Induced Pluripotent Stem Cell Differentiation and Three-Dimensional Tissue Formation Attenuate Clonal Epigenetic Differences in Trichohyalin

Anastasia Petrova,^{1–3,*} Antonio Capalbo,^{4,*} Laureen Jacquet,^{1,*} Simon Hazelwood-Smith,⁵ Dimitra Dafou,^{5,†} Carl Hobbs,⁶ Matthew Arno,⁷ Alessio Farcomeni,⁸ Liani Devito,¹ Heba Badraiq,¹ Michael Simpson,⁵ John A McGrath,² Wei-Li Di,³ Jeffrey B Cheng,⁹ Theodora M Mauro,⁹ and Dusko Ilic¹

The epigenetic background of pluripotent stem cells can influence transcriptional and functional behavior. Most of these data have been obtained in standard monolayer cell culture systems. In this study, we used exome sequencing, array comparative genomic hybridization (CGH), miRNA array, DNA methylation array, three-dimensional (3D) tissue engineering, and immunostaining to conduct a comparative analysis of two induced pluripotent stem cell (iPSC) lines used in engineering of 3D human epidermal equivalent (HEE), which more closely approximates epidermis. Exome sequencing and array CGH suggested that their genome was stable following 3 months of feeder-free culture. While the miRNAome was also not affected, $\approx 7\%$ of CpG sites were differently methylated between the two lines. Analysis of the epidermal differentiation complex, a region on chromosome 1 that contains multiple genes involved in skin barrier maturation (including trichohyalin, *TCHH*), found that in one of the iPSC clones (iKCL004), *TCHH* retained a DNA methylation signature characteristic of the original somatic cells, whereas in other iPSC line (iKCL011), the *TCHH* methylation signature matched that of the human embryonic stem cell line KCL034. The difference between the two iPSC clones in *TCHH* methylation did not have an obvious effect on its expression in 3D HEE, suggesting that differentiation and tissue formation may mitigate variations in the iPSC methylome.

Introduction

REPROGRAMMING SOMATIC CELLS back to a pluripotent state, similar to embryonic stem cells (ESCs), initially required insertion of reprogramming transgenes and their viral vectors into the host genome [1]. Previous studies demonstrated that the effect of reprogramming on terminally differentiated somatic cells was strikingly similar to that seen in neoplastic transformation. Profound epigenetic changes led to pluripotent plasticity and indefinite self-renewal of cells that were previously committed to differentiation and had a limited proliferation capacity. Although reprogramming-associated point mutations might arise during the initial stages of

conversion [2], it instead appears that these point mutations are predominantly a consequence of cloning individual cells with different mutation histories [3]. In addition, induced pluripotent stem cell (iPSC) clones might carry not only epigenetic memory of their parental somatic cell origin [4] but also have a unique pattern of DNA methylation as a result of nonuniform epigenome transformation during reprogramming [5–7]. Specific genetic and epigenetic footprints influence the molecular and functional properties of each iPSC clone. Altered gene expression and distinct differentiation propensity among early passage iPSC clones are likely to be largely attenuated through continuous passaging [7]. In contrast, human pluripotent stem cells, both human ESC (hESC) and especially iPSC, kept for

¹Assisted Conception Unit, Stem Cell Laboratory, Division of Women's Health, Women's Health Academic Centre, King's College London, London, United Kingdom.

²St John's Institute of Dermatology, King's College London, London, United Kingdom.

³Immunobiology Unit, Institute of Child Health, University College London, London, United Kingdom.

⁴GENERA, Centers for Reproductive Medicine, Rome, Italy.

⁵Division of Genetics and Molecular Medicine, King's College London, London, United Kingdom.

⁶Histology Laboratory, Wolfson Centre for Age-Related Diseases, King's College London, London, United Kingdom.

⁷Genomics Centre, King's College London, London, United Kingdom.

⁸Statistics Section, Department of Public Health and Infectious Diseases, Sapienza–University of Rome, Rome, Italy.

⁹Department of Dermatology, Veteran Affairs Medical Center, University of California San Francisco, San Francisco, California.

*These three authors contributed equally to this work.

[†]Current affiliation: Department of Genetics, Developmental and Molecular Biology, School of Biology, Aristotle University, Thessaloniki, Greece.

extended period of time in culture are prone to genomic instability [8,9]. It is unknown to what extent the cellular genetic and epigenetic makeup or culture conditions contribute to genomic instability over time.

