

High Prevalence of Evolutionarily Conserved and Species-Specific Genomic Aberrations in Mouse Pluripotent Stem Cells

URI BEN-DAVID, NISSIM BENVENISTY

Stem Cell Unit, Department of Genetics, Silberman Institute of Life Sciences, The Hebrew University, Jerusalem, Israel

Key Words. Pluripotent stem cells • Embryonic stem cells • Induced pluripotent stem cells • Genomic instability • Chromosomal aberrations

ABSTRACT

Mouse pluripotent stem cells (PSCs) are the best studied pluripotent system and are regarded as the “gold standard” to which human PSCs are compared. However, while the genomic integrity of human PSCs has recently drawn much attention, mouse PSCs have not been systematically evaluated in this regard. The genomic stability of PSCs is a matter of profound significance, as it affects their pluripotency, differentiation, and tumorigenicity. We thus performed a thorough analysis of the genomic integrity of 325 samples of mouse PSCs, including 127 induced pluripotent stem cell (iPSC) samples. We found that genomic aberrations occur frequently in mouse embryonic stem cells of various mouse strains, add in mouse iPSCs of various cell origins and derivation techniques. Four hotspots of chromosomal aberrations were detected: full trisomy 11 (with

a minimally recurrent gain in 11qE2), full trisomy 8, and deletions in chromosomes 10qB and 14qC-14qE. The most recurrent aberration in mouse PSCs, gain 11qE2, turned out to be fully syntenic to the common aberration 17q25 in human PSCs, while other recurrent aberrations were found to be species specific. Analysis of chromosomal aberrations in 74 samples of rhesus macaque PSCs revealed a gain in chromosome 16q, syntenic to the hot-spot in human 17q. Importantly, these common aberrations jeopardize the interpretation of published comparisons of PSCs, which were unintentionally conducted between normal and aberrant cells. Therefore, this work emphasizes the need to carefully monitor genomic integrity of PSCs from all species, for their proper use in biomedical research. *STEM CELLS* 2012;30:612–622

Disclosure of potential conflicts of interest is found at the end of this article.

INTRODUCTION

Since their initial derivation 30 years ago, mouse embryonic stem cells (ESCs) have been extensively used for the study of various fundamental questions in developmental and cellular biology. The salient advantages of the mouse system over other model organisms have led mouse pluripotent stem cells (PSCs) to become the most studied pluripotent system, and much of the work in this field is currently conducted using these cells. Not surprisingly, the seminal breakthrough of reprogramming mature cells into the pluripotent state was achieved for the first time in mouse [1], and mouse induced PSCs (iPSCs) are commonly regarded as the “gold standard,” to which human iPSCs are often compared. Indeed, many technical advancements and mechanistic insights have been achieved with mouse PSCs prior to their achievement with their human counterparts.

Chromosomal aberrations are known to occur in cultures of both mouse and human ESCs [2–4]. However, most of the work on genomic instability and culture adaptation of ESCs has been performed with human cells, with only few attempts to identify and characterize typical aberrations in mouse ESCs [3, 5, 6]. Moreover, while the genomic instability of human

iPSCs has recently caught much attention [7–11], chromosomal aberrations in mouse iPSCs have been reported only sporadically [12, 13], and the genome status of these cells is therefore considered to be unresolved [14].

Although mouse PSCs are irrelevant for cell therapy per se, their genomic stability in culture may have far-reaching implications: first, chromosomal aberrations might affect the differentiation capacity of mouse PSCs, similar to their effect in human PSCs [15, 16]; second, chromosomal aberrations might also influence the pluripotency of the cells, as judged by stringent pluripotency tests, such as their contribution to germline transmission and tetraploid complementation [6]; third, these aberrations are likely to increase the tumorigenicity of the cells [17], thus affecting the interpretation of in vivo experiments performed with mouse PSC-derived cells. Thus, chromosomal aberrations might jeopardize the correct interpretation of studies conducted with aberrant PSCs. Another motivation to study chromosomal aberrations in mouse PSCs would be to compare them to those that arise in human PSCs, a comparison through which novel conclusions could potentially be drawn with regard to the human cells as well.

In this study, we have performed a comprehensive analysis of chromosomal aberrations in mouse ESCs and iPSCs.

Author contributions: U.B.D.: conception and design, collection and/or assembly of data, data analysis and interpretation, and manuscript writing; N.B.: conception and design, data analysis and interpretation, manuscript writing, and final approval of manuscript.

Correspondence: Nissim Benvenisty, M.D., Ph.D., Stem Cell Unit, Department of Genetics, Silberman Institute of Life Sciences, The Hebrew University, Jerusalem, Israel. Telephone: 972-2-6586774; Fax: 972-2-6584972; e-mail: nissimb@cc.huji.ac.il Revised January 22, 2012; accepted for publication January 26, 2012; first published online in *STEM CELLS EXPRESS* February 10, 2012. © AlphaMed Press 1066-5099/2012/\$30.00/0 doi: 10.1002/stem.1057

We report the prevalence of these aberrations in mouse ESCs of various strains, and in mouse iPSCs of various cell origins and derivation techniques. Four genomic regions that recurrently acquire aberrations are characterized and compared with the ones previously identified in human PSCs.

MATERIALS AND METHODS

Gene Expression Profiles Database

Gene expression profiles from studies that involved mouse embryonic stem cells (ESCs) and mouse iPSCs, and which conducted gene expression microarray analysis using Mouse430_2 or HT_MG-430A microarray platforms (Affymetrix, CA, <http://www.affymetrix.com>), were obtained from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo>). Similarly, gene expression profiles from studies that involved rhesus macaque ESCs and iPSCs, and which conducted gene expression microarray analysis using rhesus macaque genome array (Affymetrix, CA, <http://www.affymetrix.com>), were also obtained from the GEO database (<http://www.ncbi.nlm.nih.gov/geo>). Raw .CEL files for all samples were analyzed using MAS5 probeset condensation algorithm using Expression Console (Affymetrix, CA, <http://www.affymetrix.com>). Arrays were analyzed for quality control and outliers were removed. Further outliers were removed following hierarchical clustering analysis. Gene expression profiles from studies that involved mouse epiblast stem cells (EpiSCs) and epiblast stem cell-like iPSCs (ePSCs), and which conducted gene expression microarray analysis using MouseRef-8 v2.0 microarray platform (Illumina, CA, <http://www.illumina.com>), were also obtained from the GEO database (<http://www.ncbi.nlm.nih.gov/geo>), as normalized files. Thus, the final dataset consisted of 325 samples of mouse PSCs and 74 samples of rhesus macaque PSCs (Supporting Information Table S1). Probes absent in more than 20% of the samples were discarded in the Mouse430_2, HT_MG-430A, and rhesus macaque genome array platforms. In the case of multiple probesets for any given gene, multiple instances were discarded, so that each gene would be represented by one probeset only. Whenever possible, probesets ending with “_at” were used for the analysis (otherwise probesets were randomly selected). Probesets without documented chromosomal location were also removed. Thus, separate datasets containing a single probeset for each expressed gene were generated. In order to reduce bias due to low expression levels, values under 50 (for MG430_2, HT_MG-430A, rhesus macaque genome array, and MAS5-normalized MouseRef-8 v2.0 files) or 5.5 (for Robust Multichip Average (RMA)-normalized MouseRef-8 v2.0 files) were collectively raised to this level. In order to further reduce noise in the rhesus macaque platform, the sum of squares (SSQ) of the relative expression values was calculated for each gene and highly variable genes (SSQ >50) were removed as well.

Comparative Genomic Hybridization-Piecewise Constant Fit Overexpression Analysis

For each sample, the expression value of each autosomal gene was divided by the median expression of the same gene across the entire dataset, in order to obtain a comparative value. In order to reduce possible bias from any given experiment, groups of similar samples with highly similar gene expression profiles (as judged by hierarchical clustering) were averaged for the sake of calculating the grand population median. This median then served as the baseline for examining expression bias. The data were then processed using a freely available comparative genomic hybridization (CGH) analysis software program, CGH-Explorer (<http://www.ifi.uio.no/forskning/grupper/bioinf/Papers/CGH>). Gene expression regional bias was detected using the program's piecewise constant fit algorithm, using a set of parameters as follows: least allowed deviation = 0.25–0.30; least allowed aberration size

= 30–80; Winsorize at quantile = 0.001; penalty = 10–12; threshold = 0.01. Moving-average plots of cell lines and regions of interest were drawn using the moving-average fit tool.

