Stem Cell Reports
Report
ISSCR
OPEN ACCESS

Recurrent Genetic Abnormalities in Human Pluripotent Stem Cells: Definition and Routine Detection in Culture Supernatant by Targeted Droplet Digital PCR

Said Assou,[1] Nicolas Girault,[1] Mathilde Plinet,[1] Julien Bouckenheimer,[1] Caroline Sansac,[1] Marion Combe,[1] Joffrey Mianné,[1] Chloé Bourguignon,[1] Mathieu Fieldes,[1] Engi Ahmed,[1] Thérèse Commes,[1,2] Anthony Boureux,[1,2] Jean-Marc Lemaître,[1] and John De Vos[1,3,*]
[1]IRMB, University of Montpellier, INSERM, Hôpital Saint-Eloi, 34295 Montpellier, France
[2]Institut de Biologie Computationnelle, 34095 Montpellier, France
[3]Department of Cell and Tissue Engineering, CHU Montpellier, 34295 Montpellier, France
*Correspondence: john.devos@inserm.fr
https://doi.org/10.1016/j.stemcr.2019.12.004

SUMMARY

Genomic integrity of human pluripotent stem cells (hPSCs) is essential for research and clinical applications. However, genetic abnormalities can accumulate during hPSC generation and routine culture and following gene editing. Their occurrence should be regularly monitored, but the current assays to assess hPSC genomic integrity are not fully suitable for such regular screening. To address this issue, we first carried out a large meta-analysis of all hPSC genetic abnormalities reported in more than 100 publications and identified 738 recurrent genetic abnormalities (i.e., overlapping abnormalities found in at least five distinct scientific publications). We then developed a test based on the droplet digital PCR technology that can potentially detect more than 90% of these hPSC recurrent genetic abnormalities in DNA extracted from culture supernatant samples. This test can be used to routinely screen genomic integrity in hPSCs.

INTRODUCTION

Human pluripotent stem cells (hPSCs) are a physiologically relevant cell material for research (*in vitro* modeling of human development and diseases) and regenerative medicine/cell therapies. Therefore, it is crucial that the hPSC genome remains the faithful genetic copy of the cells from which they were derived. However, genetic abnormalities (e.g., karyotype abnormalities) can arise in hPSCs, for example, during cell reprogramming, cell culture, or genome editing (Assou et al., 2018). Many of these genetic abnormalities are often recurrent. For example, gains of chromosome 12 (most frequently 12p), 17 (particularly 17q), 20, or X have been often detected using standard cytogenetic procedures (G-banding) (Lefort et al., 2009). Sub-chromosomal abnormalities, such as 20q11.21 amplification, also can be recurrent (Lefort et al., 2008). The biological significance of such recurrent abnormalities is still discussed, but they might result in a strong selective growth advantage for cultured cells, as already demonstrated for the 20q11.21 amplification (Zhang et al., 2019). Therefore, it is crucial to carefully catalog all genetic alterations found in hPSCs and identify the recurrent ones. To this aim, we carried out a meta-analysis of published genetic abnormalities found in hPSCs. We could give a precise definition of recurrent genetic abnormality and then listed all of them in a large dataset. As these recurrent genetic abnormalities are found in specific genomic regions, we developed a focused droplet digital PCR (ddPCR) approach that allows screening more than 90% of these recurrent abnormalities in DNA isolated from cell culture supernatant. This method greatly simplifies and therefore encourages the regular and systematic hPSCs screening.

RESULTS

Meta-Analysis of hPSC Genetic Abnormalities and Identification of a Recurrence Pattern

To catalog all genetic abnormalities previously detected in hPSCs using various techniques (karyotyping, fluorescence *in situ* hybridization, comparative genomic hybridization, microarray-based comparative genomic hybridization, and next-generation sequencing [NGS]), we selected primary research articles that reported genetic abnormalities in hESCs and human induced PSCs (hiPSCs), and extracted the DNA abnormality genomic coordinates as well as the experimental data to characterize these abnormalities. We collected data on 942 cell samples and on 415,750 variants and abnormalities from 107 different studies published between 2004 and 2016 (Figures 1A and 1B; Table S1). The dataset included the major publications on genetic abnormalities in hPSCs during culture and also articles that identified one or several abnormal clones in new hPSC lines. A first global analysis of all listed mutations allowed identifying genome locations where these abnormalities were more frequently localized: trisomy 12 and 12p amplification, 20q11.21 amplification, trisomy 17 and 17q amplification, chromosome 1 amplification, and trisomy X (female cell lines)

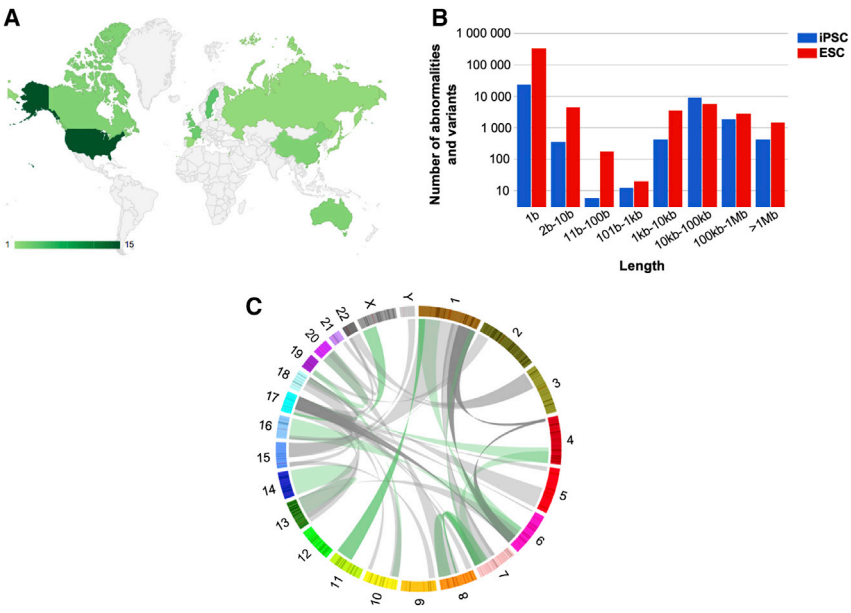


Figure 1. Description of the Genetic Abnormalities Dataset

(A) Countries contributing to the articles included in the analysis.

(B) Number of genetic abnormalities and variations collected, according to their length.

(C) Circos plot representing all translocations in this study. Numbers, chromosome; green, balanced translocations; gray, unbalanced translocations.

(Figure S1A). Abnormalities that accumulated at a specific genome location (i.e., recurrent abnormalities) were mostly aneuploidy or copy-number variations (CNV), in agreement with previous reports. No abnormality smaller than 10 base pairs (bp) displayed a recurrent profile in this large dataset (Figure S1B). We also reported 93 translocations and 20 inversions, involving all chromosomes with the exception of chromosome 12. They were mainly at chromosomes 1 and 17, but without a clear recurrent pattern (Figure 1C).

In summary, this large meta-analysis of genetic abnormalities in hPSCs confirms the recurrence of large CNVs and chromosomal abnormalities, and provides a large dataset of recurrent abnormalities.

Definition and Analysis of Recurrent Genetic Abnormalities

As no quantitative definition of a recurrent hPSC genetic abnormality exists in the literature, we wanted to establish a clear threshold for such events. Therefore, we defined a recurrent genetic abnormality as a non-polymorphic variant that overlaps with abnormalities found in other hPSC lines. The recurrence pattern most likely reflects a common functional cause that occurs in different laboratories and in different cell lines (Assou et al., 2018). We hypothesized that the abnormalities with the strongest functional impact on hPSC growth would be those that are (1) common to different hPSC lines and (2) found in different culture conditions. We estimated that a genetic abnormality met these two criteria if all/part of the altered sequence overlapped with that of other genetic abnormalities that were described in at least four other distinct scientific publications (thus, at least five articles in total) (Figure 2A). To identify recurrent genetic abnormalities based on these criteria, we analyzed all variants >10-bp-long variants that were not polymorphisms (n = 8,284). We found that recurrent abnormalities were only CNVs (including chromosomal gain or losses) (Figure 2B). By plotting the genomic coordinates of these 738 recurrent abnormalities, we found that they were mainly localized in known hotspots, such as chromosome 1, 12, 17q, 20q11.21, or X (Figure 2C). Conversely, there were no recurrent abnormalities, as defined above, in chromosome 2, 4, 10, or 21.

We then investigated the nature of these hotspot regions (Figure 2A) (Table 1 lists the 20 more frequent common regions). Specifically, four regions included more than 50% of all reported abnormal genetic abnormalities. Moreover, a limited set of common abnormal regions comprised the most recurrent genetic abnormalities. Indeed, more than 90% of recurrent abnormalities were restricted to 20 common regions. A set of probes designed to cover these regions would detect all genomic abnormalities of these regions, except balanced translocations (Figure 2D).

