

RESEARCH ARTICLE

Effects of Integrating and Non-Integrating Reprogramming Methods on Copy Number Variation and Genomic Stability of Human Induced Pluripotent Stem Cells

Xiangjin Kang[☉], Qian Yu[☉], Yuling Huang, Bing Song, Yaoyong Chen, Xingcheng Gao, Wenyin He^{*}, Xiaofang Sun^{*}, Yong Fan^{*}

Key Laboratory for Major Obstetric Diseases of Guangdong Province, Key Laboratory of Reproduction and Genetics of Guangdong Higher Education Institutes, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou, 510150, China

☉ These authors contributed equally to this work.

* 31848346@qq.com (WH); xiaofangsun@hotmail.com (XS); yongfan011@gzhmu.edu.cn (YF)



OPEN ACCESS

Citation: Kang X, Yu Q, Huang Y, Song B, Chen Y, Gao X, et al. (2015) Effects of Integrating and Non-Integrating Reprogramming Methods on Copy Number Variation and Genomic Stability of Human Induced Pluripotent Stem Cells. *PLoS ONE* 10(7): e0131128. doi:10.1371/journal.pone.0131128

Editor: Rajasingh Johnson, University of Kansas Medical Center, UNITED STATES

Received: February 23, 2015

Accepted: May 27, 2015

Published: July 1, 2015

Copyright: © 2015 Kang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are available from the Gene Expression Omnibus (GEO) via accession numbers GSM1704973–GSM1704991.

Funding: This work was supported by the National Natural Science Foundation of China (81370766, 81301184, 81401254, and U1132005), Guangdong Province Higher Education Funding (2013KJJCX0149 and Yq2013135), Guangdong Province Science and Technology Project (2014A02011029), China Postdoctoral Science Foundation (2013M542163), and the Zhujiang Science and Technology Star Project of Guangzhou (2012J2200006).

Abstract

Human-induced pluripotent stem cells (iPSCs) are derived from differentiated somatic cells using defined factors and provide a renewable source of autologous cells for cell therapy. Many reprogramming methods have been employed to generate human iPSCs, including the use of integrating vectors and non-integrating vectors. Maintenance of the genomic integrity of iPSCs is highly desirable if the cells are to be used in clinical applications. Here, using the Affymetrix Cytoscan HD array, we investigated the genomic aberration profiles of 19 human cell lines: 5 embryonic stem cell (ESC) lines, 6 iPSC lines derived using integrating vectors (“integrating iPSC lines”), 6 iPSC lines derived using non-integrating vectors (“non-integrating iPSC lines”), and the 2 parental cell lines from which the iPSCs were derived. The genome-wide copy number variation (CNV), loss of heterozygosity (LOH) and mosaicism patterns of integrating and non-integrating iPSC lines were investigated. The maximum sizes of CNVs in the genomes of the integrating iPSC lines were 20 times higher than those of the non-integrating iPSC lines. Moreover, the total number of CNVs was much higher in integrating iPSC lines than in other cell lines. The average numbers of novel CNVs with a low degree of overlap with the DGV and of likely pathogenic CNVs with a high degree of overlap with the ISCA (International Symposium on Computer Architecture) database were highest in integrating iPSC lines. Different single nucleotide polymorphisms (SNP) calls revealed that, using the parental cell genotype as a reference, integrating iPSC lines displayed more single nucleotide variations and mosaicism than did non-integrating iPSC lines. This study describes the genome stability of human iPSCs generated using either a DNA-integrating or non-integrating reprogramming method, of the corresponding somatic cells, and of hESCs. Our results highlight the importance of using a high-resolution method to monitor genomic aberrations in iPSCs intended for clinical applications to avoid any negative effects of reprogramming or cell culture.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Human-induced pluripotent stem cells are derived from differentiated somatic cells using defined factors. Like human embryonic stem cells, iPSCs are capable of unlimited proliferation and of differentiating into all cell types of the body. The generation of patient-specific iPSCs holds promise for regenerative medicine because they can provide a renewable source of autologous cells for cell therapy without the risk of immune rejection. Retrovirus- or lentivirus-based delivery systems have served as mainstream methods of generating iPSCs, although three independent groups successfully generated mouse iPSCs using tetraploid complementation to conclusively demonstrate that iPSCs are equivalent to ESCs in terms of pluripotency [1–3]. However, genomic integrations of reprogramming factors in virally generated iPSCs not only cause insertional mutagenesis but also lead to residual expression of reprogramming factors in iPSCs and their derivatives. Several recent studies reported that virally induced iPSCs harbor genetic or epigenetic and transcriptional abnormalities—including dysregulation of imprinted genes, CNVs, accumulation of point mutations, aberrant methylation patterns and other chromosomal aberrations—that are either pre-existing or generated during reprogramming [4–7]. Thus, iPSCs that maintain original genomic integrity and do not carry integrated viral vector sequences or transcription factor DNA are highly desirable for clinical applications.

Many reprogramming methods have been employed to generate human iPSCs without genome-integrating DNA elements. These methods use, for example, episomal vectors, adenoviral vectors, Sendai viral vectors, plasmids, synthetic mRNA, miRNA, protein transduction and small molecules [8–12]. All of these methods have disadvantages, such as low reprogramming efficiency, a requirement for serial transgene deliveries, or success being limited to only certain types of somatic cells, such as commonly used fibroblasts [13]. Of all the reagents used in these non-integrating reprogramming methods, episomal vectors are particularly appealing because they are easy to manipulate and because they allow a relatively high efficiency compared to the other non-integrating methods. Recently, Okita et al. modified their episomal vectors to simultaneously encode more than one reprogramming factor to generate human iPSCs efficiently [9]. Chou et al. utilized an improved episomal vector and successfully generated iPSCs from blood cells [14]. Thus, the generation of human iPSCs based on improved episomal vectors is believed to be efficient, free of genomic integration of transgenes, and representative of a step forward in the development of autologous and allogeneic stem cell therapy.