We reprogrammed human foreskin BJ fibroblasts (HFFs) into iPSC using modified synthetic mRNA. Two iPSC clones, iKCL004 and iKCL011, that we used to build human skin equivalents [human epidermal equivalent (HEE)] with a functional permeability barrier [10] showed subtle differences, which prompted us to investigate in depth the genetic and epigenetic footprint of both lines. Since the focus of our work was keratinocyte differentiation culminating in the stratum corneum derived epidermal permeability barrier, we concentrated on the epidermal differentiation complex (EDC) on chromosome 1, which contains multiple genes involved in epidermal cornification [11–19]. One of our clinical grade lines, KCL034, was used as a control in most analyses [20].

Materials and Methods

Reprogramming

Human neonatal foreskin fibroblasts BJ (ATCC, CRL-2522) were reprogrammed using modified synthetic mRNA as described [10]. At day 2–17 of reprogramming, some cells also were treated with 1 μ M pifithrin- α . Two clones with a similar growth rate, one derived in the absence (iKCL004) and one in the presence of 1 μ M pifithrin- α (iKCL011), were further characterized; pluripotency marker expression and differentiation into three germ layers in vitro and in vivo (teratomas) revealed no obvious difference between the lines [10].

Cell culture

Undifferentiated hESC line KCL034 and iPSC lines, iKCL004 and iKCL011, were cultured under standard feeder-free conditions on Matrigel (Becton Dickinson) in mTeSR1 medium (STEMCELL Technologies) under hypoxic (5% O₂) conditions. Differentiation into keratinocytes and generation of HEEs were described previously [10]. Transepithelial electrical resistance was measured with epithelial voltammeter EVOM (World Precision Instruments) as described [10,21]. Normal human keratinocytes (NHK) were cultured in EpiLife (Life Technologies). Human neonatal foreskin fibroblasts BJ were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (Hyclone).

Quantitative reverse transcriptase–polymerase chain reaction

Total RNA was extracted using the RNeasy Mini Kit (QIAGEN) following the manufacturer's instructions. A 20-min DNase I (Qiagen) treatment step was included to eliminate the potential contamination of genomic DNA. cDNA was generated by reverse transcription of total RNA (350 ng) using the RT2 First Strand Kit (Qiagen) following the manufacturer's instruction. The quantitative reverse transcriptase–polymerase chain reaction (qPCR) consisted of 3 μ L of cDNA diluted 1:10 in water, 6.25 μ L of RT2 Real Time SYBR[®] Green qPCR Master Mix (Qiagen), and 400 nM of each primer for a final volume of 12.5 μ L. qPCR cycling conditions were as follows: one cycle 95°C for 10 min; 40 cycles 95°C for 15 s, 61°C for 30 s followed by melt curve acquisition from 59°C to 95°C with

0.5°C increment. The *BAX* (Fw: CGGGTTGTCGCCCTTTTC TA; Rv: CGGGGATTGATCAGACACGTA) and *p21* (Fw: GC ACTTTGATTAGCAGCGGA; Rv: AAGACAACTACTCCC AGCCC) primers were designed using Primer3 software. The GAPDH and ACTB primers were from PrimerDesign Ltd. All samples were analyzed in triplicate, Ct values were determined, and the expression was calculated by the $2^{-\Delta\Delta C_t}$ method. GAPDH and ACTB were used for internal normalization.

Array comparative genomic hybridization

Array comparative genomic hybridization (CGH) was performed following strict protocols for each of the different steps as provided by the manufacturers of the reagents. CGH Labeling Kit (Enzo Life Sciences) was used to label 1 μ g DNA. Post labeling, DNA was purified with QIAquick PCR Purification Kit (Qiagen). Labeling efficiency and yield were assessed by spectrophotometry (NanoDrop). For hybridization, washing, and scanning of arrays, we used an Agilent 4 $\times$ 44K platform. Image quantification, array quality control (QC), and aberration detection were performed using feature extraction and DNA Analytics software packages (Agilent); 95% of array data were required to pass QC. The ADM-2 algorithm at threshold 6 (with a three-probe sliding window providing a mean detection interval of 200 kb) was used for aberration calling.