Cell Culture Supernatant as a DNA Source to Evaluate hPSC Genomic Integrity by ddPCR

We then decided to take advantage of this highly biased recurrence profile of hPSC genetic abnormalities to develop a rapid PCR-based approach to detect the most common recurrent abnormalities, including those that cannot be detected by karyotyping due to its resolution limits. We analyzed DNA extracted from different hPSC

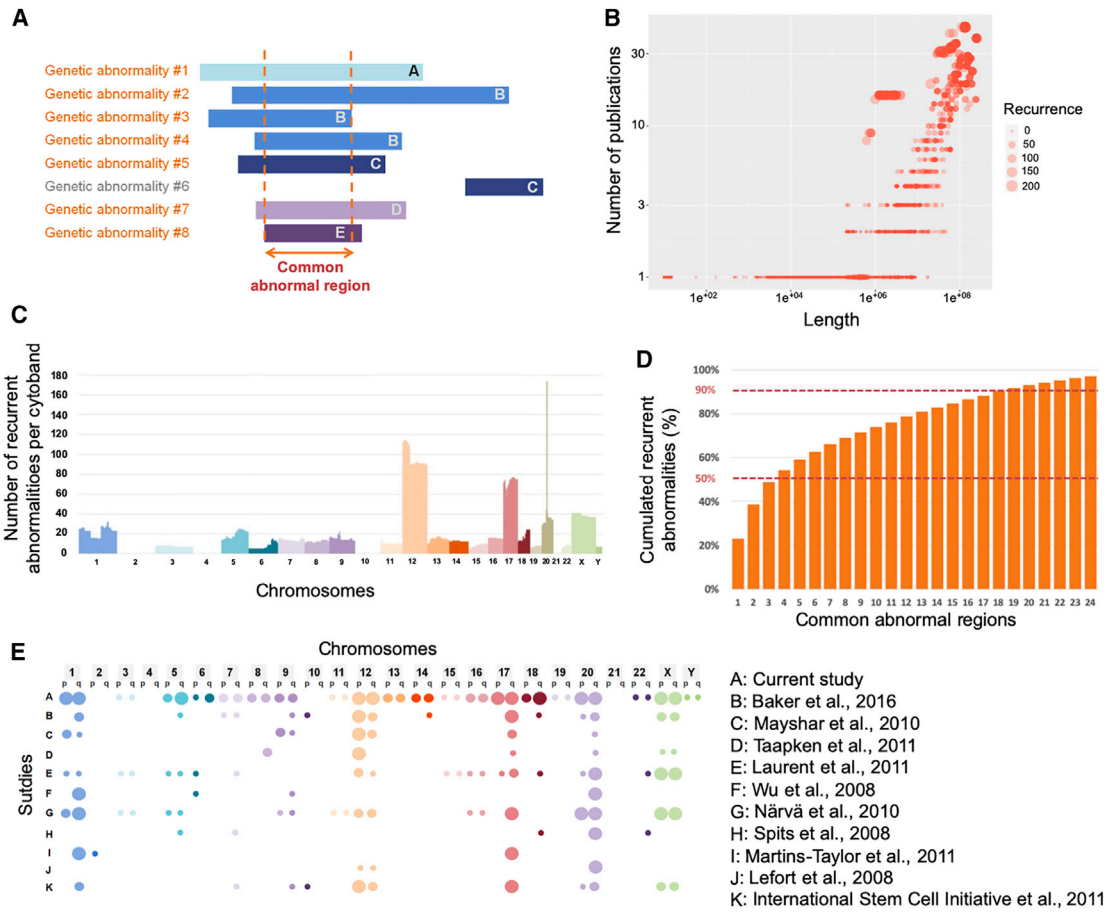


Figure 2. Definition and Analysis of Recurrent Genetic Abnormalities in hPSC

(A) Graphic representation of eight genetic abnormalities (or variants that are not polymorphic) (#1 to #8), from five different articles (A–E), one color for each article. Abnormalities in orange are recurrent genetic abnormalities because they overlapped and were from five different articles.

(B) Dot plot showing the length of all genetic abnormalities larger than 10 bp and that were not polymorphisms (n = 8,284) (x axis) versus the number of different articles that described these overlapping abnormalities (y axis).

(C) Bar plot showing the 738 recurrent genetic abnormalities according to their genomic coordinates (one color per chromosome). y axis, number of recurrent abnormalities.

(D) Percentage of cumulated recurrent abnormalities found in the 24 common abnormal regions with the most recurrent genetic abnormalities.

(E) Comparison of the recurrent abnormalities reported in this work and in a selection of other major original papers or reviews. Colors, chromosome; large bubble, high recurrence; average bubble, intermediate recurrence or no quantitative information; small bubble, low recurrence; no bubble, no recurrence. References found in Supplemental Information section.

lines (cell-DNA) without (HY03, UHOMi001-A, iCOPD2, and iCOPD9) and with genetic abnormalities (RSP4: chromosome 20 triploidy), using ddPCR and primer pairs that target chromosome X or chromosome 20. We could observe two chromosome 20 copies in DNA samples from normal hPSC lines, and three copies in the RSP4 cell-DNA sample (Figures 3A and S2). Second, to test the sensitivity of our approach, we prepared cell-DNA from UHOMi001-A cells (diploid line) mixed with increasing percentages (0%–100%) of HD291 cells (chromosome 12q trisomy). Our ddPCR approach could detect the presence of the chromosome abnormality, starting from the sample containing 10% of HD291 cells (Figure 3B).

Another major constraint to hPSC genome integrity analysis is the need to dedicate a significant part of the cell culture for this purpose. Therefore, we investigated whether genome integrity could be assessed using DNA extracted from hPSC culture supernatants. We found that when using supernatant-derived DNA

Table 1. Common Abnormal Regions

Rank	Chromosome	Start	End	n	%	Cumulated %
1	20	29 848 383	30 754 613	169	22.9%	22.9%
2	12	11 937 418	25 556 120	116	15.7%	38.6%
3	17	51 281 495	52 293 893	75	10.2%	48.8%
4	X	1	60 600 000	39	5.3%	54.1%
5	1	172 900 000	185 800 000	34	4.6%	58.7%
6	5	104 500 000	117 404 202	26	3.5%	62.2%
7	18	56 200 000	61 600 000	24	3.3%	65.4%
8	17	7 211 004	8 044 174	21	2.8%	68.3%
9	7	132 600 000	133 785 759	18	2.4%	70.7%
10	9	68 700 000	114 900 000	17	2.3%	73.0%
11	11	2 800 000	10 700 000	17	2.3%	75.3%
12	13	87 700 000	101 700 000	16	2.2%	77.5%
13	16	1	90 354 753	16	2.2%	79.7%
14	1	16 200 000	17 074 942	14	1.9%	81.6%
15	8	93 300 000	127 230 818	13	1.8%	83.3%
16	14	19 152 018	107 349 540	13	1.8%	85.1%
17	6	130 300 000	139 000 000	12	1.6%	86.7%
18	15	67 200 000	67 300 000	11	1.5%	88.2%
19	3	60 799	26 400 000	9	1.2%	89.4%
20	22	24 268 025	49 570 503	9	1.2%	90.7%

(supernatant-DNA), the concentration of DNA fragments, estimated by qPCR, increased progressively during hPSC culture and was at a highest after 7 days in culture, when cell confluence was higher than 70% (Figure 3C). To help elucidating the origin of the supernatant-DNA, we analyzed floating cells by staining with annexin V and 7-amino-actinomycin followed by flow cytometry (Figure 3D). This analysis showed that 74.5% and 56.5% of floating cells were apoptotic, 12.2% and 6.6% were dead, and 13.4% and 36.9% were viable after single-cell and mechanical passaging, respectively. We also measured the DNA integrity index by qPCR using the ALU115 primers that amplify both short (apoptotic) and long (non-apoptotic) DNA fragments, and the ALU247 primers that amplify only long non-apoptotic DNA fragments (Umetani et al., 2006). The ALU115/ALU247 ratio was higher than 2, showing that the supernatant contained a majority of apoptotic cells (Figure S3A). We also investigated the influence of the interval before DNA extraction and of the number of freeze-thaw cycles on the quantity of DNA extracted from cell culture supernatant samples (Figures S3B and S3C). These results demonstrated that stable and significant quantities of DNA can be extracted from cell culture supernatant samples.

Finally, we observed an excellent agreement between the results obtained using cell-DNA and the corresponding supernatant-DNA (supernatant collected at days 5–7 of culture) (Figure 3E). This shows that ddPCR offers the sensitivity required for evaluating genomic integrity in supernatant-DNA, and that supernatant-DNA could be used instead of cell-DNA.