Location-Enrichment Analysis

For each sample in which an aberration was detected, a list of the autosomal genes that are overexpressed (>1.5-fold, for trisomies and gains) or underexpressed (<1.5-fold, for monosomies and deletions) relative to the median expression of that gene was comprised. This list was then subjected to location-enrichment analysis, using the Expander software (<http://acgt.cs.tau.ac.il/expander>). For the rhesus macaque aberrations, the list was subjected to chi-square test contingency table analysis, using the Statistics Online Computational Resource (<http://www.socr.ucla.edu/SOCR.html>). Significance was determined as Bonferroni-corrected *p* values lower than 1.0E-4, which is the default value of the Expander program.

Synteny and Orthology Analysis

Synteny between the mouse, rhesus, and human genome was determined and drawn using the Synteny Location-Based Display of the Ensemble Genome Browser 63 (<http://www.ensembl.org>), the National Center for Biotechnology Information (NCBI) Homology Maps Page (<http://www.ncbi.nlm.nih.gov/projects/homology/maps>), and the Primate Cytogenetics database (<http://www.biologia.uniba.it/macaque/>). Orthology between the human and mouse genes that reside in the recurrently aberrant regions was examined using the HUGO Gene Nomenclature Committee (HGNC) Comparison of Orthology Predictions Search (<http://www.genenames.org/cgi-bin/hcop.pl>).

Common Fragile Sites Analysis

An updated map of the common fragile sites (CFSs) in the human genomes was obtained from Bignell et al. [18]. The syntenic regions in the mouse genome were determined and drawn using the Synteny Location-Based Display of the Ensemble Genome Browser 63 (<http://www.ensembl.org>) and the NCBI Homology Maps Page (<http://www.ncbi.nlm.nih.gov/projects/homology/maps>).

Statistical Analysis

Hierarchical clustering was performed using Partek Genomics Suite Version 6.3 (Partek, MO; <http://www.partek.com>). Gene expression values were compared between the diploid and aneuploid lines using one-tailed Student's *t* test. *p* values were corrected for multiple testing using the Bonferroni correction. For the purpose of all statistical analyses, samples of similar cells from the same study were averaged and considered as one independent sample.

RESULTS

Hotspots of Chromosomal Aberrations in Mouse PSCs

We have recently developed a methodology for detecting chromosomal aberrations in human pluripotent and multipotent stem cells, based on the gene expression patterns of the cells lines [7, 11]. Here, we applied the same methodology to perform a comprehensive analysis of chromosomal aberrations in 129 samples of mouse ESCs, 127 samples of mouse iPSCs, 25 samples of mouse ESCs that originated from iPSCs (through nuclear transfer), and 44 samples of mouse EpiSCs, from 48 different studies. Only aberrations that met the stringent criteria for statistical significance in the bioinformatic tools applied are presented and discussed (Materials and Methods). Since the exact passage number was not available for many of the analyzed samples, aberrations in different samples may sometimes reflect the same culture event. Therefore, samples of similar cells from the same study were

averaged and considered as one independent sample for statistical purposes and their visual representation. The number of independent cases of each aberration thus makes a better indicator of its prevalence than the general number of samples in which this aberration occurred.

The analysis revealed a high frequency of chromosomal aberrations in mouse ESCs from various strains. Sixty-nine aberrations were detected in 49 out of 129 ESC samples (38.0%). Some of the aberrations (for example, gain of 7qA) were identified only once, and therefore seem to have occurred sporadically (Supporting Information Fig. S1A). Other aberrations are recurrent aberrations that were detected in at least three independent cases, and these genomic loci are thus identified as hotspots of chromosomal aberrations in mouse ESCs. The analysis identified four such recurrent aberrations: full trisomies of chromosome 8 (21 samples from eight independent cases, 16.3% of the samples, Fig. 1A, 1B and Supporting Information Fig. S1B--S1H), full trisomies of chromosome 11 (10 samples from five independent cases, 7.8% of the samples, Fig. 1B and Supporting Information Fig. S1C--S1E, S1H), deletions of chromosome 10qB (17 samples from three independent cases, 13.5% of the samples, Fig. 1C and Supporting Information Fig. S1G), and deletions of chromosome 14qC-14qE (six samples from three independent cases, 4.8% of the samples, Fig. 1D and Supporting Information Fig. S1H). All mouse ESCs are susceptible to these chromosomal aberrations, as they were identified in multiple-cell lines (such as, R1, J1, D3, AB1, RW4, CGR8, and E14Tg2A) of various mouse strains (C57BL/6, 129/Sv, 129/Ola, Rosa26, ICR, and Jae as well as hybrids of these strains) (Fig. 1A--1D, Supporting Information Fig. S1A--S1I, Table 1 and Supporting Information Table S1).

We next performed the first systematic analysis of chromosomal aberrations in mouse iPSCs. Forty-five aberrations were detected in 29 out of 127 iPSC samples (22.8%). Therefore, chromosomal aberrations were detected in iPSCs at a significant proportion of the samples. The analysis identified chromosomal aberrations in iPSCs in three out of the four hotspots that were identified in ESCs: full trisomies of chromosome 8 (six samples from two independent cases, 4.7% of the samples, Fig. 1E, 1G), full trisomies of chromosome 11

(13 samples from two independent cases, 10.2% of the samples, Fig. 1F, 1J), and deletions of chromosome 14qC-14qE (four samples from two independent cases, 3.1% of the samples, Fig. 1G and Supporting Information Fig. S2A). Smaller gains in chromosome 11qD-11qE were also detected in iPSCs (three independent cases, 2.4% of the samples, Fig. 1H and Supporting Information Fig. S2A, S2B). This suggests that mouse iPSCs are prone to acquire similar characteristic chromosomal aberrations as mouse ESCs, similarly to previous reports in human PSCs [7]. Sporadic aberrations that occurred only once, such as trisomy 12, were identified as well (Supporting Information Fig. 2C).

We also included in the analysis mouse ESCs of iPSC origin. These ESC lines were isolated following the nuclear transfer of iPSCs into enucleated oocytes. As ESCs and iPSCs were found here to acquire the same chromosomal aberrations, it is not surprising that these iPSC-derived ESCs also exhibit the same common aberrations: trisomy 8 (Fig. 1I) and trisomy 11 (Fig. 1J and Supporting Information Fig. S2D). These aberrations could be acquired during the time the cells spent in culture either as ESCs or as iPSCs. Indeed, both of these possibilities could be demonstrated: in one case, it is clear that the aberration already stems from the iPSCs, as the same aberration exists both in the iPSCs of origin and in the ESCs that were derived from them (Fig. 1J); in the other case, the aberration did not exist in the iPSCs of origin, suggesting it occurred after the nuclear transfer (Supporting Information Fig. 2D).

An ideogram of the chromosomal aberrations identified in mouse ESCs, iPSCs, and iPSC-derived ESCs is presented as Figure 2, the list of recurrent aberrations in these cells is presented as Table 1, and the full list of all identified aberrations is presented as Supporting Information Table S1.