Exome sequencing

Exome sequencing was undertaken using genomic DNA extracted from BJ parental fibroblasts before and after reprogramming (two clones: iKCL004 and iKCL011). Libraries were prepared using the SureSelect Human All Exon Capture v4 (Agilent) and sequenced with 100-bp paired end reads on the HiSeq platform (Illumina).

The resulting sequence reads were aligned to the reference genome (hg19) with NovoAlign (Novocraft Technologies). Duplicate reads, resulting from PCR clonality or optical duplicates, and reads mapped to multiple locations were excluded from downstream analysis. Single-nucleotide substitutions and small insertions or deletions were identified and quality filtered within the SAMtools software package and in-house software tools. Variants were annotated with respect to genes and transcripts with the ANNOVAR tool.

To identify differentially arising new mutations, we undertook a series of pairwise comparisons with the VarScan 2 tool. Comparison was limited to protein coding regions and associated splice sites of genes defined by the GENCODE project and sites at which >20 reads were observed in all three samples (~90% of the targeted exome).

The resulting differentially called variants were filtered with the following criteria to retain high quality changes; presence of the new allele in >20% of reads in the sample in which the variant was identified, absence of any reads representing the alternative allele in the other two samples. Variants located within repeat elements or segmentally duplicated regions were excluded as likely alignment artifacts. The sequence alignments spanning each candidate variant were checked manually using the Integrative Genomics Viewer 2.3.

miRNA expression analysis

miRNA expression was evaluated in biological triplicates of undifferentiated hESC and iPSC using TaqMan Array

MicroRNA Cards (Panel A and Panel B; Applied Biosystems). The cards contain primer sets for 736 human miRNA sequences (667 unique miRNAs), two sequences from Arabidopsis, eight measurements of MammU6, and 22 RNU sequences.

The cell samples were collected in 200 μ L of total volume of lysis buffer. MicroRNA isolation and purification were conducted through the TaqMan miRNA ABC (Anti-miRNA Bead Capture) Purification Kit (Applied Biosystems). This system is designed for rapid and specific purification of human miRNAs. The Human Panel A or Panel B v3.0 Beads are superparamagnetic Dynabeads[®] covalently bound to a unique set of 377 anti-miRNA oligonucleotides for each panel. The miRNA isolation relies on hybridization of endogenous miRNAs to the corresponding anti-miRNA oligonucleotides attached to the beads. The purification protocol was followed according to the manufacturers' instructions. The bead hybridization and the miRNA elution steps were conducted on the Thermomixer[®] R (Eppendorf). All the washing steps were performed through the DynaMag[™]-2 magnet (Applied Biosystems), to clear up the solution from DNA, proteins, contaminants, and residual binding solution, while keeping the miRNAs bound to the beads. MicroRNAs were eluted in 20 μ L of final volume, to concentrate them in the smallest possible volume. The cDNA was then generated using the specific Megaplex RT Primer pools A and B v3.0 (Applied Biosystems), which include 377 primer sets for each specific miRNA isolated in the previous step, according to the manufacturer protocol. Pre-amplification was performed according to manufacturer instructions and protocol using the specific Megaplex PreAmp Primer pools A and B v3.0 (Applied Biosystems). Briefly, 25 μ L of pre-amplification products were added to 425 μ L nuclease-free water and mixed with 450 μ L TaqMan Universal Master Mix II, No UNG, then dispensed into the 384 wells by centrifugation on a Heraeus Megafuge 40 (Thermo Scientific) with the proper TaqMan Low Density Arrays card adapters. The reactions were incubated in a 384-well plate at 95°C for 10 min followed by 40 cycles of 95°C for 15 s and at 60°C for 1 min on a ViiA[™] 7 Real-Time PCR System (Applied Biosystems).

Raw data were analyzed using SDS software (Applied Biosystems), and Ct values were used as a readout. Manual inspection of all amplification plots was performed and miRNAs were excluded from the analysis if: Ct values were too high (above 35, indicating a miRNA expression too low for accurate detection) or if amplification was not achieved in at least two samples (biological replicates). Data were processed through the RealTime StatMiner 5.0 software (Integromics).