We then evaluated whether supernatant-DNA could be used to perform focused ddPCR (in culture supernatant digital PCR test, iCS-digital test) using a panel of six commercial pre-designed probes that correspond to or are close to the six most common abnormal regions (chromosomes 20q, 12, 17, X, 1, and 5). These six probes target genome regions that comprise more than 50% of all recurrent genetic abnormalities found in hPSCs (61% cumulated coverage of recurrent abnormalities). We analyzed supernatant-DNA from two hPSC lines with abnormal karyotype (HD129

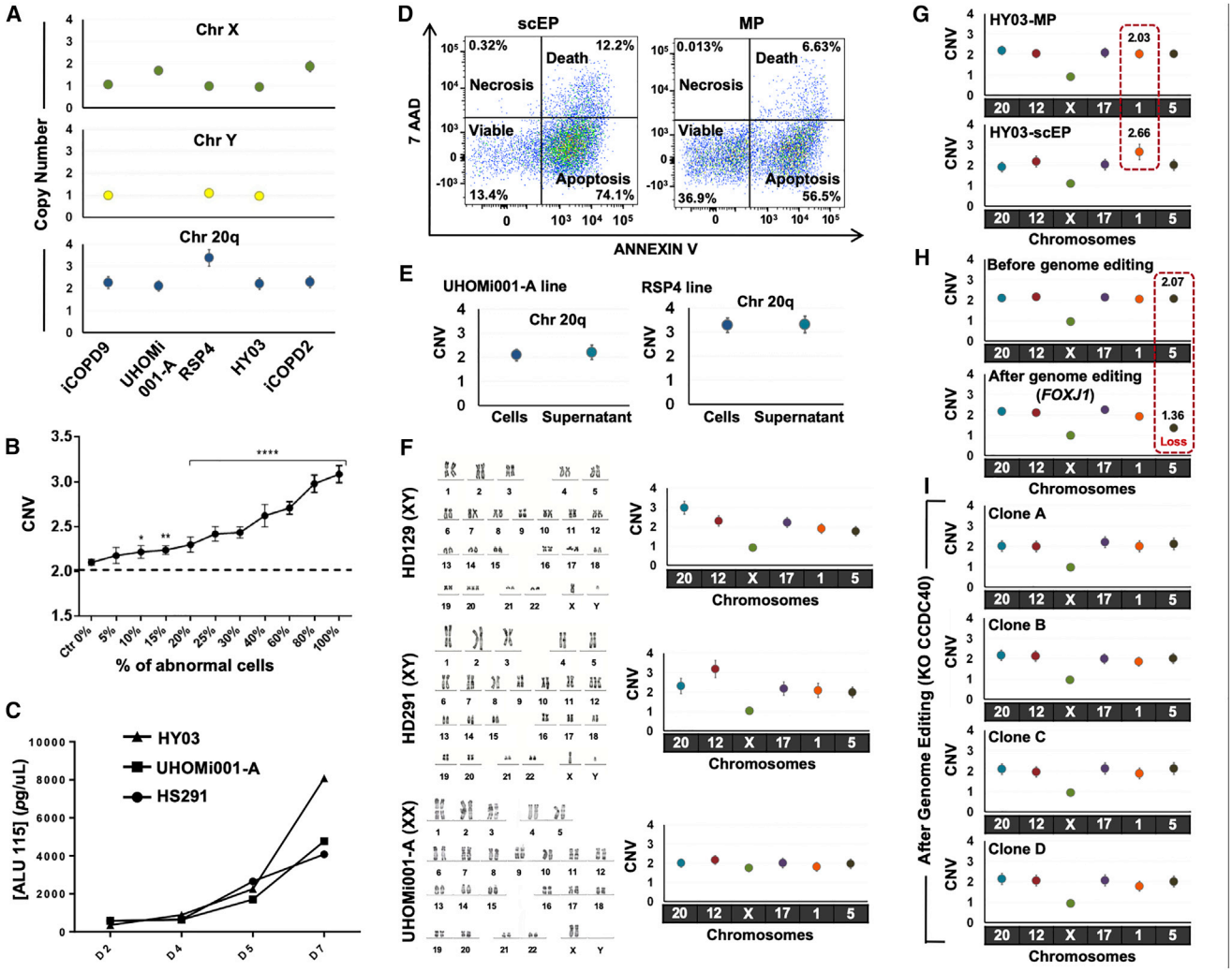


Figure 3. Detecting Recurrent Genetic Abnormalities by Focused ddPCR

Each ddPCR data point is obtained from one sample using Poisson statistics and error bars indicate the Poisson distribution at 95% confidence intervals.

(A) Copy-number variation analysis using droplet digital PCR and DNA extracted from different hPSC lines in culture.

(B) Sensitivity of the droplet digital PCR method for detecting increasing percentages (from 0% to 100%) of hPSCs harboring a trisomy 12 within a sample of euploid hPSCs. The panels represent three independent experiments ($p < 0.05$, Student's t test).

(C) Quantification of DNA in supernatant samples from one hESC line (HS291) and two hiPS lines (HY03 and UHOMi001 cells) cultured in E8 medium on Geltrex matrix. Supernatant was collected at the indicated days (D) after seeding (75,000 cells/well in a 35-mm plate). DNA was extracted from 300 μL of supernatant and quantified by ALU-qPCR with ALU115 primers.

(D) The percentage of apoptotic, necrotic, and viable cells in supernatant samples collected at day 5 was evaluated by flow cytometric analysis after staining with annexin V and 7-amino-actinomycin (7-AAD). Every dot corresponds to a single cell. ScEP, single-cell enzymatic passaging; MP, mechanical passaging.

(E) Copy number of chromosome 20q measured by ddPCR using genomic DNA from cells and supernatant as template. The error bar varies in function of the DNA source (cells or supernatant).

(F) The iCS-digital test using six probes for chromosomes 20q, 12, X, 17, 1, and 5 can identify aneuploidy. The hPSC lines HD129 (chromosome 20 triploidy), HD291 (chromosome 12 triploidy), and UHOMi001-A (euploid) were analyzed by karyotyping (classical G-banding) and with the iCS-digital test using probes targeting common abnormal regions on chromosomes 20q, 12, X, 17, 1, and 5. Karyotype images are reprinted from Stem Cells Dev. 24(5):653-62, 2015 and Stem Cell Res. 33:15-19, 2018 with permission from Elsevier.

(G) Identification of genome modifications associated with culture conditions using the iCS-digital test. After 15 passages using single-cell enzymatic passaging, the HY03 hiPSC line displayed a copy-number gain on chromosome 1 (2.66), suggesting a mosaic cell population that comprises an abnormal clone with at least three genome copies at the probe location.

(legend continued on next page)

and HD291: chromosome 20 and 12 triploidy, respectively) (Bai et al., 2015) and from one diploid line (UHOMi001-A). The iCS-digital test results overlapped with those obtained by karyotyping (Figure 3F). In conclusion, targeted ddPCR can efficiently detect CNVs and can be carried out using supernatant-DNA.

Routine Screening of hPSC Lines during Cell Culture and after CRISPR Gene Editing Using the iCS-Digital Test

The simplicity of the iCS-digital test could allow the routine screening of the most recurrent genetic abnormalities in hPSC lines, particularly when using single-cell or small-clump passaging, a major cause of genomic alterations (Bai et al., 2015). The iCS-digital test revealed that the HY03 hiPSC line, which was euploid at passage 5, harbored a chromosome 1 gain at passage 15 after single-cell passaging, but not after mechanical passaging (Figure 3G).

Cell reprogramming and genome editing using CRISPR/Cas9 require clonal expansion/selection that favors the emergence of abnormal clones. The iCS-digital test did not detect any alteration at chromosomes 20q, 12, 17, X, 1, and 5 in the hiPSC line HY03 before genome editing (Figure 3H). Conversely, after introduction of the mCherry cassette at the 3′ of the *FOXJ1* gene, or knockout (KO) of the *CCDC40* gene using the CRISPR/Cas9 methodology, the iCS-digital test showed that the single correctly edited clone (passage 7 after editing) harbored a CNV in the long arm of chromosome 5 (copy number = 1.3) (Figure 3H), while the four CCDC40-KO clones analyzed appeared euploid (six regions checked) (Figure 3I). Taken together, our results show that focused ddPCR can be used to rapidly screen iPSCs after derivation, during cell culture or amplification, and after cell cloning in settings, such as gene editing.

DISCUSSION

The present study is, to our knowledge, the largest meta-analysis of hPSC genetic abnormalities in more than 100 different research articles from many different laboratories and cell lines. This allowed us to propose a quantitative threshold to define recurrent genetic abnormalities in hPSCs and to test this threshold. Hence, a recurrent genetic abnormality is an abnormality that shares part of its abnormal sequence with other abnormalities that have been reported in at least five different publications. Our threshold should favor the detection of abnormalities that are common to different hPSC lines and different laboratories, as opposed to abnormalities that are specific to a unique cell line, a hiPSC donor cell, or to the culture conditions used in one laboratory. As expected, many regions were already reported, but some are new. Figure 2E compares our recurrent abnormalities with those reported in a selection of other major original papers or reviews. Our study adds to previous studies a rigorous definition of recurrence, and the analysis of a larger number of different samples compared with all previous works. Our analysis confirmed that structural variants, including CNVs of various lengths, are among the main recurrent genetic abnormalities in hPSC lines. It has been proposed that the high hPSC susceptibility to mitotic division errors and to the upregulation of anti-apoptotic proteins contributes to the high frequency of aneuploidy and CNVs at specific genomic locations (Zhang et al., 2019). For instance, chromosome 20q11.21 gain leads to upregulation of the anti-apoptotic protein BCL-XL (Zhang et al., 2019). Noteworthy, with the exception of the *BCL2L1* locus, we did not find convergence on any specific chromosome locus, even by using the large dataset collected for our meta-analysis. As illustrated by Figure 2C, most recurrent zones are large parts of chromosome arms. Although regions with more recurrences than others can be identified, only the *BCL2L1* locus has a pattern pointing to a precise zone. This suggests that recurrence is the consequence of several genes dispersed on a large chromosome section that must be simultaneously upregulated through copy-number gains (or downregulated for the rare cases of recurrent deletions) to promote *in vitro* culture fitness. For example, the most recurrent part of chromosome X (X:1, 60,600,000) contains several genes that are annotated as involved in regulation of cell cycle, DNA repair, apoptosis, stemness, or cancer. Consequently, two or more of these genes could be central players in the recurrence of copy-number gains at this locus (Figure S4). Recent analyses of hPSC NGS data identified recurrent small mutations in oncogenes and tumor suppressor genes, such as *TP53*, with a frequency still to be determined (Avior et al., 2019; Merkle et al., 2017). We expect that additional recurrent abnormalities will be uncovered, but this requires collecting many genome sequencing data, especially whole genome data. This will allow drawing the complete map of genetic abnormalities of hPSCs. In addition, one pitfall of our current