Rapid Acquisition of Aberrations in iPSCs of Various Cellular Origins and Derivation Methods

In several cases, chromosomal aberrations were identified in iPSCs from gene expression profiles, which were used to characterize them as pluripotent. These expression profiles most likely reflect iPSCs from early passages (<20), and the identification of so many aberrations in these lines suggests that

Figure 1. Gene expression patterns reveal hotspots of chromosomal aberrations in mouse PSCs. Moving-average plots of gene expression levels along the autosomal genome of ESCs (A--D), iPSCs (E--H), and ESCs that were derived by nuclear transfer of iPSCs (I, J). (A): Four samples of 129sv ESCs demonstrate trisomy of chromosome 8 (red lines). Three normal ESC lines from another study are presented as controls (blue lines). (B): D3 ESC sample demonstrates trisomy of chromosomes 8 and 11 (red line). Three normal ESC lines from another study are presented as controls (blue lines). (C): Nine samples of D3 ESC demonstrate deletion of chromosome 10qB (red lines). Three normal ESC lines from another study are presented as controls (blue lines). (D): Three samples of E14Tg2A ESCs demonstrate monosomy of chromosome 14 (red lines). Three normal ESC lines from another study are presented as controls (blue lines). (E): Three samples of iPSCs demonstrate trisomy of chromosome 8 (red lines). These iPSCs were derived by retroviral transduction of mouse embryonic fibroblasts with Oct4, Sox2, Klf4, and c-Myc. Three samples of normal iPSCs, from another clone reported in the same study, are presented as controls (blue lines). (F): iPSC sample demonstrates trisomy of chromosome 11 (red line). This iPSC line was derived by drug induction of peritoneal fibroblasts from transgenic mouse. Nine samples of normal iPSCs, from various clones reported in the same study, are presented as controls (blue lines). (G): Three samples of iPSCs demonstrate trisomy of chromosome 8 and deletion of chromosome 14qC-14qE (red lines). These iPSCs were derived by retroviral transduction of mouse neural stem cells with Oct4 and Klf4. Three samples of normal ESC lines from another study are presented as controls (blue lines). (H): iPSC sample demonstrates gain of chromosome 11qE2, acquired in culture within 12 passages. The cells were normal at passage 4 (blue line) and acquired the aberration by passage 16 (red line). This iPSC line was derived by drug induction of T cells from transgenic mouse. (I): iPSC-derived Rosa26 ESC sample demonstrates trisomy of chromosome 8 (red line). This ESC line was derived by nuclear transfer of iPSCs into enucleated oocytes. The iPSCs of origin were derived by adenoviral transduction of mouse tail-tip fibroblasts with the reprogramming factors. Two normal subclones of the same original cell line are presented as controls (blue lines), suggesting this aberration occurred at the ESC stage of the cells. (J): Twelve samples of iPSCs and 12 samples of iPSC-derived ESCs demonstrate trisomy of chromosome 11 (red lines). The iPSCs were derived by drug induction of MEFs from transgenic mouse and the ESCs were derived by nuclear transfer of these iPSCs into enucleated oocytes. As the iPSCs and their ESC derivatives exhibit the exact same aberration, it has most likely occurred already at the iPSC stage of the cells. Nine normal ESC samples from the same study are presented as controls (blue lines). Red, orange, and purple bars indicate gains in ESCs, iPSCs, and iPSC-derived ESCs, respectively; dark and light green bars indicate deletion in ESCs and iPSCs, respectively. Samples appear by their GSM numbers, as deposited in the GEO database. The number of normal (blue) and aberrant (red) samples in each panel appears within parentheses. Details regarding each sample are given in Table 1 and Supporting Information Table S1. Abbreviations: ES, embryonic stem; GSM, GEO sample; iPS, induced pluripotent stem.

genomic abnormalities accumulate in iPSCs already at low passages (Table 1). These aberrations most likely occur during the reprogramming process or during the first passages of the cells in culture, as was previously demonstrated for human iPSCs [7, 11]. Indeed, in one study, iPSCs were analyzed both at passage 4 and at passage 16 [19]. Our analysis revealed that one of the cell lines reported in this study exhibited normal diploid karyotype at passage 4, but by passage 16 had already acquired a gain of chromosome 11qE2 (Fig. 1H). These results demon-

strate that advantageous aberrations can take over the culture of mouse iPSCs rapidly, within several passages only.

The analysis also revealed that chromosomal aberrations arise in mouse iPSCs regardless of the cell type from which they are derived, as we could detect them in mouse iPSCs derived from mouse embryonic fibroblasts, tail-tip fibroblasts, peritoneal fibroblasts, T cells, granulocytes, and neural stem cells (Fig. 1E--1G, Supporting Information Fig. S2A--S2D, Table 1 and Supporting Information Table S1). The acquisition of

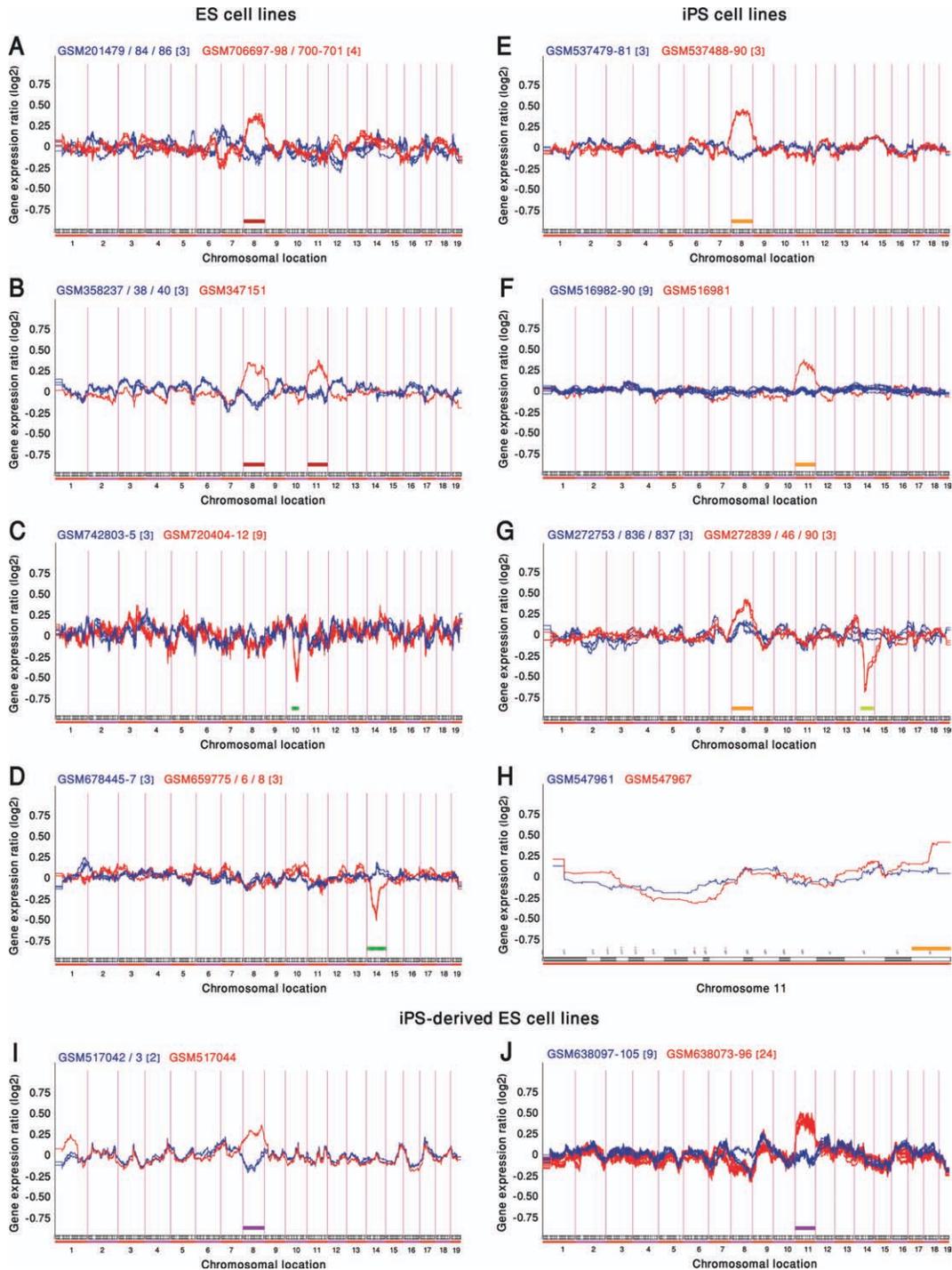


Figure 1.

