

Random Mutagenesis, Clonal Events, and Embryonic or Somatic Origin Determine the mtDNA Variant Type and Load in Human Pluripotent Stem Cells

Filippo Zambelli,^{1,2} Joke Mertens,¹ Dominika Dziedzicka,¹ Johan Sterckx,³ Christina Markouli,¹ Alexander Keller,¹ Philippe Tropel,⁴ Laura Jung,⁵ Stephane Viville,^{5,6} Hilde Van de Velde,^{1,3} Mieke Geens,¹ Sara Seneca,^{1,7} Karen Sermon,¹ and Claudia Spits^{1,*}

¹Research Group Reproduction and Genetics, Faculty of Medicine and Pharmacy, Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, Brussels 1090, Belgium

²S.I.S.Me.R. Reproductive Medicine Unit, Via Mazzini 12, Bologna 40100, Italy

³Centre for Reproductive Medicine, Universitair Ziekenhuis Brussel (UZ Brussel), Laarbeeklaan 101, Brussels, Belgium

⁴STEMCELL Technologies Inc., Vancouver, BC, Canada

⁵Institut de Parasitologie et Pathologie Tropicale, EA 7292, Fédération de Médecine Translationnelle, Université de Strasbourg, 3 rue Koeberlé, Strasbourg 67000, France

⁶Laboratoire de Diagnostic Génétique, UF3472-génétique de l'infertilité, Hôpitaux Universitaires de Strasbourg, Strasbourg 67000, France

⁷Centre for Medical Genetics, UZ Brussel, Laarbeeklaan 101, Brussels, Belgium

*Correspondence: claudia.spits@vub.be

<https://doi.org/10.1016/j.stemcr.2018.05.007>

SUMMARY

In this study, we deep-sequenced the mtDNA of human embryonic and induced pluripotent stem cells (hESCs and hiPSCs) and their source cells and found that the majority of variants pre-existed in the cells used to establish the lines. Early-passage hESCs carried few and low-load heteroplasmic variants, similar to those identified in oocytes and inner cell masses. The number and heteroplasmic loads of these variants increased with prolonged cell culture. The study of 120 individual cells of early- and late-passage hESCs revealed a significant diversity in mtDNA heteroplasmic variants at the single-cell level and that the variants that increase during time in culture are always passenger to the appearance of chromosomal abnormalities. We found that early-passage hiPSCs carry much higher loads of mtDNA variants than hESCs, which single-fibroblast sequencing proved pre-existed in the source cells. Finally, we show that these variants are stably transmitted during short-term differentiation.

INTRODUCTION

It is by now well established that human embryonic and induced pluripotent stem cells (hESCs and hiPSCs) display substantial nuclear genome instability while kept in culture (reviewed in [Lund et al., 2012](#) and in [Nguyen et al., 2013](#)). A significant amount of research has been devoted to this topic, proving that these cells undergo replicative stress ([Jacobs et al., 2014, 2016](#); [Lamm et al., 2016](#)) and the cultures are frequently taken over by recurrent chromosomal abnormalities conferring an *in vitro* survival advantage ([Amps et al., 2011](#); [Avery et al., 2013](#); [Nguyen et al., 2014](#); [Merkle et al., 2017](#)). In contrast, only a few studies provide some insight on the integrity of their mitochondrial genome, despite the important role mitochondria play in reprogramming and maintenance of the stem cell state ([Van Blerkom, 2008](#); [Lonergan et al., 2007](#)). Undifferentiated human and mouse ESCs contain few, spherical, and immature mitochondria, similar to those found in preimplantation embryos. The number and maturity of the mitochondria increase upon differentiation, concurrent with the switch from glycolysis to oxidative phosphorylation for energy production ([Facucho-Oliveira and St John, 2009](#)). Conversely, human somatic mitochondria undergo morphological and func-

tional changes during reprogramming to hiPSCs ([Suhr et al., 2010](#)), with a shift from oxidative phosphorylation to glycolysis. Furthermore, attenuating mitochondrial function in undifferentiated hESCs increases the mRNA levels of the pluripotency genes, compromises their differentiation potential, and increases the number of persisting tumorigenic cells after differentiation ([Mandal et al., 2011](#)).

Work from the field of disease modeling has provided some very interesting insight on the effect of specific mtDNA mutations on hPSC differentiation capacity, proliferation rate, and reprogramming efficiency ([Yokota et al., 2015, 2017](#)). For instance, mtDNA haplogroups appear to affect cellular function. Work on mouse ESCs has shown that in both undifferentiated and differentiating cells, the mitochondrial haplogroup has a significant impact on the expression of genes involved in pluripotency and differentiation, and does consequently influence the capacity of the cells to differentiate ([Kelly and St John, 2010](#); [Kelly et al., 2013](#)). In the human, recent work in the context of mitochondrial replacement in oocytes indicated that some haplogroups can modify the growth dynamics of hESCs, resulting in a growth advantage that can lead to a culture takeover ([Kang et al., 2016a](#)).



Maitra et al. (2005) were the first to show mtDNA changes in human pluripotent stem cells (hPSCs). They found that two out of ten hESC lines had acquired heteroplasmic single nucleotide variants (SNVs) during culture. Technical limitations at that time precluded the study of the full mitochondrial genome while simultaneously establishing the variant load. The advent of massive parallel sequencing made it possible for Prigione et al. (2011) to study four hiPSC lines in detail, and compare their full mtDNA with that of the two source cell lines. They identified a number of SNVs that significantly differed in heteroplasmic load between lines and as compared with their source cells. However, they were unable to provide an explanation for these observations. Later, our group identified by long-range PCR numerous large deletions in the mtDNA of hESCs (Van Haute et al., 2013). Most recently, two recent reports studied heteroplasmic SNVs in hiPSCs (Kang et al., 2016b; Perales-Clemente et al., 2016). Both studies found that different hiPSC lines established from the same source cells harbored different variants, frequently with a pathogenic effect, some of which could be traced back to the source cell cultures. These findings suggest that the differences among the hiPSC lines are due to their clonal nature, each line representing the mtDNA content of one individual source cell. They hypothesized that there is considerable mosaicism in the source cell cultures, and that this is related to somatic mutagenesis, and correlating to the age of the cell donor (Kang et al., 2016b).

In this study, we address the issues that were not covered by the studies discussed above. First, we aimed at thoroughly studying mtDNA variants in hESC cultures, as these were rarely investigated in the previous reports. To this aim, we carried out deep sequencing of the mtDNA of seven early-passage hESC lines. In order to identify the origin of the variants found in the hESCs, we analyzed the mtDNA in the blood of the women who donated the embryos used to derive these hESC lines. We also studied 11 human oocytes and eight inner cell masses (ICMs) from blastocysts as the source cell type of hESC. Next, we analyzed multiple passages of these hESC lines to assess the impact of long-term culture, and carried out single-cell analysis of early- and late-passage lines to better understand the dynamics of variants in culture. Secondly, we investigated the origin of the diversity in mtDNA variants in hiPSCs. We deep-sequenced 17 hiPSC lines derived from two different fibroblast lines, and by different reprogramming methods. We studied these two source fibroblasts lines as a bulk, and for one of the lines also the single cells, in order to assess the levels of mosaicism and the matching between the individual source cells and the hiPSC lines. Finally, we studied the transmission of mtDNA variants during differentiation, by differenti-

ating multiple hiPSC clones into ectoderm and definitive endoderm.

RESULTS

Early-Passage hESCs Carry Low-Frequency mtDNA Variants Similar to Those in the Embryos Used for Derivation

First, we used massive parallel sequencing to analyze the full mtDNA of seven hESC lines at early passages, i.e., under 30. This method allows the comprehensive sequencing of the full mtDNA together with the quantification of heteroplasmic sites with loads as low as 1.5%. The results are shown in Figures 1A and 1B; all details can be found in Table S3. Most of the lines displayed one to two heteroplasmic variants, none predicted to be pathogenic (Figure 1B), and with a cumulative variant load (i.e., the sum of all the frequencies of the different heteroplasmies) ranging from 2.0% to 50.8% (details in Table S3).

To investigate if these low-frequency variants were derivation-induced, inherited, or culture-induced, we first studied the mtDNA in blood samples of the women who donated the embryos for the derivation of five of the seven lines. In all cases, the homoplasmic variants were identical between the donors and the lines. However, none of the heteroplasmies identified in the hESCs was found in the blood sample of the donor. In fact, none of the blood samples contained any heteroplasmic variant (data in Table S4). This can be explained by the fact that heteroplasmic variants may be lost in rapidly proliferating tissues such as blood, and approximately 15% of variants may be germline mutations (Rebolledo-Jaramillo et al., 2014).

