the cumulative cell number declined to practically zero within a few passages (Figures 1a–c). In contrast, the MycER-expressing progenitor cells were reprogrammed into mammary stem cells, as ascertained by their ability to form mammospheres and repopulate a cleared mouse fat pad (Figure 1).

Having established that c-Myc can reprogram mammary progenitor cells into stem cells, we repeated these experiments bypassing the PKH26-sorting step. Control virus-infected mammospheres prepared from wild-type mice could be maintained only for a few passages in tissue culture, suggesting exhaustion of the stem cell population. In contrast, mammospheres infected with the virus-expressing MycER could be easily expanded with no signs of crisis or stem cell exhaustion and could also repopulate a cleared mouse fat pad (Figure 2).

The reprogramming of mammary progenitor cells into mammary stem cells described above was performed in the absence of 4-hydroxytamoxifen (TAM). When TAM was added to the media, nuclear MycER protein levels increased (Supplementary Figure 1a) and the MycER-expressing mammosphere cultures were exhausted within one passage (Figure 2a). This was associated with phosphorylation and stabilization of p53, expression of p53 target genes and apoptosis (Supplementary Figure 1b), suggesting that high levels of MycER protein induced a p53-dependent DNA damage response.[16,21–23] Consistent with this scenario, MycER-expressing mammospheres prepared from p53−/− mice could be expanded in the presence of TAM (Figure 2a). During stem cell reprogramming, MycER levels were lower and apparently insufficient to induce an overt DNA damage response (Supplementary Figure 1b), but whether they were low enough to prevent induction of genomic instability was not clear.

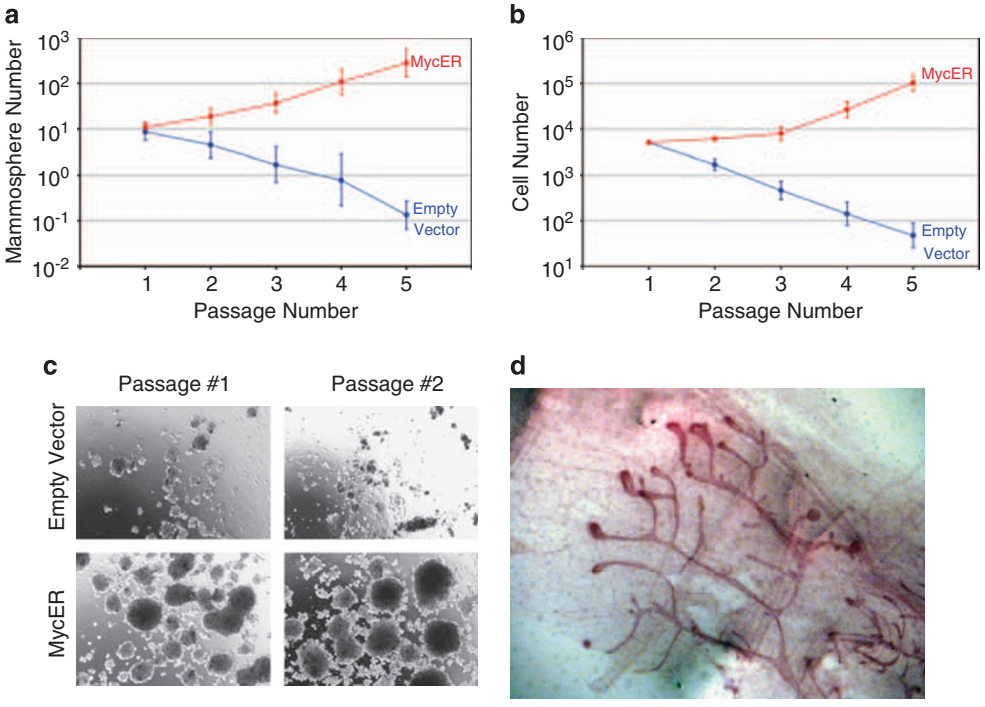


Figure 1 Reprogramming of progenitor mammary cells to stem cells by MycER. (**a** and **b**) PKH26-negative cells, which are devoid of stem cells, were isolated from secondary mammospheres and infected with a lentivirus-expressing MycER or with a control virus (Empty Vector). The cells were cultured under non-adherent conditions and mammosphere number (**a**) and cumulative cell number (**b**) were determined at each passage. Results are presented as means ± 1 S.D. (**c**) Representative images of the cultures at passages 1 and 2. (**d**) Repopulation of a cleared mouse fat pad by MycER-expressing mammospheres originally prepared from PKH26-negative cells with no evidence of tumor development

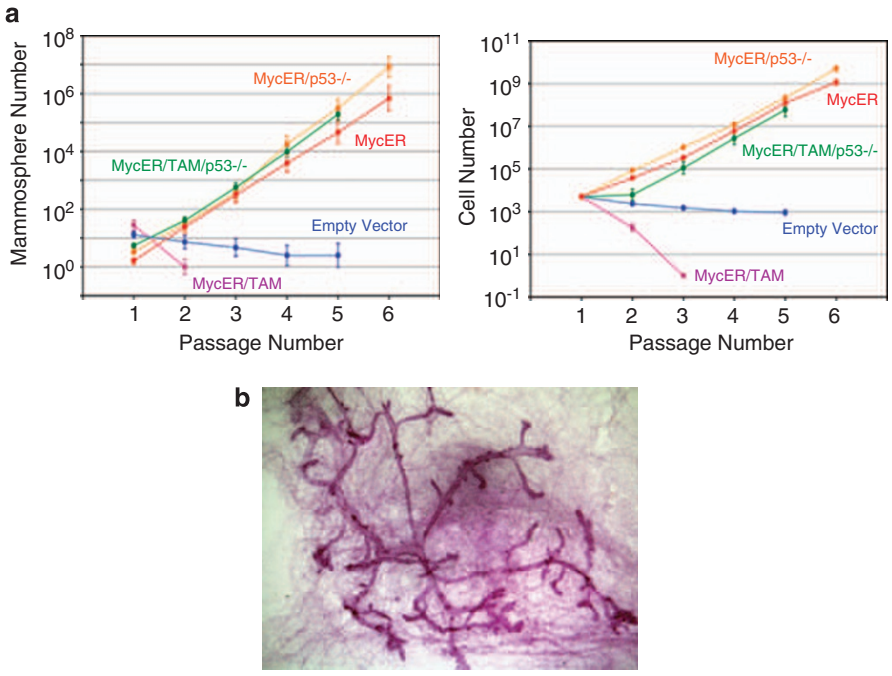


Figure 2 Stem cell reprogramming of primary mammary cells by MycER. (**a**) Primary mammary cells prepared from wild-type or *p53*−/− mice were infected with a lentivirus-expressing MycER or with a control virus (Empty Vector) and cultured under non-adherent conditions. In some cultures TAM was added to induce high levels of MycER activity. The number of mammospheres and the cumulative number of cells were determined at each passage. Results are presented as means ± 1 S.D. (**b**) Repopulation of a cleared mouse fat pad by MycER-expressing mammospheres prepared by lentiviral infection of primary mammary cells. There are no signs of tumor development in the replanted mammary gland