Table 1. Hotspots of chromosomal aberrations in mouse pluripotent stem cells

Recurrent aberration	PSC type	Mouse strain (Cell line)	Cell of origin	Passage #	Derivation method	Original study GSE #
Trisomy 8	ES	129S7/SvEvBrd (AB1)	ICM	—	Isolation from blastocysts	GSE7528
	ES	129 × 1/SvJ (RW4)	ICM	—	Isolation from blastocysts	GSE11274
	ES	129Sv/J × C57BL/6	ICM	—	Isolation from blastocysts	GSE13408
	ES	129S2/SvPas (D3)	ICM	—	Isolation from blastocysts	GSE13805
	ES	129 × 1/SvJ × 129S1/Sv (R1)	ICM	—	Isolation from blastocysts	GSE20527
	ES	129Sv	ICM	—	Isolation from blastocysts	GSE26001
	ES	129Sv	ICM	—	Isolation from blastocysts	GSE20958
	ES	129Sv	ICM	—	Isolation from blastocysts	GSE30744
	iPS-derived ES	Rosa 26	TTFs-adenoviral transduction-derived iPS cells	Early (<20)	Nuclear transfer of iPS cells	GSE20575
	iPS	Rosa 26 × ICR	NSCs	Early (<20)	Retroviral transduction with two factors (Oct4, Klf4)	GSE10806
Deletion 10B	iPS	B6 × D2	MEFS	—	Retroviral transduction with four factors (Oct4, Sox2, Klf4, and cMyc)	GSE21515
	ES	129S7/SvEvBrd (AB1)	ICM	—	Isolation from blastocysts	GSE7528
Gain 11	ES	129S2/SvPas (D3)	ICM	—	Isolation from blastocysts	GSE29072
	ES	129S2/SvPas (D3)	ICM	—	Isolation from blastocysts	GSE11628
	ES	129S4/SvJae (J1)	ICM	—	Isolation from blastocysts	GSE9978
	ES	129Sv/J × C57BL/6	ICM	—	Isolation from blastocysts	GSE13408
	ES	129S2/SvPas (D3)	ICM	—	Isolation from blastocysts	GSE13805
	ES	129 × 1/SvJ × 129S1/Sv (R1)	ICM	—	Isolation from blastocysts	GSE20527
	ES	129Sv	ICM	—	Isolation from blastocysts	GSE30744
iPS-derived ES	C57BL/6 × 129Sv/Jae F1 × 129Sv/Jae	OSKM-MEFs-drug induced iPS cells	Early (<20)	Nuclear transfer of iPS cells	GSE24705	
iPS-derived ES	Rosa 26	OSKM-Granulocytes-drug induced iPS cells	Early (<20)	Nuclear transfer of iPS cells	GSE20576	
iPS	C57BL/6 × 129Sv/Jae F1 × 129Sv/Jae	OSKM-MEFs	Early (<20)	Drug-inducible transgenic system	GSE24705	
iPS	Rosa 26	OSKM-MEFs	Early (<20)	Drug-inducible transgenic system	GSE20576	
iPS	Rosa 26	OSKM-PFs	Early (<20)	Drug-inducible transgenic system	GSE20576	
iPS	Rosa 26	OSKM-TTFs	Early (<20)	Drug-inducible transgenic system	GSE20576	
iPS	Rosa 26	OSKM-T-cells	p16	Drug-inducible transgenic system	GSE20576	
ES	129 × B6	ICM	—	Isolation from blastocysts	GSE10573	
ES	129P2/OlaHsd (CGR8)	ICM	p26	Isolation from blastocysts	GSE18660	
ES	129P2/OlaHsd (E14Tg2A)	ICM	—	Isolation from blastocysts	GSE26833	
iPS	Rosa 26	OSKM-TTFs	Early (<20)	Drug-inducible transgenic system	GSE20576	
iPS	Rosa 26 × ICR	NSCs	Early (<20)	Retroviral transduction with two factors (Oct4, Klf4)	GSE10806	

Presented are recurrent chromosomal aberrations, which have been identified in at least three independent cases, and are thus considered as hotspots. The color coding of the aberrations is identical to that of the ideogram in Figure 2. ESC and iPSC lines that were derived for the same study in which their gene expression was reported are assumed to be at an early passage (<20), if not stated otherwise. Abbreviations: ES, embryonic stem; ICM, inner cell mass; iPS, induced pluripotent stem; MEFs, mouse embryonic fibroblasts; NSCs, neural stem cells; OSKM, Oct4/Sox2/Klf4/cMyc; PFs, peritoneal fibroblasts; TTFs, tail-tip fibroblasts.

aberrations is also independent of the reprogramming method used for the iPSC derivation—aberrations were detected in iPSCs derived by retroviral transduction with 4, 3, and 2 transcription factors (Fig. 1E, 1G, Supporting Information Fig. S2C and Table 1) as well as by drug-inducible transgenic system (Fig. 1F, 1J, Supporting Information Fig. S2A, S2B and Table 1).

Relevance of Chromosomal Aberrations in PSC Comparison Studies

The prevalence of chromosomal aberrations in iPSCs, starting at early passage, and regardless of their cellular origin and derivation technique, requires caution when evaluating these cells as pluripotent and comparing them to ESCs. For example, the cell line IP20D-3 was germline transmittable, whereas the cell line IP36D-3 was not [20]. This difference was attributed solely to the different expression levels of the *Dlk1-Dio3* region between these lines [20], which is indeed a valid explanation, as the importance of the *Dlk1-Dio3* region has been demonstrated independently [21]. However, we found that IP20D-3 has a diploid karyotype, whereas IP36D-3 has an extra chromosome 8 (Fig. 1E). Trisomy 8 was demonstrated to interfere with germline transmission of ESCs [6], suggesting another plausible cause for the phenotypic difference detected between these two specific clones.

Another study has recently reported a different developmental potential of PSCs generated by different reprogramming strategies [22]. In this study, the authors compared nuclear-transfer ESCs (ntESCs), iPSCs, and iPSC-derived ESCs (iPSC-nt-ESCs) from genetically identical donor cells. While the ntESCs examined in this study could give rise to viable mice in a tetraploid complementation assay, all iPSCs and iPSC-nt-ESCs failed to generate such viable mice. The authors concluded from these results that the ground state of pluripotency was not acquired during the formation of iPSCs and iPSC-nt-ESCs. However, our analysis revealed that the iPSCs and iPSC-nt-ESCs used in this study—but not the ntESCs to which they were compared—harbor full trisomy of chromosome 11 (Fig. 1J and Supporting Information Table S1). Therefore, the difference in the developmental potential of the cells may well be attributed to this trisomy (as was shown to be the case with trisomy 8) rather than to the different reprogramming methods applied. Moreover, the overlooked trisomy may also explain why the ntESCs clustered more closely with normal ESCs than the iPSCs and iPSC-derived ESCs in an unsupervised hierarchical clustering of microarray expression data.

Caution is also bound when comparing ESC lines, as overlooked aberrations may affect results of the analyses that are based on aberrant cells. For example, a recent study demonstrated derivation of haploid mouse ESCs [23]. In this study, the haploid cells were compared with control ESCs that were considered to be normal “diploid” cells. However, our analysis revealed that all three control samples harbored trisomies of chromosomes 8 and 11, and one of them also had an extra chromosome 6 (Supporting Information Fig. 1H). Moreover, gene expression patterns were compared in this study between the haploid and diploid cells, revealing 162 genes that were downregulated in haploid versus diploid ESCs [23]. We found that 32.1% of the listed genes reside on chromosomes 8 and 11, suggesting that the comparison may have been influenced by the trisomies in the control cells.