Next, we sequenced 11 single oocytes and eight ICMs, as these cells are the closest proxy in terms of mtDNA content to the first passages after hESC derivation. The results are shown in Figures 1A and 1B, all data in Table S5. We detected heteroplasmic variants in 9 out of 11 oocytes, at cumulative loads ranging from 2.1% to 28.4%, none predicted to be pathogenic (Figure 1A, Table S5). In some cases, the same variants were found recurrently in the oocytes from the same donor, but often at different frequencies. For instance, in three out of four oocytes of donor 1 we found the variant m.150C>T, at loads of 5%, 13%, and 20%. In the ICMs, we found heteroplasmic variants in all but one, with a cumulative variant load ranging from 2.0% to 44.8%. The three ICMs donated by the same patient showed a very low number of variants and a low cumulative frequency, but no overlap was found (Table S5). Two variants were found to be likely pathogenic, with loads of 3.0% and 4.4%, respectively.

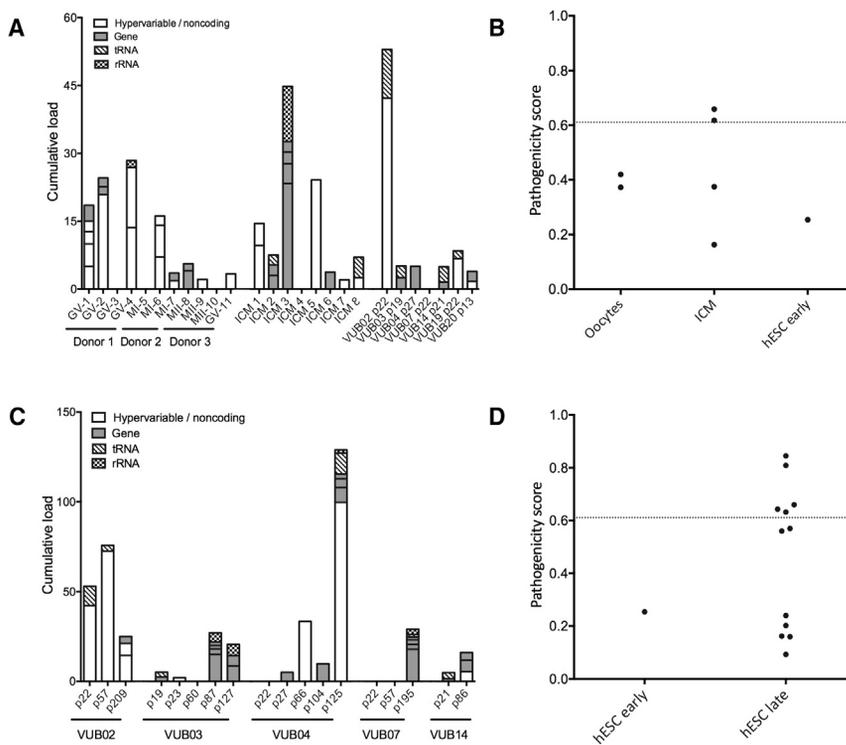


Figure 1. Heteroplasmy Levels and Pathogenicity Scores of the mtDNA Variants in Oocytes, ICMs, and hESC Lines

(A) Heteroplasmies retrieved in early hESCs compared with oocytes and ICMs. The y axis indicates the heteroplasmic load of the variants. The individual variants are defined as a single bar with the height proportional to their load, and the color pattern corresponding to its location in the mtDNA sequence, as shown in legend in the figure. The numbers after the GV, MI, and MII refer to the donor's code. All the variants from one sample are stacked to obtain the cumulative heteroplasmic load value.

(B) Pathogenicity scores obtained with Mutpred2. Variants with a score over 0.611 have a high probability of being pathogenic.

(C) Variants in different passages of VUB hESC lines, the variants are represented as in (A).

(D) Pathogenicity scores in early (passage below 30) and late (above passage 50) hESCs. All data can be found in [Tables S2–S6](#).

Overall, these results show that the type and load of the heteroplasmic variants found in early passages of hESCs and those identified in oocytes and ICMs are very similar. This suggests that the variants we identified in the hESCs were already present in the embryos at the origin of the cell lines.

An Increase in mtDNA Variants Correlates with Time in Culture and the Appearance of Chromosomal Abnormalities

To investigate the effect of prolonged *in vitro* culture, we analyzed multiple later passages of five of the hESC lines studied before, ranging from passage 57 to 209. The results are summarized in [Figures 1C](#) and [1D](#); all details can be found in [Table S3](#). While the early-passage samples showed few variants at relatively low loads, the later passages showed an overall increase in the number of variants per line (up to six in VUB07) with a higher cumulative load (ranging from 5.0% to 128.9%). In these later passages, we also found five variants predicted to be potentially pathogenic, and two others highly likely to be pathogenic. However, the variants with a high pathogenic score were present at loads below 10% ([Figure 1D](#) and [Table S3](#)).

Interestingly, the increase in the cumulative variant load in culture is not the result of the clonal expansion of pre-existing variants of the early passages. In most of the cases the variants identified in the earlier passages were not retrieved in the later ones, and vice versa ([Table S3](#)). Only one variant

of VUB02 was recurrent in all three passages analyzed (m.234A>G in passage 22, 57 and 209), although its heteroplasmic load significantly varied among passages. The same variant was observed in passages 66 and 125 of VUB04, but not in passages 22, 27, and 104. To explain these peculiar observations, we retraced the “laboratory history” of the lines and checked the karyotypes of all samples by array-based comparative genomic hybridization (CGH) (all karyotypes can be found in [Table S6](#)). In some cases, several sub-lines had clearly originated from the same line, by cycles of cryopreservation and parallel culture. In VUB02, all three samples had been collected during continuous culture, and belonged to the same chromosomally normal sub-line. This is concordant with the presence of the same variant in all samples, in which the variation in load can be caused by random drift.

Conversely, the later passages of other lines had often acquired chromosomal abnormalities. In these cases, the mtDNA genotype differed completely between sub-lines. For instance, VUB03 passage 87 and passage 127 have both a gain of 20q11.21, but both mutations occurred independently, as they have different sizes. Also, the samples were collected from cultures that were kept in the lab at different time periods. This indicates that these two samples belong to two different sub-lines, which is concordant with the difference in mtDNA variants. Another interesting case is that of VUB04, where passages 27, 66, and 125 belong to the same sub-line, which was kept in culture

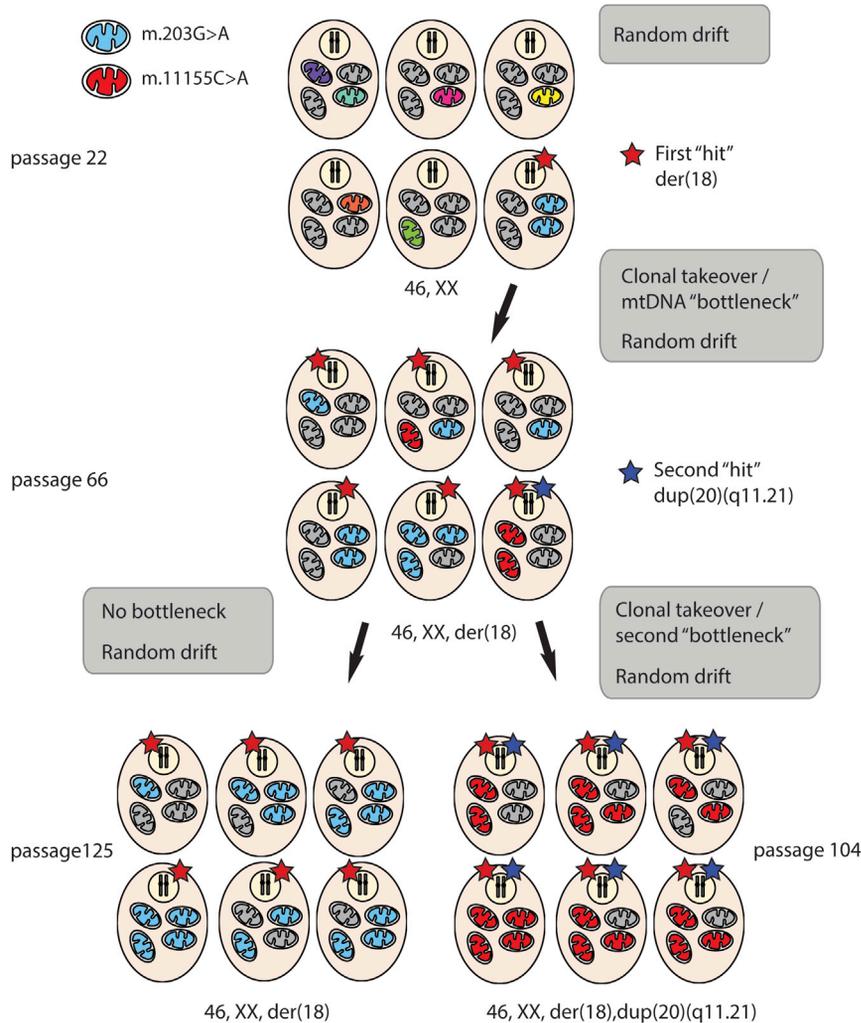


Figure 2. Schematic Representation of mtDNA Variant Transmission in Different Sub-lines of VUB04

