

Gain of 1q confers an MDM4-driven growth advantage to undifferentiated and differentiating hESC while altering their differentiation capacity

Nuša Krivec¹, Edouard Couvreur de Deckersberg¹, Yingnan Lei¹, Diana Al Delbany¹, Marius Regin¹, Stefaan Verhulst², Leo A. van Grunsven², Karen Sermon¹, Claudia Spits¹

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

²Liver Cell Biology Research Group, Faculty of Medicine and Pharmacy, Vrije Universiteit Brussel, Brussels, Laarbeeklaan 103, 1090 Brussels, Belgium

Lead contact, claudia.spits@vub.be

SUMMARY

Gains of 1q are a highly recurrent chromosomal abnormality in human pluripotent stem cells. In this work, we show that gains of 1q impact the differentiation capacity to derivatives of the three germ layers, leading to miss-specification to cranial placode and non-neural ectoderm during neuroectoderm differentiation and by poorer expression of lineage specific markers in hepatoblasts and cardiac progenitors. Competition assays show that the cells retain their selective advantage during differentiation, which is mediated by a higher expression of *MDM4*, a gene located in the common region of gain. *MDM4* drives the winner phenotype of the mutant cells in both the undifferentiated and differentiating state by reducing the cells' sensitivity to DNA-damage through decreased p53-mediated apoptosis. Finally, we find that cell density in culture plays a key role in promoting the competitive advantage of the cells by increasing DNA damage.

INTRODUCTION

At present, over 7400 hPSC lines have been registered at the hPSC lines registry (<https://hpscereg.eu>) and over 50 ongoing and completed clinical trials involve the transplantation of cells derived from hPSC^{1,2}. While the ability to differentiate is a fundamental characteristic of all hPSCs, they may differ in their differentiation capacity to specific lineages or cell types³. Causes for this variability are common genetic variation^{4,5}, epigenetic variance⁶⁻⁸, mitochondrial mutations and a broad range of environmental factors induced by differences in culture conditions³. Another significant source of variation are genetic changes acquired by the cells during in vitro culture, such as copy number variations^{9,10} and single nucleotide changes such as p53-inactivating mutations¹¹⁻¹³.

Full or segmental gains of chromosomes 1, 12, 17 and 20 are known to be highly recurrent in both human induced pluripotent stem cells (hiPSC) and human embryonic stem cells (hESC)^{9,10,14}. They confer the cells increased cloning efficiency, decreased doubling times, decreased sensitivity to apoptotic triggers, increased self-renewal and growth factor independence¹⁵⁻¹⁹. This results in a growth advantage in the undifferentiated state and a rapid culture take-over by the mutant cells^{17,20}. Presently, the mechanisms and driver gene behind this selective advantage has been fully characterized for only one of these recurrent abnormalities: the gain of 20q11.21 results in a decreased sensitivity to apoptosis due to higher *BCL2L1* expression^{16,17}. Whether this selective advantage is maintained during differentiation remains unanswered. Further, there is increasing evidence that these recurrent abnormalities affect the differentiation capacity of hPSC, often in a cell-lineage specific manner^{15,18,21}. Gains of 20q11.21 and losses of 18q lead to impaired neuroectoderm commitment^{22,23} and gain in chromosome 12p results in large foci of residual undifferentiated cells persisting after hepatoblast differentiation²⁴.

In this study, we focused on the gain of 1q, which is one of the most commonly acquired genetic abnormalities in hPSC worldwide⁹, while little is known on its functional effects. Human ESC with a complex karyotype, including a gain of 1p, have an altered gene expression profile with an activation of WNT signaling and deregulation of tumor suppressors and oncogenes, as well as a differentiation bias towards ectoderm²¹ and gain of 1q was found to recurrently appear during in vitro neural differentiation, suggesting a selective advantage during differentiation²⁵. In this work, we found that gain of 1q alters the differentiation capacity of hESC and established that *MDM4*, a regulator of p53 located in the minimal region of gain, drives their competitive advantage both in the undifferentiated state and during differentiation by reducing their sensitivity to DNA-damage induced p53-mediated apoptosis.

RESULTS

Figure 1 shows the overall setup of this study. All experiments were carried out on the hESC lines VUB03 and VUB19^{26,27}. In vitro culture led to three independent events of gain in chromosome 1. VUB19 and one subline of VUB03 acquired a gain of the entire q arm of

chromosome 1 (termed VUB19^{1q21.1qter} and VUB03^{1q21.1qter}), another subline of VUB03 gained a smaller region spanning 3.3 Mb in 1q32.1 (VUB03^{1q32.1}).

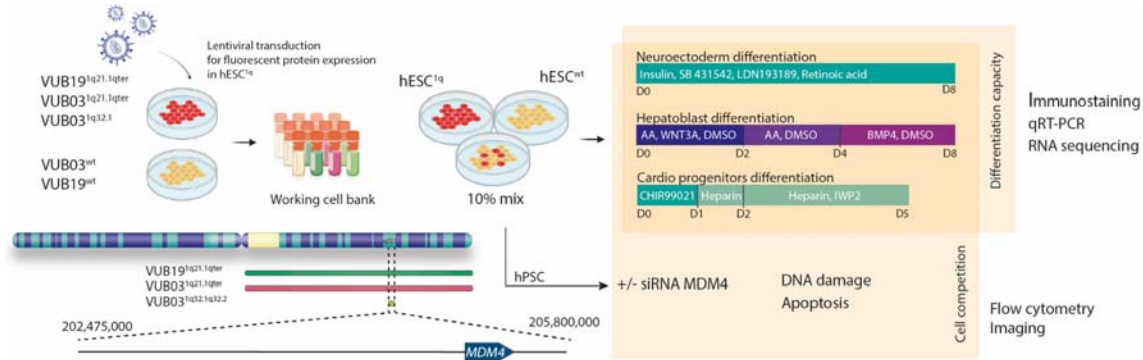


Figure 1: Schematic representation of experimental setup of this study. Lentiviral transduction was used to induce stable expression of fluorescent proteins in hESC^{1q}. Working cell banks were established comprising mutant labeled cell lines and their isogenic counterparts. Pure wild-type and 1q lines were utilized, along with mixes containing 10% of 1q cells, for differentiation to three lineages: neuroectoderm, hepatoblasts and cardiac progenitors. Differentiation capacity was assessed through immunostaining, qPCR, and RNA sequencing. Cell competition results were measured by analyzing the ratio of mutant and wild-type cells before and after differentiation using flow cytometry and immunostaining. To study the role of *MDM4* in the growth advantage of hESC^{1q}, *MDM4* was downregulated by siRNA prior to the competition assays. To study the mechanisms by which *MDM4* provides the selective advantage, DNA damage was induced by bleomycin and apoptosis levels were evaluated using Annexin V staining and flow cytometry, while DNA damage levels were studied through immunostaining for gamma-H2AX.

hESC^{1q} mis-specify to placode and non-neural ectoderm during neuroectoderm differentiation and differentiate to immature hepatoblasts and cardiac progenitors than hESC^{wt}

We investigated the impact of the gain of 1q on trilineage differentiation by subjecting hESC^{wt} and hESC^{1q} to neuroectoderm (NE), hepatoblast (HEP) and cardiac progenitor (CP) differentiation (VUB03^{1q21.1qter}, VUB19^{1q21.1qter}, VUB03^{1q32.1}, VUB03^{wt} and VUB19^{wt}, all lines differentiated at least in triplicate, Figure 1). We measured the expression of six lineage-specific markers and of *NANOG* and *POU5F1* to evaluate differentiation efficiency and to test for residual undifferentiated cells (Figure 2A and 2C and Figure S1). All differentiated cells had almost undetectable levels of *POU5F1* and *NANOG* mRNA, and no POU5F1-positive cells appeared in the immunostainings (Figure 2A), indicating that all cells exited the undifferentiated state (Figure 2C and Figure S1).

NE^{1q} showed significantly lower mRNA levels of *PAX6* and *SOX1* compared to the levels in NE^{wt} (Figure 2C $p^{PAX6}=0.0323$ and $p^{SOX1}=0.0002$, unpaired t-test, N=15), indicating a decreased neuroectodermal differentiation efficiency. In line with this, immunostaining showed a lower percentage of PAX6-positive cells in differentiated hESC^{1q} than in their isogenic hESC^{wt} counterparts, but without reaching statistical significance (Figure 2B,

NE^{1q}=32.15%, NE^{wt}=61.24%, $p=0.0974$, unpaired t-test, N=3). For the HEP differentiation, we found no significant difference between HEP^{wt} and HEP^{1q} in the mRNA expression levels of *HNF4A* while *AFP* had a significantly lower expression in HEP^{1q} (Figure 2C, $p^{\text{HNF4A}}=0.2247$, $p^{\text{AFP}}=0.0142$, unpaired t-test, N=9) and immunostaining for HNF4A showed similar percentages of HNF4A positive cells in both groups (Figure 2B, HEP^{1q}=41.60%, HEP^{wt}=29.79%, $p=0.4514$, unpaired t-test, N=3). Similarly, the mRNA levels of the CP marker *GATA4* was not significantly different between CP^{wt} and CP^{1q} while *NKX2-5* was significantly lower in CP^{1q} (Figure 2C, $p^{\text{GATA4}}=0.1537$, $p^{\text{NKX2.5}}=0.0398$, unpaired t-test, N=18), which was confirmed by the immunostaining for GATA4 (Figure 2B, CP^{1q}=29.56%, CP^{wt}=22.74%, $p=0.4390$, unpaired t-test).

Taken together, these results show that while hESC^{1q} show a decreased differentiation efficiency into neuroectoderm, they do not remain undifferentiated, suggesting that part of the cells miss-specify to alternative cell fates. For the mesendoderm lineages, hESC^{1q} appear to commit equally efficiently to this germ layer than their genetically balanced isogenic counterpart, as reflected by the similar expression for the early markers of endoderm and mesoderm induction, while showing a decreased expression of the markers for further cell-type commitment.

To further characterize the differentiated cells, we carried out bulk mRNA sequencing of 43 samples: 15 samples of neuroectoderm (NE^{wt}=6, NE^{1q}=9, 14 of cardiac progenitors (CP^{wt}=6, CP^{1q}=8) and 14 of hepatoblasts (HEP^{wt}=6, HEP^{1q}=8). Differential gene expression analysis showed that NE^{1q} differentially expressed 1603 genes as compared to NE^{wt}, while HEP^{1q} and CP^{1q} differentially expressed 189 and 241 genes respectively as compared to HEP^{wt} and CP^{wt} (Figure S2, $|\log_2\text{fold-change}|>1.0$ and false discovery rate (FDR)<0.05). We surmised that this significant difference in the number of differentially expressed genes was caused by miss-specification of hESC^{1q} to neuroectoderm and were rather yielding a mixed cell population, which did not occur in the mesendoderm lineages.