Syntenic Aberrations in Mouse and Human PSCs

Next, we examined whether the chromosomal aberrations that arise in mouse and human PSCs are evolutionarily conserved. For that aim, we compared the hotspots of chromosomal aberrations

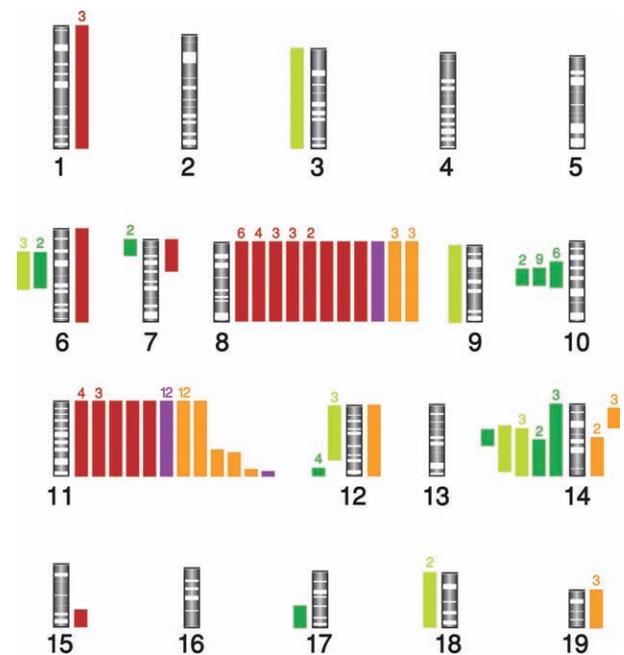


Figure 2. Ideogram of chromosomal aberrations in mouse pluripotent stem cells (PSCs). This ideogram represents the chromosomal aberrations identified in the autosomal genomes of mouse embryonic stem cells (ESCs), induced PSCs (iPSCs), and iPSC-derived ESCs. Bars to the right of the chromosome represent gains, and bars to the left of the chromosome represent deletions. Red, orange, and purple bars indicate gains in ESCs, iPSCs, and iPSC-derived ESCs, respectively; dark and light green bars indicate deletions in ESCs and iPSCs, respectively. Chromosomal aberrations in samples from similar cells from the same study are shown and considered as a single aberration, for the purpose of all statistical analyses and graphic presentations. Whenever a bar represents more than one sample, the number of represented samples is indicated above it.

in mouse PSCs, identified here, with the known common aberrations in human PSCs [2, 7, 11]. Interestingly, human chromosome 17 is completely syntenic to the distal half of mouse chromosome 11 [24]. Gains in chromosome 17 are the second most common aberrations in human PSCs [2, 7, 11], while gains in chromosome 11 were found here to be the most common aberrations in mouse PSCs (Fig. 2 and Table 1). The relatively high resolution of our methodology enabled us to narrow-down the aberrant regions in these chromosomes and to define minimal recurrently aberrant regions of ~10 Mb in both species (17q25 in human and 11qE2 in mouse) (Fig. 3A). Importantly, we found that these minimal recurrent regions are fully syntenic between mouse and human, suggesting that these aberrations are evolutionarily conserved (Fig. 3A).

BIRC5 (Survivin) is an antiapoptotic gene that is important for the survival and the teratoma-formation capacity of human ESCs. Ablation of this gene induces apoptosis in human ESCs in vitro and in teratomas in vivo [25]. Human ESCs with trisomy 17 were shown to generate more aggressive teratomas than diploid ESCs [26], and *BIRC5* was suggested as a candidate gene that might be responsible for the advantageous growth that trisomy 17 confers [2]. Here, we report that *BIRC5* resides inside the minimal recurrently aberrant region of human chromosome 17q25, and its mouse ortholog *Birc5* resides inside the syntenic recurrently aberrant region of mouse chromosome 11qE2 (Fig. 3A). Moreover, *Birc5* is significantly upregulated in all the aberrant mouse cell lines (average fold change = 1.34, p value = 3.5E-4) and