The two variants represented by blue (m.203G>A) and red (m.11155C>A) mitochondria are passenger mutations of cells that acquired a chromosomal abnormality conferring growth advantage and resulted in clonal culture takeover. For the variant m.203G>A, we observe how one bottleneck (first "hit," red star) causes the appearance of the variant in passage 66 and an expansion until a homoplasmic level at passage 125. During *in vitro* culture, spontaneous mutation and random drift resulted in cells carrying new mtDNA variants. An independently grown dish of VUB04 acquired a second chromosomal abnormality (second "hit," blue star), generating a sub-line not carrying the m.203G>A variant but showing another variant, m.11155C>A, passenger of the cell acquiring the second chromosomal aberration. der(18) = der(18)t(18,5)(q21.2; q14.2).

continuously. During that time span, the line acquired a derivative chromosome 18 (der[18]t[18,5] [q21.2; q14.2]) conferring growth advantage, which at passage 66 was still mosaic and expanding, and at passage 125 was present in all cells. This is concomitant with a progressive increase in the load of m.203G>A. Additionally, during culture, a second sub-line from VUB04 appeared, with a second abnormality (a gain of 20q11.21) that took over the culture, resulting in a different heteroplasmy pattern at passage 104, as compared with the other VUB04 sub-line.

These observations can be explained by assuming that the cultures contain a significant cellular diversity in terms of mtDNA variants. When a cell acquires a chromosomal abnormality that results in a selective advantage, the culture goes through a genetic bottleneck, and after some time, all the cells of the sub-line will be descendants of that one mutant cell. In this scenario, any mtDNA variants in that founder cell will be passenger mutations that will also take over the culture, although without necessarily

having any functional benefit. Figure 2 illustrates how this hypothesis would apply to VUB04.

mtDNA Cellular Mosaicism in Early- and Late-Passage hESCs

The observations described above led to two questions: how much cellular diversity is there in terms of mtDNA variants, and does it fluctuate with time in culture? To test this, we collected single cells from VUB03, VUB04, and VUB14, at early passages (p23, p22, and p21, respectively) and late passages (p87, p104, and p86, respectively). For each passage, we collected 20 single cells, which were analyzed using only one of the two different long-range PCRs for mtDNA enrichment (set 1 or set 2). This means that ten cells were studied individually for each specific region; the results have not been pooled. The data are summarized in Figure 3, all data can be found in Tables S7–S9.

Overall, the results show a remarkable level of diversity in the mtDNA variants among individual cells. Most of the

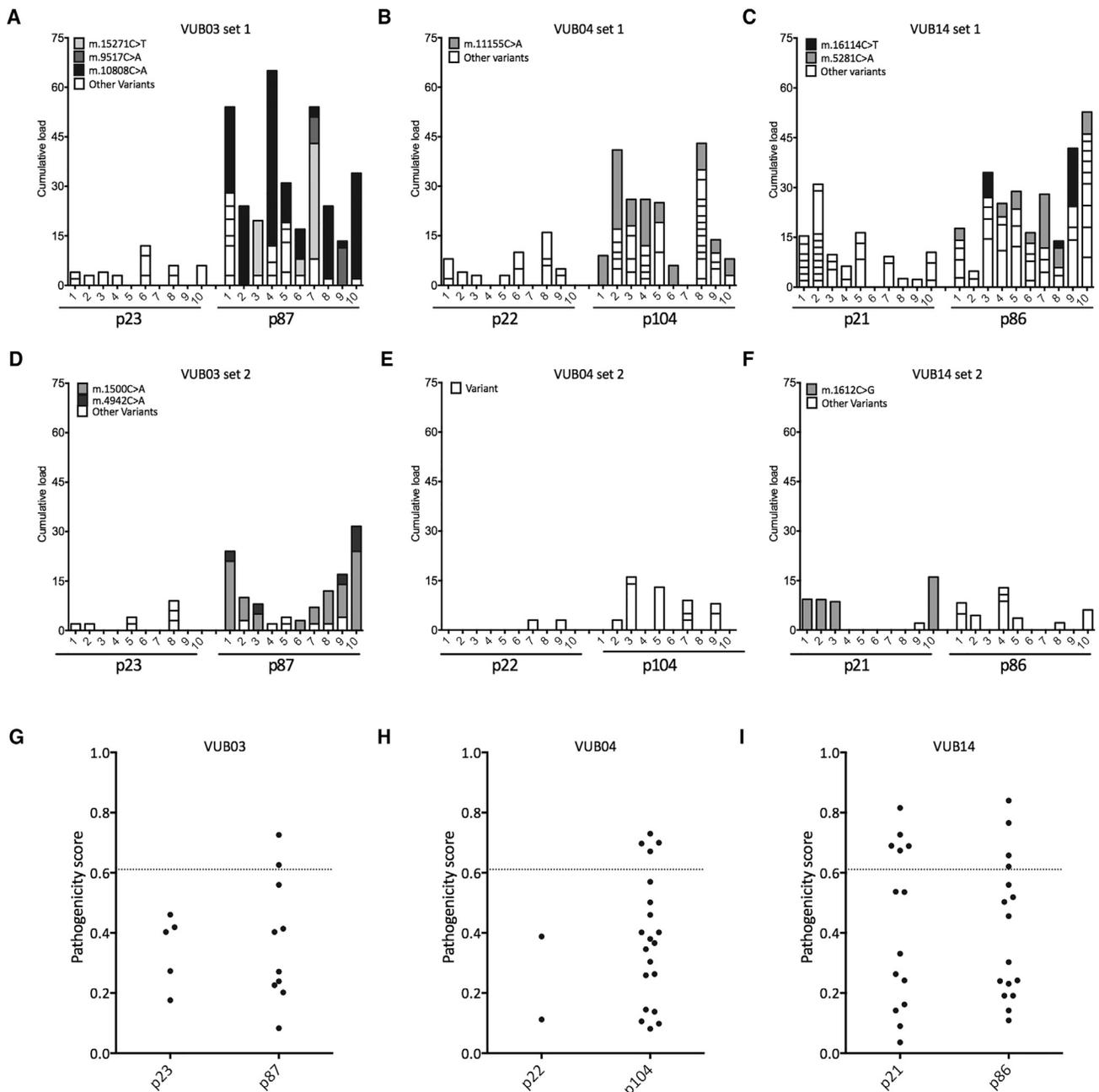


Figure 3. Heteroplasmies Observed in Single Cells of Early and Late hESC Lines

Single cells from three VUB lines (VUB03, VUB04 and VUB14) were sequenced after mtDNA PCR amplification using set 1 (A–C) or set 2 (D–F). The y axis represents the load of the variants observed; the single variants are defined as a single bar with the height proportional to their frequency. All the variants from one sample are stacked to obtain the cumulative load. The variants in white are unique to one cell, while the ones recurring in more than one cell are color-coded as indicated in the legends. For every line, the pathogenicity results are reported (G–I); a threshold of 0.611 defines variants with a high probability of being pathogenic. All data can be found in [Tables S7–S9](#).

variants are unique to each cell, especially in early passages, with heteroplasmic loads between 2% (the lower detection limit) and 35% (Figures 3A–3F, white bars). Variants found in more than one cell were also consistently identified in

the analysis of the DNA sample extracted from the bulk culture dish used to collect the single cells. Recurrent variants were found in all late-passage samples, and in the early passage of VUB14 (Figures 3A–3F, colored bars). Importantly,



the pathogenicity scores of the variants in the later passages were higher as compared with the early ones (Figures 3G–3I). These recurrent variants possibly have two different origins, in line with the results described in the previous sections. It is most likely that the variant in early-passage VUB14 pre-existed in the embryo that was used for derivation, while the recurrent variants in the later passages appear because of the culture takeover by a chromosomally abnormal cell.