To address the issue of genomic instability, mammary cells, reprogrammed into stem cells by infecting them with the lentivirus-expressing MycER, were passaged for 9 weeks in the absence of tamoxifen. Then, the cells were serially diluted in 48-well plates, to obtain single stem cell clones, which were expanded for 3–6 weeks, again in the absence of tamoxifen, before preparing genomic DNA (Figure 3a). Control genomic DNA was prepared from non-infected primary mammospheres (passage 5). Eight randomly selected reprogrammed stem cell clones were subjected to cGH analysis using high-density arrays covering chromosomes 10–13 in their entirety and part of chromosomes 9 and 14, corresponding in total to a quarter of the mouse genome.

Four out of the eight examined clones had focal copy number changes (CNCs) (Figure 3b). The first clone had a deletion of about 100 kbp within the retinoic acid receptor-related orphan receptor A (*Rora*) gene in chromosome 9. In humans, the *RORA* gene maps to the common fragile site (CFS) FRA15A.[24,25] The second reprogrammed clone had a deletion of about 250 kbp within the phosphodiesterase 4D (*Pde4D*) gene in chromosome 13. Like *Rora*, *Pde4D* is a very large gene. The third clone had a deletion of about 200 kbp within the protein tyrosine phosphatase receptor type G (*Ptprg*) gene in chromosome 14. In humans, *PTPRG* maps within the CFS FRA3B,[25] although most deletions targeting FRA3B involve the adjacent fragile histidine triad (*FHIT*) gene. The fourth clone had an amplification of the jumonji AT rich interactive domain 2 (*Jarid2*) and dystrobrevin-binding protein 1 (*Dtnbp1*) genes in chromosome 13. The protein product of *Jarid2* associates with the Polycomb repressive complex 2 and regulates the self-renewal of embryonic stem cells,[26–28] suggesting that the amplification of *Jarid2* may have been selected in this clone. The observed frequency of genomic aberrations within just a quarter of the mouse genome indicates that reacquisition of stemness features in the reprogramming protocol is associated with the occurrence of genomic rearrangements.

Genomic instability in iPS cells. The analysis of the genomes of induced mammary stem cells, described above, prompted us to explore the presence of genomic instability in iPS cells prepared by standard protocols.[1–5] We first focused on iPS cells reprogrammed using three factors. Cells prepared from six independent clones (from two independent experiments, three clones per experiment) expressed pluripotency markers and could be induced to differentiate and form embryoid bodies *in vitro* (Supplementary Figure 2). Further, all these clones were diploid, as determined by karyotype analysis (data not shown). cGH was performed for chromosomes 5–13 in their entirety and for part of chromosomes 4 and 14, corresponding in total to half the mouse genome. Only two clones showed signs of genomic instability. One clone had a small gain within chromosome 8 targeting the odd Oz/ten-m homolog 3 (*Odz3*) gene; a second clone also had a small gain, again within chromosome 8, targeting the cadherin 13 (*Cdh13*) gene (Figure 4). Both *Cdh13* and *Odz3* are large genes.

We subsequently examined iPS cells that had been reprogrammed using four factors. Pluripotency was assessed by expression of pluripotency markers (including the