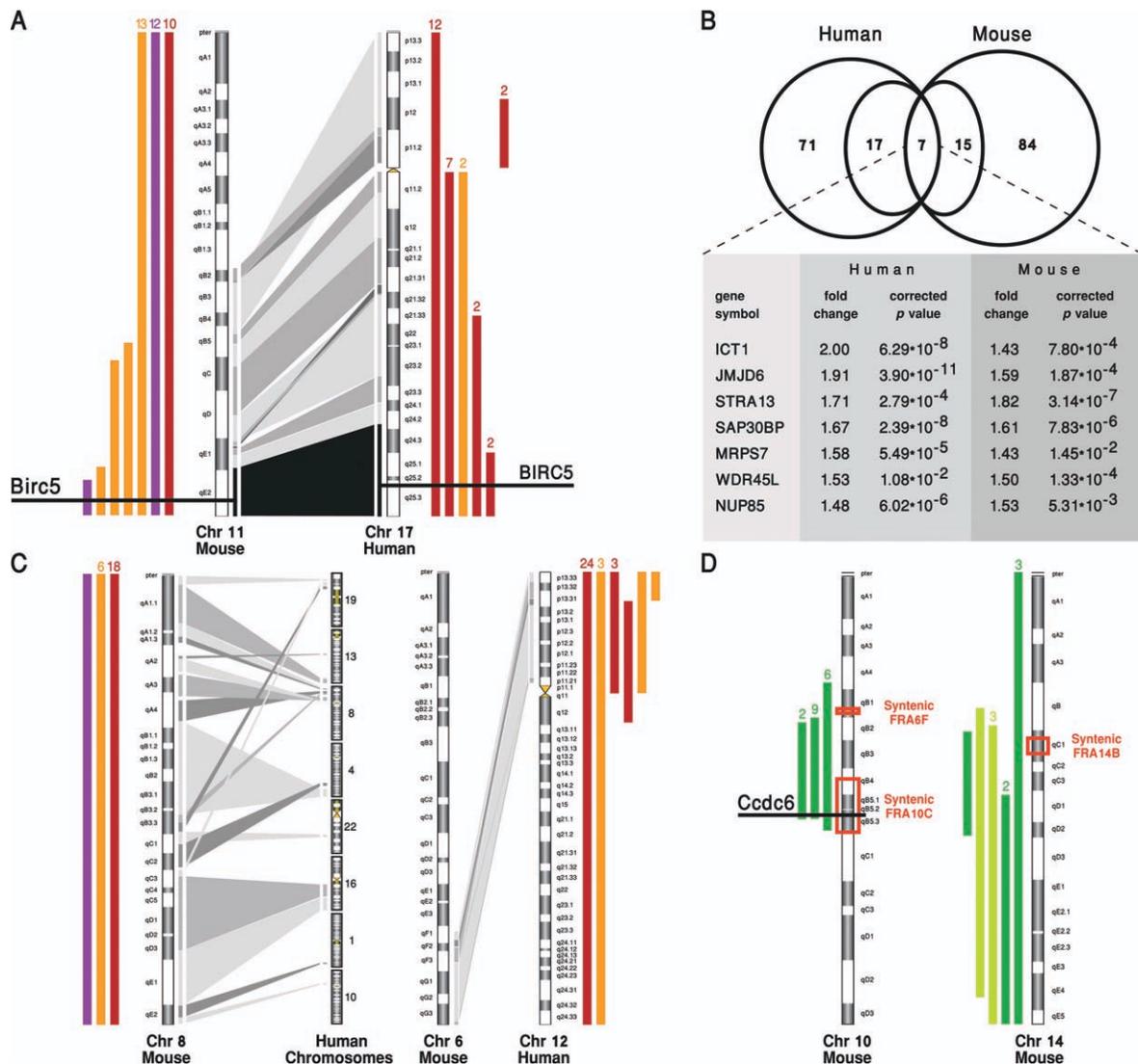


Figure 3. Syntenic and species-specific recurrent aberrations in mouse pluripotent stem cells (PSCs). (A): The recurrently aberrant region, mouse 11qE2, is completely syntenic to the recurrently aberrant region 17q25. The synteny between mouse chromosome 11 and human chromosome 17 is presented in the middle. The recurrent chromosomal aberrations identified in these chromosomes in PSCs are presented as bars to the sides of the chromosomes. Red, orange, and purple bars indicate gains in embryonic stem cells (ESCs), induced PSCs (iPSCs), and iPSC-derived ESCs, respectively. The chromosomal aberrations in mouse PSCs were identified in this study; the aberrations in human PSCs were previously reported by Baker et al. [2], Mayshar et al. [7], and Ben-David et al. [11]. The numbers above the bars represent the number of cases in which each aberration was identified. The chromosomal locations of BIRC5/Birc5 are indicated with a black line. (B): A Venn diagram of the genes that reside inside the minimally recurrent aberration in human 17q25 and mouse 11qE2. 95 genes are expressed in human PSCs from this region, 24 of which are significantly overexpressed in the cell lines that harbor the chromosomal gain. One hundred six genes are expressed in mouse PSCs from this region, 22 of which are significantly overexpressed in the cell lines that harbor the chromosomal gain. Only seven orthologous genes are overexpressed both in mouse and in human aberrant cells. These genes are listed in the table below the Venn diagram, with their average fold change and its statistical significance (corrected for multiple testing by Bonferroni correction). (C): Species-specific recurrently aberrant regions in mouse and human PSCs. Left: the recurrently aberrant region, mouse chromosome 8, is syntenic to multiple human chromosomes. However, while trisomy 8 is very frequent in mouse PSCs, none of the syntenic human regions is recurrently aberrant in human PSCs. Right: the recurrently aberrant region in human PSCs, human 12p, is syntenic to mouse 6qF-6qG. However, while trisomy 12p is the most frequent aberration in human PSCs, trisomy 6 is very rare in mouse PSCs. The recurrent aberrations identified in these chromosomes in PSCs are presented as bars to the sides of the chromosomes. Red, orange, and purple bars indicate gains in ESCs, iPSCs, and iPSC-derived ESCs, respectively. The chromosomal aberrations in mouse PSCs were identified this study; the aberrations in human PSCs were previously reported by Baker et al. [2], Mayshar et al. [7], and Ben-David et al. [11]. The numbers above the bars represent the number of cases in which each aberration was identified. (D): Deletions in 10qB and 14qC-14qE are unique to mouse PSCs and have not been detected in the syntenic human regions. The edges of these recurrent aberrations correspond to genomic regions that are syntenic to common fragile sites (CFSs) in human. Dark and light green bars indicate deletions in ESCs and iPSCs, respectively. Similar aberrations from the same study are presented as a single aberration, and the number of represented samples is indicated above the bars. Genomic regions that are syntenic to human CFSs are delineated in red. The chromosomal location of *Ccdc6* is indicated with a black line. Abbreviations: ICT1, immature colon carcinoma transcript 1; JMJD6, jumonji domain containing 6; MRPS7, mitochondrial ribosomal protein S7; NUP85, nucleoporin 85kDa; STRA13, stimulated by retinoic acid 13 homolog; SAP30BP, SAP30 binding protein; WDR45L, WDR45-like.

in most of the aberrant human cell lines (average fold change = 1.39, p value = .017). These results strongly support the notion that *BIRC5/Birc5* plays an important role in the recurrence of this aberration, both in mouse and in human.

One of the advantages of the methodology applied here is that once a recurrently aberrant region is identified, it is readily accessible to an expression analysis of the genes that reside inside it. We therefore decided to compare the gene expression levels between the syntenic aberrant regions in mouse and human. The minimal recurrently aberrant region in human 17q25 contains 95 genes that are expressed in human PSCs; the minimal recurrently aberrant region in mouse 11qE2 contains 106 genes that are expressed in mouse PSCs. Sixty-two of these genes are orthologs, and are thus expressed in both species from the same syntenic region. A comparison of the human PSCs that harbor the aberration to those that do not revealed 24 genes, which are significantly overexpressed. The same comparison between mouse PSCs that harbor this aberration to those that do not identified 22 significantly overexpressed genes. The intersection between these groups, that is, the list of genes that are significantly overexpressed both in mouse and in human aberrant cell lines (after applying a stringent correction for multiple testing), comprises merely seven genes (Fig. 3B). We present these genes as novel putative candidates that might be functionally relevant for this aberration, both in mouse and in human.

The most upregulated gene in this group (Fig. 3B) is immature colon carcinoma transcript 1 (*ICT1*). This gene codes for a mitochondrial protein, whose function was shown to be essential for cell vitality and mitochondrial function in HeLa cells [27]. *ICT1* is more highly expressed in undifferentiated than in differentiated colon carcinoma cells [28], and it is highly expressed in both human and mouse PSCs [29, 30]. Genome-wide TF-binding studies suggest its promoter is bound by multiple pluripotency factors, including *Oct4* and *Nanog*, in mouse ESCs [31]; genome-wide RNAi knockdown studies showed that downregulation of *SOX2/Sox2*, *OCT4/Oct4*, and/or *NANOG/Nanog* significantly downregulates the expression of *ICT1/Ict1* in both human and mouse ESCs as well as in human embryonal carcinoma cells [31]. Lastly, in a recent genome-wide RNAi screen, knockdown of *ICT1* itself in hESCs significantly reduced the fraction of undifferentiated cells in culture [32], suggesting it is important for self-renewal in these cells. Future studies should show whether this gene is indeed important for the recurrent gains in 17q25 and 11qE2 in human and mouse PSCs.

Species-Specific Aberrations in Mouse and Human PSCs

The comparison of the recurrent chromosomal aberrations between mouse and human revealed that apart from the syntenic aberration discussed above, the other recurrent aberrations are not evolutionarily conserved. In human PSCs, the most common autosomal aberrations, besides gains of chromosome 17q, are gains of chromosome 12p [2, 7] and 20q11.21 [4, 33, 34]; we did not detect any gains in the syntenic regions (in chromosomes 6 and 2, respectively) in mouse PSCs, apart from a single trisomy of chromosome 6 (one out of 325 samples, 0.3% of the samples) (Fig. 3C). On the other hand, the human syntenic regions to the mouse hotspots of chromosomal aberrations—chromosomes 8, 10qB, and 14qC-14qE—do not exhibit recurrent aberrations in human PSCs (Fig. 3C). We conclude that most of the chromosomal aberrations in PSC cultures are species-specific and are not conserved between mouse and human.

As it has been suggested that mouse and human PSCs represent different developmental stages [35, 36], the different chromosomal aberrations that occur in these cells may result from their distinct developmental stages rather than being species specific. In order to determine between these possibilities, we gathered data and analyzed the genomic integrity of 36 samples of mouse EpiSCs and eight ePSCs, from seven different studies. Two samples of trisomy 3 (in one cell line), and two samples of trisomy 10 (in another cell line) were detected (Supporting Information Fig. 3A, 3B and Supporting Information Table S1). Thus, we could detect in these cells neither the characteristic mouse PSC aberrations nor aberrations that were syntenic to the hotspots of human PSC aberrations. The fact that not even a single aberration in chromosomes 8 and 11 could be detected in these cells implies that EpiSCs do not tend to acquire the typical aberrations of PSCs (either ESCs or iPSCs). However, future studies of chromosomal aberrations in larger datasets of EpiSCs, as well as in “naïve” human PSCs, are needed in order to resolve this issue.

Analysis of Rhesus Macaque PSCs

In order to further examine the evolutionary conservation of the detected chromosomal aberrations, we analyzed the genomic integrity of 71 samples of rhesus macaque ESCs and three samples of rhesus macaque iPSCs. The genomic integrity of such cells has never been evaluated before. Although the rhesus macaque genome has been sequenced [37], the cytogenetic organization of its chromosomes is not yet fully resolved. Thus, we limited the rhesus analysis to identification of gains and deletions at the resolution of chromosome arms or whole chromosomes. Four aberrations were detected in four out of 74 samples (two independent cases, 5.4% of the samples, Supporting Information Fig. 4A–4C and Supporting Information Table S1). This low frequency of aberrations should probably be attributed to the low passage number of the cell lines examined, since expression arrays were almost always performed right after the derivation of the rhesus PSCs.