For instance, in the early passage of VUB03, 11 out of 20 cells carry mtDNA variants, but all of them are unique. At this point, the line was chromosomally normal. At passage 87, the line acquired a gain of 20q11.21, an event that occurred sometime after passage 60, when the line was still genetically normal. By passage 87, all the cells carried mtDNA variants, of which five were recurrent and two were present in most cells. This strongly suggests that the founder cell that originally acquired the gain of 20q11.21 also carried those five variants. Similarly, in VUB04, 9 out of 20 of the cells from the early, chromosomally normal passage carry diverse and unique variants. In the late passage, when the line had become genetically abnormal, carrying a derivative chromosome 18 and a gain of 20q11.21, 15 out of 20 cells carried mtDNA variants. Most of these variants were different from the variants found at passage 22, but were still unique to one cell, except variant m.11155C>A, which is found in nine out of ten cells. It can be assumed that the cell that acquired the gain 20q11.21 also carried the m.11155C>A variant.

It is noteworthy that, when removing the recurrent variants, assumed to be already present in the founder cell that took over the culture, the late hESC passages still show an increased number of variants per cell, and at higher individual loads. The maximum cumulative load for set 1 in VUB03 was 12.5% in the early and 23.3% in the late passage, in VUB04 it was 9.7% in the early and 35.4% in the late passage, and in VUB14 it was 30.5% in the early and 78.1% in the late passage.

High-Load mtDNA Variants in hiPSCs Originate from the Source Fibroblasts and Stably Transmit during Differentiation

We studied the mitochondrial genome of early-passage clonal hiPSC lines derived from two different fibroblast cultures at two different institutions, named STBG-iPSCs (nine clones) and VUBi004 (eight clones). The STBG-iPSC lines were generated with different reprogramming methods, which did not appear to have a clear differential impact on the mtDNA integrity. For VUBi004, we sequenced passage 5 of all the lines and passage 10 of four clones. Data are shown in Figure 4 and the complete data in Tables S10 and S11.

The first striking point is that (low-passage) hiPSCs very frequently harbored variants with high cumulative loads (up to 172.6% in the STBG-iPSC lines and 239.6% in VUBi004 clones), and predominantly located in coding regions. This is in stark contrast with hESCs, which even in the late passages did not accumulate comparable loads. Furthermore, potentially pathogenic variants, which in late-passage hESCs never showed loads above 10%, can reach near homoplasmy in hiPSCs. The variants retrieved in clones originating from the same source fibroblasts are rarely the same, with only two lines displaying the same set of variants at slightly different loads, while sequencing the same clonal line at passage 5 and passage 10 yielded the same exact heteroplasmies, with very small fluctuations in their frequency (Table S11).

This might be again due to significant cellular diversity in the fibroblast culture. When looking into the bulk source fibroblasts of the STBG-iPSC lines, we found a cumulative variant load of 15.3%, resulting from seven variants, but none of these was the same as those identified in the different STBG-iPSC clones. Conversely, the source fibroblasts of VUBi004 carried 16 variants, with a cumulative load of 64.6%. Seven of these variants were also found in the iPSC lines, with much lower heteroplasmic loads in the fibroblast culture compared to the iPSC clones. During the technical setup and validation of the sequencing method used in this paper, we sequenced 11 single cells from the VUBi004 fibroblast culture (part of the data shown in Table 1, all results in Table S12 and in Zambelli et al., 2017). We found that some of the individual fibroblast cells carry at high loads the variants that are identified at lower loads in the DNA extracted from whole culture dishes. Therefore, the heteroplasmic load measured in the DNA extracted from whole culture dishes is in fact the average of a mosaic culture. For instance, m.12071T>C is present at a frequency of 9.9% in the bulk DNA sample, but is nearly homoplasmic in 2 out of the 11 tested fibroblast cells, while the rest of the cells are only carrying wild-type sequence. In line with this, when comparing these results with those of the VUBi004 clones, we find very similar patterns. For instance, the variant m.2760A>G has a load of 5.8% in the bulk fibroblast culture, but is present in one of the single fibroblast at 95.8% and in one of the clonal hiPSC lines at 72.2%.

Finally, we assessed the transmission and maintenance of the heteroplasmic load of these variants during short-term differentiation. For this, we differentiated six hiPSC clones into definitive endoderm and ectoderm. We characterized the cells by immunostaining for SOX17 and PAX6, and real-time PCR for SOX17, FOXA2, PAX6, SOX1, and POU5F1 (part of the data is shown in Figures 5A and 5B, the rest can be found in Figure S1). We sequenced the mtDNA of the cells prior to the start of the differentiation,

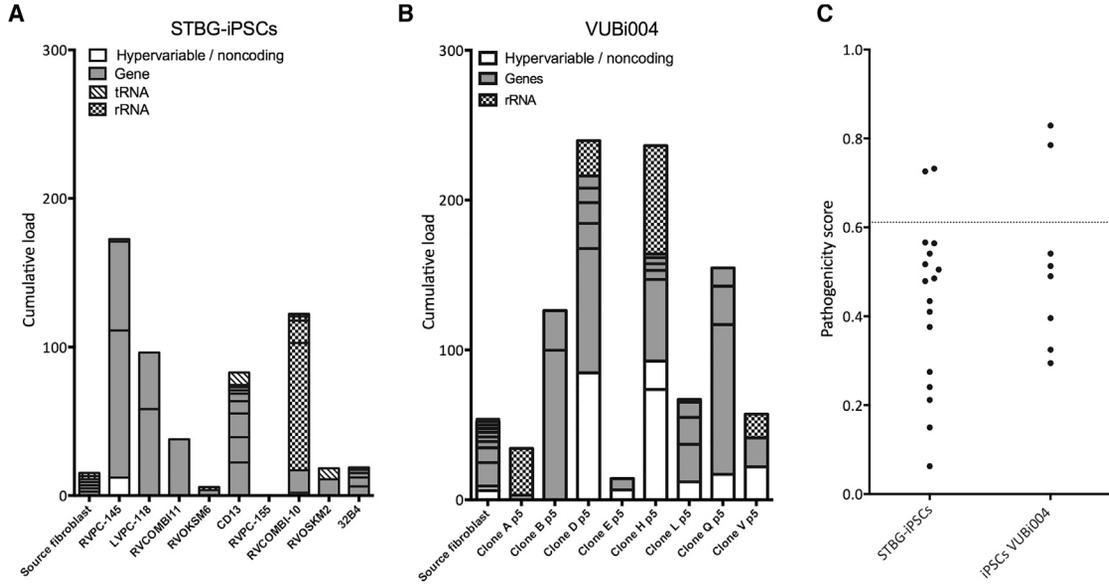


Figure 4. Heteroplasmies Observed in hiPSCs Lines and Their Source Cells

(A) Heteroplasmies found in multiple lines derived with different reprogramming methods from individual 1. (B) Heteroplasmies found in hiPSC lines derived from individual 2, all with the same reprogramming method. The y axis represents the frequency of the variants observed; the single variants are defined as a single bar with the height proportional to their frequency, and the color pattern corresponding to the location in the mtDNA, as shown in legend in the figure. All the variants from one sample are stacked to obtain the cumulative load value. (C) Pathogenicity scores obtained with Mutpred2; a threshold of 0.611 defines variants with a high probability of being pathogenic. All data can be found in [Tables S10](#) and [S11](#).

and at the endpoint. The results show that the heteroplasmic variants, both synonymous and non-synonymous with pathogenic potential, are preserved at near identical loads throughout two independent processes of differentiation ([Figure 5C](#), pathogenicity scores in [Table S11](#)).