Genomic stability is critical for clinical applications of human iPSCs. Because genetic aberrations have been strongly associated with cancer, it is important that iPSCs prepared for clinical use are free from cancer-associated genomic alterations. Recently, several studies have identified genomic abnormalities such as chromosomal aneuploidy and translocations, megabase-scale duplications and deletions, and point mutations in iPSCs. Laurent et al. have reported increased subchromosomal CNVs in pluripotent cell samples, with the enriched CNVs located in specific genomic regions [5]. They also found increased numbers of deletions in human iPSCs associated with tumor-suppressor genes, whereas duplications of oncogenes are found in iPSCs that have been cultured for extended periods of time. Gore et al. have reported that point mutations accumulate particularly in oncogenic pathways in otherwise karyotypically normal human iPSCs [15]. Hussein et al. found that early-passage human iPSCs have more CNVs than intermediate-passage iPSCs or ESCs [7]. However, most of the iPSCs described in these studies have genomic abnormalities generated from integrating reprogramming methods. In this study, we performed high-resolution HD genotyping on 10 iPSC lines, the two matched fibroblast and AF lines from which they came, and four hESC lines. The iPSC lines were reprogrammed in our lab using an integrating method (four-factor lentiviral) and a

non-integrating method (episomal vector). We found that the iPSC lines generated by non-integrating methods have lower incidences of genomic aberrations than those generated by integrating methods. Our results suggest that a high-resolution method such as HD genotyping should be strongly recommended for monitoring genomic aberrations throughout iPSC preparation to minimize any effects of acquired genomic aberrations in clinical applications.

Materials and Methods

Cell culture

All experiments were approved by the ethical committee of The Third Affiliated Hospital of Guangzhou Medical University. The tissue donors provided their written informed consent for participation. Briefly, the donors were informed that their tissues would be used for scientific research and that the research results would be published in a scientific research journal but that their names will not be listed in the paper. The donors were not compensated for their donation. Human amniotic fluid (AF) cell lines, fetal fibroblast (FF) cell lines and human ES cell lines were independently established in our laboratory [16,17]. AF cells were cultured in AmnioMAX-II Complete Medium (Invitrogen, Carlsbad, CA, USA). Fetal fibroblast (FF) cell lines were cultured in fibroblast medium (Dulbecco's Modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) (Hyclone, Logan, UT, USA), 1 mM glutamine (Gibco), 1% non-essential amino acids (NEAA) (Gibco), and 100 IU/ml penicillin/streptomycin (Gibco)). Cells were infected with STEMCCA lentiviral supernatants generated by the transfection of 293T packaging cells as previously described [18]. To generate non-integrating iPSCs, AF cells or FF cells were transfected with Episomal iPSC Reprogramming Vectors (Invitrogen, A14703) by electroporation or by transfection with the CytoTune-iPS 2.0 Sendai Reprogramming Kit (Life Technologies, A16518). Transfected cells were then plated onto vitronectin-coated culture dishes according to the manufacturer's instructions.

Human ESC and iPSC lines were cultured on Matrigel-coated tissue culture dishes (ES qualified; BD Biosciences) with mTeSR1 (STEMCELL Technologies, Vancouver, BC, Canada) at 37°C and 5% CO₂ in a 100% humidified atmosphere incubator. The culture medium was refreshed daily until the cells were ready to passage or harvest. Cells were passaged using 1 mg/ml dispase (Gibco) every 3–4 days.

DNA purification

Genomic DNA was extracted and purified with QIAamp DNA Blood Mini Kit (QIAGEN, Germany) according to the manufacturer's instructions. RNA digestion was performed at 37°C for 1 hour using RNase (QIAGEN). The concentration and quality of the DNA samples were determined using a spectrophotometer (Nanodrop 2000, Thermo Scientific, USA) and 1% agarose gel electrophoresis.

CNV and SNP genotyping

The Affymetrix Cytoscan HD array (Affymetrix, Santa Clara, CA, USA), which interrogates 2,696,550 copy number markers across the human genome, was used to characterize cell lines of interest to determine whether there are different associations with de novo CNVs generated using different reprogramming methods. For each sample, 250 ng of input genomic DNA was amplified and labeled according to the manufacturer's instructions [19]. Hybridization of the labeled product using a GeneChip Hybridization Oven 645 (Affymetrix) was followed by washing using the GeneChip Fluidics Station 450 (Affymetrix) and scanning using the GeneChip Scanner 3000 7G (Affymetrix). Data processing was performed in Chromosome Analysis Suite

2.0 (CHAS2.0) software (Affymetrix), via the standard reference library provided by the manufacturer based on the Hidden Markov Model. The segments filter was set to 300 kb and 50 markers for CNV and to 3 Mb and 50 markers for LOH.

Validation of CNV

qRT-PCR was performed to validate CNV calls. SYBR Green assays were performed in triplicate and the predicted copy number was calculated based on the equation: $CN = 2^{(2^{-\Delta\Delta Ct})}$.

Accession numbers

The Gene Expression Omnibus (GEO) accession number for the Affymetrix Cytoscan HD array data which has been submitted to Gene Expression Omnibus in this paper is GSM1704973–GSM1704991.