Presence/absence indicators were compared among groups through a logistic mixed model, with a slide specific intercept to capture any overdispersion. The *P*-values arising from Wald tests on the group indicator coefficients were then adjusted using Benjamini–Hochberg (BH) correction. The global false discovery rate (FDR) was guaranteed below the rate of 5%. We then proceeded to compare expression levels after Loess normalization as this normalization approach has been shown to effectively reduce standard deviations, increase sensitivity and improve accuracy of differential miRNA expression detection, and increase interplatform concordance [22]. Once again we controlled for multiplicity using BH correction. Effect sizes of miRNAs relative concentration were measured by means of fold changes. Pearson's correlation was used to generate a heatmap for intragroup correlation plot among all biological samples.

Hierarchical clustering analysis exploiting complete linkage clustering method and Euclidean distance similarity measure was performed on fold changes using KCL034 cell lines as a calibrator to compare similarity of miRNA expression profile between iPSC and hESC lines. A Volcano plot was generated using significance versus fold change on the *y*- and *x*-axes, respectively.

Whole epigenome methylation array

Epigenome-wide methylation (>480,000 CpG sites) was analyzed using Infinium Human Methylation 450 BeadChip Kit (Illumina) in biological triplicates of 500 ng bisulfite treated DNA for each: BJ parental fibroblasts (negative control), KCL034 (positive control), iKCL004, and iKCL011 line. GenoSplice Technology performed QC of the data, their processing, and further analyses. Microarray data have been deposited in the NCBI Gene Expression Omnibus (GEO) with the accession number GSE55006.

Whole genome gene-expression array

Total RNA from undifferentiated hESC and iPSC, as well as control primary (NHK) and hESC/iPSC-differentiated keratinocytes (NHK), was reverse transcribed into cDNA using SuperScript III Reverse Transcriptase (Life Technologies). A total amount of 750 ng of biotin-labeled aRNA in a 5 μ L volume was then used for HumanHT-12 Expression BeadChip whole-genome gene expression direct hybridization assay system (Illumina) according to manufacturer's instructions, run on iScan system (Illumina) and analyzed as described [10]. All samples were analyzed as biological replicates from three independent experiments. Microarray data have been deposited in the NCBI GEO with the accession number GSE55898.

Immunostaining

Immunostaining has been performed as previously described [23,24]. The following antibodies were used: rabbit anti-KRT14 (Covance), -Oct4 (Santa Cruz Biotechnology), -Sox2 (GeneTex), -LCE2B (Bioss), -p63 (Abcam), mouse anti-cMyc (Santa Cruz Biotechnology), -Klf4 (ATGen), -TCHH (ImmuQuest), and goat anti-LIN28 (R&D Systems). Secondary antibodies were purchased from Jackson ImmunoResearch. Actin stress fibers were visualized with rhodamine-conjugated phalloidin and DNA with a nuclear stain Hoechst 33342, both from Life Technologies.

Results

Reprogramming had no effect on mutation rate and genomic stability of iKCL004 and iKCL011 lines

We first reprogrammed HFFs using synthetic modified mRNA and following a modified version of the protocol originally described by Warren et al. [25]. mRNA transfection efficiency was on average >75% (Fig. 1A, B). We increased the B18R concentration from 200 to 300 ng/mL to further suppress the innate immune response triggered with the exogenous mRNA cocktail. We also added B18R to the medium 4 h prior the first transfection to precondition the cells. Regardless, a regime of daily transfection had a deleterious effect on the cells; the vast majority did not survive (Fig. 1C). At day 4 of reprogramming, we assessed the