DISCUSSION

Taken together, our data reveal significant levels of mtDNA mosaicism and diversity in hPSC cultures, with significant differences between hESCs and hiPSCs. While early passages of hESCs display few variants at a low heteroplasmic level, hiPSCs and late-passage hESCs show a drastic shift in heteroplasmy and in type of variants, with an increase in their pathogenicity scores. These significant differences are, in the case of hiPSCs, due to the generation of clonal lines from a strongly mosaic somatic cell culture, while for hESCs, it is due to another type of clonal event: the culture takeover by one cell, driven by a chromosomal abnormality conferring growth advantage.

The results of early hESC passages show that the lines carry variants very similar to those we identified in human oocytes and ICMs, and reflect normal human germline variance. In this sense, our data are in line to that reported

by two groups, which recently observed the presence of few mtDNA variants with relatively low pathogenicity in human oocytes ([Boucret et al., 2017](#); [Kang et al., 2016b](#)). Also, the study of [Kang et al. \(2016b\)](#) included data from human somatic cell nuclear transfer hESCs. Our data are similar in terms of the type and loads of mtDNA variants detected, with generally few variants per line and mainly in noncoding regions.

In contrast, late-passage hESCs carry more and higher load variants, which do not appear to necessarily be the result of a progressive increase in load of the variants detected in the early passages. Our results suggest that this is mainly caused by the clonal expansion of one cell in culture. The single-cell sequencing data show that even at early passage, there is considerable diversity in mtDNA genotypes, likely originating from random drift and low levels of spontaneous mutagenesis. Culture takeover by a single stem cell carrying an acquired chromosomal abnormality conferring a selective advantage, results also in a “takeover” by an mtDNA genotype as a passenger variant (as illustrated in [Figure 2](#)). In this manner, the clonal expansion results in a previously undetectable variant (mainly because it is only present in one or few cells) quickly rising in frequency until taking over the culture. In the pioneering work of [Maitra et al. \(2005\)](#), the



Table 1. Comparison of Sequencing Data of Bulk Fibroblast Culture, Single Fibroblasts, and hiPSC Clonal Lines from Individual 2

Variants	Bulk DNA	Single Fibroblasts					hiPSCs		
		Cell 1	Cell 4	Cell 8	Cell 9	Cell 11	Clone A	Clone H	Clone V
m.414T>G	6.2	0	0	0	–	–	0	73.7	0
m.2760A>G	5.8	–	–	–	95.8	0	0	72.2	0
m.2947T>C	0	–	–	–	0	0	31.2	0	15.8
m.3563G>A	1.6	–	–	–	0	31.5	0	0	0
m.9276G>A	3.1	0	15	0	–	–	0	0	0
m.12071T>C	9.9	98	0	100	–	–	0	0	0
m.12850A>G	15.6	0	100	0	–	–	0	0	0
m.15258T>C	0	0	0	0	–	–	3.2	0	17.4
m.15617G>A	4.2	0	0	41	–	–	0	0	0
m.16166A>G	3	0	0	0	–	–	0	18.9	0

Both single cells and clonal hiPSCs are characterized by the presence of variants at relatively high heteroplasmic levels. In most cases, these variants are present in the bulk DNA of the fibroblast culture but always at lower levels. In one case, the same variant is found in a single cell, in the bulk culture and in one hiPSC clone (m.2760A>G). For clarity, only the samples with matching variants are indicated, all data can be found in [Table S11](#) (VUBi004) and [Table S12](#) (single fibroblasts). – indicates no results for this region, this depending on the primer set used for mtDNA amplification.

authors found that two out of ten hESC lines acquired heteroplasmic mtDNA variants during long time in culture (45 additional passages in one case and 175 in the other). Interestingly, the lines that acquired these variants also had acquired a chromosomal abnormality, while none of the five karyotypically normal lines showed *de novo* mtDNA variants. These observations can be explained by our model, and are fully in line with the findings on our hESC lines.

In the case of hiPSCs, the main cause for the increase in heteroplasmic variants seems to be the clonal expansion of a single reprogrammed somatic cell from a culture with significant cellular diversity in terms of mtDNA. This diversity in mtDNA variants among lines derived from the same source was also reported in the works of [Prigione et al. \(2011\)](#), [Kang et al. \(2016b\)](#), and [Perales-Clemente et al. \(2016\)](#), all of them suggesting that the variants were already present in the fibroblast population. In our work, we show that this is indeed correct. The single-cell sequencing data show that every fibroblast from the source cell culture displays a specific subset of mtDNA variants, which are often personal, and at near homoplasmic loads. Furthermore, the identification of an individual fibroblast matching one of the hiPSC lines further strengthens this point. Also, we did not find significant differences in the type or load of the variants between the different lines from the same source but established using different reprogramming methods. This is in line with the results of [Prigione et al. \(2011\)](#), who also found no difference in the type and number of variants detected in cells reprog-

rammed with two different methods. Finally, we found that the variants are stably transmitted to the differentiated progeny. In this work, we used short differentiation protocols, but it is unlikely that the stability of the variants is related to the differentiation time, since a report from [Hämäläinen et al. \(2013\)](#) showed stable levels of heteroplasmy in hiPSCs after longer differentiation protocols, up to 16 weeks for neurons and 7 weeks for teratomas. Nevertheless, we cannot exclude that specific variants might influence the growth and/or differentiation capacity of the cells, leading to a shift in the heteroplasmy levels in the differentiated cells.

Another important aspect of our findings is the difference in pathogenic potential seen in variants present in hiPSCs compared with hESCs. The higher pathogenic potential of the variants retrieved in the hiPSCs shows that the source cells, despite being derived from healthy individuals, can carry a certain degree of dysfunctional, heavily mutated cells. Furthermore, it appears that pathogenic variants are not necessarily selected against during reprogramming or during hPSC culture. It is possible that this is because the metabolic needs of pluripotent cells greatly rely on glycolysis rather than oxidative phosphorylation. Conversely, once the hPSCs are differentiated, this may result in metabolic dysfunction, as elegantly demonstrated by [Kang et al. \(2016b\)](#). This underscores the importance of determining the mtDNA variant load in hPSC lines before their use in a clinical setting.

Finally, our data on early-passage hESCs and on short-term culture of hiPSCs suggests that, if no culture takeover

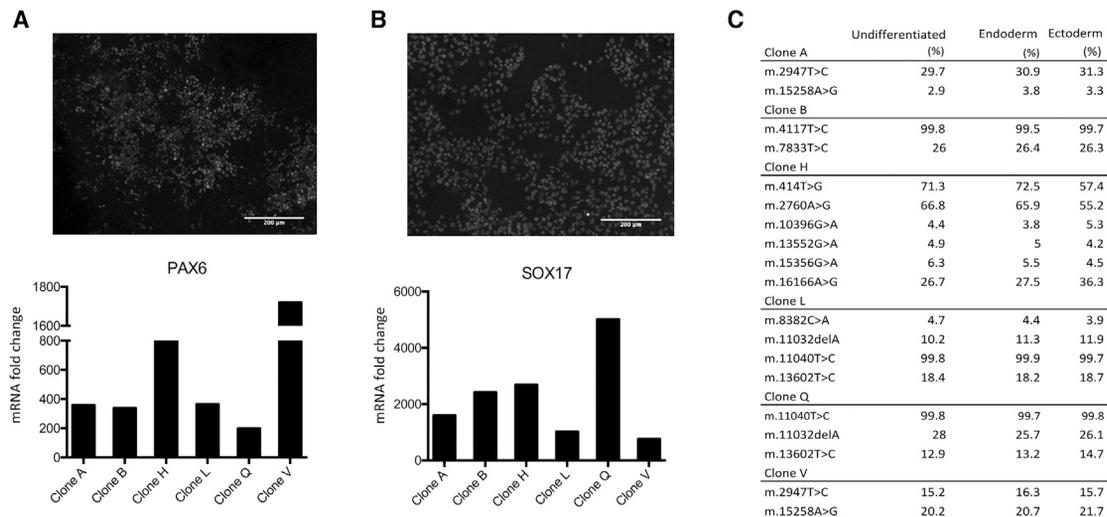


Figure 5. Transmission of mtDNA Heteroplasmies during hiPSC Differentiation

(A and B) All cell lines show abundant expression of (A) ectodermal (PAX6) and (B) endodermal (SOX17) markers, proving the successful differentiation of the six clones of VUBi004 into the two different lineages. One example of immunohistochemistry is shown in the figure, the remaining data can be found in Figure S1. The quantitative real-time PCR data shown in the plots is represented as fold-change increase relative to the undifferentiated cells.