Results

Previous studies have shown that CNVs are generated during the reprogramming process. To investigate the effect of integrating and non-integrating reprogramming methods on genomic integrity of iPSCs, high-resolution CNV and SNP genotyping (Affymetrix Cytoscan HD array) was performed on 19 samples comprising 2 somatic cell lines, 5 hESC lines, 6 iPSC lines generated by the integrating method and 6 iPSC lines generated by the non-integrating method (Table 1). These cells clearly showed hESCs characteristics (S1 Fig). Fig 1A shows a map of the areas of CNVs identified in all the samples.

Comparison of CNVs in somatic cells, hESCs and iPSCs

Large deletions (>1 Mb) of chromosome 13 were observed in the integrating line iPS-FF-2. In lines iPS-FF-g and iPS-FF-h, 7 large deletions were identified in chromosomes 1, 3, 8, 10, 16, 17 and 22. The line iPS-FF-h had five large duplications of the short arm of chromosome 2. A large deletion of chromosome 13 was observed in iPS-AF-3 (Fig 1A). The maximum sizes of the duplication and deletion were 23116.3 kb and 26586.3 kb, respectively, and were both found in the integrating iPSC lines (Fig 1B). Moreover, the number of total CNVs was greater in integrating iPSC lines than in the other cell lines (Fig 1B). The average length of 3-copy duplications was significantly higher in the integrating iPSC lines (826.1 kb) than in the other lines. The median number of CNVs (>100 kb) in the non-integrating iPSC lines (5.7) was approximately the same as that in hESCs (3.2) and somatic cells (6.0) but six-fold lower than in the integrating iPSC lines (33.8) (Fig 1C). Despite the considerable variation in the number of CNVs among these samples, the average numbers of aberrations in the integrating iPSC lines were overall significantly higher than that in the somatic cells, hESCs and non-integrating iPSC lines (Fig 1B). Among the pluripotent stem cell lines, the genomic aberrations were evenly distributed throughout the genomes in iPSC lines. In contrast, the distribution of CNVs in the four hESCs samples was highly skewed; CNVs were detected only in chromosomes 3, 7, 8, 10, 14 and X (Fig 2A) in hESCs. These results suggest that there had been a positive selection in culture for iPSCs with large aberrations that greatly changed their genomic stability from that of their parental cell lines (Fig 1C).

In addition to these large deletions and duplications, recurrent CNVs associated with each reprogramming method were individually identified in the human pluripotent stem cell samples. We had previously observed that the validation rate for small deletions, which are usually detected at a higher frequency in pluripotent cells, was significantly lower than that for small

Table 1. A list of human cells analyzed for a genomic integrity state in this study.

Cell ID	Cell Type	Reprogramming Method	Ability of Differentiation
AF	Human amniotic fluid cells(P3)	None	None
FF	Human fetal skin fibroblast cells(P5)	None	None
HES17	Embryonic Stem Cells(P29)	None	Pluripotent
HES20	Embryonic Stem Cells(P25)	None	Pluripotent
HES26	Embryonic Stem Cells(P25)	None	Pluripotent
HES29	Embryonic Stem Cells(P24)	None	Pluripotent
HES40	Embryonic Stem Cells(P26)	None	Pluripotent
iPS-AF-35	Induced Pluripotent Cells (P17)	Integrating	Pluripotent
iPS-AF-411	Induced Pluripotent Cells (P15)	Integrating	Pluripotent
iPS-FF-g	Induced Pluripotent Cells (P12)	Integrating	Pluripotent
iPS-FF-h	Induced Pluripotent Cells (P12)	Integrating	Pluripotent
iPS-AF-2	Induced Pluripotent Cells (P15)	Integrating	Pluripotent
iPS-AF-3	Induced Pluripotent Cells (P15)	Integrating	Pluripotent
iPS-AF-6	Induced Pluripotent Cells (P11)	Non- Integrating	Pluripotent
iPS-AF-7	Induced Pluripotent Cells (P10)	Non- Integrating	Pluripotent
iPS-AF-8	Induced Pluripotent Cells(P12)	Non- Integrating	Pluripotent
iPS-AF-9	Induced Pluripotent Cells (P11)	Non- Integrating	Pluripotent
iPS-AF-11	Induced Pluripotent Cells (P12)	Non- Integrating	Pluripotent
iPS-FF-c	Induced Pluripotent Cells (P15)	Non- Integrating	Pluripotent

Numbers in parenthesis with P indicate passage in culture on the cells used in the genomic integrity analysis.

doi:10.1371/journal.pone.0131128.t001

duplications. Therefore, for subsequent analyses, we focused on recurrent de novo duplications relative to the parental cell line. In both the lines iPS-FF-g and iPS-FF-h, we saw duplications of approximately 100 kb located at 7q21.11 (this duplications contain an OMIM gene: SEAM3E) as well as 7q31.32 (this duplications contain an OMIM gene: CADPS2) (Fig 2A). These duplications were found neither in the parental somatic cell samples nor in the non-integrating iPSCs, implying they might be specifically induced by the integrating reprogramming method.

Catalogue of the identified CNVs in diverse cell lines

On the other hand, no significant differences in the total number of CNVs were observed in the pluripotent cells derived from the two somatic cell sources (Fig 2A). No aberration was recurrently located on a particular chromosome in any somatic cell-derived iPSC line.