To determine the alternate cell fate acquired by hESC^{1q} upon neuroectoderm differentiation, we carried out differential gene expression analysis of the NE^{1q} and the NE^{wt} relative to bulk RNA sequencing data of undifferentiated hESC (N=38 samples, archived lab data). First, we tested the expression of neuroectoderm markers, which we found to be less induced in both NE^{1q} than in NE^{wt} (Figure 2D). We then queried the data for the expression of different sets of markers for embryonic and extra-embryonic lineages that appear in early human development. Table S1 shows the lists of gene sets we curated from published data²⁸⁻⁴⁴. We found NE^{1q} showed high levels of expression of markers of non-central nervous system (non-CNS) ectodermal lineages such as non-neural ectoderm and of cranial placode (Figure 2E and Figure 2F, Figure S3). To confirm this, we double-stained NE^{wt} and NE^{1q} for the non-CNS marker TFAP2A³⁹ and the NE marker PAX6 and found that while hESC^{wt} efficiently differentiate to a homogeneous neuroectoderm cell-population, hESC^{1q} produce a mix of neuroectoderm and non-CNS cells (Figure 2G).

Next, we used this same approach to study the cell types obtained from the HEP and CP samples. While hESC^{1q} equally induce some of the early HEP and CP differentiation markers as hESC^{wt}, 68.6% (24/35) and 63.2% (12/19) of HEP and CP lineage-specific markers have a lower expression in the mutant cells (Figure 2H and Figure 2K). HNF4 and GATA4 stainings (Figure 1A, Figure S1) and RNA sequencing data (Figure S4) did not suggest misspecification in hESC^{1q}. While differentiating to HEP and CP, we analysed the expression of markers specific to different stages of the differentiation of these cell types.

Human ESC-derived HEP have a profile between hepatoblast stage 2 and fetal hepatocyte 2 (Figure 2D), and cells with a gain of 1q show overall lower expression of 64% of markers (16/25). Gene-set enrichment analysis of the differential gene expression analysis between HEP^{wt} and HEP^{1q} showed that 145 of the 162 significantly enriched gene-sets in the canonical pathways gene set had negative enrichment scores and that they were frequently related to processes key to hepatocyte and liver function, including cholesterol metabolism, ferroptosis, transsulfuration, plasma lipoprotein remodelling, folate metabolism, selenium micronutrient network and metabolism of steroids (FDR<0.05, Table S2). Gene ontology enrichment analysis showed that 1574 of the 1913 gene sets had negative enrichment scores, including amino acid metabolic process, fatty acid metabolic process and steroid metabolic process (Figure 2J, Table S2).

In the case of CP, CP^{1q} and CP^{wt} equally express a profile between cardiac progenitors and cardiomyocytes, where CP^{1q} have lower expression of genes marking the later stages of differentiation (Figure 2L). Gene set enrichment analysis of the CP samples showed that CAR^{1q} have negative enrichment scores for 157 of the 178 significantly enriched canonical pathway gene-sets, including sets related to dilated cardiomyopathy, folding of actin, striated muscle contraction and hypertrophic cardiomyopathy (Table S2). These genes are key to correct heart contraction functions. Gene ontology enrichment analysis identified negative enrichment scores for 1146 of the 1786 significantly enriched set, including terms, such as I band, Z disc, sarcomere, myofibril, contractile fiber, muscle system process and muscle development (Figure 2M, Table S2).

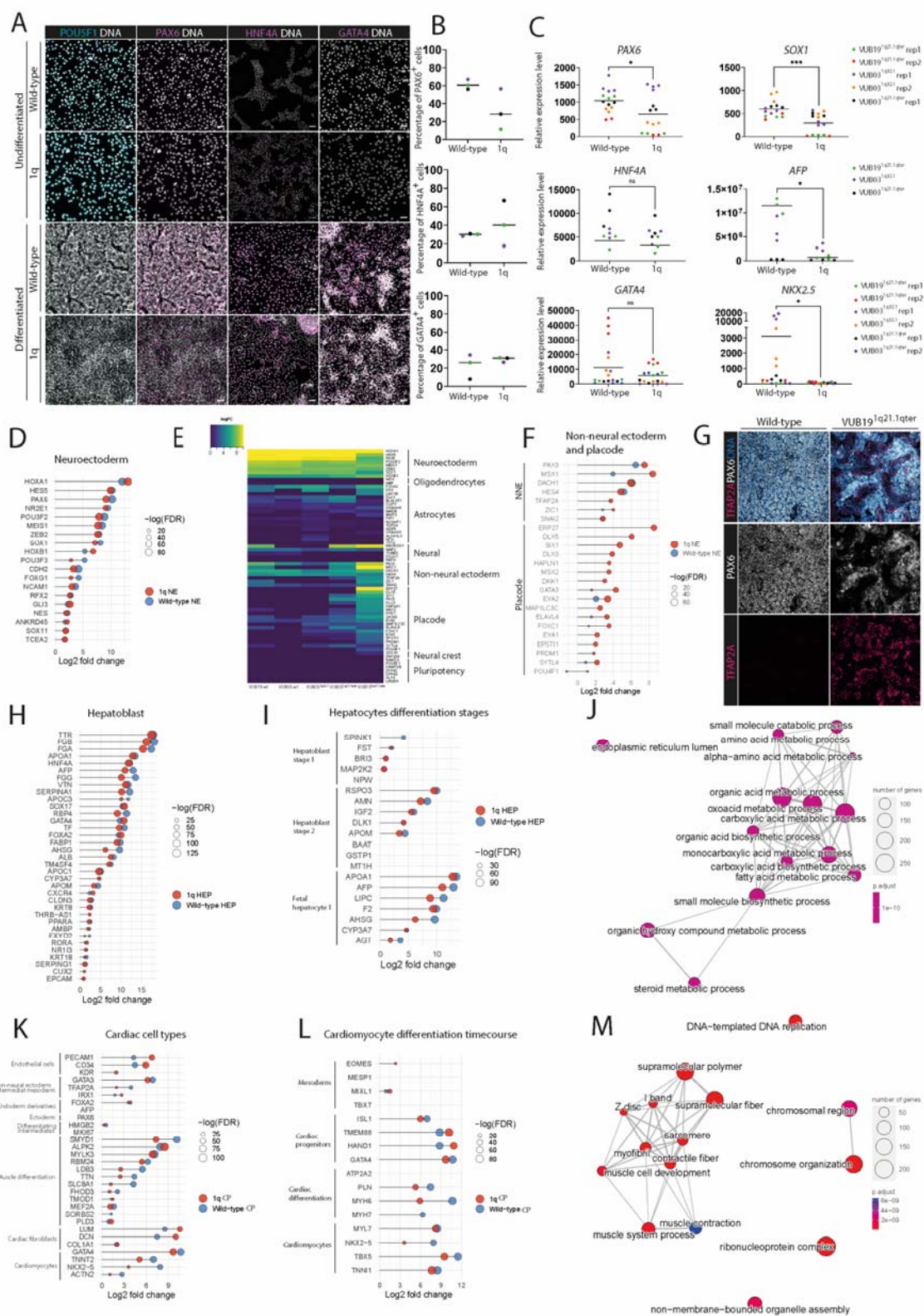


Figure 2. hESC^{1q} miss-specify to placode and non-neural ectoderm during neuroectoderm differentiation and differentiate to more immature hepatoblasts and cardiac progenitors than hESC^{wt}.

A Examples of immunostaining for POU5F1, PAX6, HNF4A and GATA4 of hESC^{wt} and hESC^{1q} before and after the completion of induction of differentiation to neuroectoderm (NE, day 8), hepatoblast (HEP, day 8), cardiac progenitors (CP, day 5) of VUB19^{1q21.1qter}. Figure 1S shows the images for cell lines VUB03^{1q21.1qter} and VUB03^{1q32.1} included in the study.

B Quantification of the number of PAX6, HNF4A and GATA4-positive cells in the immunostainings shown in A and in Figure 1S. No statistically significant differences were found.

C mRNA quantification by quantitative real-time PCR of *PAX6* and *SOX1* in NE^{wt} and NE^{1q}, *HNF4A* and *GATA4* in HEP^{wt} and HEP^{1q}, and *GATA4* and *NKX2.5* in CP^{wt} and CP^{1q}; * p<0.05, *** p<0.001. Results for *POU5F1* and *NANOG* are shown in Figure 1S. Each differentiation was carried out in three technical replicates at the same time. Different hESC lines and independent experimental replicates (rep) are coded in different colors.

D Lollipop diagrams representing the differential gene expression of neuroectoderm markers in NE^{wt} and NE^{1q} vs undifferentiated cells. All genes shown in this panel and in E, H, I, K, and L are induced with a log₂fold-change over 0 and FDR<0.05.

E Lollipop diagrams representing the differential gene expression of placode markers in NE^{wt} and NE^{1q} vs undifferentiated cells.

F Heatmap representing the induction of expression of genes marking different ectodermal cell types and the undifferentiated hESC state in NE obtained from the five lines included in this study.

G Representative images of the immunostaining of NE^{wt} and NE^{1q} for PAX6 and TFAP2A.

H Lollipop diagrams representing the differential gene expression of hepatoblast markers in HEP^{wt} and HEP^{1q} vs undifferentiated cells.

I Lollipop diagrams representing the differential gene expression of markers of different stages of hepatocyte differentiation in HEP^{wt} and HEP^{1q} vs undifferentiated cells.

J Enrichment map showing the top-15 deregulated pathways from gene ontology gene set enrichment analysis in HEP^{wt} vs HEP^{1q}. The size of nodes indicates the number of genes in each pathway and the color represents adjusted p-value. Pathways that cluster together have overlapping gene sets.

K Lollipop diagrams representing the differential gene expression of genes marking different cardiac cell types in CP^{wt} and CP^{1q} vs undifferentiated cells.

L Lollipop diagrams representing the differential gene expression of genes marking different stages of cardiomyocyte differentiation in CP^{wt} and CP^{1q} vs undifferentiated cells.

M Enrichment map showing the top-15 deregulated pathways from gene ontology gene set enrichment analysis in CP^{wt} vs CP^{1q}.