(H) Analysis of genome stability using the iCS-digital test in the male hiPSC line HY03 before (mechanical passage M53, clumps passage Cl2, single-cell passage SC11) and seven passages after genome editing.
(I) CNV characterization of chromosome 20q, 12, X, 17, 1, and 5 in four CCDC40_KO HY03 clones obtained using CRISPR/Cas9 technology. No abnormality was detected in any of the four clones.

analysis is that samples used in our work were analyzed using methods which have biases, and this could have influenced the recurrence score. Whole-genome NGS would provide the unbiased data necessary for the comprehensive identification of anomaly recurrence in PSC. Finally, the comparison with parental samples (in the case of iPSC lines), and the comparison with data from a large number of healthy volunteer DNA samples is necessary to rigorously distinguish recurrence from polymorphisms because the databases routinely used to identify polymorphisms are neither comprehensive nor error-free.

Nonetheless, the recurrence distribution is biased, and this allows studying common abnormal regions by targeted PCR. For example, by targeting only the four most common abnormal regions, more than 50% of all recurrent genetic abnormalities are covered, and more than 90% by targeting 20 regions (Table 1). Screening of recurrent genetic abnormalities is paramount to claim that hPSCs are normal (Bai et al., 2013; International Stem Cell Initiative et al., 2011). Although a targeted approach is by definition not exhaustive, it is an effective strategy to rule out the most frequent and functionally damaging abnormalities found in hPSC lines. We demonstrated that this strategy can be carried out simply by using DNA from culture supernatant using ddPCR. Other competitive technologies exist, but none can meet all the following requirements: range, resolution, sensitivity, low cost, data analysis workload, and complex result interpretation. For example, DNA microarrays, and whole-exome and whole-genome sequencing can be used for large-scale hPSC genome integrity analyses, but are limited by the long time required for sample processing and their high cost. Karyotyping is still the gold standard for hPSC analysis, but it is time-consuming. Rapid tests for the routine screening of cultured cells are mostly provided by PCR-based technologies (Baker et al., 2016). The advantages and disadvantages of the techniques used to assess hPSC genome integrity are listed in Table S2. Ultimately, we anticipate that NGS will become the primary technique for assessing hPSC quality because sequencing costs continue to decrease.

Genomic integrity of hPSCs in culture should be frequently assessed. We recently noted in a series of 25 consecutive studies on hiPSCs that the current genomic screening practices were unsatisfactory because no genomic integrity follow-up was carried out for any of the hiPSC lines (Assou et al., 2018). This could be explained by the labor and costs involved in the implementation of classical screening techniques, such as karyotyping (see also Table S2). Therefore, a simple test that can rapidly rule out the most frequent recurrent genomic abnormalities might promote adhesion to good practices for hPSC genomic integrity screening. Moreover, karyotyping can miss abnormalities that are smaller than 5–10 Mb. For instance, we found that among all 170 recurrent genomic abnormalities on chromosome 20, 168 overlapped with 20q11.21, and among them 135 were smaller than 5 Mb. By contrast, the probes of the iCS-digital test can detect all 168 20q11.21 abnormalities. Therefore, PCR-based tests are appropriate for early detection of genetic abnormalities that arise during cell culture to avoid performing experiments with compromised hPSC lines.

In conclusion, we used a large dataset of hPSC genomic abnormalities based on more than 100 publications to define recurrent genetic abnormalities, including a set of common abnormal genomic regions that involve more than 90% of all recurrent abnormalities. This offered the opportunity to develop a targeted ddPCR approach that could be performed on culture supernatant. Therefore, we propose a simple test based on supernatant-DNA (iCS-digital test) that could be used to routinely screen cultured hPSC lines for excluding the presence of common CNVs.

EXPERIMENTAL PROCEDURES

Recurrence Analysis

All genetic abnormalities were converted using their GRCh37A genome coordinates. The complete pipeline can be found in Supplemental Experimental Procedures. In brief, to define recurrent genetic abnormalities, each abnormality >10 bp and that was not a polymorphism (n = 8,284 abnormalities in total) was compared with all the others, and abnormalities with a reciprocal overlap >0.33 and larger than 0.2 Mb, and the number of publications from which these abnormalities originated were counted, using R. Abnormalities that overlapped with other abnormalities that came from at least four other publications (number of total publications ≥ 5) were defined as recurrent and kept. This process was carried out iteratively and rapidly converged. Loss and gain were considered indiscriminately because the aim was to identify common abnormal regions.

Cell Culture

All PSC were derived with the ethical oversight of a Comité de protection des Personnes (CPP). PSC lines were maintained on Geltrex matrix (Thermo Fisher Scientific) in Essential 8 Medium (Thermo Fisher Scientific) or on Matrigel (BD Biosciences) in mTeSR1medium (STEMCELL Technologies), according to the manufacturer's instructions. Cells were grown in 35-mm dishes at 37°C and were passaged either mechanically or dissociated into single cells or small clumps every week. More details for cell reprogramming and cell passaging can be found in Supplemental Experimental Procedures.

Culture Medium Collection, Nucleic Acid Extraction, Quantification of Supernatant-DNA, Flow Cytometry, ddPCR, and Genome Editing using CRISPR/Cas9

See Supplemental Experimental Procedures for details.

Statistical Analysis

For ddPCR, absolute quantification was based on the number of positive droplets and Poisson sampling statistics, as follows: $\lambda = -\ln(1 - k/n)$, where k is the number of positive droplets, n is the total number of droplets, and λ is the mean copies per droplet. To test the statistical difference of the ddPCR results between samples with different percentages of trisomic cells, first the values were normalized to those of the 0% sample within each biological replicate. Poisson statistics and Student's t test were applied; $p < 0.05$ defined statistical significance.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.stemcr.2019.12.004.

AUTHOR CONTRIBUTIONS

S.A., N.G., and J.D.V. designed the study and analyzed data. S.A., M.C., J.M., and E.A., performed the experiments. M.P., J.B., C.S., C.B., M.F., and E.A., collected and analyzed data. T.C. and A.B. analyzed and interpreted the data. S.A. and J.D.V. wrote the paper. J.-M.L. revised the manuscript. All authors approved the final version prior to submission.

CONFLICT OF INTERESTS

S.A. and J.D.V. filed a patent entitled "Non-invasive methods for assessing genetic integrity of pluripotent stem cells" (priority number: EP20150306389) and are co-founder, shareholder and scientific advisor of Stem Genomics SAS that acquired the exploitation rights of this patent.

ACKNOWLEDGMENTS

We thank Camille Novo for contributing to the collection of PSC genomic abnormalities, Elena Hauser for technical help and Guilhem Requirand for his expertise in flow cytometry. This study was funded by INSERM, France, ANR (INGESTEM), France and the SATT AxLR, France.

Received: May 25, 2019
Revised: December 2, 2019
Accepted: December 4, 2019
Published: January 02, 2020

REFERENCES

Assou, S., Bouckenheimer, J., and De Vos, J. (2018). Concise review: assessing the genome integrity of human induced pluripotent stem cells: what quality control metrics? Stem Cells *36*, 814–821.

Avior, Y., Eggan, K., and Benvenisty, N. (2019). Cancer-related mutations identified in primed and naive human pluripotent stem cells. Cell Stem Cell *25*, 456–461.

Bai, Q., Desprat, R., Klein, B., Lemaître, J.-M., and De Vos, J. (2013). Embryonic stem cells or induced pluripotent stem cells? A DNA integrity perspective. Curr. Gene Ther. *13*, 93–98.

Bai, Q., Ramirez, J.-M., Becker, F., Pantesco, V., Lavabre-Bertrand, T., Hovatta, O., Lemaître, J.-M., Pellestor, F., and De Vos, J. (2015). Temporal analysis of genome alterations induced by single-cell passaging in human embryonic stem cells. Stem Cells Dev. *24*, 653–662.

Baker, D., Hirst, A.J., Gokhale, P.J., Juarez, M.A., Williams, S., Wheeler, M., Bean, K., Allison, T.F., Moore, H.D., Andrews, P.W., et al. (2016). Detecting genetic mosaicism in cultures of human pluripotent stem cells. Stem Cell Reports *7*, 998–1012.