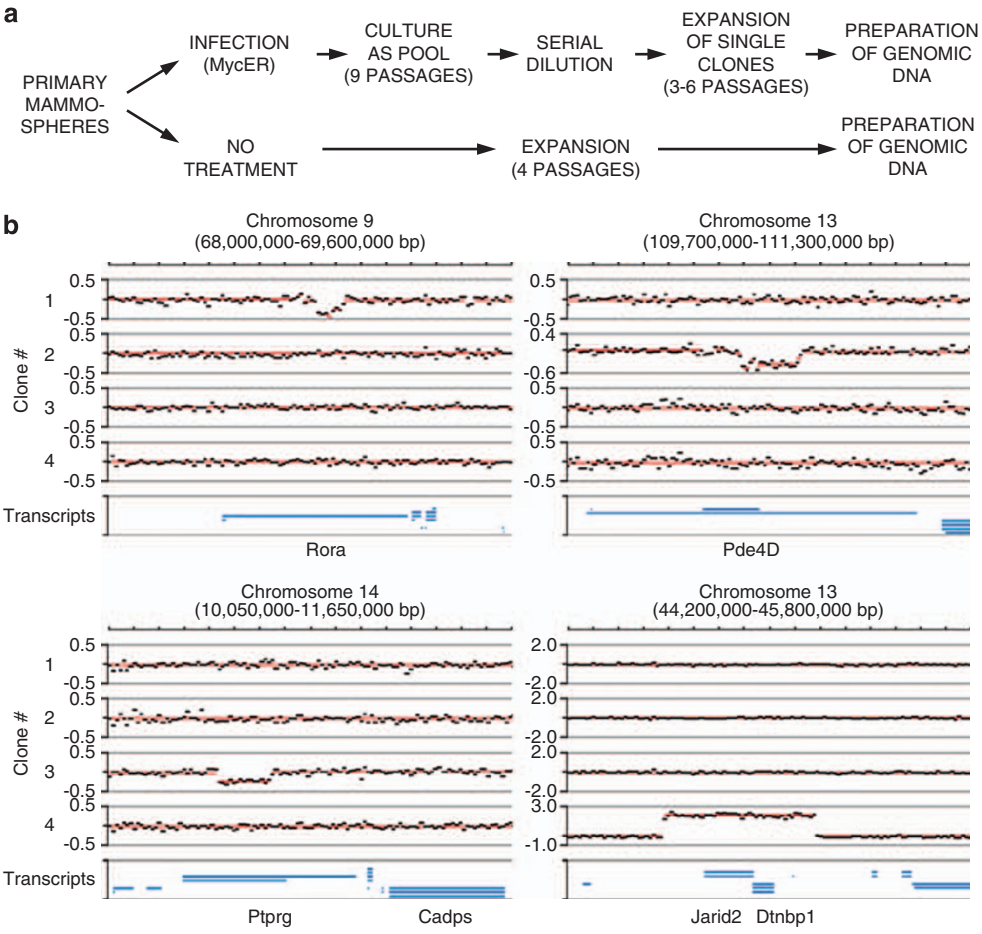


Figure 3 Genomic instability in MycER-induced mammary stem cells. (**a**) Protocol for isolation of genomic DNA from wild-type mammary cells and from MycER-reprogrammed stem cell clones. (**b**) cGH analysis of four MycER-reprogrammed stem cell clones. Selected regions of mouse chromosomes 9, 13 and 14 are indicated. The results are expressed as $\log_2$ ratios of the intensity signals of the stem cell clone DNA to the control DNA. Each black bar represents the average of 10 probes with 1388-bp median spacing. The red lines represent the statistical average and show discontinuities, when a series of averaged probe data deviates in a statistically significant manner from its neighbors. Transcripts are indicated by blue lines and the names of select genes are indicated

endogenous *Oct4* locus and alkaline phosphatase), the ability to grow in the presence of ERK and GSK3 kinase inhibitors (the most stringent conditions for the propagation of embryonic stem cells) and the ability to differentiate into embryoid bodies (Figure 5). The karyotypes of cells from four independent clones were examined; two clones were diploid and two aneuploid (data not shown). Cells from the diploid clones were subjected to cGH for half the mouse genome, as described above for the OSK iPS cells (Figure 6a). Cells from the first clone had an amplification in chromosome 7 that involved >20 genes, as well as a deletion in chromosome 8 targeting the fat mass and obesity associated (*Fto*) gene (Figure 6b), a large gene. In contrast, we could not identify any focal CNCs in cells from the second clone.

Genomic lesions induced by DNA replication stress. Given the conservation of CFS in the mouse and human genomes,[29–31] it appears that some of the CNCs identified in the induced stem cells map to CFS. Specifically, the deletions in the *Rora* and *Ptprg* genes map to the human CFS FRA15A and FRA3B, respectively. Several of the remaining identified CNCs targeted large genes, again suggesting DNA replication stress as the culprit.[30,31] To explore the spectrum of CNCs generated by DNA replication stress, we treated cells for 4 weeks with low doses of aphidicolin, a prototypical agent for inducing DNA replication stress, and then monitored CNCs by cGH (Figure 7a). For this experiment, we employed a mouse–human hybrid cell line, GM11713A, which contains a single copy of human chromosome 3, thus, facilitating the detection of deletions. This cell line was previously used to map aphidicolin-induced deletions within the *FHIT* gene.[32]

Six aphidicolin-treated clones were isolated for cGH analysis spanning the entire chromosome 3. Two clones were essentially identical, except for one small deletion, suggesting that they arose from a common ancestor cell. The concordance of the findings for these two clones provided a validation of the quality of the cGH analysis, but reduced the number of independent clones from six to five. In addition to deletions targeting the *FHIT* gene in FRA3B in four out of the five clones, two clones had deletions within the *PTPRG* gene, also mapping to the FRA3B CFS (Figure 7b). Deletions were further found in the genes dedicator of cytokinesis 3 (*DOCK3*), cell adhesion molecule 2 (*CADM2*), limbic system-associated

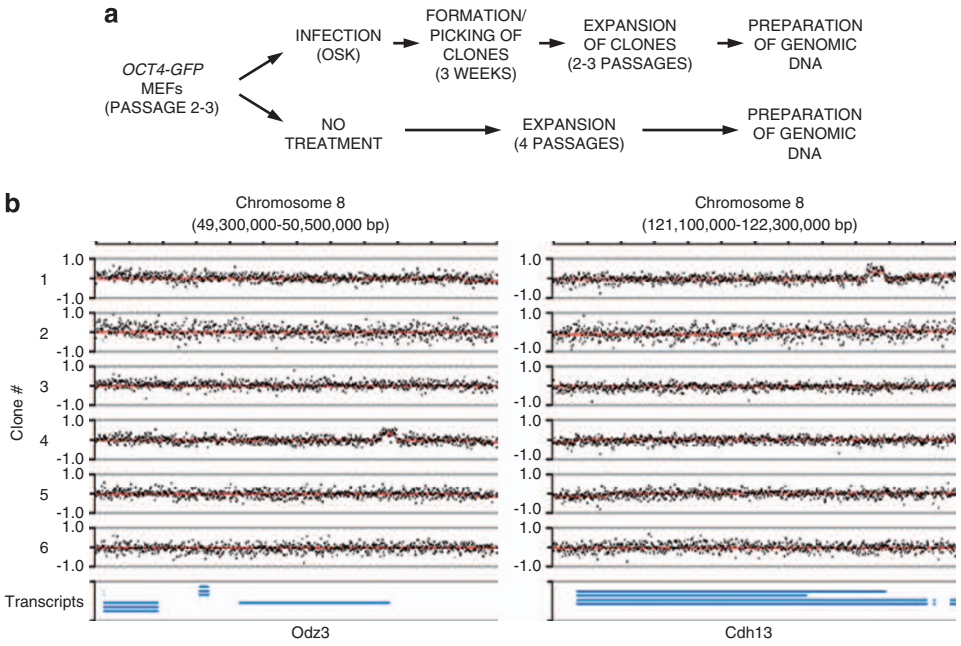


Figure 4 Genomic instability in iPS cells induced by OSK. (**a**) Protocol for isolation of genomic DNA from iPS cells and from non-infected MEFs. (**b**) cGH analysis of six iPS cell clones. Two regions of mouse chromosome 8 are indicated. The results are expressed as $\log_2$ ratios of the intensity signals of the iPS cell DNA to the control DNA. Each black bar represents a single probe with 1388-bp median spacing. The red lines represent the statistical average and show discontinuities, when a series of averaged probe data deviates in a statistically significant manner from its neighbors. Transcripts are indicated by blue lines and the names of select genes are indicated