The two identified aberrations are a duplication of chromosome 16q and a full trisomy of chromosome 17 (Supporting Information Fig. S4A–S4C). Interestingly, rhesus macaque chromosome 16 is completely syntenic to human chromosome 17 [37], and the duplication occurred in the region that is syntenic to the chromosome arm 17q (Supporting Information Fig. S4D), which is the arm usually duplicated in human PSCs [2, 7, 11]. Similar to the upregulation of *Birc5* in the aberrant mouse and human PSCs, this gene is also overexpressed in the rhesus cell line with a gain of 16q. In contrast, rhesus macaque chromosome 17 is syntenic to human chromosome 13 [37], which exhibits genomic instability only very rarely [2, 7, 11]. Although more research is needed before these results can be generalized, they do support the notion that the common aberration in human chromosome 17q (corresponding to mouse chromosome 11q and rhesus chromosome 16) is evolutionarily conserved, while other aberrations are species specific.

DISCUSSION

Here, we performed a large-scale analysis of mouse PSCs and revealed high frequency of aberrations in mouse ESCs of various mouse strains as well as in mouse iPSCs of various cell origins and derivation techniques. The accumulation of aberrations in iPSCs begins already at early passages, and these aberrations can take over the culture within several passages. Importantly, all of the aberrations identified in this study are

presumed to have overtaken the culture, as our expression-based approach cannot detect aberrations that are present only at a small subset of the population [7].

The frequency of aberrations in iPSCs was found to be somewhat lower than that in ESCs (23.8% vs. 38.0%, respectively). This difference may result from the longer time that ESCs were grown in culture at the time of analysis, providing them with more time and replication cycles for culture adaptation. For most of the analyzed samples, the exact passages of the ESCs and iPSCs were not specifically mentioned in the original studies, preventing their direct comparison. However, most of the iPSCs were derived especially for their respective studies and were reported in these studies for the first time, whereas most of the ESCs were cell lines cultured routinely at the laboratories and used as controls (Table 1).

Besides the difference in the general frequency of aberrations between ESCs and iPSCs, the relative incidence of each of these aberrations also seems somewhat different: trisomy 8 is more prevalent than trisomy 11 in ESCs (eight vs. five independent cases; 16.3% vs. 7.8% of the samples, respectively), in line with previous reports [3, 6], whereas an opposite trend exists in iPSCs (two independent cases, 4.7% of the samples, for trisomy 8; five independent cases, 12.6% of the samples, for gains in chromosome 11) (Fig. 2 and Table 1). A similar phenomenon was recently reported in human PSCs: both human ESCs and iPSCs acquire gains in chromosomes 12 and 17 [2, 11]; however, gains in chromosome 17 are much more common in human ESCs than they are in human iPSCs [11, 38]. Interestingly, all the ESCs that harbored trisomy 11—but none of the iPSCs that harbored this trisomy—also harbored trisomy 8 (Fig. 1B, Supporting Information Fig. S1C–S1E, Table 1 and Supporting Information Table S1), further supporting a potential difference in the importance of these aberrations between these cell types or between their growth conditions in culture.

Four hotspots of chromosomal aberrations characterize mouse PSCs: full trisomy of chromosome 11 (with a minimal recurrent gain in 11qE2), full trisomy of chromosome 8, and smaller deletions in chromosomes 10qB and 14qC–14qE. The most common aberration in mouse PSCs, gain of 11qE2, is fully syntenic to the second most frequent aberration in human PSCs, gain of 17q25. The antiapoptotic gene *BIRC5/Birc5* resides inside this region and is upregulated in the aberrant lines, supporting its importance in these aberrations. A comparison of the expression of orthologous genes in this region proposes further genes that may be involved in the selection advantage conferred by this gain in both species, the most promising of which seems to be *ICT1/Ict1*. Notably, *BIRC5/Birc5* does not appear in this list of candidate genes since its overexpression in human PSCs is not statistically significant after applying a stringent multiple-testing correction for the unbiased comparison of orthologous genes. This may reflect the ability of multiple genes within this region to drive the selection advantage, thus smaller gains would need to be identified in order to unequivocally determine the relative importance of each gene.

Interestingly, the syntenic regions to the other three identified hotspots in mouse PSCs are not known hotspots of chromosomal aberrations in human PSCs, and other common aberrations in human PSCs were not detected in the mouse syntenic regions. Analysis of mouse EpiSCs also failed to detect aberrations in syntenic regions to the characteristic aberrations of human PSCs. Together, these results imply that most of the chromosomal aberrations in PSC cultures are species-specific and are not conserved between mouse and human. This specificity may result from the different culture

conditions and signaling requirements of mouse and human PSCs; from the unique organization of the chromosomes in each species, which may turn an advantageous aberration into disadvantageous under different chromosomal contexts; or from other inherent species differences. In line with this observation, a recent large-scale analysis of the genomic stability of human ESCs suggested that the structural rearrangement of 20q11.21, common in human ESCs, may be a relatively species-specific event, presumably resulting from a recent pericentric inversion shared only by gorilla, chimp, and human [4]. Importantly, our results from rhesus macaque support both the evolutionary conservation of the common gain in mouse chromosome 11/human chromosome 17 and the species-specificity of other aberrations.

PSCs are rapidly proliferating cells with a unique cell cycle [39]. The most probable mechanism for the occurrence of full trisomies and monosomies in PSCs is a defective mitosis, characterized by supernumerary centrosomes [40], non-disjunction of sister chromatids [41], and the “uncoupling” of the spindle checkpoint from apoptosis [42]. However, small copy number variations have to occur through alternative mechanisms. CFSs are late-replicating genomic loci that are predisposed to spontaneous or induced DNA breakage [43]. CFSs are preferred targets for chromosomal rearrangements at the early stages of tumorigenesis [44, 45]. A recent study of human iPSCs reported that deletions recurred more frequently within CFS regions compared with the whole genome [9]. We therefore examined whether the smaller regions in chromosomes 11 and 14, which we identified as hotspots of deletions in mouse PSCs, reside inside or nearby CFSs.

Although only few CFSs have been accurately mapped in the mouse genome, CFSs are considered to be highly conserved between mouse and human [46, 47]. We thus used a recently published map of CFSs in the human genome [18] and compared their syntenic regions in the mouse genome to the regions of recurrent deletions in mouse PSCs (namely, the deletions in chromosomes 10qB and 14qC–14qE). Interestingly, the fragile site FRA6F, which is accurately mapped in the human genome to a ~1 Mb long region, correlates quite accurately to the beginning of the deletion in chromosome 10 (Fig. 3D). This region is deleted in many human tumors [48]. Another CFS, FRA10C, which is also known to be involved in cancer rearrangements [49], matches exactly the other end of the common deletion of 10qB (Fig. 3D). Moreover, the gene *Ccdc6*, which was shown to reside inside human FRA10C and to undergo DNA breakage after exposure to fragile site-inducing chemicals [49], is located at the end of the aberration (Fig. 3D). *Ccdc6* is indeed significantly downregulated in the aberrant lines (fold change = 0.56, *p* value = .01). Lastly, part of a third CFS, FRA14B, matches the beginning of the deletion in Chr14 (although this CFS has been mapped only at the band resolution, enabling a rather general comparison) (Fig. 3D). These findings suggest that fragile genomic regions may be major targets of genomic instability in mouse PSCs, potentially due to replication stress these cells experience during the reprogramming process or during their expansion in culture.

CONCLUSIONS

Taken together, our results reveal both similarities and differences between chromosomal aberrations in mouse and human PSCs and emphasize the need to carefully and frequently monitor their genomic integrity for their proper use in

biological research. Specifically, when comparing the quality of iPSCs of various origins and derivation methods, or when comparing iPSCs to ESCs, it is crucial to assess the chromosomal aberrations these cells harbor, as these aberrations can interfere with the properties of the cells and thus jeopardize the results of such comparisons.

ACKNOWLEDGMENTS

We thank Tamar Golan-Lev for her assistance with the graphic design. N.B. is the Herbert Cohn Chair in Cancer Research.

This research was supported by The Legacy Heritage Biomedical Science Partnership Program of the Israel Science Foundation (Grant No. 943/09) and by the Centers of Excellence Legacy Heritage Biomedical Science Partnership (Grant No. 1801/10).

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

The authors indicate no potential conflicts of interest.