(C) Comparison of the heteroplasmy levels of mtDNA variants in hiPSC in their undifferentiated and differentiated state. In all cases but one, the variants are stably transmitted through short-term differentiation. The ectoderm obtained from clone H shows a small fluctuation in the heteroplasmic load of some of the variants.

is involved, the mtDNA variant load remains grossly stable during at least 20 passages, although individual cells may be slowly adding up spontaneous mutations. On the other hand, it was already shown that older individuals carry increased mtDNA variant levels in different tissues (Kang et al., 2016b), this clearly having a significant impact on the obtained hiPSCs. This factor should probably be taken into account when identifying cell donors for reprogramming, and can complicate the identification of suitable autologous hiPSC lines for regenerative medicine in aged individuals. Nevertheless, despite the high load of mtDNA variants we found in most of the hiPSCs analyzed, in both our sources we were also able to obtain lines with a low variant load, with one line carrying no heteroplasmies. Although this type of screening is cumbersome, it is important to bear in mind the significant impact mitochondria have on the functionality of the differentiated progeny.

EXPERIMENTAL PROCEDURES

Ethical Approval

The Institutional Medical Ethics Committee of the University Hospital UZ Brussel and the Vrije Universiteit Brussel approved all experiments in this study. Experiments involving human embryos were also approved by the Belgian Federal Committee on medical and scientific research on human embryos *in vitro*.

All individuals involved in the study gave signed informed consent.

Cell Culture

The VUB hESC lines were derived and characterized as previously described (Mateizel et al., 2006, 2010). The lines were grown either on inactivated mouse embryonic fibroblast feeder layers using a standard hESC medium as described in (Mateizel et al., 2006, 2010), or on culture dishes coated with 5 µg/mL of recombinant Laminin-521 (Biolamina) in Nutristem (Biological Industries) with 0.5% 100 U/mL penicillin/streptomycin (Thermo Fisher Scientific), as described in (Dziedzicka et al., 2016).

The VUBi004 iPSC lines were obtained by transducing fibroblasts with p4ORF-dTomato as described in Warlich et al. (2011). Clonal iPSC lines were manually isolated at day 30 post-transduction, and further cultured individually on Laminin-521 and Nutristem. The iPSC lines derived in the laboratory of S. Viville, named STBG-iPSCs, were established using different reprogramming methods, as listed in the Table S1 and as described in Jung et al. (2014).

The fibroblast line was established from a skin biopsy from a 53-year-old female donor. The cells were cultured in F-12 Nutrient Mix Ham (Life Technologies) with 20% fetal calf serum (Thermo Fisher Scientific), 0.5% penicillin/streptomycin (Thermo Fisher Scientific), and 1% glutamine (Thermo Fisher Scientific). The cells were expanded up to passage 4.

hiPSC Differentiation

The protocol for differentiation to definitive endoderm was adapted from Sui et al. (2012). hiPSCs were seeded on Laminin-521



coated dishes at a density of 30,000 cells/cm². The day after plating, cells were treated with a medium containing RPMI 1640 supplemented by GlutaMAX (Thermo Fisher Scientific), 0.5% B27 supplement (Thermo Fisher Scientific), 100 ng Recombinant Human/Mouse/Rat Activin A (R&D Systems), and 3 μ M CHIR99021 (Stemgent). The second day, cells were treated with the same medium without CHIR99021. Cells were cultured for 2 additional days in RPMI 1640.

The ectoderm differentiation protocol was adapted from [Chetty et al. \(2013\)](#) and [Chambers et al. \(2009\)](#). hiPSCs were seeded on Laminin-521 coated dishes at a density of 50,000 cells/cm², and treated the day after with a medium consisting of KnockOut D-MEM (Thermo Fisher Scientific), 10% KnockOut Serum Replacement (Thermo Fisher Scientific), 500 ng/mL Recombinant Human Noggin Protein (R&D Systems), and 10 μ M SB431542 (Tocris). Differentiation medium was changed daily and cells treated for a total of 4 days.

Oocytes and Embryos

Oocytes were retrieved after conventional ovarian stimulation. In this study, we used oocytes found to be immature after removal of the granulosa cells and thus not useful for the patient's treatment. We analyzed 11 oocytes obtained from three donors at different stages of maturation: 5 at the germinal vesicle stage, 3 at metaphase I, and 3 at metaphase II matured *in vitro* (details can be found in [Table S2](#)). The ICMs were isolated from blastocysts shown to be affected with a monogenic disease after preimplantation genetic diagnosis. Blastocysts were vitrified after trophectoderm biopsy on day 5/6, warmed ([Van Landuyt et al., 2015](#)), and cultured overnight in blastocyst medium (ORIGIO, CooperSurgical) to day 6. We studied eight ICMs coming from six donors (details in [Table S2](#)).

Bulk DNA Extraction and Collection of ICMs and Single Cells

DNA was isolated from whole culture dishes by proteinase K-SDS lysis, followed by phenol-chloroform extraction and ethanol precipitation.

All single cells (hESCs, fibroblasts, and oocytes) and ICMs were collected by mouth-controlled pipetting as previously described ([Spits et al., 2006](#)). Briefly, all cells were sequentially washed in three individual droplets of calcium- and magnesium-free medium, and collected in sterile PCR tubes containing 2.5 μ L lysis solution (200 mM NaOH and 50 mM DTT). The samples were stored at -20°C until further processing. Prior to PCR amplification, the samples were lysed by 10-min incubation at 65°C .

The ICMs were isolated from blastocyst stage human embryos. The ICM was separated from the trophectoderm by mechanical dissociation using a 1,480-nm diode laser (Saturn 5 Laser, Research Instruments, UK) as described in [Capalbo et al. \(2013\)](#). The zona pellucida was removed from the oocytes by incubation with acidic Tyrode's solution prior to collection. For single hESC collection, the hESCs were dissociated into single-cell suspension using TrypLE Express (Thermo Fisher Scientific).

mRNA Quantification

Total RNA was extracted using the RNeasy Micro Kit (Qiagen), and reverse transcription was performed using the First-Strand cDNA

Synthesis Kit (GE Healthcare) following the manufacturer's instructions. Quantitative real-time PCR was performed using the ViiA 7 thermocycler (Thermo Fisher Scientific) and ViiA 7 software v1.2 (Thermo Fisher Scientific), using the standard settings. Quantitative real-time PCR was carried out on 40 ng of cDNA, using the qPCR MasterMix Plus Low ROX (Eurogentec), and TaqMan gene expression assays for *SOX17*, *FOXA2* (endoderm), *SOX1*, *PAX6* (ectoderm), and *GUSB* (Housekeeping) or 1.8 μ M primer mix (IDT) and 250 nM probe (Thermo Fisher Scientific) for *POU5F1* (pluripotency) and *UBC* (housekeeping). Fold changes were calculated with the ddCt method and *UBC* and *GUSB* were used as endogenous control.

Immunostainings

Definitive endoderm was fixed with 4% paraformaldehyde (Sigma-Aldrich), permeabilized with 0.1% Triton X-100 (Sigma-Aldrich) and blocked with 3% BSA (Sigma-Aldrich). Primary antibodies were diluted in 1.5% BSA at a final concentration of 2 μ g/ μ L for *SOX17* (Goat Polyclonal IgG; R&D Systems, Cat# AF1924) and 1 μ g/ μ L for OCT-3/4 (Mouse Monoclonal IgG, Santa Cruz Biotechnology, Cat# SC-5279) and incubated overnight. Donkey anti-Goat IgG (H + L) Alexa Fluor 488 (Thermo Fischer Scientific, Cat# A-11055) and donkey anti-mouse IgG (H + L) Alexa Fluor 594 (Thermo Fischer Scientific, Cat# R37115) were used as secondary antibodies. Ectoderm cells were fixed with 4% paraformaldehyde (Sigma-Aldrich), permeabilized with 100% methanol (Sigma-Aldrich), and blocked with 10% fetal bovine serum (Thermo Fischer Scientific). Primary antibodies were diluted in the blocking solution in a final concentration of 2 μ g/ μ L and 1 μ g/ μ L for *PAX6* (Mouse Monoclonal IgG, Abcam, Cat# ab78545) and OCT3A (Rabbit Monoclonal IgG, Cell Signaling, Cat# C30A3), respectively, and incubated overnight. Donkey anti-rabbit (H + L) Alexa Fluor 488 (Thermo Fischer Scientific) and donkey anti-mouse (H + L) Alexa Fluor 594 (Thermo Fischer Scientific) were used as secondary antibodies. For both endoderm and ectoderm, nuclear staining was performed with Hoechst 33342 (Thermo Fischer Scientific).

mtDNA Massive Parallel Sequencing

We have previously developed a well-validated method for a massive parallel sequencing-based approach to study the mtDNA, both in DNA samples and in single cells. This method can reliably detect SNVs with loads $\geq 1.5\%$ in DNA samples and oocytes, and $\geq 2\%$ in single cells and ICMs ([Zambelli et al., 2017](#)).