membrane protein (*LSAMP*) and *N*-acetylated alpha-linked acidic dipeptidase-like 2 (*NAALADL2*), none of which map within established CFS (Supplementary Figure 3). Finally, a gain was detected within the ELKS–RAB6-interacting/CAST family member 2 (*ERC2*) gene (Figure 7b). An interesting feature of this analysis is that all the identified CNCs mapped to large genes. In fact, the CNCs targeted 7 of the 17 largest genes of chromosome 3, whereas no CNCs were detected in any of the remaining annotated genes of chromosome 3 (Figure 7c). Further, while DNA replication stress is usually associated with deletions, one of the observed CNCs was a gain. Similar findings have been obtained by another group using a similar experimental approach.[32,33] Thus, based on the spectrum of CNCs induced by aphidicolin, we propose that of the eight CNCs observed in the induced stem cells (mammary stem cells and iPS cells), six may be linked to DNA replication stress.

Discussion

Previous analyses of mouse and human iPS cells has demonstrated the presence of aneuploid karyotypes in some clones.[3,34] However, it was assumed that the clones that had diploid karyotypes did not have genomic aberrations.[3,35] Our analysis and data accumulating in the literature suggests that this assumption was incorrect.[36–38] Although it is difficult to be conclusive, the most likely mechanism underlying the observed genomic instability in induced stem cells is oncogene-induced DNA replication stress.[6,39–42] The presence of CNCs targeting very large genes and the lower frequency of genomic aberrations in stem cells induced without *c-Myc* support this tentative conclusion. It is important to note that the mechanisms by which c-Myc and other oncogenes induce DNA replication stress and genomic instability are not well established. Yet, we know that genomic instability is not directly linked to the number of cell divisions or to proliferation rate.[39] In some settings, oncogenes induce genomic instability within one cell cycle.[43] Importantly, the induced stem cells we examined were all early passage and were all derived from very early passage mammary cells or MEFs.

A key question is whether a few genomic aberrations compromise the function and utility of reprogrammed stem cells. To answer this question, it will be important to determine whether the observed genomic aberrations are due to transient genomic instability during the reprogramming process or whether the genomic instability persists even after reprogramming. Our study does not address this point. We note that the induced mammary stem cells described here, despite harboring CNCs, were capable of repopulating cleared mouse fat pads and none of the mammary glands derived from these cells have become cancerous so far (in 20 reconstituted mice; albeit within the short time frame of 7 months after transplantation; data not shown). Also, it has been established that OSK and OSKC-iPS cells can form viable fertile mice; yet, the efficiency is variable, possibly reflecting the presence of genomic aberrations in the iPS cells.[2,3,5,10] Finally, mice derived from iPS cells often develop tumors and many of them die *in utero*, suggesting developmental abnormalities.[2,10]

As considerable effort is being placed to develop more efficient protocols for inducing stem cells, care must be

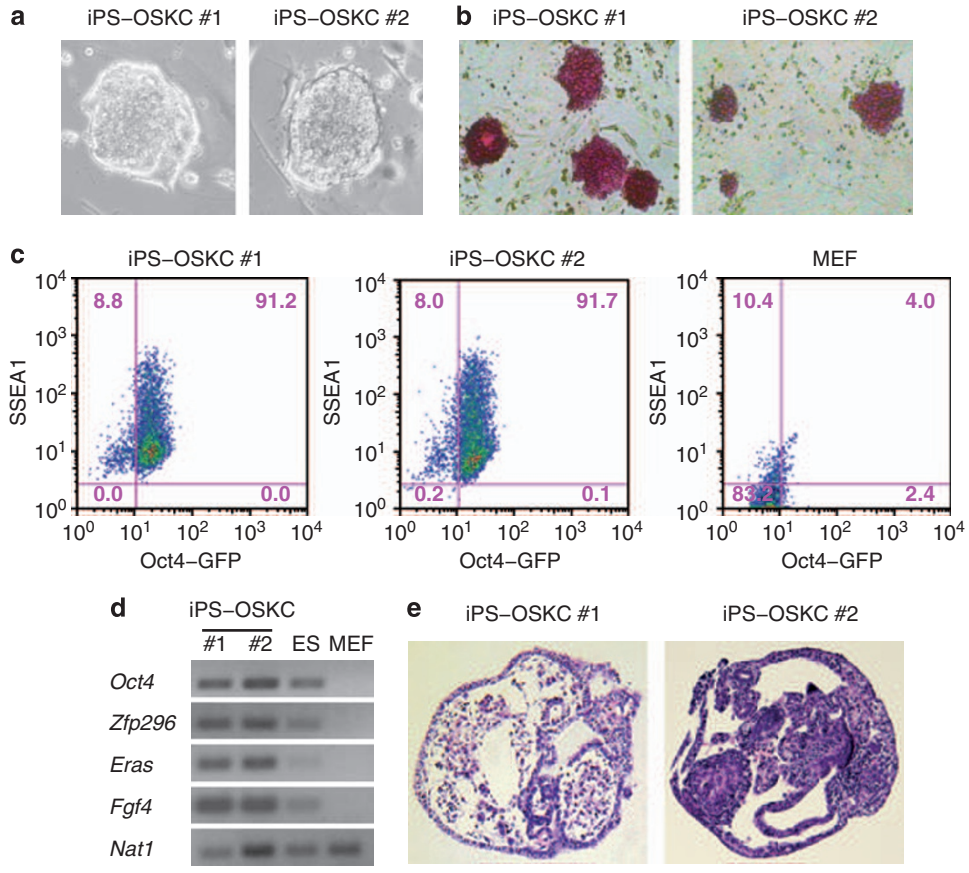


Figure 5 Characterization of iPS cells induced by four factors (OSKC). (**a**) Morphology of iPS-OSKC clones 1 and 2 by phase-contrast microscopy. (**b**) Staining of iPS-OSKC clones 1 and 2 for alkaline phosphatase activity. (**c**) Expression of pluripotency markers in iPS-OSKC cells. Expression levels of SSEA-1 and Oct4-GFP in iPS-OSK clone 1 and 2 cells and in the MEFs from which these iPS cells were derived, as determined by flow cytometry. The percentages of cells expressing high or low levels of SSEA-1 and Oct4-GFP are indicated by purple-colored numbers. (**d**) Expression of pluripotency genes in iPS-OSKC cells. Levels of *Oct4*, *Zfp296*, *Eras* and *Fgf4* in iPS-OSKC clones 1 and 2, in embryonic stem (ES) cells and in MEFs were determined by PCR. *Nat1* expression serves as a standard. (**e**) Morphology of embryoid bodies formed after differentiation of iPS-OSKC clones 1 and 2 for 7 days *in vitro*, as revealed by hematoxylin–eosin staining