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

Lefort, N., Feyeux, M., Bas, C., Féraud, O., Bennaceur-Griscelli, A., Tachdjian, G., Peschanski, M., and Perrier, A.L. (2008). Human embryonic stem cells reveal recurrent genomic instability at 20q11.21. Nat. Biotechnol. *26*, 1364–1366.

Lefort, N., Perrier, A.L., Laâbi, Y., Varela, C., and Peschanski, M. (2009). Human embryonic stem cells and genomic instability. Regen. Med. *4*, 899–909.

Merkle, F.T., Ghosh, S., Kamitaki, N., Mitchell, J., Avior, Y., Mello, C., Kashin, S., Mekhoubad, S., Ilic, D., Charlton, M., et al. (2017). Human pluripotent stem cells recurrently acquire and expand dominant negative P53 mutations. Nature *545*, 229–233.

Umetani, N., Kim, J., Hiramatsu, S., Reber, H.A., Hines, O.J., Bilchik, A.J., and Hoon, D.S.B. (2006). Increased integrity of free circulating DNA in sera of patients with colorectal or periampullary cancer: direct quantitative PCR for ALU repeats. Clin. Chem. *52*, 1062–1069.

Zhang, J., Hirst, A.J., Duan, F., Qiu, H., Huang, R., Ji, Y., Bai, L., Zhang, F., Robinson, D., Jones, M., et al. (2019). Anti-apoptotic mutations desensitize human pluripotent stem cells to mitotic stress and enable aneuploid cell survival. Stem Cell Reports *12*, 557–571.

Stem Cell Reports, Volume *14*

Supplemental Information

Recurrent Genetic Abnormalities in Human Pluripotent Stem Cells: Definition and Routine Detection in Culture Supernatant by Targeted Droplet Digital PCR

Said Assou, Nicolas Girault, Mathilde Plinet, Julien Bouckenheimer, Caroline Sansac, Marion Combe, Joffrey Mianné, Chloé Bourguignon, Mathieu Fieldes, Engi Ahmed, Thérèse Commes, Anthony Boureux, Jean-Marc Lemaître, and John De Vos

Figure S1

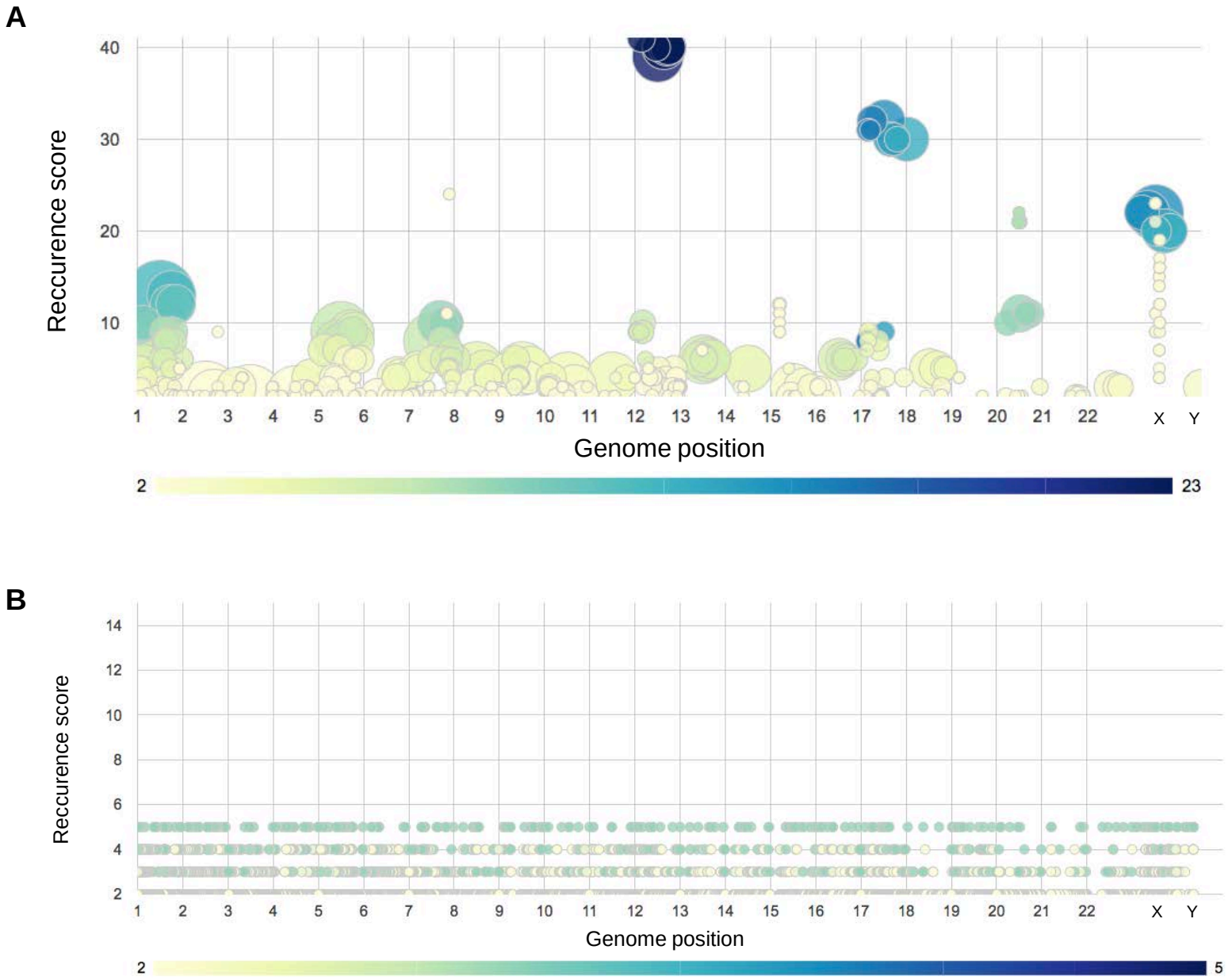

A
Reccurence score
40
30
20
10
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
Genome position
2
23
B
Reccurence score
14
12
10
8
6
4
2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
Genome position
2
5

Figure S2

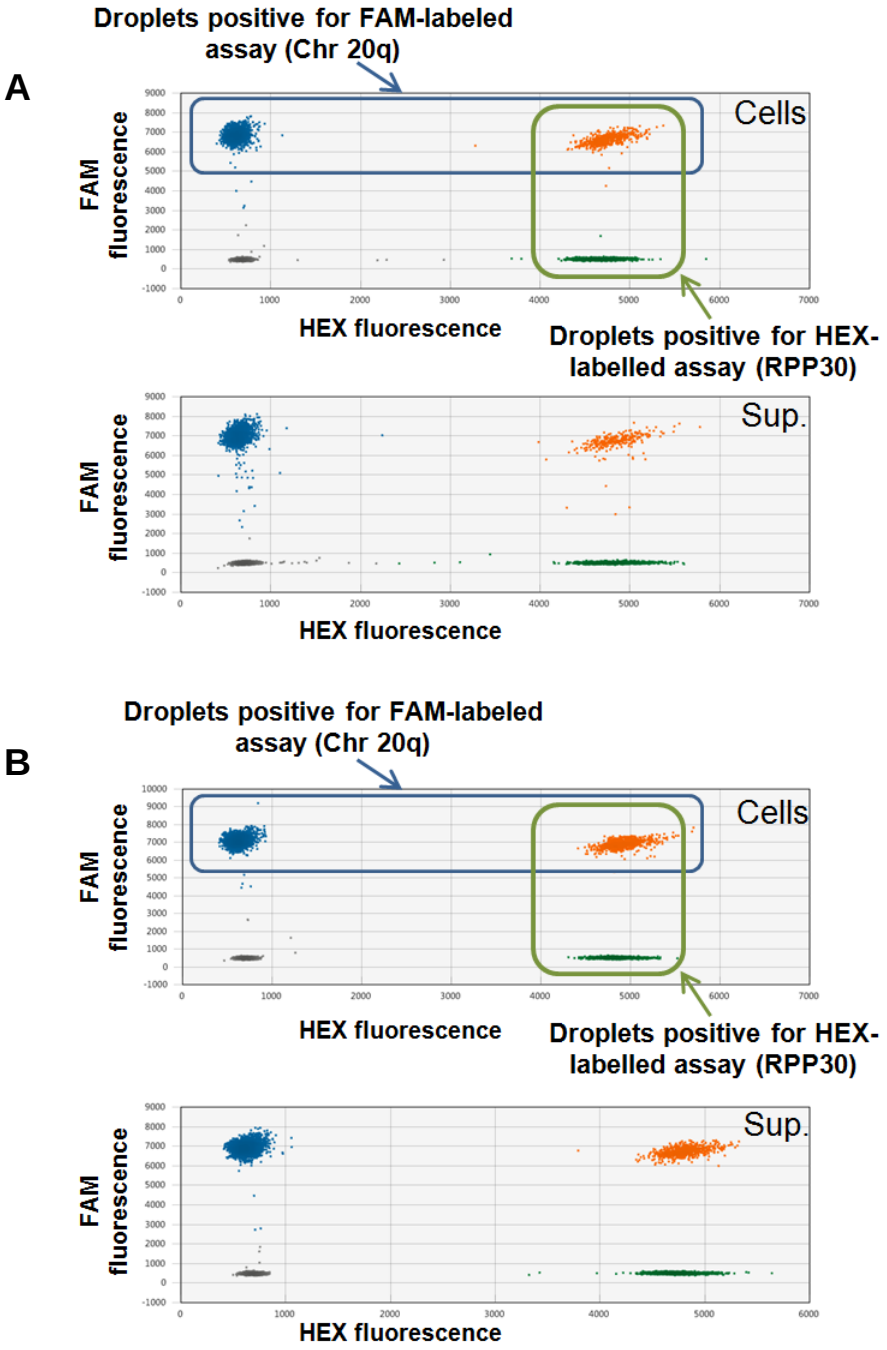

A
Droplets positive for FAM-labeled assay (Chr 20q)
Cells
FAM fluorescence
HEX fluorescence
Droplets positive for HEX-labelled assay (RPP30)
Sup.
FAM fluorescence
HEX fluorescence
B
Droplets positive for FAM-labeled assay (Chr 20q)
Cells
FAM fluorescence
HEX fluorescence
Droplets positive for HEX-labelled assay (RPP30)
Sup.
FAM fluorescence
HEX fluorescence