We generated a list of 33,506 common CNVs identified in 1,038 healthy individuals and of DGVs. These common CNVs could be considered to be mostly functionally neutral. We designated as “novel CNVs” the set of CNVs that were identified in our study and overlapped to a low extent (<10%) with common CNVs. The average number of novel CNVs was 0.5 in somatic cells, 1.0 in human ESCs, 1.0 in non-integrating iPSCs, and 9.3 in integrating iPSCs. The novel CNV fraction was significantly higher in early-passage integrated iPSC lines, while the fraction in non-integrating iPSC lines was as low as the fraction in somatic cell lines (Fig 2B). A set of CNVs from our study that had high coverage rates in the ISCA database (>90%) were designated as ISCA-related CNVs, the average number of which in integrating iPSCs was high. The average number of ISCA-related CNVs were 3.0, 1.8 and 2.0, respectively in somatic cells, hESCs and non-integrating iPSCs, respectively, and was 8.5 in integrating iPSCs. The novel CNV fraction was significantly higher in early-passage integrating iPSCs (Fig 2B).

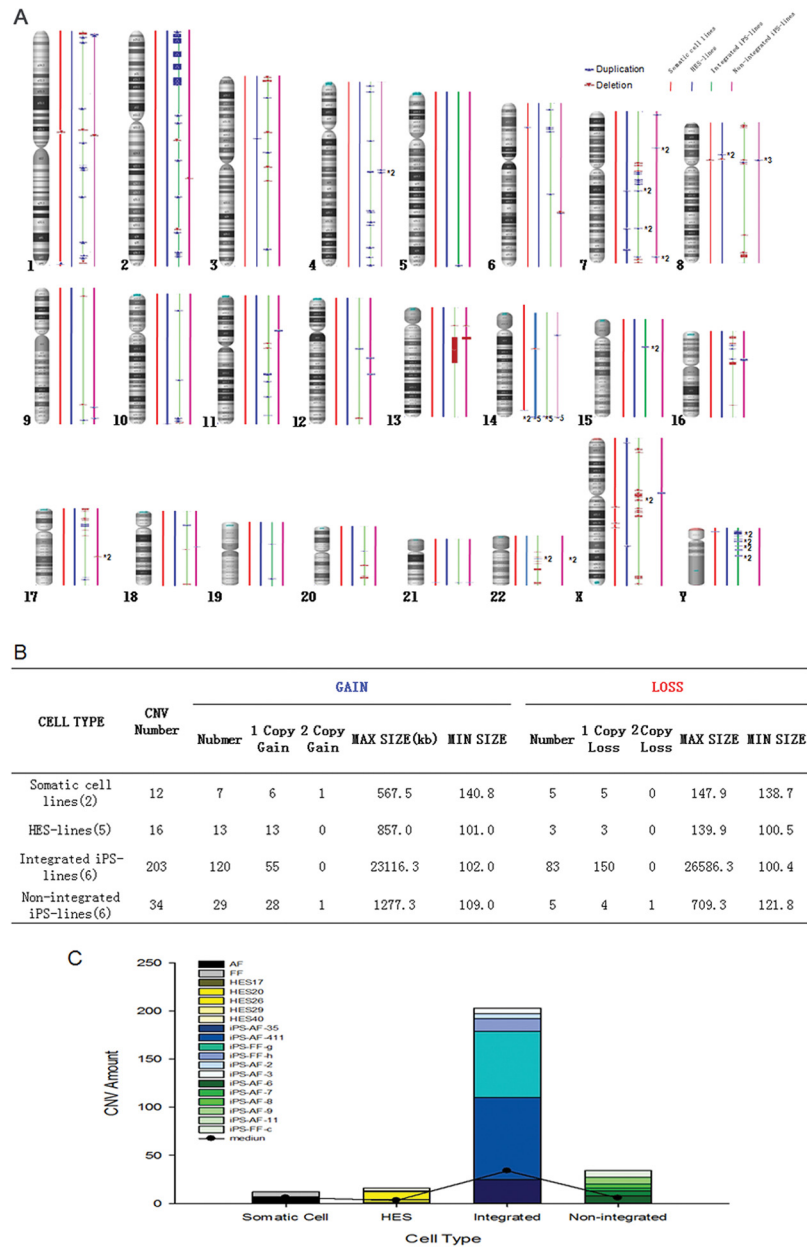


Fig 1. CNV analysis in somatic cells, hESCs and hiPSCs. A) Genomic representation demonstrating the sharp difference in the number of somatic cells, hESCs and hiPSCs. Blue boxes represent amplifications, and red boxes represent deletions, with sizes varying with CNV coverage. B) Comparison of total numbers and average sizes of CNVs between different cell lines. C) Detailed numbers of CNV in each line and the average value of particular cell types.

doi:10.1371/journal.pone.0131128.g001

In AF cell lines and AF-derived iPSCs, 33 CNVs were detected that overlapped with CNVs in integrating iPSCs, while eight other CNVs were detected that only overlapped with CNVs in non-integrating iPSCs (Fig 2C). Gene ontology (GO) analysis revealed that genes containing CNVs particularly induced by the integrating reprogramming method were related mainly to protein transport, plasma membrane and ion binding (Fig 2D).