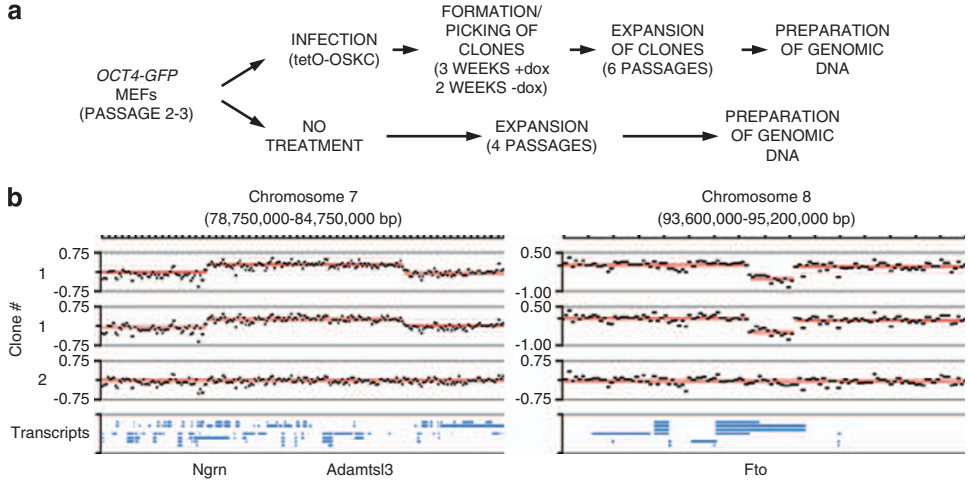
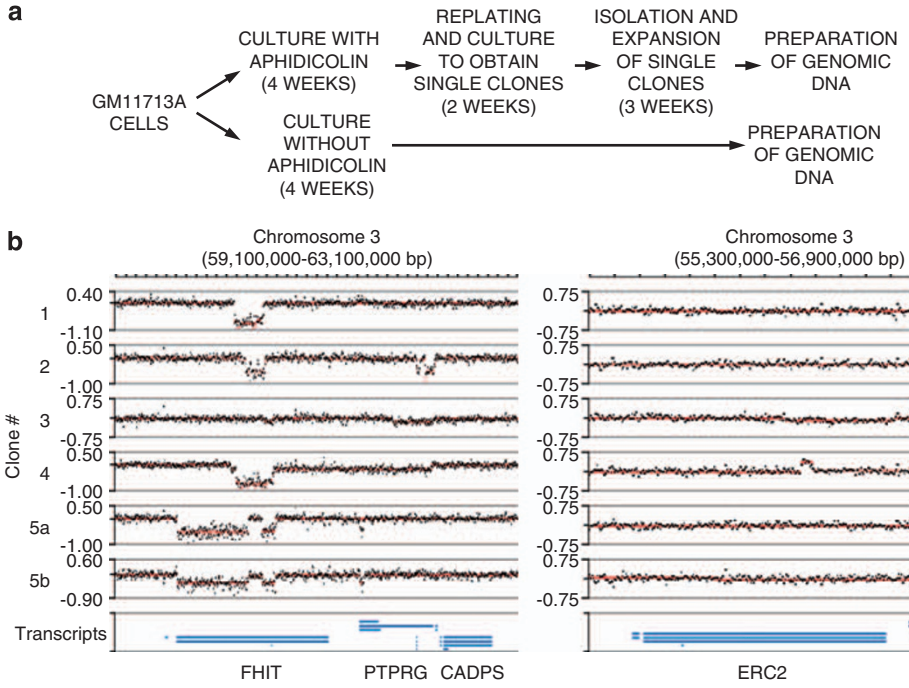


Figure 6 Genomic instability in iPS cells induced by OSKC. (**a**) Protocol for isolation of genomic DNA from iPS cells and from non-infected MEFs. (**b**) cGH analysis of two iPS cell clones. The array hybridization for clone 1 was performed twice and both replicates are shown. Selected regions of mouse chromosomes 7 and 8 are indicated. The results are expressed as $\log_2$ ratios of the intensity signals of the iPS cell DNA to the control DNA. Each black bar represents the average of 10 probes with 1388-bp median spacing. The red lines represent the statistical average and show discontinuities, when a series of averaged probe data deviates in a statistically significant manner from its neighbors. Transcripts are indicated by blue lines and the names of select genes are indicated



c

Gene	Size (bp)	Status	Gene	Size (bp)	Status
FHIT	1,502,089	4 - del	ZBTB20	808,611	0
IGSF4D	1,109,815	1 - del	PTPRG	732,428	2 - del
ROBO1	992,672	0	RBMS3	723,678	0
ERC2	960,056	1 - gain	DOCK3	708,994	1 - del
CNTN4	957,000	0	MAGI1	684,604	0
CACNA2D3	951,095	0	LPP	666,721	0
NAALADL2	946,318	1 - del	ULK4	666,595	0
NLGN1	884,873	0	LSAMP	635,237	1 - del
GRM7	880,291	0			