FIG. 1. Reprogramming HFFs using synthetic modified mRNA. **(A)** Synthetic modified mRNAs are translated into proteins. Cells transfected with mRNA cocktail expressing reprogramming factors Klf4, cMyc, Oct4, LIN28, and Sox2. The 15-h time point was chosen for analysis as it corresponds to the average peak expression time post-transfection of the different reprogramming factors [25]. **(B)** More than 75% cells express reprogramming factors 15 h after mRNA transfection. The cells were fixed and immunostained for Oct4, Klf4, Sox2, c-Myc, and LIN28 15 h post-transfection with mRNA reprogramming cocktail. Nuclei were visualized with Hoechst 33342. Total cell number was calculated. Data are presented as the mean of cells positive for Oct4, Klf4, Sox2, c-Myc, and LIN28 $\pm$ standard error of the mean ($n=3$). **(C)** Daily transfection had a deleterious effect on the cells; vast majority did not survive. Day 0: the cells before the first transfection demonstrate normal morphology and density. Day 3: decreased cell density and cell rounding indicate a high apoptotic rate. Day 6: massive apoptosis is seen. Surviving cells start undergoing mesenchymal epithelial transformation and adopt an epithelial morphology. Day 9: surviving cells continue to transition to an epithelial morphology. Day 12: first signs of compaction and colony formation are noted. Day 15: increased compaction. Cells become smaller and more hESC like. **(D)** Inhibitory effect of pifithrin- α on p53-mediated transcription. qPCR expression analyses of the two p53 target genes, *BAX*, and *p21* at Day 4 of reprogramming showed decrease in *BAX*, but not *p21* mRNA levels ($n=3$ biological replicates). Transfected samples are compared with the negative control (untransfected BJ cells). Data are given by normalized mean $\pm$ standard error of the mean ($n=3$); statistical significance was calculated by two-way analysis of variance with Tukey post hoc test. **** $P<0.0001$. **(E)** An average reprogramming efficiency from two independent experiments was calculated as number of trypan blue positive colonies per total number of cells plated per well of six-well dish. HFF, human foreskin fibroblast; hESC, human embryonic stem cell; qPCR, quantitative reverse transcriptase-polymerase chain reaction.

inhibitory effect of pifithrin- α on p53-mediated transcription using qPCR. Transcription of two p53 target genes BAX and CDKN1A (known as p21) was elevated > sixfold in cells subjected to daily mRNA transfections, suggesting that p53 activation is induced by reprogramming itself. Pifithrin- α treated cells had reduced levels of BAX, but no effect on p21 transcription was observed (Fig. 1D). Daily mRNA transfections caused massive cell death in both conditions, although cell survival seemed to be somewhat higher in the presence of pifithrin- α . Although reprogramming efficiency was about 30% higher in the presence of 1 μ M pifithrin- α (0.86% vs. 0.66%), in our hands, it was still lower than the

reprogramming efficiency reported by Warren et al. [25] regardless of the p53 inhibition (Fig. 1E). Harsh conditions in cell culture might result in survival of clones that carry growth-promoting mutation(s) associated with tumorigenesis. To assess the genomic integrity of our lines, we subjected iKCL004 and iKCL011 to array CGH at two different time points, one at early passages (<10) and the other about 3 months after the lines were adapted to the feeder-free culture. Even though it is widely known that extended periods of cell culture under feeder-free conditions are linked to higher rates of chromosomal abnormalities in pluripotent stem cells, especially iPSC [8,9], we did not

FIG. 2. hESC/iPSC miRNA profiling. (A) Volcano plot showing no significant miRNA differentially expressed between two iPSC lines ($n = 3$ biological replicates per group). Significance versus fold change on the y- and x-axes, respectively, is plotted. Horizontal black lines represent the statistical significance threshold. Green and red lines represent the upper and lower fold changes thresholds, respectively. (B) Correlation heatmap among the profiles of miRNA expression of all the samples analyzed (three biological replicates per each of two iPSC and one hESC line). Superior to the heatmap and on the left of it three biological replicates and their relative grouping are reported. The color intensity varies in function of the Pearson correlation index ranging from green to red for a lower or a higher correlation, respectively. The proper numerical value of correlation between each couple of sample is reported within the squares at the intersections. (C) Volcano plot showing no significant miRNA differentially expressed between two iPSC lines, iKCL004 and iKCL011, and a control hESC line KCL034. (D) Hierarchical clustering of fold-change values obtained from the comparison of the two iPSC lines against hESC line used as a calibrator. The analysis was performed exploiting complete linkage, Euclidean distance, and z-score normalization, and the heatmaps, beside reporting four clusters of genes showing a similar behavior (highlighted by the four groups of dendograms with the same colored halo below the heatmaps), do not reveal significant differences among the samples analyzed. (E) Hierarchical clustering of fold-change values obtained from the comparison of the two iPSC lines against hESC line used as a calibrator. The analysis was performed exploiting complete linkage, Euclidean distance, and z-score normalization and the heatmaps, beside reporting four clusters of genes showing a similar behavior (highlighted by the four groups of dendograms with the same colored halo below the heatmaps), do not reveal significant differences among the samples analyzed. iPSC, induced pluripotent stem cell.