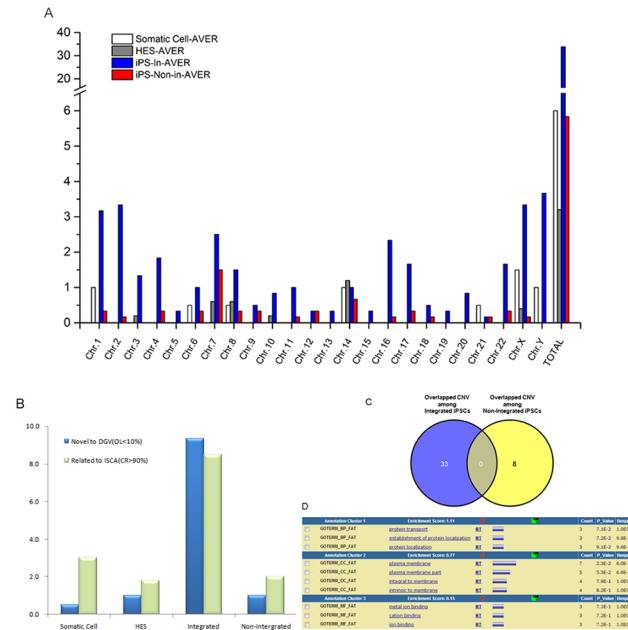


Fig 2. Comparison of CNVs induced by different reprogramming methods. A) Analysis of CNV distribution among chromosomes in somatic cells, hESCs and hiPSCs. B) Average number of novel CNVs with low degrees of overlap with DGVs (blue bar) and likely pathogenic CNVs with high degrees of overlap with the ISCA database (green bar) in somatic cells, hESCs and hiPSCs. C) Venn diagrams for total CNVs acquired after integrating or non-integrating reprogramming of hiPSCs. Total numbers of overlapping CNVs ($n > 2$) in integrating (blue) and non-integrating (yellow) iPSC lines are displayed for representative hiPSC lines. The overlapping areas indicate the number of CNVs that are common between cell lines generated by the two methods. D) GO analysis of the overlapping CNVs among the integrating hiPSC lines.

doi:10.1371/journal.pone.0131128.g002

LOH and SNP in human iPSCs

In both integrating and non-integrating iPSC lines, 72.5% of 2-copy LOH fragments (>1500 kb, >50 marker count) originated from the parental cells and were stably maintained without any recovery of heterozygosity. LOH fragments were recurrently found in the X chromosome, which accounted for 25.4% of the total number (Fig 3A). The integrating line iPS-AF-3 had the exact same LOH pattern as that of its parental cells. Both the lines iPS-AF-9 and iPS-AF-11 had induced loss of heterozygosity in Xp22.2 and Xq23 relative to the AF cells. Because LOH was calculated based on the call for the SNP marker, these data support the hypothesis that a significant number of mutations occur during or shortly after reprogramming and then were selected during the expansion of iPSCs. Further analysis based on SNP markers revealed that, compared to the parental cells, integrating iPSC lines displayed more different SNP calls than did non-integrating iPSC lines (Fig 3B).

Genetic mosaicism in human iPSCs

Genetic mosaicism resulted from the de novo generation of CNVs, followed by selection for less damaged cells during propagation. From the microarray data, no mosaicism was detected in non-integrating iPSC lines, while mosaicism was clearly evident in approximately 38%-48% of integrating iPSC lines (Fig 4). Moreover, in the tested cell lines, mosaicism was occasionally found in human AF cells but not in hESCs or human fibroblasts.

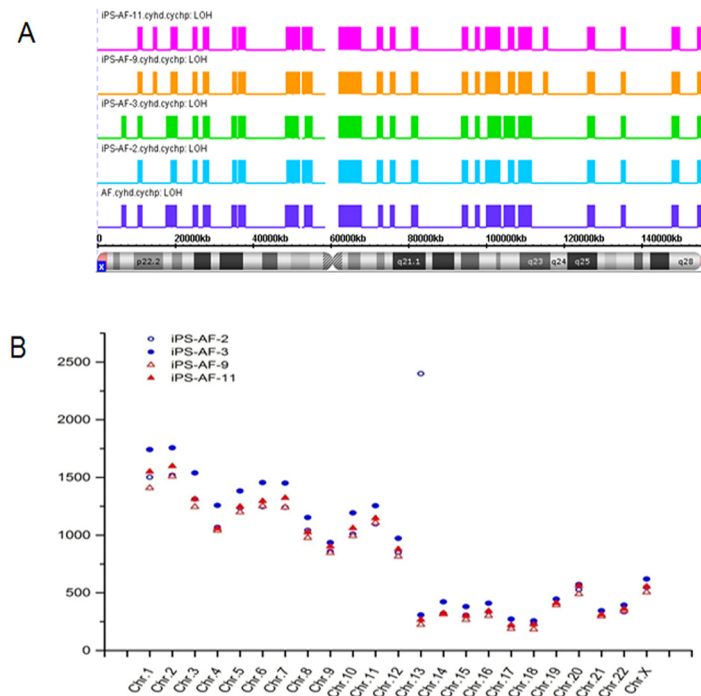


Fig 3. Analysis of LOH and SNVs in the genomes of integrating hiPSCs and non-integrating hiPSCs derived from the same parental cell line. A) LOH changes in the X chromosome caused by the reprogramming method. B) Analysis of SNV distribution among chromosomes.

doi:10.1371/journal.pone.0131128.g003

Discussion

iPSCs present an ideal system for distinguishing between the effects of reprogramming and those of cell passaging on genome stability [7,20]. This study is unique in combining non-pluripotent and pluripotent samples to detect cell type-specific as well as reprogramming method-specific genomic aberrations in a high-resolution microarray platform that enables the detection of kilobase-length aberrations, which are not detectable by karyotyping or low-resolution microarrays. We mainly aimed to evaluate the effects of iPSCs generated by integrating and non-integrating reprogramming methods on genomic stability. Four hESC lines were used as controls, although some of the abnormalities might have been present in the embryos from