Figure 7 Genomic instability induced by chronic DNA replication stress. (**a**) Protocol for isolation of genomic DNA from aphidicolin-treated GM11713A cells and untreated cells. (**b**) cGH analysis of six aphidicolin-treated clones. Selected regions of human chromosome 3 are indicated. The results are expressed as $\log_2$ ratios of the intensity signals of the aphidicolin-treated cell DNA to the control DNA. Each black bar represents the average of 10 probes with 475-bp median spacing. The red lines represent the statistical average and show discontinuities, when a series of averaged probe data deviates in a statistically significant manner from its neighbors. Transcripts are indicated by blue lines and the names of select genes are indicated. The fifth and six clones are related to each other and are referred to as subclones 5a and 5b. (**c**) List of the 17 largest genes of human chromosome 3, according to their size, indicating also the presence of identified genomic aberrations in the five aphidicolin-treated clones. The status column indicates the number of clones with a genomic aberration and the type of aberration. del, deletion

taken that increased efficiency is not achieved at the expense of genomic stability. For example, *c-Myc* enhances the efficiency of reprogramming,[1,4,5] but in our study it promoted genomic instability. Inactivation of *p53*, which also increases the efficiency of reprogramming,[11–14] probably does so by allowing genomically unstable cells to escape apoptosis; thus explaining why iPS cells derived from *p53*-deficient fibroblasts form malignant tumors in mice.[17] Finally, some of the chemicals used for reprogramming, such as 5-azacytidine, are known DNA damage-inducing agents.[44] On a brighter note, if oncogene-induced DNA replication stress is a significant contributor to the genomic instability observed in induced stem cells, then modifications to the reprogramming protocols that mitigate oncogenic stress should improve the quality of the generated cells. Nevertheless, at this time, our results and the results of others,[36–38] suggest that great caution should be exercised when planning human therapies using induced stem cells.

Materials and Methods

Induced mammary stem cells. To prepare mammosphere cultures, mammary tissues from 5-month-old virgin female FVB mice (Harlan) or *p53*−/− mice (in the C57/BL6J background, back-crossed from a 129v background) were collected and dissociated mechanically. Disaggregation of the tissues was completed by enzymatic digestion in DMEM supplemented with 100 U/ml Hyaluronidase (Sigma, St. Louis, MO, USA) and 200 U/ml Collagenase (Sigma) for approximately 3 h at 37 °C. The digested material was then centrifuged and filtered through 100, 70, 40 and 20 μm meshes. Red blood cells were lysed by incubating the cell suspension in 0.2% NaCl. The resulting cells were plated on ultralow adhesion plates (Falcon, BD Biosciences, San Jose, CA, USA) at 100 000 cells/ml in MEBM medium (BioWhittaker, Walkersville, MD, USA) supplemented with 5 μg/ml insulin, 0.5μg/ml hydrocortisone, 2% B27 (Invitrogen, Carlsbad, CA, USA), 20 ng/ml EGF and bFGF (BD Biosciences, San Jose, CA, USA) and 4 μg/ml heparin (Sigma). Primary mammospheres were allowed to form for 6 days. At passaging, the number of mammospheres was counted, the mammospheres were then dissociated in single cells and the number of cells counted. A total of 5000 cells were then plated in 24-well plates. At each passage, the number of mammospheres reflects the number of stem cells that had been plated, because a stem cell can form a mammosphere, whereas, by definition, progenitor cells cannot form mammospheres.

For distinguishing native stem cells from progenitor cells, primary mammary cells were incubated with the PKH26 dye (Sigma) for 5 min, as previously described.[19]

The reaction was blocked in 1% BSA and the cells were plated and passaged twice to obtain secondary mammospheres. Single cells obtained from disaggragated secondary mammospheres were sorted by flow cytometry on the basis of PKH26 fluorescence (FACS Vantage SE flow cytometer, Becton and Dickinson, Franklin Lakes, NJ, USA). The PKH26-high population was isolated as the most epifluorescent of the total population (about 0.3% of the cells). PKH26-negative cells were gated at 10 times less fluorescence units with respect to the PKH26-high population (about 30% of the cells). Since stem cells divide once or just a few times during each passage, they retain high levels of the dye, whereas the progenitor cells that divide many times, become PKH-negative. For non-infected mammospheres obtained from wild-type mice, the frequency of stem cells in the PKH26-high population is about 1 : 3, whereas the corresponding frequency in the PKH26-negative population is $<1:80\,000$.[19]

Lentiviral infections were performed using the lentiviral vector pWPI (Addgene, Cambridge, MA, USA) carrying the GFP reporter gene and either no other insert (Empty Vector) or MycER as an insert. Disaggregated primary mammospheres or PKH26-negative cells were plated in Phoenix-generated viral supernatants and infections were carried out in three cycles of 6 h each. After 6 days in culture, secondary mammospheres were collected, disaggregated and FACS sorted for the expression of GFP. For serial passage experiments, 5000 cells from disaggregated mammospheres were plated in 24-well plates and, after 6 days, counted, and re-plated at the same density. Treatment with 500 nM TAM (Sigma) was performed continuously for 7 days during mammosphere formation or once on replating of disaggregated mammospheres on day 1 of culture.

For transplantation experiments, mammosphere cell suspensions were pelleted into Eppendorf benchtop microfuge tubes, counted and resuspended in PBS. In all, 30 μl of cell suspensions were transplanted in the cleared fat pad of 3-week-old virgin female FVB mice. The presence of positive outgrowths was evaluated 2 months later by whole mount analysis.

Immunoblot analysis of protein extracts from mammosphere cultures was performed using anti-vinculin (Sigma), anti-p53 (clone Al25, gift from K Helin), anti-phospho-serine15-p53 (Santa Cruz Biotechnology, Santa Cruz, CA, USA), anti-p21 (clone F5, Santa Cruz Biotechnology) and anti-cleaved caspase3 (clone D175, Cell Signaling Technology, Danvers, MA, USA) primary antibodies, HRP-conjugated secondary antibodies (Sigma) and the SuperSignal West Pico Substrate detection kit (Pierce, Rockford, IL, USA). For immunofluorescence, single cell suspensions were fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X-100 and blocked with 3% BSA. Cells were stained with anti-Myc antibody (provided by S Hann, Vanderbilt University School of Medicine, Nashville, TN, USA).

iPS cells induced by OSK. MEFs containing an *Oct4-GFP* transgene[45] were infected with a lentiviral vector-expressing human *OCT4*, *SOX2* and *KLF4* (OSK) from an SFFV promoter as a single transcript with self-cleaving 2A sequences separating each gene. Colonies of reprogrammed cells were picked and expanded after 20–25 days. iPS cells were cultured on feeder layers in embryonic stem cell (ESC) medium (DMEM, 15% knock-out serum replacement supplement (Invitrogen), L-glutamine, penicillin–streptomycin, nonessential amino acids, β-mercaptoethanol, and 1000 U/ml LIF).