The mtDNA fraction was enriched through long-range (LR) PCR (two amplicons, generated by primer set 1 and primer set 2). The primer sequences and cycling protocol are described in [Supplemental Experimental Procedures](#).

The LR-PCR products were pooled, sheared with a Covaris M220 sonicator (Life Technologies) and size-selected for fragments from 200 to 900 bp. The fragmented products underwent end repair, adenylation, and paired-end adapters ligation with the Illumina TruSeq DNA PCR Free HT sample preparation kit. Sequencing was carried out on a HiSeq 1500 (Illumina) platform with an average sequencing depth of 7500 \times .

The files were first aligned to the mitochondrial revised Cambridge Reference sequence (NC 012920.1) using BWA-MEM and sorted ([Li, 2013](#); [McKenna et al., 2010](#)). This was followed by



GATK realignment around indels and recalibration. Finally, variant calling itself was done using Mutect v2.0 (Cibulskis et al., 2013), with a lower frequency threshold of 1.5% for bulk DNA and oocytes and 2% for single cells and ICMs. The data were also analyzed using the mtDNA server (Weissensteiner et al., 2016), and only SNVs identified by both methods were considered. Variants found recurrently in all the samples or in long homopolymeric stretches were not considered. (Excluded variants and further details can be found in Supplemental Experimental Procedures.) In case of a discordant heteroplasmy level, the frequency retrieved with mtDNA server was considered. Pathogenic potential of non-synonymous changes was assessed using Mutpred2 (Pejaver et al., 2017)

4 × 44K Human Genome Array-CGH

We used the protocol for oligo-arrays as implemented in the ISO 15189 laboratory of the Center for Medical Genetics Universitair Ziekenhuis Brussel (UZ Brussel, Belgium), based on the protocol provided by Agilent Technologies, as previously described in Jacobs et al. (2014). In brief, the DNA was purified using Amicon Ultra 30K centrifugal filter tubes (Merck Chemicals); 400 ng of DNA was labeled using the BioPrime aCGH Labeling module (Life Technologies). Blocking and hybridization was carried out using Cot-1 DNA (Life Technologies) and Blocking Agent and hybridization buffer (Agilent Technologies). The samples are hybridized on the microarray slide (4 × 44K Human Genome CGH Microarray, Agilent Technologies) for 24 hr. Slides were washed and dried with Oligo Wash Stabilization and Drying solution following the manufacturer's protocol (Agilent Technologies). Scanning of the slides was done using an Agilent dual laser DNA microarray scanner G2566AA (Agilent Technologies). The data were extracted using Agilent Feature Extraction software and analyzed with the "ArrayCGHbase" analysis platform (<http://medgen.ugent.be/arrayCGHbasecmgg/>) (Menten et al., 2005).

ACCESSION NUMBERS

All the sequencing data have been submitted to the Sequence Read Archive with accession code SRA: SRP125058.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, 1 figure, and 12 tables and can be found with this article online at <https://doi.org/10.1016/j.stemcr.2018.05.007>.

AUTHOR CONTRIBUTIONS

F.Z. co-wrote the manuscript and performed all experiments except when stated otherwise; J.M. performed single-cell collection and assisted with long-range PCR; J.S. and H.V.D.V. performed ICM isolation and oocyte collection; D.D., C.M., A.K., and M.G. provided cell lines and assisted with hESC culture and differentiation experiments; P.T., L.J., and S.V. derived STBG-hiPSCs lines; S.S. assisted with data analysis; K.S. critically revised the manuscript; and C.S. designed the study and co-wrote the manuscript. All co-authors proofread the manuscript.

ACKNOWLEDGMENTS

The authors would like to acknowledge Lise Barbé for providing DNA samples from the VUBi004 lines. The work was funded by the Wetenschappelijke Fonds Willy Gepts (UZ Brussel), the FWO-Vlaanderen (1506616N), and the Methusalem grant of the Vrije Universiteit Brussel to K.S.

Received: December 13, 2017

Revised: May 10, 2018

Accepted: May 15, 2018

Published: June 14, 2018

REFERENCES

- 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.
- Avery, S., Hirst, A.J., Baker, D., Lim, C.Y., Alagaratnam, S., Skotheim, R.I., Lothe, R.A., Pera, M.F., Colman, A., Robson, P., et al. (2013). BCL-XL mediates the strong selective advantage of a 20q11.21 amplification commonly found in human embryonic stem cell cultures. *Stem Cell Rep.* *1*, 379–386.
- Boucret, L., Bris, C., Seegers, V., Goudenège, D., Desquret-Dumas, V., Domin-Bernhard, M., Ferré-L'Hotellier, V., Bouet, P.E., Descamps, P., Reynier, P., et al. (2017). Deep sequencing shows that oocytes are not prone to accumulate mtDNA heteroplasmic mutations during ovarian ageing. *Hum. Reprod.* *32*, 2101–2109.
- Capalbo, A., Wright, G., Elliott, T., Ubaldi, F.M., Rienzi, L., and Nagy, Z.P. (2013). FISH reanalysis of inner cell mass and trophectoderm samples of previously array-CGH screened blastocysts shows high accuracy of diagnosis and no major diagnostic impact of mosaicism at the blastocyst stage. *Hum. Reprod.* *28*, 2298–2307.
- Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol.* *27*, 275–280.
- Chetty, S., Pagliuca, F.W., Honore, C., Kweudjeu, A., Rezanian, A., and Melton, D.A. (2013). A simple tool to improve pluripotent stem cell differentiation. *Nat. Methods* *10*, 553–556.
- Cibulskis, K., Lawrence, M.S., Carter, S.L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E.S., and Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* *31*, 213–219.
- Dziedzicka, D., Markouli, C., Barbé, L., Spits, C., Sermon, K., and Geens, M. (2016). A high proliferation rate is critical for reproducible and standardized embryoid body formation from laminin-521-based human pluripotent stem cell cultures. *Stem Cell Rev.* *12*, 721–730.
- Facucho-Oliveira, J.M., and St John, J.C. (2009). The relationship between pluripotency and mitochondrial DNA proliferation during early embryo development and embryonic stem cell differentiation. *Stem Cell Rev.* *5*, 140–158.