REFERENCES

- Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006;126:663–676.
- Baker DE, Harrison NJ, Maltby E et al. Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. *Nat Biotechnol* 2007;25:207–215.
- Sugawara A, Goto K, Sotomaru Y et al. Current status of chromosomal abnormalities in mouse embryonic stem cell lines used in Japan. *Comp Med* 2006;56:31–34.
- Amps K, Andrews PW, Anyfantis G et al. Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nat Biotechnol* 2011;29:1132–1144.
- Rebuzzini P, Neri T, Mazzini G et al. Karyotype analysis of the euploid cell population of a mouse embryonic stem cell line revealed a high incidence of chromosome abnormalities that varied during culture. *Cytogenet Genome Res* 2008;121:18–24.
- Liu X, Wu H, Loring J et al. Trisomy eight in ES cells is a common potential problem in gene targeting and interferes with germ line transmission. *Dev Dyn* 1997;209:85–91.
- Mayshar Y, Ben-David U, Lavon N et al. Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell* 2010;7:521–531.
- Laurent LC, Ulitsky I, Slaviv I 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.
- Hussein SM, Batada NN, Vuoristo S et al. Copy number variation and selection during reprogramming to pluripotency. *Nature* 2011;471:58–62.
- Gore A, Fung HL, Young J et al. Somatic coding mutations in human induced pluripotent stem cells. *Nature* 2011;471:63–67.
- Ben-David U, Mayshar Y, Benvenisty N. Large-scale analysis reveals acquisition of lineage-specific chromosomal aberrations in human adult stem cells. *Cell Stem Cell* 2011;9:97–102.
- Pasi CE, Dereli-Oz A, Negrini S et al. Genomic instability in induced stem cells. *Cell Death Differ* 2011;18:745–753.
- Sommer CA, Sommer AG, Longmire TA et al. Excision of reprogramming transgenes improves the differentiation potential of iPSCs generated with a single excisable vector. *Stem Cells* 2010;28:64–74.
- Yeo JC, Ng HH. Transcriptomic analysis of pluripotent stem cells: Insights into health and disease. *Genome Med* 2011;3:68.
- Yang S, Lin G, Tan YQ et al. Differences between karyotypically normal and abnormal human embryonic stem cells. *Cell Prolif* 2010;43:195–206.
- Moon SH, Kim JS, Park SJ et al. Effect of chromosome instability on the maintenance and differentiation of human embryonic stem cells in vitro and in vivo. *Stem Cell Res* 2011;6:50–59.
- Ben-David U, Benvenisty N. The tumorigenicity of human embryonic and induced pluripotent stem cells. *Nat Rev Cancer* 2011;11:268–277.
- Bignell GR, Greenman CD, Davies H et al. Signatures of mutation and selection in the cancer genome. *Nature* 2010;463:893–898.
- Polo JM, Liu S, Figueroa ME et al. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat Biotechnol* 2010;28:848–855.
- Liu L, Luo GZ, Yang W et al. Activation of the imprinted Dlk1–Dio3 region correlates with pluripotency levels of mouse stem cells. *J Biol Chem* 2010;285:19483–19490.
- Stadtfeld M, Apostolou E, Akutsu H et al. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* 2010;465:175–181.
- Jiang J, Ding G, Lin J et al. Different developmental potential of pluripotent stem cells generated by different reprogramming strategies. *J Mol Cell Biol* 2011;3:197–199.
- Leeb M, Wutz A. Derivation of haploid embryonic stem cells from mouse embryos. *Nature* 2011;479:131–134.
- Zody MC, Garber M, Adams DJ et al. DNA sequence of human chromosome 17 and analysis of rearrangement in the human lineage. *Nature* 2006;440:1045–1049.
- Blum B, Bar-Nur O, Golan-Lev T et al. The anti-apoptotic gene survivin contributes to teratoma formation by human embryonic stem cells. *Nat Biotechnol* 2009;27:281–287.
- Blum B, Benvenisty N. The tumorigenicity of diploid and aneuploid human pluripotent stem cells. *Cell Cycle* 2009;8:3822–3830.
- Handa Y, Hikawa Y, Tochio N et al. Solution structure of the catalytic domain of the mitochondrial protein ICT1 that is essential for cell vitality. *J Mol Biol* 2010;404:260–273.
- van Belzen N, Diesveld MP, van der Made AC et al. Identification of mRNAs that show modulated expression during colon carcinoma cell differentiation. *Eur J Biochem* 1995;234:843–848.
- Assou S, Le Carrouer T, Tondeur S et al. A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. *Stem Cells* 2007;25:961–973.
- Wu C, Orozco C, Boyer J et al. BioGPS: An extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol* 2009;10:R130.
- Jung M, Peterson H, Chavez L et al. A data integration approach to mapping OCT4 gene regulatory networks operative in embryonic stem cells and embryonal carcinoma cells. *PLoS One* 2010;5:e10709.
- Chia NY, Chan YS, Feng B et al. A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* 2010;468:316–320.
- Lefort N, Feyeux M, Bas C et al. Human embryonic stem cells reveal recurrent genomic instability at 20q11.21. *Nat Biotechnol* 2008;26:1364–1366.
- Spits C, Mateizel I, Geens M et al. Recurrent chromosomal abnormalities in human embryonic stem cells. *Nat Biotechnol* 2008;26:1361–1363.
- Hanna J, Cheng AW, Saha K et al. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc Natl Acad Sci USA* 2010;107:9222–9227.
- Buecker C, Geijsen N. Different flavors of pluripotency, molecular mechanisms, and practical implications. *Cell Stem Cell* 2010;7:559–564.
- Gibbs RA, Rogers J, Katze MG et al. Evolutionary and biomedical insights from the rhesus macaque genome. *Science* 2007;316:222–234.
- Taapken SM, Nisler BS, Newton MA et al. Karyotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. *Nat Biotechnol* 2011;29:313–314.
- Singh AM, Dalton S. The cell cycle and Myc intersect with mechanisms that regulate pluripotency and reprogramming. *Cell Stem Cell* 2009;5:141–149.
- Holubcova Z, Matula P, Sedlackova M et al. Human embryonic stem cells suffer from centrosomal amplification. *Stem Cells* 2011;29:46–56.
- Cervantes RB, Stringer JR, Shao C et al. Embryonic stem cells and somatic cells differ in mutation frequency and type. *Proc Natl Acad Sci USA* 2002;99:3586–3590.
- Mantel C, Guo Y, Lee MR et al. Checkpoint-apoptosis uncoupling in human and mouse embryonic stem cells: A source of karyotypic instability. *Blood* 2007;109:4518–4527.

- 43 Schwartz M, Zlotorynski E, Kerem B. The molecular basis of common and rare fragile sites. *Cancer Lett* 2006;232:13–26.
- 44 Hellman A, Zlotorynski E, Scherer SW et al. A role for common fragile site induction in amplification of human oncogenes. *Cancer Cell* 2002;1:89–97.
- 45 Bester AC, Roniger M, Oren YS et al. Nucleotide deficiency promotes genomic instability in early stages of cancer development. *Cell* 2011;145:435–446.
- 46 Helmrich A, Stout-Weider K, Hermann K et al. Common fragile sites are conserved features of human and mouse chromosomes and relate to large active genes. *Genome Res* 2006;16:1222–1230.
- 47 Matsuyama A, Shiraishi T, Trapasso F et al. Fragile site orthologs FHIT/FRA3B and Fhit/Fra14A2: Evolutionarily conserved but highly recombinogenic. *Proc Natl Acad Sci USA* 2003;100:14988–14993.
- 48 Morelli C, Karayianni E, Magnanini C et al. Cloning and characterization of the common fragile site FRA6F harboring a replicative senescence gene and frequently deleted in human tumors. *Oncogene* 2002;21:7266–7276.
- 49 Gandhi M, Dillon LW, Pramanik S et al. DNA breaks at fragile sites generate oncogenic RET/PTC rearrangements in human thyroid cells. *Oncogene* 2010;29:2272–2280.



See www.StemCells.com for supporting information available online.