Human ESC^{1q} have an *MDM4*-driven competitive advantage that is retained during differentiation

Human ESC^{1q} have a well-established competitive advantage over their genetically balanced counterparts in the undifferentiated state^{9,11,14}. We next aimed at determining whether the cells retain this selective advantage during differentiation, and which gene is driving their winning phenotype.

We first looked at the 53 commonly deregulated genes in our HEP, CP and NE samples, and found that *MDM4*, a regulator of p53 activity that is located in the common region of gain and which has been previously suggested as a key gene for the gain of 1q¹¹, is consistently upregulated in all samples (Figure 3A). Gene set enrichment analysis of the differentially expressed genes in NE, HEP and CP from 1q and wt cells for the Reactome pathways related to p53 signaling indeed shows that the transcriptional regulation by p53, the regulation of p53 activity, the G1/S damage checkpoint and the p53 dependent responses to DNA-damage in G1 and S are all significantly negatively enriched in cells carrying a gain of 1q (Figure 3B, Table S2, Table S3). This led us to the hypothesis that the higher expression of *MDM4* in cells with a gain of 1q results in the inhibition of p53-mediated transcriptional activation upon DNA damage, leading to a decreased induction of apoptosis upon DNA damage, thus providing a competitive advantage to the mutant cells (Figure 3C).

To determine if cells with a gain of 1q retained their ability to take over the culture during differentiation, we carried out competition assays during NE, HEP and CP induction. For this, 10% of hESC^{1q} stably expressing a fluorescent protein was introduced into a unlabeled hESC^{wt} culture. Differentiation was initiated the next day, and was controlled by immunostaining for PAX6, HNF4A and GATA4 (Figure 3D). To measure culture-take over, the proportion of hESC^{1q} was determined by flow cytometry at the onset and at the end of differentiation.

We found that in all three differentiations, the cells with a gain of 1q outcompeted wild type cells (Figure 3E). During the NE induction, the proportion of 1q cells increased in average 33.9%±2.7% during the 8-day differentiation, from a mean of 10.9% at the onset to 44.9% at day 8 ($p < 0.0001$, unpaired t-test, N=18-21). This was similar for the 8-day differentiation to HEP, where the 1q cells increased in average 49.3%±4.9% (mean at day 0=11.6%, mean at day 8=60.9%, $p < 0.0001$, unpaired t-test, N=9). This increase was less pronounced during the 5-day CP differentiation, with an average 22.7%±3.2% increase (mean at day 0=10.9%, mean at day 5=33.5%, $p < 0.0001$, unpaired t-test, N=18- 21), which may be attributable to the shorter time span of the CP differentiation as compared to NE and HEP.

Next, we tested the role of *MDM4* in the competitive advantage of cells with a 1q. For the undifferentiated cells, we downregulated *MDM4* for 24h by siRNA in hESC^{1q} on the day before to start of the competition assay. The cells were imaged daily (Figure 3F), and the proportion of hESC^{1q} was measured with an automated cell counter. On average, hESC^{1q} increased from 14.0 to 33.7% in three days in the untreated condition, while no increase was observed after siRNA treatment (N=3, $p^{1q}=0.0028$, $p^{si1q}=0.7905$, 2-way ANOVA, Figure 3G). For the competition assays during differentiation, the hESC^{1q} were treated for 24h with the siRNA and mixed at a 1:9 ratio, and differentiation was initiated the next day.

Differentiation was shortened to 4 days, as this was the time previous work in the lab showed that a single siRNA transfection could reliably sustain a gene's downregulation²². Figure 3H shows the effect of the siRNA on the expression of *MDM4* and of downstream targets of the p53 signaling pathway in the three hESC^{1q}. In the untreated conditions of the competition assays, the fraction of cells with a gain of 1q became significantly larger in all three differentiations (Figure 3I and 3J). The increase was most pronounced after NE induction, with an average increase of 30.0% ($p^{1q}=0.0053$, 2-way ANOVA). In HEP and CP, the mean increases were 11.3% ($p^{1q}=0.0131$, 2-way ANOVA) and 11.67 ($p^{1q}=0.0149$, 2-way ANOVA), respectively. Conversely, in the siRNA-treated competition assays, the fraction of 1q cells remained unchanged (Figure 3I and 3J). Taken together, these results show that reducing the levels of *MDM4* in the mutant cells abolishes their competitive advantage both in the undifferentiated state and during differentiation.

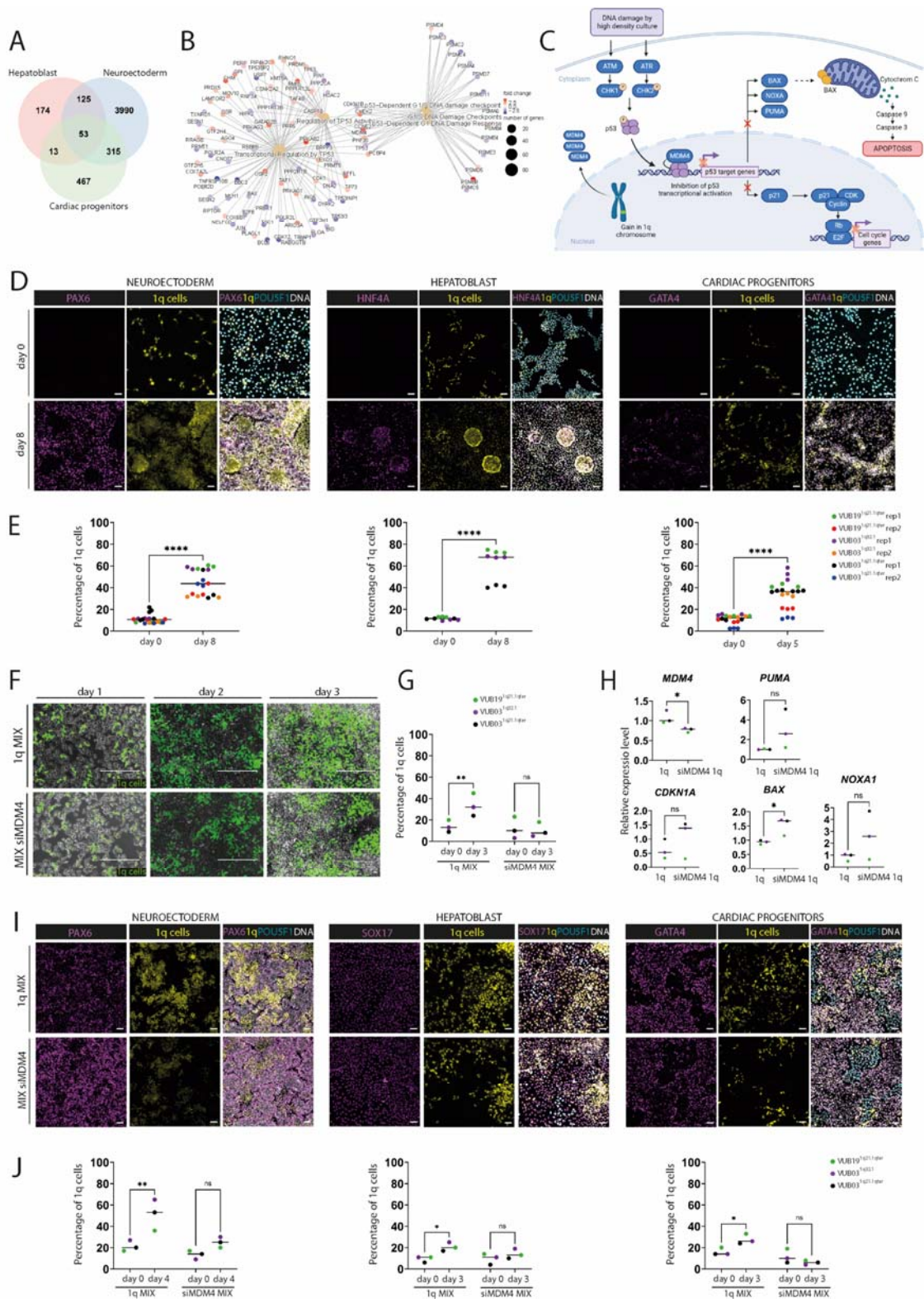


Figure 3. $hESC^{1q}$ have an *MDM4*-driven competitive advantage that is retained during differentiation

A Venn diagram showing the overlaps in differentially expressed genes between control and 1q-cells, for the three cell types with \log_2 -fold-change over 0 and $FDR < 0.05$. *MDM4* is part of the core of 53 commonly deregulated genes.

B Reactome pathway gene set enrichment analysis of the C2 library canonical pathways related to p53 signaling in NE, HEP and CP from 1q and wt cells.

C Graphical summary of the model in which *MDM4* has a higher expression because of the gain of 1q, resulting in an inhibition of the p53 signaling pathway and in turn decreasing the sensitivity of the cells to DNA-damage.

D Examples of immunostaining for PAX6, HNF4A and GATA4 after induction cell competition during the differentiation to neuroectoderm, hepatoblast and cardiac progenitors.

E Quantification of the numbers of 1q-cells at the start and end of differentiation.

F Competition assays in the undifferentiated state. $hESC^{1q}$ are mixed in at a 1:9 ratio with $hESC^{wt}$ and traced by their fluorophore expression. In the siMDM4 condition, the $hESC^{1q}$ have been treated with siRNA against *MDM4* prior to mixing.

G Quantification of the numbers of $hESC^{1q}$ at the start of the cell competition (day 0) and at day 3 of in vitro culture.

H mRNA expression of *MDM4* and downstream p53 targets in siMDM4 treated and untreated $hESC^{1q}$

I Examples of immunostaining for PAX6, SOX17 and GATA4 after induction cell competition during the differentiation to neuroectoderm, hepatoblast and cardiac progenitors, with and without treatment of the $hESC^{1q}$ with siRNA against *MDM4*.

J Quantification of the numbers of 1q-cells at the start of the cell competition during differentiation (day 0) and at day 3 or 4, with and without siRNA against *MDM4*.

* $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$

Higher *MDM4* expression in $hESC^{1q}$ results in a decreased sensitivity to DNA-damage induced apoptosis

Next, we sought to elucidate by which mechanisms higher expression of *MDM4* confers the competitive advantage to the cells. We first tested the hypothesis that the higher expression of *MDM4* by cells carrying a gain of 1q leads to a decreased p53-mediated apoptosis in response to DNA damage. For this, we induced DNA damage in $hESC^{1q}$ and $hESC^{wt}$ using bleomycin and carried out a time-course measurement of apoptosis and cell death.