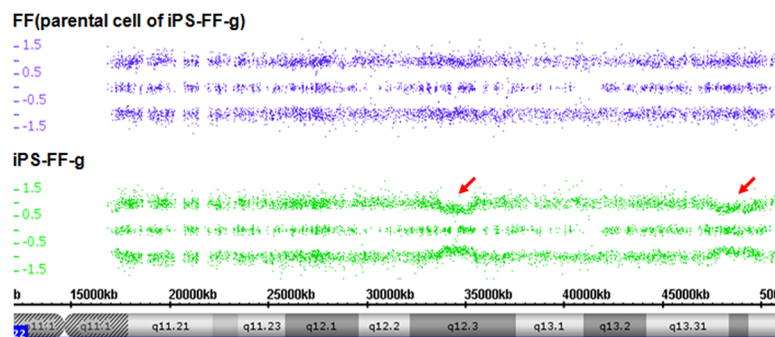


Fig 4. Chas data browser representation of the mosaicism of chromosomes in integrating hiPSCs compared to its parental cells.

doi:10.1371/journal.pone.0131128.g004

which the cells were derived. The parental cells from which iPSCs were derived were also analyzed, because the difference between the pluripotent cells and cultured somatic cells cannot be ignored.

Schlaeger et al performed a small-scale genome-wide aCGH analysis and found that the majority of CNVs were preexisting and that the frequency of possible de novo aberrations was uniformly low among representative patient-derived human iPSC lines obtained by diverse methods[13]. In this study, we intended to identify method-specific differences in CNV loads via microarray analysis with a high probe density. The results presented here indicate that the integrating iPSC lines contain numerous genomic aberrations. Interestingly, non-integrating iPSCs and hESCs were found to contain more duplications, while integrating iPSCs were found to contain more deletions, than somatic non-pluripotent cells. These findings further substantiate the persistence of the effects of these two reprogramming methods of inducing pluripotent cells as well as of the subsequent selection forces[13]. According to our results, early-passage non-integrating iPSCs generated by the episomal or Sendai-virus (SeV) method were pluripotent and had fewer genomic anomalies compared to those generated by the virus method, which suggests that the non-integrating method should be considered more reliable, resulting in fewer genomic rearrangements in iPSC lines[13]. Episomal and SeV methods are the most commonly practiced reprogramming strategies to date[21,22]. These methods are based on plasmid DNA that can be easily generated according to cGMP-compatible processes[23]. Because SeV reprogramming is efficient with a low workload, increasing method-adoption rates have been reported for SeV[13]. Though it has been reported that SeV viral sequences are completely absent in most lines at higher passages, it is worth investigating in detail its dynamic effects on genomic stability.

Hussein et al. concluded that the reprogramming process is associated with a high mutation rate in the resultant early-passage human iPSC lines: novel CNVs in human iPSCs were mostly generated during the reprogramming process[7]. Early-passage cells then endured strong selection pressure and lost the majority of their de novo mutations[5]. Although our study was based mainly on a microarray approach, the high density and coverage of SNP and CNV probes enabled us to identify copy number changes, site mutations, and genetic mosaicism in tested cell lines. These mutations might confer a growth or survival disadvantage to the cells and therefore be selected against, eventually resulting in a CNV profile similar to those found in hESCs. Given that early-passage iPSCs were collected to focus on the influence of the reprogramming procedure, some genetic changes appear to result generally from selection pressures in the continuous cell culture, rather than specifically from the reprogramming. We extended the analysis by asking whether all the CNVs observed in iPSCs in this study pre-existed in the parental somatic populations at low frequencies or were introduced during reprogramming and the subsequent culturing. We compared the loads of CNVs and the mosaicism in the parental cell lines from which all tested iPSCs were derived and found more copy number variants (>50 kb) and mosaicism (>10%) in AF cells than in fibroblast cells. However, as regards iPSCs derived from these two types of cells, no significant differences in either CNV load or mosaicism were detected in this study. This observation suggests that aberrations in iPSCs are induced mainly by reprogramming and are not selected from pre-existing aberrations in somatic cells during the process.

Differential expression and DNA methylation in somatic cells, hESCs and iPSCs have already been reported in previous studies[6,17,24]. All reprogramming methods resulted in similar expression and DNA methylation patterns, with a small portion of genes showing method-specific expression[25]. A similar observation was made from a DNA methylation analysis that showed that only 0.14% of CpG sites were differentially methylated in episomally derived iPSC lines compared with virally derived ones[17]. These results show that diverse

reprogramming strategies have major differences in their effects on genome structure but not on gene regulation.

The possible mechanisms responsible for reprogramming-induced CNVs are still under exploration. A series of recent studies investigated mutations reminiscent of those induced by replication stress and found CNVs located in regions of genomic fragility, such as common fragile sites and subtelomeric regions[4,26]. Unlike the results of other previous studies, no recurrent regions of duplication on chromosomes 12 and 20, which lie in close proximity to known pluripotency genes, were observed in our study. One explanation could be that the iPSC samples were of an early passage, although low passage number does not in itself ensure genetic integrity. A higher level of novel CNVs were detected in integrating iPSCs, suggesting the presence of a greater replication stress in those lines. Meanwhile, previous studies showed that the level of reactive oxygen species increased during reprogramming, which might contribute to the incidence of mutations in the genome[27]. The gene expression profiles of iPSC lines from both integrating and non-integrating methods will be further investigated to identify any differences in enzymatic activities relevant to ROS.

Genomic stability is critical for the clinical use of induced pluripotent stem cells and should be assessed frequently at the endpoint of every treatment. The results of our present study highlight the importance of determining the molecular mechanisms involved in reprogramming somatic cells to be pluripotent, as well as the necessity for optimizing reprogramming and culture conditions to improve genetic stability of iPSCs and their safety for clinical cell therapy.