Expression of endogenous pluripotency genes, such as *Nanog*, *Oct4* and *Sox2*, was measured in iPS cells and matched MEFs in triplicate by quantitative PCR. Expression levels were normalized to the levels present in embryonic stem (ES) cells. For embryoid body differentiation, the iPS cells were washed in IMDM medium supplemented with FBS, L-Glu and MonoThioGlycerol to remove LIF. Cells were then resuspended in methylcellulose (40%) containing IMDM medium and plated. GFP expression from the *Oct4-GFP* reporter was monitored for 10 days.

iPS cells induced by OSKC. MEFs containing an *Oct4-GFP* knock-in reporter allele[46] were infected with a lentiviral vector-expressing mouse *Oct4*, *Sox2*, *Klf4* and *c-Myc* (OSKC) from a doxycycline-regulated TetO-mini CMV promoter as a single transcript with self-cleaving 2A sequences separating each gene.[47] After lentiviral infection, the cells were cultured in ESC medium and treated with doxycycline (1 μg/ml) for 23 days. iPS cell colonies were isolated 12 days after doxycycline withdrawal. Genomic DNA was extracted from passage 6 iPS cells.

To monitor expression of pluripotency markers, iPS cells were washed in PBS, stained with anti-SSEA-1 phycoerythrin-conjugated antibody (eBioscience, San Diego, CA, USA) and examined for SSEA1 and Oct4-GFP expression by flow cytometry (FACS Calibur flow cytometer, Becton and Dickinson). In addition, expression of the endogenous genes *Oct4*, *Zfp296*, *Eras* and *Fgf4* was monitored in the iPS cells, matched MEFs and ES cells by PCR, as previously described.[48] Embryoid body differentiation was induced, as described above for the iPS-OSK cells. 21 days later, embryoid bodies were collected, fixed and stained with hematoxylin and eosin.

Aphidicolin-induced DNA replication stress. The GM11713A cell line, a mouse–human hybrid cell line containing a single wild-type copy of human chromosome 3 (Coriell Cell Repository, Camden, NJ, USA), was subjected to DNA replication stress over a period of 30 days, by adding every 4 days to the tissue culture media 0.4 μM aphidicolin. At the end of the 30-day period, the cells were trypsinized and hundred cells were plated on a 10 cm diameter plate. Two weeks later, single colonies were isolated and expanded for 3 weeks, at which time genomic DNA was prepared.

cGH analysis. Control DNA was labeled with Cy5, whereas DNA from the induced stem cells or aphidicolin-treated cells was labeled with Cy3. Equal amounts of Cy5- and Cy3-labeled DNA were hybridized against high-density tiling microarrays. For analysis of mouse genomic DNA the tiling arrays contained probes covering chromosomes 5–8 and parts of chromosomes 4 and 9 (Roche Nimblegen, Madison, WI, USA, array MM8 WG CGH2; Build 36) and/or chromosomes 10–13 and parts of chromosomes 9 and 14 (Roche Nimblegen array MM8 WG CGH3; Build 36). The median probe density on these arrays was 1388 bp. For analysis of human genomic DNA, the tiling arrays contained probes covering chromosome 3 (Roche Nimblegen array HG18 CHR3 FT B3734001-00-01; Build 36). The median probe density on this array was 475 bp.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements. Financial support for this project was provided by the Swiss National Foundation to TDH and DT, the National Institutes of Health, USA to TDH, the European Commission Seventh Framework Programme (GENICA) to TDH and PGP, and the Italian Ministry of Health, Project Giovanni Ricercatori to SC and GT.

1. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006; **126**: 663–676.
2. Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature* 2007; **448**: 313–317.
3. Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K *et al. In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 2007; **448**: 318–324.
4. Wernig M, Meissner A, Cassady JP, Jaenisch R. c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* 2008; **2**: 10–12.
5. Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol* 2008; **26**: 101–106.
6. Halazonetis TD, Gorgoulis VG, Bartek J. An oncogene-induced DNA damage model for cancer development. *Science* 2008; **319**: 1352–1355.
7. Maherali N, Hochedlinger K. Guidelines and techniques for the generation of induced pluripotent stem cells. *Cell Stem Cell* 2008; **3**: 595–605.
8. Ellis J, Bruneau BG, Keller G, Lemischka IR, Nagy A, Rossant J *et al.* Alternative induced pluripotent stem cell characterization criteria for *in vitro* applications. *Cell Stem Cell* 2009; **4**: 198–199.
9. Daley GQ, Lensch MW, Jaenisch R, Meissner A, Plath K, Yamanaka S. Broader implications of defining standards for the pluripotency of iPSCs. *Cell Stem Cell* 2009; **4**: 200–201.
10. Zhao XY, Li W, Lv Z, Liu L, Tong M, Hai T *et al.* Viable fertile mice generated from fully pluripotent iPS cells derived from adult somatic cells. *Stem Cell Rev* 2010; **6**: 390–397.
11. Hong H, Takahashi K, Ichisaka T, Aoi T, Kanagawa O, Nakagawa M *et al.* Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. *Nature* 2009; **460**: 1132–1135.
12. Kawamura T, Suzuki J, Wang YV, Menendez S, Morera LB, Raya A *et al.* Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature* 2009; **460**: 1140–1144.
13. Marion RM, Strati K, Li H, Murga M, Blanco R, Ortega S *et al.* A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* 2009; **460**: 1149–1153.
14. Utikal J, Polo JM, Stadtfeld M, Maherali N, Kulalert W, Walsh RM *et al.* Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. *Nature* 2009; **460**: 1145–1148.
15. Hanna J, Saha K, Pando B, van Zon J, Lengner CJ, Creyghton MP *et al.* Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* 2009; **462**: 595–601.