Figure 4A shows the percentages of live and of apoptotic and dead cells for $hESC^{wt}$ (N=3), $hESC^{1q}$ (N=3) and $hESC^{1q}$ treated with siRNA against *MDM4* (N=3), at the start of bleomycin treatment and at the subsequent 2, 4 and 6-hour time-points. $hESC^{wt}$ start undergoing apoptosis at 2h after exposure to bleomycin, followed by a rapid decrease in the numbers of live cells. In contrast, apoptotic cells start appearing in $hESC^{1q}$ as from 4h of exposure and reach 41.9% of apoptotic cells at 6h, as compared to 78.6 % in $hESC^{wt}$ (unpaired t-test, $p = 0.0169$). Treating $hESC^{1q}$ with siMDM4 significantly increases their sensitivity to DNA

damage, with apoptosis initiating at 2h, and reaching 49.4% at 4h. siMDM4-treated hESC^{1q} cells do not reach same levels of dead cells at 6h as in hESC^{wt} cells, although the differences are not statistically significant (54.9% vs 78.6%, unpaired t-test, p=0.1530, Figure 4A). This may be explained by an incomplete transfection of the cells, or to overall insufficiently stable downregulation of *MDM4* to the levels of hESC^{wt}.

This raised the question of how a decrease of sensitivity to DNA-damage could provide a competitive advantage to hESC and differentiating cells. We observed during daily monitoring of the competition assays that hESC^{1q} started outcompeting their genetically balanced counterparts once the cultures became confluent. Previous work has indicated that hPSC are prone to replication stress and DNA damage^{45,46}, which can be mitigated by addition of nucleosides to the medium⁴⁷ and exacerbated in higher cell density culture due to medium acidification^{48,49}. We therefore hypothesized that higher culture cell density generates the conditions for a strong competitive advantage of 1q gains, by increasing the levels of DNA damage.

We studied the cell proliferation dynamics of hESC^{wt} and hESC^{1q} over time in culture in pure form and mixed at a 1:9 ratio (N=3 for each condition). Daily cell numbers count showed that the numbers of hESC^{wt} and hESC^{1q} increased similarly until day 7, suggesting that they have similar cell doubling times. As from day 8, hESC^{1q} continue proliferating even when they have reached a very high density (Figure 4B). Daily analysis of the ratio between hESC^{wt} and hESC^{1q} showed no statistically significant changes until day 3, after which the proportion of hESC^{1q} started steadily increasing (Figure 4C, N=3). This coincided with culture dishes at day 3 still showing empty areas, whereas at day 4 the cells had reached confluence, suggesting that this is a flipping point for 1q gains to start providing a competitive advantage (Figure 4D). To investigate the relationship between cell culture density and DNA damage, we measured DNA damage by γ H2AX staining in hESC^{1q} and hESC^{wt}, in low and high cell density cultures (mean of 7,547 cells/cm² and 200,050cm/cm² respectively, representative images are shown in Figure 4E and cell counts shown in Figure 4F), and found that cells grown at low density had significantly lower numbers of γ H2AX foci than when grown in high density (Figure 4G, wt_{low}:11.38, wt_{high}:17.95, 1q_{low}:13.71, 1q_{high}:26.22, p<0.0001, one-way ANOVA). While there were no differences in DNA-damage between in hESC^{1q} and hESC^{wt} grown at low density (Figure 4G, 1.2-fold change, p=0.2314, one-way ANOVA), at high density, hESC^{1q} showed higher γ H2AX foci counts than hESC^{wt} (Figure 4G, 1.5-fold change, p<0.0001, one-way ANOVA). Lastly, we counted the numbers of cells in the cell cultures until day 5 and calculated the densities. The mean densities in days 1 and 2 are in the range of the 'low density' group in the DNA damage staining, with a mean of 33,711 cells/cm² at day 1 and 87,854 cells/cm² at day 2. From day 3, the densities are in the range and above of the 'high density' group, with 265,224cells/cm² at day 3, 331,680 cells/cm² at day 4 and 351,475 cells/cm² at day 5. This indicates that at this point, the cells start undergoing DNA damage more frequently, reaching the condition when a decreased

p53-mediated induction of apoptosis starts providing a selective advantage, so cells with a gain of 1q will start outcompeting their genetically balanced counterparts.

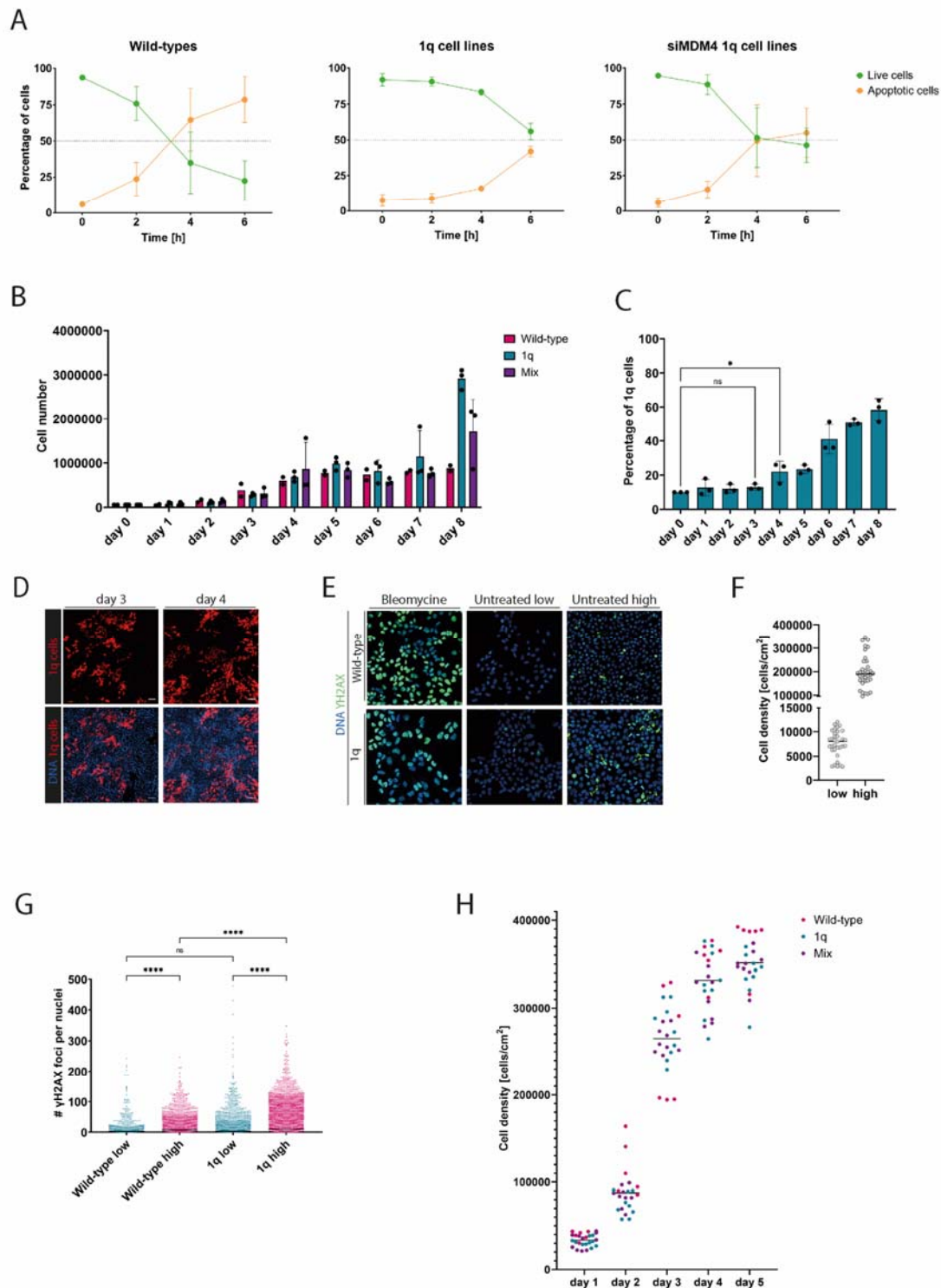


Figure 4. Higher MDM4 expression in hESC^{1q} results in a decreased sensitivity to DNA-damage induced apoptosis

A Percentages of apoptotic and live cells in hESC^{1q} and hESC^{wt} upon 0h, 2h, 4h and 6h of exposure to bleomycin.

B Evolution of the numbers of cells in one culture dish from day 0 to day 8, for hESC^{1q}, hESC^{wt} and a 1:9 mix of hESC^{1q} and hESC^{wt}.

C Percentage of hESC^{1q} over the course of 8 days in culture, starting from a 1:9 mix of hESC^{1q} and hESC^{wt}. The proportion of hESC^{1q} becomes significantly increased as from day 4. * p<0.05.

D Representative images illustrating the cell density at days 3 and 4 of the time-course experiment. Confluency is reached at day 4 of this experiment.

E Examples of the staining for gammaH2AX in hESC^{1q} and hESC^{wt}, upon exposure to bleomycin and in low- and high-density cultures.

F Cell densities for the low- and high-density culture conditions shown in E and G. The dots represent the counts made in different images of dishes of different cell lines. The line represents the mean of the dots.

G Counts of gammaH2AX foci per nucleus in hESC^{1q} and hESC^{wt} at low and high-density culture conditions. Each dot represents a single nucleus, hESC^{wt-low} N=976, hESC^{wt-high} N=2879, hESC^{1q-low} N=2995, hESC^{1q-high} N=7158. **** p<0.0001, one-way ANOVA and Šidák correction.