Figure S3

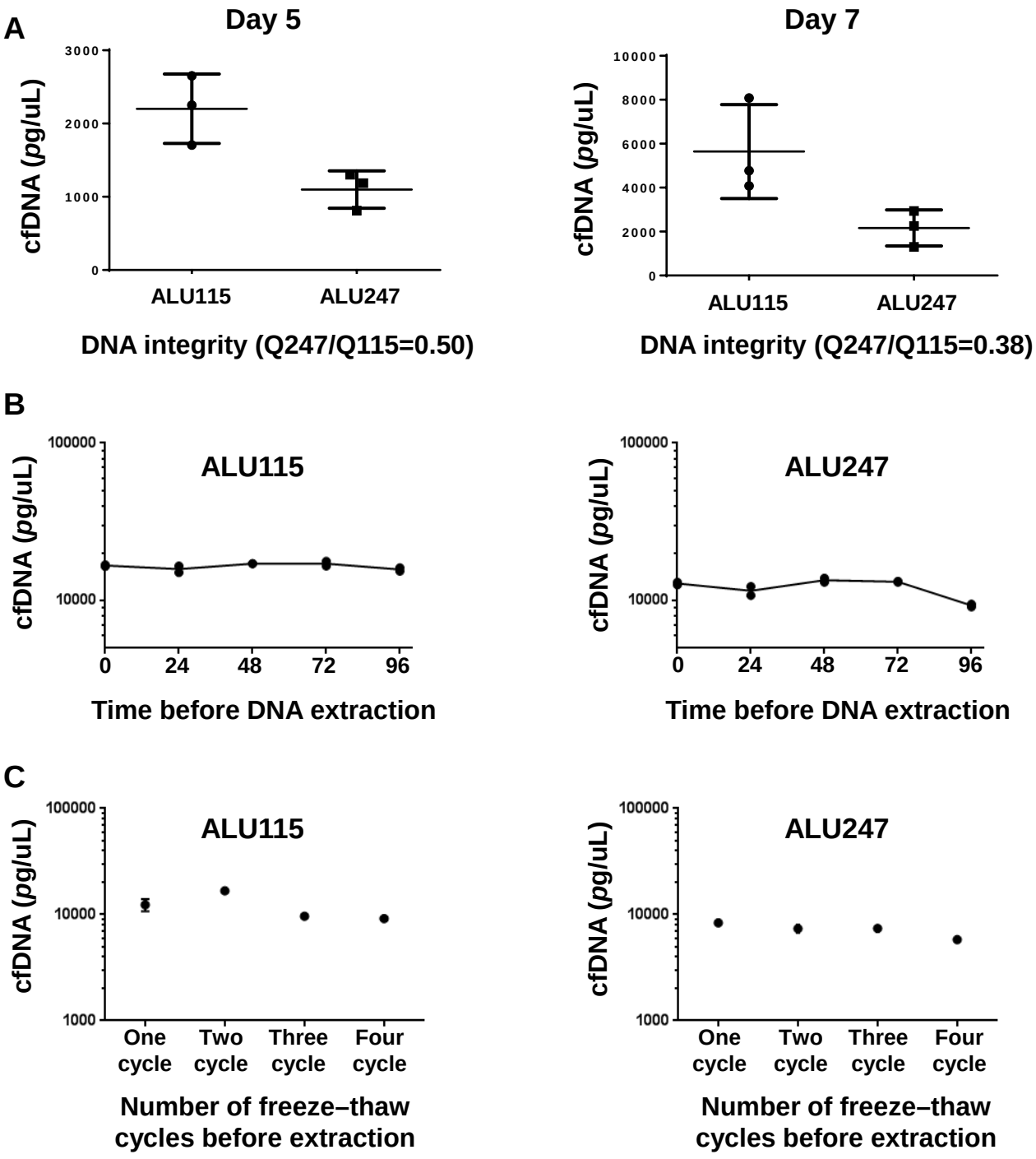
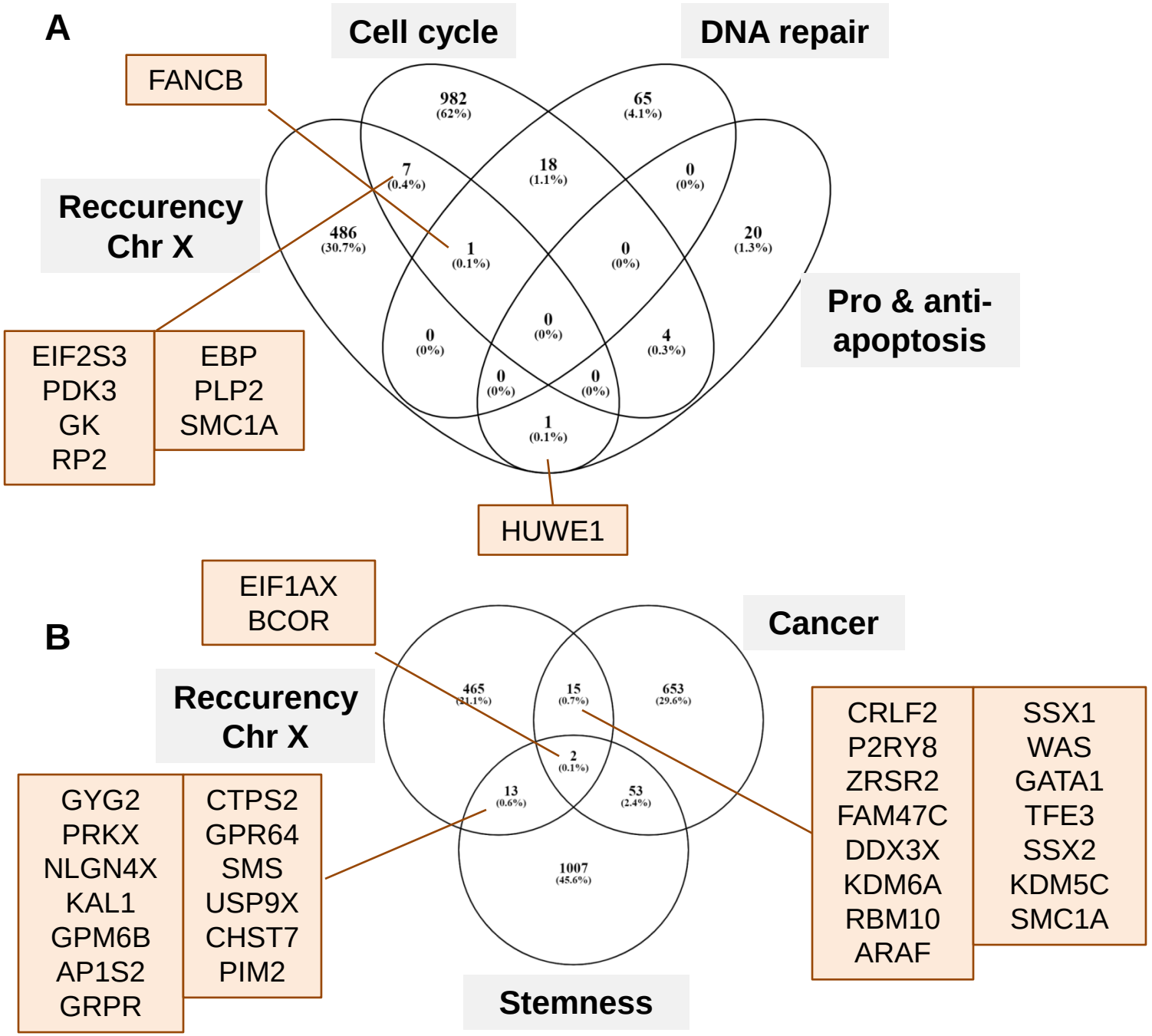