Supporting Information

S1 Fig. Immunohistochemistry of the stem cell-specific surface antigens Oct4, Nanog and Tra-1-60 in AF-iPSCs and FF-iPSCs and teratoma formation of those iPSCs by subcutaneous implantation into NOD/SCID mice. The iPSCs differentiated into various tissues, including ectoderm (neural tissues), mesoderm (cartilage) and endoderm (glandular tissues). (TIF)

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (81370766, 81301184, 81401254, and U1132005), Guangdong Province Higher Education Funding (2013KJJCX0149 and Yq2013135), Guangdong Province Science and Technology Project (2014A02011029), China Postdoctoral Science Foundation (2013M542163) and the Zhujiang Science and Technology Star Project of Guangzhou (2012J2200006).

Author Contributions

Conceived and designed the experiments: YF WH XS. Performed the experiments: XK QY YH. Analyzed the data: WH YF. Contributed reagents/materials/analysis tools: BS YC XG. Wrote the paper: WH YF.

References

1. Zhao XY, Li W, Lv Z, Liu L, Tong M, Hai T, et al. (2009) iPS cells produce viable mice through tetraploid complementation. *Nature* 461: 86–90. doi: [10.1038/nature08267](https://doi.org/10.1038/nature08267) PMID: [19672241](https://pubmed.ncbi.nlm.nih.gov/19672241/)
2. Kang L, Wang J, Zhang Y, Kou Z, Gao S (2009) iPS cells can support full-term development of tetraploid blastocyst-complemented embryos. *Cell Stem Cell* 5: 135–138. doi: [10.1016/j.stem.2009.07.001](https://doi.org/10.1016/j.stem.2009.07.001) PMID: [19631602](https://pubmed.ncbi.nlm.nih.gov/19631602/)
3. Boland MJ, Hazen JL, Nazor KL, Rodriguez AR, Gifford W, Martin G, et al. (2009) Adult mice generated from induced pluripotent stem cells. *Nature* 461: 91–94. doi: [10.1038/nature08310](https://doi.org/10.1038/nature08310) PMID: [19672243](https://pubmed.ncbi.nlm.nih.gov/19672243/)