DISCUSSION

Despite that the appearance of abnormal karyotypes in hESC was first noticed nearly 20 years ago⁵⁰ we still have limited insight onto their mechanisms of origin, their driver genes and their functional consequences. In this study, we find that hESC^{1q} have an abnormal response to the differentiation cues to neuroectoderm, with a part of the population misspecifying to non-CNS cells. Also, while entering the mesendodermal lineage equally efficiently as hESC^{wt}, hESC^{1q} generate hepatoblasts and cardiac progenitors that have gene-expression profiles suggestive of a poor cell maturation. It is interesting to note that hESC with a gain of 20q11.21^{22,23}, an isochromosome 20⁵¹ and with a loss of 18q⁵² all display an abnormal response to the neuroectoderm differentiation signaling cues. This common response suggests a particular vulnerability of the neuroectodermal lineage to genetic abnormalities, with potentially common signaling pathways disrupting the differentiation process. In this study, we have identified the alternative cell fate as non-CNS ectoderm, the cells expressing markers of cranial placode and non-neural ectoderm. These cell types are not commonly obtained in the standard NE differentiation by dual-SMAD inhibition, but it has been shown that the addition of different concentrations of BMP4 and modulation the WNT pathway will drive the cells towards these fates³⁹. Further, it has been shown that DNA damage-induced stabilization of p53 in hPSC results in an activation of TGF-beta signaling and a downregulation of BMP4 pathway genes⁵³, suggesting that deregulation of these pathways specifically in hESC^{1q} during the differentiation process may be at the basis of this alternate fate acquisition.

The differentiation of mesendoderm cell types yielded similar proportions of correctly specified cells, but gene-expression analysis revealed that the hepatoblasts and cardiac

progenitors obtained from hESC^{1q} had profiles suggestive of a more immature cell types. Whether this reflects delayed or impaired maturation and whether this extends to other mesendoderm derivatives remains to be elucidated, but it further supports the notion that genetic abnormalities do affect the quality of the cells obtained after differentiation, which can negatively impact both the clinical translation of hPSC as well as the reproducibility of results when using these cells in research⁵⁴. Further, here we also show that the genetically variant cells retain their ability to take over a culture during differentiation. Considering that 10 to 20% of cells in an hPSC culture carry a variety of copy number variations⁵⁵⁻⁵⁷, and 10% of hESC lines carry mosaic large structural variants¹¹, it is possible that the differentiation of a seemingly genetically normal line yields a mixed population of euploid and aneuploid cells with varying degrees of correct cell specification. This lays a further layer of complexity into controlling the outcomes of differentiation, as current methods for genetic screening will easily miss low-grade mosaicism.

In this work, we also identify *MDM4* as the driver gene behind the competitive advantage of the cells with a gain of 1q, and find that this advantage is based on a lower sensitivity to DNA damage-induced apoptosis by the suppression of p53 activity by *MDM4*. An interesting aspect of our findings is that cell density in culture plays a key role in generating the conditions under which hESC^{1q} can outcompete their wild-type counterparts. This finding opens the possibility of tailoring culture conditions to suppress mutant cell take-over, such as by controlling cell density or decreasing DNA damage by regulating the pH of the culture medium^{49,56} or by addition of factors such as nucleosides⁵⁸.

Finally, *MDM4* has recently been shown to also be the driver of the addiction of cancers to chromosome 1q gains to support their malignant growth⁵⁹. The similarities between the genetic abnormalities found in cancers and hPSC have been already noted in the past⁶⁰⁻⁶². Gains of 20q11.21, one of the most common acquired abnormalities in hPSC⁹, appear in over 80% of pancreatic and colorectal carcinomas^{63,64} and gains of 1q appears in over half of the cases of hepatocellular carcinoma⁶⁵ and have been recently identified as the first copy number alterations occurring in breast cancer and melanoma evolution⁵⁹. These parallels, together with the high incidence of the gain of 1q in culture, the fact that the variant cells can take over also during differentiation and that the gain of 1q impacts the final cell product, stress the importance of factoring in the genetic integrity of the cells not only in a clinical setting, but also in a research context to ensure reliable and reproducible science^{66,68}.

MATERIALS AND METHODS

hESCs lines, cell culture and banking and transgenic modification

All the hESC lines were derived from human embryos and characterized as described in ^{26,27}. They are registered in the EU hPSC registry (<https://hpscereg.eu/>). VUB03 and VUB19 remained genetically balanced up to passage 20 and 24, respectively. All stocks of hESC lines are preserved in liquid nitrogen in 90 % Knock-out serum (Thermo Fisher Scientific) and 10% sterile DMSO. hESC were cultured in dishes coated with human recombinant Laminin-521 (Biolamina). Laminin-521 was diluted in phosphate buffered saline (PBS) (Thermo Fisher) with calcium and magnesium. Coated dishes were stored overnight at 4°C before the use. NutriStem™ (Biological Industries) medium with 10mM penicillin/streptomycin was changed daily. hESC cultures were kept in the incubator at 37°C and 5% CO₂. For passaging, cells were washed with PBS and incubated with recombinant enzymes TrypLE (Thermo Fisher) for 10 minutes at 37°C for single cell dissociation. TrypLE was deactivated with NutriStem™ medium and cells were centrifuged at 1000 rpm for 5 min. Pellet was resuspended in 1 ml of medium and counted with image-based cytometer Tali™ (Invitrogen) if needed or transferred in a new dish in ratio between 1:20 to 1:50.

Labelled hESC1q cells were generated by infecting hESC1q with lentiviral particles expressing fluorescent protein. HEK293T cells were transfected with pMDG (VSV.G encoding plasmid), pCMVΔR8.9 (gag---pol encoding plasmid) as packing vectors and three different vectors each with their own fluorescent protein (LeGO---EF1a---V2---Puro (Venus fluorescent protein), LeGO---EF1a---C2---Puro (mCherry) and pCDH---EF1---MCS---pA---PGK---copGFP---T2A---Puro (Green fluorescent protein). Vectors with mCherry and Venus fluorescent protein were kindly provided by Kristoffer Riecken⁶⁷. Vector with GFP was a gift from VUB Laboratory for molecular and cellular therapy. Vector was purchased from Addgene (<https://www.addgene.org/17448/>) and cloned CMV promoter to EF1a. Lentiviruses were produced by transfection of HEK 293T cells with VSV.G and gag-pol plasmids together with the plasmid of interest and PEI (2ug per ug of DNA, Polysciences Inc) in Opti-MEM medium (Thermo Fisher Scientific). After 4h of transfection the transfection cocktail was replaced by complete medium. Supernatant containing lentiviral particles was collected after 48 and 72h and stored at -80C. A day before hPSC were seeded at the density 50.00cells/cm². hESC were transduced with 1:1 mix of Nutristem and complete medium containing lentivirus together with 1:1000 protamine sulfate (LEO Pharma; 10mg/ml). Cells were incubated with transduction cocktail for 4h. Next, cells were washed 5 times with PBS and refreshed with Nutristem medium. Selection of successfully transduced cells was done with FACS.

Working cell banks comprising labeled 1q cell lines and their isogenic counterparts were established, which were additionally controlled for cell identity and karyotype. We utilized the working cell bank for subsequent experiments and the cells were not kept in culture for more than 5 passages after thawing to prevent them from genetic drift.

Genome characterization

Genomic DNA was extracted with DNeasy Blood & Tissue Kit (Qiagen) following producer's instructions. Concentration was assessed with NanoDrop 1000 (Thermo Fisher Scientific). DNA was stored at +4°C. The genetic content of the hESCs was assessed through shallow whole-genome sequencing by the BRIGHTcore of UZ Brussels, Belgium, as previously described⁶⁹. Copy-number assays were used to control for gains of 1q, 12p and 20q11.21 at regular intervals. Real time polymerase chain reaction was performed on ViiA™ 7 system (Applied Biosystems). Total volume of qPCR reaction was 20µl and it contained 10µl of 2x qPCR Master Mix Low ROX (Eurogentec), 1µl of 20x TaqMan Copy Number Assay (Life Technologies), 1µl of nuclease-free water and 40 ng of DNA. TaqMan™ Copy Number Reference Assay RNase P (Applied Biosystems, 4403328) was added instead of nuclease-free water. Genomic DNA from leukocytes of healthy donor was used as a control. Non template control was tested in each experiment. Cycling parameters are 2min at 50°C, 40 cycles with 10min at 95°C and 15 sec at 95°C and 1 min at 60°C. Each sample was tested in triplicates. Copy number assays used were KIF14 (Thermo Fisher, Hs_cn), NANOG (Thermo Fisher, Hs_cn) and ID1(Thermo Fisher, Hs_cn). Experiments were run as either comparative Ct (cycle threshold) or relative standard curve analysis. Data analysis was performed by ViiATM 7 v2.0 software or Copy Caller v2.1 (Applied Biosystems).

Neuroectoderm differentiation

The protocol for neuroectoderm induction was adapted from Douvaras et al., 2015⁷⁰. We seeded 22.500 cells/cm² on dishes coated with Laminin-521 and 1:100 RevitaCell (Thermo Fisher Scientific). After 24h differentiation was induced with freshly prepared neural induction medium containing from DMEM/F12 (Thermo Fisher Scientific), 1x NEAA (TFS), 1x GlutaMAX (TFS), 1x 2-mercaptoethanol (TFS), 25 ug/ml insulin (Sigma-Aldrich) and 1x penicillin/streptomycin (TFS). The medium was supplemented with 10µM SB431542 (Tocris), 250 nM LDN193189 (STEMCELL Technologies) and 100 nM RA (Sigma-Aldrich). The medium was refreshed every day for 8 days.

Cardiac progenitor differentiation

The protocol for cardiac progenitor differentiation was modified from Lin et al., 2020⁷¹. hESC were seeded at the density of 22.500 cells/cm² on Laminin-521 coated plates. RevitaCell was added to the cell suspension in 1:100 dilution before the seeding. When cells reached 90% confluency, differentiation was initiated with cardiomyocyte differentiation basal medium (CDBM) supplemented with 5µM CHIR99021 (Tocris). After 24h treatment with CHIR99021, medium was changed with fresh CDBM with addition of 0,6U/ml heparin (Sigma-Aldrich). For the next 3 days CDBM with 0,6U/ml heparin and 3mM IWP2 (Tocris) was refreshed daily. CDBM medium was composed of DMEM/F12 (Thermo Fisher Scientific), 64 mg/L L-ascorbic acid (Sigma-Aldrich), 13.6 ug/L sodium selenium (Sigma-Aldrich), 10ug/ml transferrin (Sigma-Aldrich) and 1x chemically defined lipid concentrate (Thermo Fisher Scientific).