A
Day 5
cfDNA (pg/uL)
3000
2000
1000
0
ALU115
ALU247
DNA integrity (Q247/Q115=0.50)
Day 7
cfDNA (pg/uL)
10000
8000
6000
4000
2000
0
ALU115
ALU247
DNA integrity (Q247/Q115=0.38)
B
ALU115
cfDNA (pg/uL)
100000
10000
0
24
48
72
96
Time before DNA extraction
ALU247
cfDNA (pg/uL)
100000
10000
0
24
48
72
96
Time before DNA extraction
C
ALU115
cfDNA (pg/uL)
100000
10000
1000
One cycle
Two cycle
Three cycle
Four cycle
Number of freeze–thaw cycles before extraction
ALU247
cfDNA (pg/uL)
100000
10000
1000
One cycle
Two cycle
Three cycle
Four cycle
Number of freeze–thaw cycles before extraction

Figure S4


A
Cell cycle
DNA repair
FANCB
Reccurency Chr X
982 (62%)
65 (4.1%)
7 (0.4%)
18 (1.1%)
0 (0%)
486 (30.7%)
1 (0.1%)
0 (0%)
20 (1.3%)
0 (0%)
0 (0%)
4 (0.3%)
0 (0%)
0 (0%)
1 (0.1%)
Pro & anti-apoptosis
EIF2S3
PDK3
GK
RP2
EBP
PLP2
SMC1A
HUWE1
B
EIF1AX
BCOR
Cancer
Reccurency Chr X
465 (21.1%)
15 (0.7%)
653 (29.6%)
2 (0.1%)
13 (0.6%)
53 (2.4%)
1007 (45.6%)
CRLF2
P2RY8
ZRSR2
FAM47C
DDX3X
KDM6A
RBM10
ARAF
SSX1
WAS
GATA1
TFE3
SSX2
KDM5C
SMC1A
GYG2
PRKX
NLGN4X
KAL1
GPM6B
AP1S2
GRPR
CTPS2
GPR64
SMS
USP9X
CHST7
PIM2
Stemness

Supplemental figure legends

Figure S1

A. Bubble plot showing recurrent DNA abnormalities >10bp in hPSC lines. Each bubble represents a region where several DNA abnormalities from several publications are found. The bubble size is proportional to the region length and the color indicates the number of publications that reported abnormalities in that region. The horizontal axis corresponds to the genome position and the vertical axis corresponds to the recurrence score (see Material and Methods).

B. Same as in A, but for abnormalities and variants ≤10bp.

Figure S2

A. 2D plot showing a typical result obtained from quantifying chromosome 20q CNV in a euploid hPSC line (UHOMi001-A) using DNA from cells and supernatant (Sup.), as indicated. Blue: droplets positive for the target CNV (Chr20q); green: droplets positive for the reference gene (*RPP30*); grey: negative droplets (containing no target or reference genes); orange: droplets positive for both the reference and target genes.

B. FACS-like plot showing a typical result obtained from quantifying chromosome 20q CNV in an aneuploid hPSC line (RSP4) using DNA from cells and supernatant (Sup.), as indicated. Same color code as in A.

Figure S3

A. Concentration of DNA in supernatant samples and effects of various pre-analytical conditions. Quantification of supernatant-DNA by quantitative PCR with two sets of ALU primers (115 and 247 bp) that amplify DNA fragments of different lengths in supernatant samples collected from three hPSC lines at day 5 and day 7. The ALU 115 and ALU 247 values are significantly different ($p<0.05$). The Q247/Q115 ratio indicates the DNA integrity value (Q115 corresponds to the DNA concentration obtained using the ALU 115 primers and Q247 to the concentration obtained with the ALU 247 primers). The mean

Q247/Q115 ratio in supernatant samples collected at day 5 and day 7 was 0.50 and 0.38 respectively, suggesting that the DNA released in the supernatant originates mostly from apoptotic rather than necrotic cells.

B. Comparison of supernatant-DNA amount when supernatant was stored at room temperature for different times (24h, 48h, 72h and 96h) before extraction. DNA concentrations (pg/uL) were determined using the ALU115 and ALU247 primers. DNA concentration was not affected by keeping supernatant at room temperature.

C. Comparison of supernatant-DNA amount according to the number of freeze–thaw cycles before extraction. DNA concentration slightly decreased after the supernatants underwent 3-4 freeze–thaw cycles before extraction but the DNA quality was not affected.

Figure S4

A. Intersection of genes in the recurrent regions with cell cycle genes (signature obtained by comparison of samples with high proliferation index, such as rapidly dividing early CD71+erythroid progenitors and CD105+ endothelial cells, with somatic samples, (Assou et al., 2009), DNA repair genes (Wood et al., 2005), pro-and anti-apoptotic genes (BCL2 family members).

B. Intersection with pluripotency-associated genes (previously published data set with a consensus PSC stemness gene list, (Assou et al., 2007) and cancer genes (Cancer Gene Census, http://www.sanger.ac.uk/genetics/CGP/Census/). Venn diagrams show the number of genes in each comparison and the genes shared.

Supplemental Tables

Table S2: The most frequently used methods for genetic stability assessment with their main advantages and disadvantages.

	Cells		**DNA**		
	G-banding Karyotype	**FISH**	**Microarray**	**Whole genome/ exome sequencing**	**PCR/ddPCR**
Resolution	Poor	Excellent			
Quantity	Several cells in metaphase	Several cells	>1000 ng of DNA		>1ng of DNA
Sensitivity	≥10%	1%	10 - 20%	10 - 20%	10 - 20%
Price (dollars)	400-600	200-300	500	1500	100-200
Results	Require specialist staff		Need for a bio-informatician for analysis		Easy (software)
Timing	1-2 weeks				1 day
Advantages	- Gold standard for the detection of aneuploidy, polyploidy, and other large chromosomal imbalances	- High sensitivity and reproducibility	- Provide information on DNA regions with gains or losses	- Very high and scalable throughput, sensitivity and accuracy - Assess the whole genome at single-base resolution	- High precision for the CNV and SNV detection at a reasonable cost
Disadvantages	- High number of metaphases are needed - Cannot detect sub-karyotypic variants	- Does not allowthe comprehensive screening of chromosomal aberrations	- Cannot detect balanced rearrangements, such as inversions.	- Demanding computational power - Huge data analysis workload - Complex result interpretation -Expensive	- Balanced rearrangements not detected - Does not allow the comprehensive screening of chromosomal aberrations

Table S3: List Bio-Rad ddPCR assay ID

Chromosome	**Unique Assay ID (Bio-Rad)**	**Gene symbol**	**Locus**
20	#dHsaCP2506319	ID1	20q11
12	#dHsaCP1000374	NCAPD2	12p13
X	#dHsaCP2506654	STS	Xp22

17	# dHsaCP1000054	RPS6KB1	17q23
1	#dHsaCP1000482	SOAT1	1q25
5	#dHsaCNS501868922	PITX1	5q31
Reference	# dHsaCP2500350	RPP30	10q23

Supplemental Experimental Procedures

Recurrence scores

A first analysis was carried out using a recurrence score (RS) for data split in two datasets (>10 bp and ≤10 bp) from which polymorphic data (sequences present in dbSNP or DGV) were removed. RS of the first analysis was computed by comparing each abnormality to all the others and by identifying abnormalities with a reciprocal overlap of at least 0.2. Regions with a reciprocal overlap higher than 80% were merged. For each overlap, RS was computed as follows: RS = a * s, where (a) is the number of abnormalities that contributed to define this overlap (identical abnormalities from the same cell line in the same study were counted only once) and (s) the number of different studies from which these overlapping abnormalities came from.

To define recurrent genetic abnormalitiesn for each abnormality >10 bp we computed an overlap 'Ov' as follow: Ov = End Ov – Start Ov, where End Ov is the minimal value of the ends of both abnormalities that are compared, and Start Ov is the maximal of their starts. Equilibrated translocations were excluded.

Bedtools (Quinlan and Hall, 2010) was used to identify common abnormal regions. The common regions that cover the highest number of genetic abnormalities were calculated in a spreadsheet.

Cell reprogramming and cell passaging

The human hESC lines HD129 and HD291 were derived in our laboratory (Bai et al., 2015). The hiPSC lines UHOMi001-A (Ahmed et al., 2018), RSP4, iCOPD2A1, iCOPD9A2, and HY03 were reprogrammed using the Sendai virus and the CytoTune®-iPS 2.0 Sendai Reprogramming Kit (Thermo Fisher Scientific), and

display all the PSC features: grow as typical PSCs, are positive for pluripotency markers (*OCT4*, *NANOG*, *SOX2*, *TRA1-60*, *TRA1-81*, *SSEA3*, *SSEA4*) and for phosphatase alkaline activity, and can differentiate into cells of all three germ layers. Mechanical passaging was carried out under an inverted microscope in a hood using scalpels. For single-cell enzymatic passaging and clump passaging, colonies were pre-incubated with the Rho-associated protein kinase (ROCK) inhibitor Y-27632 for 1h, and then dissociated with TrypLE™ Select (Invitrogen) or EDTA (Versene Solution, Thermo Fisher Scientific) at 37°C for 10min.