4. Martins-Taylor K, Nisler BS, Taapken SM, Compton T, Crandall L, et al. (2011) Recurrent copy number variations in human induced pluripotent stem cells. *Nat Biotechnol* 29: 488–491. doi: [10.1038/nbt.1890](https://doi.org/10.1038/nbt.1890) PMID: [21654665](https://pubmed.ncbi.nlm.nih.gov/21654665/)
5. Laurent LC, Ulitsky I, Slavin I, Tran H, Schork A, Montgomery KD, 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. doi: [10.1016/j.stem.2010.12.003](https://doi.org/10.1016/j.stem.2010.12.003) PMID: [21211785](https://pubmed.ncbi.nlm.nih.gov/21211785/)
6. Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, et al. (2011) Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* 471: 68–73. doi: [10.1038/nature09798](https://doi.org/10.1038/nature09798) PMID: [21289626](https://pubmed.ncbi.nlm.nih.gov/21289626/)
7. Hussein SM, Batada NN, Vuoristo S, Ching RW, Autio R, Närvä E, et al. (2011) Copy number variation and selection during reprogramming to pluripotency. *Nature* 471: 58–62. doi: [10.1038/nature09871](https://doi.org/10.1038/nature09871) PMID: [21368824](https://pubmed.ncbi.nlm.nih.gov/21368824/)
8. Goh PA, Caxaria S, Casper C, Rosales C, Warner TT, Coffey PJ, et al. (2013) A systematic evaluation of integration free reprogramming methods for deriving clinically relevant patient specific induced pluripotent stem (iPS) cells. *PLoS One* 8: e81622. doi: [10.1371/journal.pone.0081622](https://doi.org/10.1371/journal.pone.0081622) PMID: [24303062](https://pubmed.ncbi.nlm.nih.gov/24303062/)
9. Okita K, Matsumura Y, Sato Y, Okada A, Morizane A, Okamoto S, et al. (2011) A more efficient method to generate integration-free human iPS cells. *Nat Methods* 8: 409–412. doi: [10.1038/nmeth.1591](https://doi.org/10.1038/nmeth.1591) PMID: [21460823](https://pubmed.ncbi.nlm.nih.gov/21460823/)
10. Warren L, Manos PD, Ahfeldt T, Loh YH, Li H, Lau F, et al. (2010) Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* 7: 618–630. doi: [10.1016/j.stem.2010.08.012](https://doi.org/10.1016/j.stem.2010.08.012) PMID: [20888316](https://pubmed.ncbi.nlm.nih.gov/20888316/)
11. Yu J, Chau KF, Vodyanik MA, Jiang J, Jiang Y (2011) Efficient feeder-free episomal reprogramming with small molecules. *PLoS One* 6: e17557. doi: [10.1371/journal.pone.0017557](https://doi.org/10.1371/journal.pone.0017557) PMID: [21390254](https://pubmed.ncbi.nlm.nih.gov/21390254/)
12. Hou P, Li Y, Zhang X, Liu C, Guan J, Li H, et al. (2013) Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* 341: 651–654. doi: [10.1126/science.1239278](https://doi.org/10.1126/science.1239278) PMID: [23868920](https://pubmed.ncbi.nlm.nih.gov/23868920/)
13. Schlaeger TM, Daheron L, Brickler TR, Entwisle S, Chan K, Cianci A, et al. (2015) A comparison of non-integrating reprogramming methods. *Nat Biotechnol* 33: 58–63. doi: [10.1038/nbt.3070](https://doi.org/10.1038/nbt.3070) PMID: [25437882](https://pubmed.ncbi.nlm.nih.gov/25437882/)
14. Chou BK, Mali P, Huang X, Ye Z, Doney SN, Resar LM, et al. (2011) Efficient human iPS cell derivation by a non-integrating plasmid from blood cells with unique epigenetic and gene expression signatures. *Cell Res* 21: 518–529. doi: [10.1038/cr.2011.12](https://doi.org/10.1038/cr.2011.12) PMID: [21243013](https://pubmed.ncbi.nlm.nih.gov/21243013/)
15. Gore A, Li Z, Fung HL, Young JE, Agarwal S, Antosiewicz-Bourget J, et al. (2011) Somatic coding mutations in human induced pluripotent stem cells. *Nature* 471: 63–67. doi: [10.1038/nature09805](https://doi.org/10.1038/nature09805) PMID: [21368825](https://pubmed.ncbi.nlm.nih.gov/21368825/)
16. Fan Y, Luo Y, Chen X, Sun X (2010) A modified culture medium increases blastocyst formation and the efficiency of human embryonic stem cell derivation from poor-quality embryos. *J Reprod Dev* 56: 533–539. PMID: [20657158](https://pubmed.ncbi.nlm.nih.gov/20657158/)
17. He W, Kang X, Du H, Song B, Lu Z, Huang Y, et al. (2014) Defining differentially methylated regions specific for the acquisition of pluripotency and maintenance in human pluripotent stem cells via microarray. *PLoS One* 9: e108350. doi: [10.1371/journal.pone.0108350](https://doi.org/10.1371/journal.pone.0108350) PMID: [25250679](https://pubmed.ncbi.nlm.nih.gov/25250679/)
18. Fan Y, Luo Y, Chen X, Li Q, Sun X (2012) Generation of human beta-thalassemia induced pluripotent stem cells from amniotic fluid cells using a single excisable lentiviral stem cell cassette. *J Reprod Dev* 58: 404–409. PMID: [22498813](https://pubmed.ncbi.nlm.nih.gov/22498813/)
19. Luo Y, Li J, Zhu D, Fan Y, Li S, Sun X (2014) High-resolution chromosomal microarray analysis of early-stage human embryonic stem cells reveals an association between X chromosome instability and skewed X inactivation. *Cell Biosci* 4: 74. doi: [10.1186/2045-3701-4-74](https://doi.org/10.1186/2045-3701-4-74) PMID: [25506417](https://pubmed.ncbi.nlm.nih.gov/25506417/)
20. Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, et al. (2011) Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* 144: 439–452. doi: [10.1016/j.cell.2010.12.032](https://doi.org/10.1016/j.cell.2010.12.032) PMID: [21295703](https://pubmed.ncbi.nlm.nih.gov/21295703/)
21. Park TS, Huo JS, Peters A, Talbot CC Jr., Verma K, Zimmerlin L, et al. (2012) Growth factor-activated stem cell circuits and stromal signals cooperatively accelerate non-integrated iPSC reprogramming of human myeloid progenitors. *PLoS One* 7: e42838. doi: [10.1371/journal.pone.0042838](https://doi.org/10.1371/journal.pone.0042838) PMID: [22905176](https://pubmed.ncbi.nlm.nih.gov/22905176/)
22. Macarthur CC, Fontes A, Ravinder N, Kuninger D, Kaur J, Bailey M, et al. (2012) Generation of human-induced pluripotent stem cells by a nonintegrating RNA Sendai virus vector in feeder-free or xeno-free conditions. *Stem Cells Int* 2012: 564612. doi: [10.1155/2012/564612](https://doi.org/10.1155/2012/564612) PMID: [22550511](https://pubmed.ncbi.nlm.nih.gov/22550511/)

23. Okita K, Yamakawa T, Matsumura Y, Sato Y, Amano N, Watanabe A, et al. (2013) An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells. *Stem Cells* 31: 458–466. doi: [10.1002/stem.1293](https://doi.org/10.1002/stem.1293) PMID: [23193063](https://pubmed.ncbi.nlm.nih.gov/23193063/)
24. Newman AM, Cooper JB (2010) Lab-specific gene expression signatures in pluripotent stem cells. *Cell Stem Cell* 7: 258–262. doi: [10.1016/j.stem.2010.06.016](https://doi.org/10.1016/j.stem.2010.06.016) PMID: [20682451](https://pubmed.ncbi.nlm.nih.gov/20682451/)
25. Guenther MG, Frampton GM, Soldner F, Hockemeyer D, Mitalipova M, Jaenisch R, et al. (2010) Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell* 7: 249–257. doi: [10.1016/j.stem.2010.06.015](https://doi.org/10.1016/j.stem.2010.06.015) PMID: [20682450](https://pubmed.ncbi.nlm.nih.gov/20682450/)
26. Agarwal S, Daley GQ (2011) Telomere dynamics in dyskeratosis congenita: the long and the short of iPS. *Cell Res* 21: 1157–1160. doi: [10.1038/cr.2011.120](https://doi.org/10.1038/cr.2011.120) PMID: [21788988](https://pubmed.ncbi.nlm.nih.gov/21788988/)
27. Asensi KD, Fortunato RS, dos Santos DS, Pacheco TS, de Rezende DF, Rodrigues DC, et al. (2014) Reprogramming to a pluripotent state modifies mesenchymal stem cell resistance to oxidative stress. *J Cell Mol Med* 18: 824–831. doi: [10.1111/jcmm.12226](https://doi.org/10.1111/jcmm.12226) PMID: [24528612](https://pubmed.ncbi.nlm.nih.gov/24528612/)