- Hämäläinen, R.H., Manninen, T., Koivumäki, H., Kislin, M., Otonkoski, T., and Suomalainen, A. (2013). Tissue- and cell-type-specific manifestations of heteroplasmic mtDNA 3243A>G mutation in human induced pluripotent stem cell-derived disease model. *Proc. Natl. Acad. Sci. USA* *38*, E3622–E3630.
- Jacobs, K., Mertzaniidou, A., Geens, M., Thi Nguyen, H., Staessen, C., and Spits, C. (2014). Low-grade chromosomal mosaicism in human somatic and embryonic stem cell populations. *Nat. Commun.* *5*, 4227.
- Jacobs, K., Zambelli, F., Mertzaniidou, A., Smolders, I., Geens, M., Nguyen, H.T., Barbé, L., Sermon, K., and Spits, C. (2016). Higher-density culture in human embryonic stem cells results in DNA damage and genome instability. *Stem Cell Rep.* *6*, 330–341.
- Jung, L., Tropel, P., Moal, Y., Teletin, M., Jeandidier, E., Gayon, R., Himmelsbach, C., Bello, F., André, C., Tosch, A., et al. (2014). ONSL and OSKM cocktails act synergistically in reprogramming human somatic cells into induced pluripotent stem cells. *Mol. Hum. Reprod.* *20*, 538–549.
- Kang, E., Wu, J., Gutierrez, N.M., Koski, A., Tippner-hedges, R., Agaronyan, K., Platero-luengo, A., Martinez-redondo, P., Ma, H., Lee, Y., et al. (2016a). Mitochondrial replacement in human oocytes carrying pathogenic mitochondrial DNA mutations. *Nature* *540*, 270–275.
- Kang, E., Wang, X., Tippner-Hedges, R., Ma, H., Folmes, C.D.L., Gutierrez, N.M., Lee, Y., Van Dyken, C., Ahmed, R., Li, Y., et al. (2016b). Age-related accumulation of somatic mitochondrial DNA mutations in adult-derived human iPSCs. *Cell Stem Cell* *18*, 625–636.
- Kelly, R.D.W., and St John, J.C. (2010). Role of mitochondrial DNA replication during differentiation of reprogrammed stem cells. *Int. J. Dev. Biol.* *54*, 1659–1670.
- Kelly, R.D.W., Sumer, H., McKenzie, M., Facucho-Oliveira, J., Trounce, I.A., Verma, P.J., and St John, J.C. (2013). The effects of nuclear reprogramming on mitochondrial DNA replication. *Stem Cell Rev.* *9*, 1–15.
- Lamm, N., Ben-David, U., Golan-Lev, T., Storchová, Z., Benvenisty, N., and Kerem, B. (2016). Genomic instability in human pluripotent stem cells arises from replicative stress and chromosome condensation defects. *Cell Stem Cell* *18*, 253–261.
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*, 1303.3997 [q-GN].
- Loneragan, T., Bavister, B., and Brenner, C. (2007). Mitochondria in stem cells. *Mitochondrion* *7*, 289–296.
- Lund, R.J., Närvä, E., and Lahesmaa, R. (2012). Genetic and epigenetic stability of human pluripotent stem cells. *Nat. Rev. Genet.* *13*, 732–744.
- Maitra, A., Arking, D.E., Shivapurkar, N., Ikeda, M., Stastny, V., Kasauaei, K., Sui, G., Cutler, D.J., Liu, Y., Brimble, S.N., et al. (2005). Genomic alterations in cultured human embryonic stem cells. *Nat. Genet.* *37*, 1099–1103.
- Mandal, S., Lindgren, A.G., Srivastava, A.S., Clark, A.T., and Banerjee, U. (2011). Mitochondrial function controls proliferation and early differentiation potential of embryonic stem cells. *Stem Cells* *29*, 486–495.
- Mateizel, I., De Temmerman, N., Ullmann, U., Cauffman, G., Sermon, K., Van de Velde, H., De Rycke, M., Degreef, E., Devroey, P., Liebaers, I., et al. (2006). Derivation of human embryonic stem cell lines from embryos obtained after IVF and after PGD for monogenic disorders. *Hum. Reprod.* *21*, 503–511.
- Mateizel, I., Spits, C., De Rycke, M., Liebaers, I., and Sermon, K. (2010). Derivation, culture, and characterization of VUB hESC lines. *In Vitro Cell. Dev. Biol. Anim.* *46*, 300–308.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* *20*, 1297–1303.
- Menten, B., Pattyn, F., de Preter, K., Robbrecht, P., Michels, E., Buysse, K., Mortier, G., De Paepe, A., van Vooren, S., Vermeesch, J., et al. (2005). arrayCGHbase: an analysis platform for comparative genomic hybridization microarrays. *BMC Bioinformatics* *6*, 124.
- 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.
- Nguyen, H.T., Geens, M., and Spits, C. (2013). Genetic and epigenetic instability in human pluripotent stem cells. *Hum. Reprod. Update* *19*, 187–205.
- Nguyen, H.T., Geens, M., Mertzaniidou, A., Jacobs, K., Heirman, C., Breckpot, K., and Spits, C. (2014). Gain of 20q11.21 in human embryonic stem cells improves cell survival by increased expression of Bcl-xL. *Mol. Hum. Reprod.* *20*, 168–177.
- Pejaver, V., Urresti, J., Lugo-martinez, J., Kymberleigh, A., Lin, G.N., Nam, H., Mort, M., David, N., Sebat, J., Iakoucheva, L.M., et al. (2017). MutPred2: inferring the molecular and phenotypic impact of amino acid variants. *bioRxiv* <https://doi.org/10.1101/134981>.
- Perales-Clemente, E., Cook, A.N., Evans, J.M., Roellinger, S., Secreto, F., Emmanuele, V., Oglesbee, D., Mootha, V.K., Hirano, M., Schon, E.A., et al. (2016). Natural underlying mtDNA heteroplasmy as a potential source of intra-person hiPSC variability. *EMBO J.* *35*, e201694892.
- Prigione, A., Lichtner, B., Kuhl, H., Struys, E., Wamelink, M., Lehrach, H., Ralser, M., Timmermann, B., and Adjaye, J. (2011). Human induced pluripotent stem cells harbor homoplasmic and heteroplasmic mitochondrial DNA mutations while maintaining human embryonic stem cell-like metabolic reprogramming. *Stem Cells* *29*, 1338–1348.
- Rebolledo-Jaramillo, B., Su, M.S.-W., Stoler, N., McElhoe, J.A., Dickins, B., Blankenberg, D., Korneliusen, T.S., Chiaromonte, F., Nielsen, R., Holland, M.M., et al. (2014). Maternal age effect and severe germ-line bottleneck in the inheritance of human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* *111*, 15474–15479.
- Spits, C., Le Caignec, C., De Rycke, M., Van Haute, L., Van Steirteghem, A., Liebaers, I., and Sermon, K. (2006). Whole-genome multiple displacement amplification from single cells. *Nat. Protoc.* *1*, 1965–1970.



- Suhr, S.T., Chang, E.A., Tjong, J., Alcasid, N., Perkins, G.A., Goissis, M.D., Ellisman, M.H., Perez, G.I., and Cibelli, J.B. (2010). Mitochondrial rejuvenation after induced pluripotency. *PLoS One* *5*, e14095.
- Sui, L., Mfopou, J.K., Geens, M., Sermon, K., and Bouwens, L. (2012). FGF signaling via MAPK is required early and improves Activin A-induced definitive endoderm formation from human embryonic stem cells. *Biochem. Biophys. Res. Commun.* *426*, 380–385.
- Van Blerkom, J. (2008). Mitochondria as regulatory forces in oocytes, preimplantation embryos and stem cells. *Reprod. Biomed. Online* *16*, 553–569.
- Van Haute, L., Spits, C., Geens, M., Seneca, S., and Sermon, K. (2013). Human embryonic stem cells commonly display large mitochondrial DNA deletions. *Nat. Biotechnol.* *31*, 20–23.
- Van Landuyt, L., Polyzos, N.P., De Munck, N., Blockeel, C., Van de Velde, H., and Verheyen, G. (2015). A prospective randomized controlled trial investigating the effect of artificial shrinkage (collapse) on the implantation potential of vitrified blastocysts. *Hum. Reprod.* *30*, 2509–2518.
- Warlich, E., Kuehle, J., Cantz, T., Brugman, M.H., Maetzig, T., Galla, M., Filipczyk, A.A., Halle, S., Klump, H., Schöler, H.R., et al. (2011). Lentiviral vector design and imaging approaches to visualize the early stages of cellular reprogramming. *Mol. Ther.* *19*, 782–789.
- Weissensteiner, H., Forer, L., Fuchsberger, C., Schöpf, B., Kloss-Brandstätter, A., Specht, G., Kronenberg, F., and Schönherr, S. (2016). mtDNA-Server: next-generation sequencing data analysis of human mitochondrial DNA in the cloud. *Nucleic Acids Res.* *44*, W64–W69.
- Yokota, M., Hatakeyama, H., Okabe, S., Ono, Y., and Goto, Y.I. (2015). Mitochondrial respiratory dysfunction caused by a heteroplasmic mitochondrial DNA mutation blocks cellular reprogramming. *Hum. Mol. Genet.* *24*, 4698–4709.
- Yokota, M., Hatakeyama, H., Ono, Y., Kanazawa, M., and Goto, Y. (2017). Mitochondrial respiratory dysfunction disturbs neuronal and cardiac lineage commitment of human iPSCs. *Cell Death Dis.* *8*, e2551.
- Zambelli, F., Vancampenhout, K., Daneels, D., Brown, D., Mertens, J., Van Dooren, S., Caljon, B., Gianaroli, L., Sermon, K., Voet, T., et al. (2017). Accurate and comprehensive analysis of single nucleotide variants and large deletions of the human mitochondrial genome in DNA and single cells. *Eur. J. Hum. Genet.* *25*, 1229–1236.