Culture medium collection and nucleic acid extraction

Before passaging, supernatant (1.5 mL) was collected into a safe-lock tube (DNase-free) from cell cultures that were at least 70% confluent. DNA was extracted from 200 µL of supernatant using the QIAmp DNA Mini Blood Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Briefly, 20 µL proteinase K and 200 µL Buffer AL were added to each supernatant sample. After pulse vortexing for 15s, the lysis mixture was incubated in an Eppendorf tube (1.5 mL) at 56°C for 10min. The highly denaturing conditions and elevated temperatures favored the complete release of DNA from any bound proteins. After adding 200 µL of cold ethanol (100%) to the lysates, samples were transferred in QIAamp Mini columns and centrifuged at 6000g for 1min followed by two wash steps (in Buffer AW1 and Buffer AW2) to eliminate contaminants. Then, supernatant-DNA was eluted in 60 µL Buffer AE and stored at -20°C.

Quantification of supernatant-DNA by ALU-qPCR and QuBit

Supernatant-DNA was analyzed by qPCR (LC480, Roche)using the ALU 115 and ALU 247 primers, as previously described in (Umetani et al., 2006). The sequences of the ALU115 primers were: forward, 5′-CCTGAGGTCAGGAGTTCGAG-3′; reverse, 5′-CCCGAGTAGCTGGGATTACA-3′. The ALU247 primers were: forward, 5′-GTGGCTCACGCCTGTAATC-3′; reverse, 5′-CAGGCTGGAGTGCAGTGG-3′. One µL of each eluted supernatant-DNA sample was added to a reaction mixture containing 2X LightCycler480 SYBR Green I master mix (Roche Applied Science, Germany) and 0.25 µM of forward and reverse primers (ALU-115 and ALU-247) as described in (Umetani et al., 2006) in a total volume of 10 µL. Reactions were carried out in 96-well white plates using an EpMotion 5070 Liquid Handling Workstation (Eppendorf). All reactions were performed in triplicate. A negative control (RNAse/DNAse-free water) was included in each run. The supernatant-DNA concentration was determined using a standard curve obtained by successive dilutions of a commercial human genomic DNA sample. DNA integrity was calculated as the ratio of the qPCR results with the two primer sets (ALU115 and ALU247). The ALU115 set amplifies smaller fragments that result from apoptosis and the ALU247 set amplifies only larger fragments that result from necrosis. Supernatant-DNA concentration was quantified using the QuBit dsDNA HS Assay Kit and a Qubit 2.0 fluorometer following the manufacturer's instructions (Life Technologies).

Flow cytometric detection of apoptosis and necrosis using the Annexin-V and 7-AAD assay

The PE Annexin-V Apoptosis Detection Kit I (BD Pharmingen, Ref: 559763) was used to quantify the percentage of apoptotic and necrotic cells in supernatant samples. Briefly, samples were incubated with PE Annexin-V in buffer containing 7-

Amino-Actinomycin (7-AAD) according to the kit protocol (http://www.bdbiosciences.com/ds/pm/tds/559763.pdf), and analyzed by flow cytometry at the Montpellier Resources Imaging (MRI) facility (https://www.mri.cnrs.fr/en). HiPS cells are used as controls for FACS gating.

Digital droplet PCR (ddPCR)

The ddPCR workflow was performed according the Bio-Rad instructions (Bio-Rad QX200 system). Briefly, reactions were set up using one primer pair that targets the region of interest (for instance: CNV-chr20) and a second primer pair that targets the reference gene (*RPP30*). The two primer sets were labeled with different fluorophores (FAM and HEX). DNA (amount) from each sample was added to the TaqMan PCR reaction mixture (final volume of 20 µL) that included 2XSupermix No dUTP (Bio-Rad, Ref: 1863023) and the primer sets. Each reaction mixture was loaded in a disposable plastic cartridge (Bio-Rad) with 70 µL of droplet generation oil (Bio-Rad) and placed in the droplet generator (Bio-Rad). The cartridge was removed from the droplet generator, and the droplets collected in the droplet well were then manually transferred with a multichannel pipette to a 96-well PCR plate. The PCR amplification conditions were: 94°C for 10min, 40 cycles of 94°C for 30s, and 60°C for 1min, followed by 98°C for 10min and ending at 4°C. After amplification, the plate was loaded into the QX200 Droplet Reader (Bio-Rad). Copy number was assessed using the Quantasoft software. For testing the ddPCR sensitivity in detecting a CNV-12q gain, the UHOMi001-A diploid and the HD291 aneuploid line were used. Cells were grown on Geltrex matrix in E8 Medium prior to the experiment and then dissociated using trypsin and counted. After mixing the two cell lines to obtain increasing concentrations (from 0% to 100%) of abnormal cells within the diploid

population, each mixed sample was processed for genomic DNA extraction using the QIAmp DNA Mini Blood Kit (Qiagen, Hilden, Germany) and for ddPCR analysis. Reference for the designed BioRad ddPCR probes can be found in Table S3.

To establish the sensitivity of the droplet digital PCR method, we used sampleswith increasing percentages (from 0 to 100%) of hPSCs harboring a trisomy 12 within a sample of euploid hPSCs. A significant difference in the CNV copy number compared with control (0%) was observed in samples with at least 10% of abnormal cells (*p*-value <0.05, Student's t test). The panels represent three biological replicates. The error bars (generated by the QuantaSoft software) for each well represent the 95% confidence intervals using Poisson statistics with the total number of droplets. All values of copy number were corrected by adding -0.083 because more than 50 experiments on samples with a normal count of chromosome 12 have shown a bias with this probe with a median value of 2.083.

Generation of the FOXJ1_mCherry and CCDC40_KO iPSC lines using CRISPR/Cas9

A stock of HY03 (75k) M53Cl2SC6 cells was made and after thawing their euploidy was confirmed using the iCS-digital test for detection of CNV anomalies at mechanical passage M53, clumps passage Cl2, single cell passage SC11 for FOXJ1_mCherry tagging, and at mechanical passage M53, clumps passage Cl2, single cell passage SC17 for CCDC40_KO cells respectively. The day before transfection, 25000 HY03 (75k) M53Cl2SC7 cells per cm^2 were plated in a 6-well plate coated with Geltrex matrix and with E8 supplemented with Y-27632 (10µM). The day of transfection, medium was refreshed using the Lipofectamine Stem transfection reagent (Invitrogen) at least 2 hours before transfection, following the

manufacturer's instructions. Briefly, 2 µg of pSpCas9(BB)-2A-GFP (PX458) (gift from Feng Zhang, Addgene plasmid #48138) containing the FOXJ1 targeting sgRNA sequence 5'-GGGCCTTCTTGTAAGAGGCC-3' or the CCDC40 targeting sgRNA sequence 5'-CTCCTCGTTGGCGGCTGCGC-3' with 1 µg of MBX plasmid (gift from Linzhao Cheng, Addgene plasmid #64122) and 1 µg of homemade donor plasmid pUC19_FOXJ1_mCherry_cNEO for FOXJ1_mCherry tagging were mixed with 4 µL of Lipofectamine and left at room temperature for 10min to form complexes. The Lipofectamine-DNA complexes were added on top of the cells, distributed by gently swirling the plate, and incubated at 37˚C, 5% CO_2. The following day, the medium was changed with fresh E8 medium supplemented with G418 (200 µg/ml) for the FOXJ1_mCherry cells, and then changed daily for 6 days. Colonies were manually picked and transferred into a 96-well plate for amplification. At confluence, clones were passaged to a 24-well plate, and then to a 6-well plate. DNA was collected to screen clones by bridge-PCR, transgene copy counting, and Sanger sequencing for FOXJ1_mCherry tagging or by high resolution melt analysis (HRMA) followed by Sanger sequencing for CCDC40_KO. Finally, the presence of genomic abnormalities was checked using the iCS-digital test.

Supplemental references

Ahmed, E., Sansac, C., Fieldes, M., Bergougnoux, A., Bourguignon, C., Mianné, J., Arnould, C., Vachier, I., Assou, S., Bourdin, A., et al. (2018). Generation of the induced pluripotent stem cell line UHOMi001-A from a patient with mutations in CCDC40 gene causing Primary Ciliary Dyskinesia (PCD). Stem Cell Res. 33, 15–19.

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

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

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

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinforma. Oxf. Engl. 26, 841–842.

Närvä, E., Autio, R., Rahkonen, N., Kong, L., Harrison, N., Kitsberg, D., Borghese, L., Itskovitz-Eldor, J., Rasool, O., Dvorak, P., et al. (2010). High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity. Nat. Biotechnol. 28, 371–377.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinforma. Oxf. Engl. 26, 841–842.

Spits, C., Mateizel, I., Geens, M., Mertzanidou, A., Staessen, C., Vandeskelde, Y., Van der Elst, J., Liebaers, I., and Sermon, K. (2008). Recurrent chromosomal abnormalities in human embryonic stem cells. Nat. Biotechnol. 26, 1361–1363.

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

Wu, H., Kim, K.J., Mehta, K., Paxia, S., Sundstrom, A., Anantharaman, T., Kuraishy, A.I., Doan, T., Ghosh, J., Pyle, A.D., et al. (2008). Copy number variant analysis of human embryonic stem cells. Stem Cells 26, 1484–1489.

Wood, R.D., Mitchell, M., and Lindahl, T. (2005). Human DNA repair genes, 2005. Mutat. Res. 577, 275–283.