Hepatoblast differentiation

The hepatoblast differentiation protocol was based on protocol from Boon et al. 2020⁷². hESC were seeded with 1:100 RevitaCell (Thermo Fisher Scientific) at the density of 35.000cells/cm² on Laminin-521 coated plates. Next day Nutristem medium was refreshed. Differentiation was initiated the next day with liver differentiation medium (LDM) supplemented with 50 ng/ml Activin A (STEMCELL Technologies), 50ng/ WNT3A (PreproTech) and 6ul/ml DMSO (Sigma-Aldrich). After 48h, medium was refreshed without WNT3A factor. On days 4 and 6 LDM medium with added 50ng/ml BMP4 (STEMCELL Technologies) and 6ul/ml DMSO was used. LDM medium was composed of MCDB 201 medium with pH 7,2 (Sigma-Aldrich), DMEM High Glucose medium (Westburg Life Sciences), L-Ascorbic Acid (Sigma-Aldrich), Insuli-Transferrin-Selenium (ITS-G, Thermo Fisher Scientific), Linoleic Acid-Albumin (LA-BSA, Sigma-Aldrich), 2-Mercaptoethanol (Thermo Fisher Scientific) and Dexamethasone (Sigma-Aldrich).

Total RNA isolation, cDNA synthesis and quantitative real time PCR

RNA was extracted from cell pellets with RNeasy Mini Kit (Qiagen) following the producer's protocol. Concentration of obtained RNA was measured with NanoDrop 1000 (Thermo Fisher Scientific). RNA was stored at -80°C. Reverse transcription of RNA to cDNA was performed with First-Strand cDNA Synthesis Kit (GE Healthcare) following producer's instructions. cDNA was stored at -20°C.

Real-time qPCR was performed on ViiA™ 7 system (Applied Biosystems). Total volume of qPCR reaction was 20µl and it contained 10µl of 2x qPCR Master Mix Low ROX (Eurogentec), 1µl of 20x TaqMan Gene Expression Assay (Life Technologies), 1µl of nuclease-free water and 40 ng of cDNA. Each sample was tested in triplicates, GUSB was used as a house-keeping gene. Genes and assays are listed in Table S4. We used the standard cycling parameters of the ViiA7 instrument. Experiments were run as comparative Ct (cycle threshold) analysis. Data analysis was performed by ViiATM 7 v2.0 software.

Immunofluorescent Staining

Cells were washed 3 times with PBS and incubated for 15 min with 3,6% paraformaldehyde at room temperature for fixation. Cells were washed again 3 times with PBS. Blocking was performed with 1-2 hour incubation with 10% fetal bovine serum (FBS). Primary antibodies were diluted in specific ratio in 10 % FBS and incubated with cells for 1 hour in the dark at room temperature. After washing 3 times with PBS, secondary antibodies were diluted in FBS together with 1:2000 dilution of Hoechst 33342 (Life Technologies) and incubated for 2 hours at the room temperature. Cells were again washed three times with PBS and stored at 4°C until analyzed with confocal microscope Zeiss. List of antibodies is shown in Table S4.

Flow cytometry and Annexin V staining

The hESC were harvested by incubating them 10 min at 37°C with 1 ml of TrypLE (Thermo Fisher). The cell suspension was gently pipetted up and down with 2 ml of medium and strained through 20 µm cell strainer (pluriStrainer) to eliminate any cell clumps. Single cell suspension was centrifuged at 1000 rpm for 5 min. Medium was aspirated, cells were resuspended in 1 ml of PBS and spun down at 1000 rpm for 5 min. If needed, the cells were incubated with Alexa Fluor® 647 Annexin V (BioLegend) based on the manufacturer's instructions. Next, the cells were washed with 1mL of PBS and centrifuged. Incubation with Live/dead stain (Thermo Fisher Scientific) was performed following the producer's protocol. Cells were washed again and fixed with 3,6% paraformaldehyde at room temperature for 10 min. After centrifugation (1000 rpm for 5 min) the paraformaldehyde was aspirated, the cells were washed with 1 ml of PBS and spun down. The fixed cells were resuspended in PBS and stored at 4°C until the analysis with Flow cytometer Aria III (BD).

RNA Sequencing

RNA-seq library preparation was performed using QuantSeq 3² mRNA-Seq Library Prep Kits (Lexogen) following Illumina protocols. Sequencing was performed on a high-throughput Illumina NextSeq 500 flow cell. The FastQC algorithm⁷³ was used to perform quality control on the raw sequence reads prior to the downstream analysis. The raw reads were aligned to the new version of the human Ensembl reference genome (GRCh38.p13) with Ensembl (GRCh38.83.gtf) annotation using STAR version 2.5.3 in 2-pass mode⁷⁴. The aligned reads were then quantified, and transcript abundances were estimated using RNA-seq by expectation maximization (RSEM, version 1.3.3)⁷⁵.

The count matrices were imported into R software (version 3.3.2) for further processing. The edgeR⁷⁶ package was utilized to identify differentially expressed genes (DEGs) between groups. Transcripts with a count per million (cpm) greater than 1 in at least two samples were considered for the downstream analysis. Genes with a log₂-fold change greater than 1 or less than -1 and a false discovery rate (FDR)-adjusted P-value less than 0.05 were considered significantly differentially expressed. Volcano plots of DEGs were generated using the ggplot2⁷⁷ package in R.

Principal component analysis (PCA) and heatmap clustering were performed using normalized counts and R packages. The heatmap was generated using the heatmap.2 function. PCA was performed using the prcomp function and plotted using ggplot2. Gene set enrichment analysis (GSEA), gene ontology enrichment analysis and reactome pathway gene set enrichment analysis was applied to detect the enrichment of pathways using the fgseaMultilevel, gseGO and enrichPathway function in R. Values are ranked by $\text{sign}(\log_{2}FC) * (-\log_{10}(FDR))$. $|\text{NES}| > 0$ and adjusted p-value < 0.05 were considered the thresholds for significance for the gene sets.

Downregulation of MDM4 gene with siRNA

SMARTpool siRNA for MDM4 was purchased from Horizon Discovery. 1q hESC were seeded at the density 40.000 cells/cm². After 24h cells were transfected with 50nM SMARTpool siRNA in Nutristem medium without antibiotics together with RNAiMAX (Thermo Fisher Scientific), according to manufacturer's instructions. Transfection cocktail was added to the cells 24h prior to seeding for differentiation.

Statistics

All differentiation experiments were carried out in at least triplicate ($n \geq 3$). All data are presented as the mean \pm standard error of the mean (SEM) or standard deviation (SD). Statistical evaluation of differences between 2 groups was performed using unpaired two-tailed t tests, one-way or two-way ANOVA in GraphPad Prism9 software, with $p < 0.05$ determined to indicate significance.

Data availability

The RNA sequencing counts per million tables are provided in the supplementary material. The raw sequencing data are available upon request. Source data can be retrieved from osf.io/qzyp4.

Author contributions

N.K. carried out all of the experiments and bioinformatics analysis unless stated otherwise and co-wrote the manuscript. E.C.D.D. assisted with the bioinformatics analysis. Y.L., and D.AD. assisted with the cell culture. S.V and L.A.vG assisted with the hepatoblast differentiation and the flow cytometry K.S. edited the paper. C.S. cowrote the manuscript and designed and supervised the experimental work. All co-authors provided scientific input and proof-read the manuscript.

Competing interests

The Authors declare no Competing Financial or Non-Financial Interests.

Acknowledgments

The authors wish to thank Kristoffer Riecken of the University of Hamburg for kindly sharing the lentiviral constructs for the overexpression of the fluorescent proteins mCherry and Venus. Thanks also go to Yannick De Vlaeminck for the assistance with production of lentivirus and kindly sharing the modified GFP plasmid. Y.L. is a predoctoral fellow supported by the China Scholarship Council (CSC), and N.K., M.R. and E.C.D.D. are predoctoral fellows supported by the Fonds voor Wetenschappelijk Onderzoek Vlaanderen (FWO). This research was supported by the Methusalem Grant to Karen Sermon (Vrije Universitet Brussel).

REFERENCES

1. Kobold, S., Guhr, A., Mah, N., Bultjer, N., Seltmann, S., Seiler Wulczyn, A.E.M., Stacey, G., Jie, H., Liu, W., Löser, P., et al. (2020). A Manually Curated Database on

- Clinical Studies Involving Cell Products Derived from Human Pluripotent Stem Cells. *Stem Cell Reports* 15, 546–555. 10.1016/j.stemcr.2020.06.014.
2. Kurtz, A., Mah, N., Chen, Y., Fuhr, A., Kobold, S., Seltsmann, S., and Müller, S.C. (2022). Human pluripotent stem cell registry: Operations, role and current directions. *Cell Prolif* 55, e13238. 10.1111/CPR.13238.
 3. Keller, A., Dzedzicka, D., Zambelli, F., Markouli, C., Sermon, K., Spits, C., and Geens, M. (2018). Genetic and epigenetic factors which modulate differentiation propensity in human pluripotent stem cells. *Hum Reprod Update* 24, 162–175. 10.1093/humupd/dmx042.
 4. Kilpinen, H., Goncalves, A., Leha, A., Afzal, V., Alasoo, K., Ashford, S., Bala, S., Bensaddek, D., Casale, F.P., Culley, O.J., et al. (2017). Common genetic variation drives molecular heterogeneity in human iPSCs. *Nature* 546, 370–375. 10.1038/nature22403.
 5. Rouhani, F.J., Zou, X., Danecek, P., Badja, C., Amarante, T.D., Koh, G., Wu, Q., Memari, Y., Durbin, R., Martincorena, I., et al. (2022). Substantial somatic genomic variation and selection for BCOR mutations in human induced pluripotent stem cells. *Nat Genet* 54, 1406. 10.1038/S41588-022-01147-3.
 6. Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., Ziller, M., Croft, G.F., Amoroso, M.W., Oakley, D.H., et al. (2011). Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* 144, 439–452. 10.1016/j.cell.2010.12.032.
 7. Motosugi, N., Sugiyama, A., Okada, C., Otomo, A., Umezawa, A., Akutsu, H., Hadano, S., and Fukuda, A. (2022). De-erosion of X chromosome dosage compensation by the editing of XIST regulatory regions restores the differentiation potential in hPSCs. *Cell reports methods* 2. 10.1016/J.CRMETH.2022.100352.
 8. Sekiya, A., Takasawa, K., Arai, Y., Horike, S. ichi, Akutsu, H., Umezawa, A., and Nishino, K. (2022). Variation of DNA methylation on the IRX1/2 genes is responsible for the neural differentiation propensity in human induced pluripotent stem cells. *Regen Ther* 21, 620–630. 10.1016/J.RETH.2022.11.007.
 9. 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. 10.1038/nbt.2051.
 10. 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. 10.1016/j.stemcr.2016.10.003.
 11. Merkle, F.T., Ghosh, S., Genovese, G., Handsaker, R.E., Kashin, S., Meyer, D., Karczewski, K.J., O’Dushlaine, C., Pato, C., Pato, M., et al. (2022). Whole-genome analysis of human embryonic stem cells enables rational line selection based on genetic variation. *Cell Stem Cell* 29, 472–486. 10.1016/j.stem.2022.01.011.
 12. 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, 1–11. 10.1038/nature22312.
 13. Avior, Y., Lezmi, E., Eggan, K., and Benvenisty, N. (2021). Cancer-Related Mutations Identified in Primed Human Pluripotent Stem Cells. *Cell Stem Cell* 28, 10–11. 10.1016/j.stem.2020.11.013.
 14. Andrews, P.W., Barbaric, I., Benvenisty, N., Draper, J.S., Ludwig, T., Merkle, F.T., Sato, Y., Spits, C., Stacey, G.N., Wang, H., et al. (2022). The consequences of