16. Kuerbitz SJ, Plunkett BS, Walsh WV, Kastan MB. Wild-type p53 is a cell cycle checkpoint determinant following irradiation. *Proc Natl Acad Sci USA* 1992; **89**: 7491–7495.
17. Sarig R, Rivlin N, Brosh R, Bornstein C, Kamer I, Ezra O *et al.* Mutant p53 facilitates somatic cell reprogramming and augments the malignant potential of reprogrammed cells. *J Exp Med* 2010; **207**: 2127–2140.
18. Dontu G, Abdallah WM, Foley JM, Jackson KW, Clarke MF, Kawamura MJ *et al. In vitro* propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev* 2003; **17**: 1253–1270.
19. Cicalese A, Bonizzi G, Pasi CE, Faretta M, Ronzoni S, Giulini B *et al.* The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells. *Cell* 2009; **138**: 1083–1095.
20. Littlewood TD, Hancock DC, Danielian PS, Parker MG, Evan GI. A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. *Nucl Acids Res* 1995; **23**: 1686–1690.
21. El-Deiry WS, Tokino T, Velculescu VE, Levy DB, Parsons R, Trent JM *et al.* WAF1, a potential mediator of p53 tumor suppression. *Cell* 1993; **75**: 817–825.
22. Mai S, Fluri M, Siwarski D, Huppi K. Genomic instability in MycER-activated Rat1A-MycER cells. *Chrom Res* 1996; **4**: 365–371.
23. Ray S, Atkuri KR, Deb-Basu D, Adler AS, Chang HY, Herzenberg LA *et al.* MYC can induce DNA breaks *in vivo* and *in vitro* independent of reactive oxygen species. *Cancer Res* 2006; **66**: 6598–6605.
24. Zhu Y, McAvoy S, Kuhn R, Smith DI. RORA, a large common fragile site gene, is involved in cellular stress response. *Oncogene* 2006; **25**: 2901–2908.
25. Bignell GR, Greenman CD, Davies H, Butler AP, Edkins S, Andrews JM *et al.* Signatures of mutation and selection in the cancer genome. *Nature* 2010; **463**: 893–898.
26. Assou S, Cerecedo D, Tondeur S, Pantesco V, Hovatta O, Klein B *et al.* A gene expression signature shared by human mature oocytes and embryonic stem cells. *BMC Genomics* 2009; **10**: 10.
27. Peng JC, Valouev A, Swigut T, Zhang J, Zhao Y, Sidow A *et al.* Jarid2/jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. *Cell* 2009; **139**: 1290–1302.
28. Shen X, Kim W, Fujiwara Y, Simon MD, Liu Y, Mysliwiec MR *et al.* Jumonji modulates polycomb activity and self-renewal *versus* differentiation of stem cells. *Cell* 2009; **139**: 1303–1314.
29. Glover TW, Berger C, Coyle J, Echo B. DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum Genet* 1984; **67**: 136–142.
30. Helmrich A, Stout-Weider K, Hermann K, Schrock E, Heiden T. Common fragile sites are conserved features of human and mouse chromosomes and relate to large active genes. *Genome Res* 2006; **16**: 1222–1230.
31. Smith DI, Zhu Y, McAvoy S, Kuhn R. Common fragile sites, extremely large genes, neural development and cancer. *Cancer Lett* 2006; **232**: 48–57.
32. Durkin SG, Ragland RL, Arlt MF, Mulle JG, Warren ST, Glover TW. Replication stress induces tumor-like microdeletions in FHIT/FRA3B. *Proc Natl Acad Sci USA* 2008; **105**: 246–251.
33. Arlt MF, Mulle JG, Schaibley VM, Ragland RL, Durkin SG, Warren ST *et al.* Replication stress induces genome-wide copy number changes in human cells that resemble polymorphic and pathogenic variants. *Am J Hum Genet* 2009; **84**: 339–350.
34. Mayshar Y, Ben-David U, Lavon N, Biancotti JC, Yakir B, Clark AT *et al.* Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell* 2010; **7**: 521–531.
35. Yu J, Hu K, Smuga-Otto K, Tian S, Stewart R, Slukvin II *et al.* Human induced pluripotent stem cells free of vector and transgene sequences. *Science* 2009; **324**: 797–801.
36. Yu J, Hu K, Smuga-Otto K, Tian S, Stewart R, Slukvin II *et al.* Correction: human induced pluripotent stem cells free of vector and transgene sequences. *Science* 2009; **324**: 1266.
37. Chin MH, Mason MJ, Xie W, Volinia S, Singer M, Peterson C *et al.* Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 2009; **5**: 111–123.
38. Laurent LC, Ulitsky I, Slavin I, Tran H, Schork A, Morey R *et al.* 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* 2011; **8**: 106–118.
39. Gorgoulis VG, Vassiliou LV, Karakaidos P, Zacharatos P, Kotsinas A, Liloglou T *et al.* Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* 2005; **434**: 907–913.
40. Bartkova J, Horejsi Z, Koed K, Kramer A, Tort F, Zieger K *et al.* DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature* 2005; **434**: 864–870.
41. Bartkova J, Rezaei N, Liontos M, Karakaidos P, Kletsas D, Issaeva N *et al.* Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature* 2006; **444**: 633–637.
42. Di Micco R, Fumagalli M, Cicalese A, Piccinin S, Gasparini P, Luise C *et al.* Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. *Nature* 2006; **444**: 638–642.
43. Denko NC, Giaccia AJ, Stringer JR, Stambrook PJ. The human Ha-ras oncogene induces genomic instability in murine fibroblasts within one cell cycle. *Proc Natl Acad Sci USA* 1994; **91**: 5124–5128.
44. Covey JM, D'Incalci M, Tilchen EJ, Zaharko DS, Kohn KW. Differences in DNA damage produced by incorporation of 5-aza-2′-deoxycytidine or 5,6-dihydro-5-azacytidine into DNA of mammalian cells. *Cancer Res* 1986; **46**: 5511–5517.
45. Boiani M, Kehler J, Scholer HR. Activity of the germline-specific Oct4-GFP transgene in normal and clone mouse embryos. *Methods Mol Biol* 2004; **254**: 1–34.
46. Lengner CJ, Camargo FD, Hochedlinger K, Welstead GG, Zaidi S, Gokhale S *et al.* Oct4 expression is not required for mouse somatic stem cell self-renewal. *Cell Stem Cell* 2007; **1**: 403–415.
47. Sommer CA, Stadtfeld M, Murphy GJ, Hochedlinger K, Kotton DN, Mostoslavsky G. Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* 2009; **27**: 543–549.
48. Takahashi K, Okita K, Nakagawa M, Yamanaka S. Induction of pluripotent stem cells from fibroblast cultures. *Nat Protoc* 2007; **2**: 3081–3089.

Supplementary Information accompanies the paper on Cell Death and Differentiation website (http://www.nature.com/cdd)