- recurrent genetic and epigenetic variants in human pluripotent stem cells. *Cell Stem Cell* 29, 1624–1636. 10.1016/J.STEM.2022.11.006.
15. Werbowetski-Ogilvie, T.E., Bossé, M., Stewart, M., Schnerch, A., Ramos-Mejia, V., Rouleau, A., Wynder, T., Smith, M.-J., Dingwall, S., Carter, T., et al. (2009). Characterization of human embryonic stem cells with features of neoplastic progression. *Nat Biotechnol* 27, 91–97. 10.1038/nbt.1516.
 16. 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. 10.1093/molehr/gat077.
 17. 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 Reports* 1, 1–8. 10.1016/j.stemcr.2013.10.005.
 18. Herszfeld, D., Wolvetang, E., Langton-Bunker, E., Chung, T.-L., Filipczyk, A. a, Houssami, S., Jamshidi, P., Koh, K., Laslett, A.L., Michalska, A., et al. (2006). CD30 is a survival factor and a biomarker for transformed human pluripotent stem cells. *Nat Biotechnol* 24, 351–357. 10.1038/nbt1197.
 19. Ben-David, U., Arad, G., Weissbein, U., Mandefro, B., Maimon, A., Golan-Lev, T., Narwani, K., Clark, A.T., Andrews, P.W., Benvenisty, N., et al. (2014). Aneuploidy induces profound changes in gene expression, proliferation and tumorigenicity of human pluripotent stem cells. *Nat Commun* 5, 4825. 10.1038/ncomms5825.
 20. Price, C.J., Stavish, D., Gokhale, P.J., Stevenson, B.A., Sargeant, S., Lacey, J., Rodriguez, T.A., and Barbaric, I. (2021). Genetically variant human pluripotent stem cells selectively eliminate wild-type counterparts through YAP-mediated cell competition. *Dev Cell* 56, 2455–2470.e10. 10.1016/j.devcel.2021.07.019.
 21. Yang, S., Lin, G., Tan, Y.-Q., Deng, L.-Y., Yuan, D., and Lu, G.-X. (2010). Differences between karyotypically normal and abnormal human embryonic stem cells. *Cell Prolif* 43, 195–206. 10.1111/j.1365-2184.2010.00669.x.
 22. Markouli, C., De Deckersberg, E.C., Regin, M., Nguyen, H.T.T., Zambelli, F., Keller, A., Dziedzicka, D., De Kock, J., Tilleman, L., Van Nieuwerburgh, F., et al. (2019). Gain of 20q11.21 in Human Pluripotent Stem Cells Impairs TGF- β -Dependent Neuroectodermal Commitment. *Stem Cell Reports* 13, 163–176. 10.1016/j.stemcr.2019.05.005.
 23. Jo, H.Y., Lee, Y., Ahn, H., Han, H.J., Kwon, A., Kim, B.Y., Ha, H.Y., Kim, S.C., Kim, J.H., Kim, Y.O., et al. (2020). Functional in vivo and in vitro effects of 20q11.21 genetic aberrations on hPSC differentiation. *Sci Rep* 10, 1–15. 10.1038/s41598-020-75657-7.
 24. Keller, A., Lei, Y., Krivec, N., Couvreur de Deckersberg, E., Dziedzicka, D., Markouli, C., Sermon, K., Geens, M., and Spits, C. (2021). Gains of 12p13.31 Delay WNT-Mediated Initiation of hPSC Differentiation and Promote Residual Pluripotency in a Cell Cycle Dependent Manner. *bioRxiv*, <https://doi.org/10.1101/2021.05.22.445238>. <https://doi.org/10.1101/2021.05.22.445238>.
 25. Varela, C., Denis, J.A., Polentes, J., Feyeux, M., Aubert, S., Champon, B., Piétu, G., Peschanski, M., and Lefort, N. (2012). Recurrent genomic instability of chromosome 1q in neural derivatives of human embryonic stem cells. *Journal of Clinical Investigation* 122, 569–574. 10.1172/JCI46268.

26. 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. 10.1007/s11626-010-9284-4.
27. 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. 10.1093/humrep/dei345.
28. Wesley, B.T., Ross, A.D.B., Muraro, D., Miao, Z., Saxton, S., Tomaz, R.A., Morell, C.M., Ridley, K., Zacharis, E.D., Petrus-Reurer, S., et al. (2022). Single-cell atlas of human liver development reveals pathways directing hepatic cell fates. *Nat Cell Biol* *24*, 1487–1498. 10.1038/S41556-022-00989-7.
29. Cheng, S., Brenière-Letuffe, D., Ahola, V., Wong, A.O.T., Keung, H.Y., Gurung, B., Zheng, Z., Costa, K.D., Lieu, D.K., Keung, W., et al. (2023). Single-cell RNA sequencing reveals maturation trajectory in human pluripotent stem cell-derived cardiomyocytes in engineered tissues. *iScience* *26*. 10.1016/J.ISCI.2023.106302.
30. Saeki, T., Yoshimatsu, S., Ishikawa, M., Hon, C.C., Koya, I., Shibata, S., Hosoya, M., Saegusa, C., Ogawa, K., Shin, J.W., et al. (2022). Critical roles of FGF, RA, and WNT signalling in the development of the human otic placode and subsequent lineages in a dish. *Regen Ther* *20*, 165–186. 10.1016/J.RETH.2022.04.008.
31. Dincer, Z., Piao, J., Niu, L., Ganat, Y., Kriks, S., Zimmer, B., Shi, S.H., Tabar, V., and Studer, L. (2013). Specification of functional cranial placode derivatives from human pluripotent stem cells. *Cell Rep* *5*, 1387–1402. 10.1016/J.CELREP.2013.10.048.
32. Spreng, A.S., Brüll, M., Leisner, H., Suci, I., and Leist, M. (2022). Distinct and Dynamic Transcriptome Adaptations of iPSC-Generated Astrocytes after Cytokine Stimulation. *Cells* *11*. 10.3390/CELLS11172644.
33. Jovanovic, V.M., Weber, C., Slamecka, J., Ryu, S., Chu, P.-H., Sen, C., Inman, J., De Sousa, J.F., Barnaeva, E., Hirst, M., et al. (2023). A defined roadmap of radial glia and astrocyte differentiation from human pluripotent stem cells. *Stem Cell Reports* *18*. 10.1016/J.STEMCR.2023.06.007.
34. Friedman, C.E., Nguyen, Q., Lukowski, S.W., Helfer, A., Chiu, H.S., Miklas, J., Levy, S., Suo, S., Han, J.D.J., Osteil, P., et al. (2018). Single-Cell Transcriptomic Analysis of Cardiac Differentiation from Human PSCs Reveals HOPX-Dependent Cardiomyocyte Maturation. *Cell Stem Cell* *23*, 586–598.e8. 10.1016/J.STEM.2018.09.009.
35. Saeki, T., Yoshimatsu, S., Ishikawa, M., Hon, C.C., Koya, I., Shibata, S., Hosoya, M., Saegusa, C., Ogawa, K., Shin, J.W., et al. (2022). Critical roles of FGF, RA, and WNT signalling in the development of the human otic placode and subsequent lineages in a dish. *Regen Ther* *20*, 165–186. 10.1016/J.RETH.2022.04.008.
36. Ealy, M., Ellwanger, D.C., Kosaric, N., Stapper, A.P., and Heller, S. (2016). Single-cell analysis delineates a trajectory toward the human early otic lineage. *Proc Natl Acad Sci U S A* *113*, 8508–8513. 10.1073/PNAS.1605537113.
37. Seal, S., and Monsoro-Burq, A.H. (2020). Insights Into the Early Gene Regulatory Network Controlling Neural Crest and Placode Fate Choices at the Neural Border. *Front Physiol* *11*. 10.3389/FPHYS.2020.608812.
38. Gupta, S., Polit, L.D., Fitzgerald, M., Rowland, H.A., Murali, D., Buckley, N.J., and Subramaniam, S. (2023). Temporal transcriptional control of neural induction in human induced pluripotent stem cells. *Front Mol Neurosci* *16*. 10.3389/FNMOL.2023.1139287.

39. Tchieu, J., Zimmer, B., Fattahi, F., Amin, S., Zeltner, N., Chen, S., and Studer, L. (2017). A Modular Platform for Differentiation of Human PSCs into All Major Ectodermal Lineages. *Cell Stem Cell* 21, 399–410.e7. 10.1016/J.STEM.2017.08.015.
40. Russo, L., Sladitschek, H.L., and Neveu, P.A. (2022). Multi-layered regulation of neuroectoderm differentiation by retinoic acid in a primitive streak-like context. *Stem Cell Reports* 17, 231–244. 10.1016/J.STEMCR.2021.12.014.
41. Yilmaz, A., Braverman-Gross, C., Bialer-Tsypin, A., Peretz, M., and Benvenisty, N. (2020). Mapping Gene Circuits Essential for Germ Layer Differentiation via Loss-of-Function Screens in Haploid Human Embryonic Stem Cells. *Cell Stem Cell* 27, 679–691.e6. 10.1016/J.STEM.2020.06.023.
42. Vallier, L., Touboul, T., Chng, Z., Brimpari, M., Hannan, N., Millan, E., Smithers, L.E., Trotter, M., Rugg-Gunn, P., Weber, A., et al. (2009). Early cell fate decisions of human embryonic stem cells and mouse epiblast stem cells are controlled by the same signalling pathways. *PLoS One* 4. 10.1371/JOURNAL.PONE.0006082.
43. Shang, Z., Chen, D., Wang, Q., Wang, S., Deng, Q., Wu, L., Liu, C., Ding, X., Wang, S., Zhong, J., et al. (2018). Single-cell RNA-seq reveals dynamic transcriptome profiling in human early neural differentiation. *Gigascience* 7, 1–19. 10.1093/GIGASCIENCE/GIY117.
44. Barry, C., Schmitz, M.T., Jiang, P., Schwartz, M.P., Duffin, B.M., Swanson, S., Bacher, R., Bolin, J.M., Elwell, A.L., McIntosh, B.E., et al. (2017). Species-specific developmental timing is maintained by pluripotent stem cells ex utero. *Dev Biol* 423, 101–110. 10.1016/J.YDBIO.2017.02.002.
45. Vallabhaneni, H., Lynch, P.J., Chen, G., Park, K., Liu, Y., Goehle, R., Mallon, B.S., Boehm, M., and Hursh, D.A. (2018). High Basal Levels of γ H2AX in Human Induced Pluripotent Stem Cells Are Linked to Replication-Associated DNA Damage and Repair. *Stem Cells* 36, 1501–1513. 10.1002/stem.2861.
46. 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. 10.1016/j.stem.2015.11.003.
47. Halliwell, J.A., Frith, T.J.R., Laing, O., Price, C.J., Bower, O.J., Stavish, D., Gokhale, P.J., Hewitt, Z., El-Khamisy, S.F., Barbaric, I., et al. (2020). Nucleosides Rescue Replication-Mediated Genome Instability of Human Pluripotent Stem Cells. *Stem Cell Reports* 14, 1009–1017. 10.1016/j.stemcr.2020.04.004.
48. 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 Reports* 6, 330–341. 10.1016/j.stemcr.2016.01.015.
49. Liu, W., Ren, Z., Lu, K., Song, C., Cheung, E.C.W., Zhou, Z., and Chen, G. (2018). The suppression of medium acidosis improves the maintenance and differentiation of human pluripotent stem cells at high density in defined cell culture medium. *Int J Biol Sci* 14, 485–496. 10.7150/ijbs.24681.
50. Draper, J.S., Smith, K., Gokhale, P., Moore, H.D., Maltby, E., Johnson, J., Meisner, L., Zwaka, T.P., Thomson, J.A., and Andrews, P.W. (2004). Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. *Nat Biotechnol* 22, 53–54. 10.1038/nbt922.
51. Vitillo, L., Anjum, F., Hewitt, Z., Stavish, D., Laing, O., Baker, D., Barbaric, I., and Coffey, P. (2023). The isochromosome 20q abnormality of pluripotent cells interrupts germ layer differentiation. *Stem Cell Reports*. 10.1016/J.STEMCR.2023.01.007.

52. Lei, Y., Al Delbany, D., Krivec, N., Regin, M., Couvreur De 3 Deckersberg, E., Janssens, C., Ghosh, M., Sermon, K., and Spits, C. (2023). SALL3 mediates the loss of neuroectodermal differentiation potential in human embryonic stem cells with chromosome 18q loss. *bioRxiv*, <https://doi.org/10.1101/2023.06.26.546513>. <https://doi.org/10.1101/2023.06.26.546513>.
53. Gonzales, K.A.U., Liang, H., Lim, Y.-S., Chan, Y.-S., Yeo, J.-C., Tan, C.-P., Gao, B., Le, B., Tan, Z.-Y., Low, K.-Y., et al. (2015). Deterministic Restriction on Pluripotent State Dissolution by Cell-Cycle Pathways. *Cell* *162*, 564–579. [10.1016/j.cell.2015.07.001](https://doi.org/10.1016/j.cell.2015.07.001).
54. Andrews, P.W., Barbaric, I., Benvenisty, N., Draper, J.S., Ludwig, T., Merkle, F.T., Sato, Y., Spits, C., Stacey, G.N., Wang, H., et al. (2022). The consequences of recurrent genetic and epigenetic variants in human pluripotent stem cells. *Cell Stem Cell* *29*, 1624–1636. [10.1016/j.stem.2022.11.006](https://doi.org/10.1016/j.stem.2022.11.006).
55. 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. [10.1038/ncomms5227](https://doi.org/10.1038/ncomms5227).
56. Jacobs, K., Zambelli, F., Mertzaniidou, A., Smolders, I., Geens, M., Nguyen, H.T., Sermon, K., and Spits, C. (2016). Higher-Density Culture in Human Embryonic Stem Cells Results in DNA Damage and Genome Instability. *Stem Cell Reports* *6*, 330–341. [10.1016/j.stemcr.2016.01.015](https://doi.org/10.1016/j.stemcr.2016.01.015).
57. Keller, A., Tilleman, L., Dziedzicka, D., Zambelli, F., Sermon, K., Van Nieuwerburgh, F., Spits, C., and Geens, M. (2019). Uncovering low-level mosaicism in human embryonic stem cells using high throughput single cell shallow sequencing. *Sci Rep* *9*, 14844. [10.1038/s41598-019-51314-6](https://doi.org/10.1038/s41598-019-51314-6).
58. Halliwell, J.A., Frith, T.J.R., Laing, O., Price, C.J., Bower, O.J., Stavish, D., Gokhale, P.J., Hewitt, Z., El-Khamisy, S.F., Barbaric, I., et al. (2020). Nucleosides Rescue Replication-Mediated Genome Instability of Human Pluripotent Stem Cells. *Stem Cell Reports* *14*, 1009–1017. [10.1016/j.stemcr.2020.04.004](https://doi.org/10.1016/j.stemcr.2020.04.004).
59. Girish, V., Lakhani, A.A., Thompson, S.L., Scaduto, C.M., Brown, L.M., Hagenson, R.A., Sausville, E.L., Mendelson, B.E., Kandikuppa, P.K., Lukow, D.A., et al. (2023). Oncogene-like addiction to aneuploidy in human cancers. *Science* (1979) *381*, 2023.01.09.523344. [10.1126/science.adg4521](https://doi.org/10.1126/science.adg4521).
60. Na, J., Baker, D., Zhang, J., Andrews, P.W., and Barbaric, I. (2014). Aneuploidy in pluripotent stem cells and implications for cancerous transformation. *Protein Cell*. [10.1007/s13238-014-0073-9](https://doi.org/10.1007/s13238-014-0073-9).
61. Baker, D.E.C., Harrison, N.J., Maltby, E., Smith, K., Moore, H.D., Shaw, P.J., Heath, P.R., Holden, H., and Andrews, P.W. (2007). Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. *Nat Biotechnol* *25*, 207–215. [10.1038/nbt1285](https://doi.org/10.1038/nbt1285).
62. Ben-David, U., and Benvenisty, N. (2011). The tumorigenicity of human embryonic and induced pluripotent stem cells. *Nat Rev Cancer* *11*, 268–277. [10.1038/nrc3034](https://doi.org/10.1038/nrc3034).
63. Voutsadakis, I.A. (2021). Chromosome 20q11.21 Amplifications in Colorectal Cancer. *Cancer Genomics - Proteomics* *18*, 487–496. [10.21873/cgp.20274](https://doi.org/10.21873/cgp.20274).
64. Mahlamäki, E.H., Höglund, M., Gorunova, L., Karhu, R., Dawiskiba, S., André-Sandberg, Å., Kallioniemi, O.-P., and Johansson, B. (1997). Comparative genomic hybridization reveals frequent gains of 20q, 8q, 11q, 12p, and 17q, and losses of 18q, 9p, and 15q in pancreatic cancer. *Genes Chromosomes Cancer* *20*, 383–391. [10.1002/\(SICI\)1098-2264\(199712\)20:4<383::AID-GCC10>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1098-2264(199712)20:4<383::AID-GCC10>3.0.CO;2-O).
65. Guichard, C., Amaddeo, G., Imbeaud, S., Ladeiro, Y., Pelletier, L., Maad, I. Ben, Calderaro, J., Bioulac-Sage, P., Letexier, M., Degos, F., et al. (2012). Integrated

- analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet* *44*, 694–698. 10.1038/ng.2256.
66. ISSCR Standards for human stem cell use in research (2023). <https://www.isscr.org/standards-document>.
 67. Weber, K., Thomaschewski, M., Warlich, M., Volz, T., Cornils, K., Niebuhr, B., Täger, M., Lütgehetmann, M., Pollok, J.M., Stocking, C., et al. (2011). RGB marking facilitates multicolor clonal cell tracking. *Nat Med* *17*, 504–509. 10.1038/NM.2338.
 68. Ludwig, T.E., Andrews, P.W., Barbaric, I., Benvenisty, N., Bhattacharyya, A., Crook, J.M., Daheron, L.M., Draper, J.S., Healy, L.E., Huch, M., et al. (2023). ISSCR standards for the use of human stem cells in basic research. *Stem Cell Reports* *18*, 1744–1752. 10.1016/J.STEMCR.2023.08.003.
 69. Bayindir, B., Dehaspe, L., Brison, N., Brady, P., Ardui, S., Kammoun, M., Van der Veken, L., Lichtenbelt, K., Van den Bogaert, K., Van Houdt, J., et al. (2015). Noninvasive prenatal testing using a novel analysis pipeline to screen for all autosomal fetal aneuploidies improves pregnancy management. *Eur J Hum Genet* *23*, 1286–1293. 10.1038/ejhg.2014.282.
 70. Douvaras, P., and Fossati, V. (2015). Generation and isolation of oligodendrocyte progenitor cells from human pluripotent stem cells. *Nat Protoc* *10*, 1143–1154. 10.1038/nprot.2015.075.
 71. Lin, Y., and Zou, J. (2020). Differentiation of Cardiomyocytes from Human Pluripotent Stem Cells in Fully Chemically Defined Conditions. *STAR Protoc* *1*, 100015. 10.1016/j.xpro.2020.100015.
 72. Boon, R., Kumar, M., Tricot, T., Elia, I., Ordovas, L., Jacobs, F., One, J., De Smedt, J., Eelen, G., Bird, M., et al. (2020). Amino acid levels determine metabolism and CYP450 function of hepatocytes and hepatoma cell lines. *Nat Commun* *11*, 1393. 10.1038/s41467-020-15058-6.
 73. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* *15*, 550. 10.1186/s13059-014-0550-8.
 74. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* *29*, 15–21. 10.1093/bioinformatics/bts635.
 75. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* *12*, 323. 10.1186/1471-2105-12-323.
 76. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* *26*, 139–140. 10.1093/bioinformatics/btp616.
 77. Wickham, H. (2016). ggplot2 (Springer International Publishing) 10.1007/978-3-319-24277-4.