FIG. 3. Whole epigenome methylation profile. **(A)** Comparative analysis of iKCL004 and iKCL011 methylome found that 35,550 (7%) of CpGs are differentially methylated between the two lines. Nearly a half of them were hypermethylated (18,815 or 53%) and the other half was hypomethylated (16,737 or 47%). Nearly twice more hypomethylated sites (3,516 or 18%) were found in body than hypermethylated sites (1,924 or 9%). Both hyper- and hypomethylated CpG sites were predominantly distributed in islands, shores, and shelves, across all chromosomes. **(B)** Manhattan plot, color-coded by chromosomes.

detect any copy number changes using array CGH in either iKCL004 or iKCL011 over 3 months of culture on Matrigel in mTeSR1 medium [10], suggesting that the genome of both lines is relatively stable. Next, we subjected iKCL004, iKCL011, and parental BJ fibroblast to exome sequencing. The comparative analysis did not identify any alleles that had arisen during the reprogramming, which suggested that the rate of mutation was not substantially affected during reprogramming of BJ fibroblasts.

miRNAome is similar in both iPSC lines

Recently, it has been demonstrated that miRNAs can contribute to iPSC somatic donor memory [26]. Therefore, we evaluated miRNA expression in biological triplicates of iKCL004 and iKCL011 using TaqMan Array MicroRNA Cards containing primer sets for 736 human miRNA sequences. This miRNA profiling technology is a medium throughput approach that holds the advantage of high

accuracy thanks to the use of single assay TaqMan chemistry [27]. After filtering, as described in “Materials and Methods” section, and Loess-based normalization, we compared 383 miRNA expressed in iKCL004 and iKCL011. We found that the average fold change was 0.9998, with a confidence interval 0.995–1.004, $P=0.702$. Kendall’s Tau rank correlation coefficient was 0.860, $P<0.0001$ (if the two rankings are the same, the value is 1 and if they are completely independent, the value is 0). No miRNA was flagged as differentially expressed after multiplicity correction with Benjamini–Hochberg FDR method at the 5% level (Fig. 2A).

Considering that even minimal cellular variation of miRNA concentration can be biologically significant, we next repeated every single miRNA analysis at the 10% and 20% levels, which are less stringent. Even with these larger values for the Type I error, no miRNA was selected as statistically significant in any analysis.

We then analyzed the relative expression miRNA profiles of iKCL004 and iKCL011 against the hESC line KCL034 to see if one of the two iPSC lines was more similar to hESC. Pearson’s correlation showed high similarity between all cell samples analyzed (>0.78 ; Fig. 2B). No significance was found between miRNA signature of KCL034 compared with iKCL011 and iKCL004 (Fig. 2C, E).

iKCL004 and iKCL011 have a different methylation signature

During reprogramming, to acquire pluripotency potential, the DNA of somatic cells undergoes a global epigenetic transformation. Variability of methylation in iPSC lines very often influences the molecular and functional properties of iPSC lines [2,7,28]. Although epigenome-wide methylation analysis suggested that both iKCL004 and iKCL011 display the methylation signature of pluripotent stem cells [10], about 7% of CpG sites (35,550 out of $>480,000$) were differentially methylated (difference >0.2) between iKCL004 and iKCL011, suggesting that these iPSC lines are likely to harbor spots of aberrant epigenetic reprogramming (Fig. 3A, B). Although the differences in DNA methylation signature appeared to have no effect on miRNA expression profile (Fig. 2), they may influence the differentiation propensity of these two iPSC lines.

Since our group is primarily interested in differentiation of hESC and hiPSC into keratinocytes, especially the formation of the epidermal permeability barrier, we next focused on the EDC on chromosome 1. EDC is a ≈ 2 -Mb region at chromosome band 1q21, which contains multiple conserved gene encoding stratum corneum proteins [11–19] (Fig. 4A). We found no difference in DNA methylation of regions associated with either late cornified envelope or small proline-rich region (SPRR) families, as well as with filaggrin, filaggrin 2, involucrin, and loricrin (Supplementary Table S1; Supplementary Figs. S1–S3;

FIG. 5. hESC/iPSC-derived keratinocytes. (A, B) Heatmap and principal component analysis of EDC genes in undifferentiated (T0) hESC (KCL034) and iPSC (iKCL004 and iKCL011), hESC/iPSC-derived keratinocytes (T3), and NHK and fibroblasts (BJ) parental to iPSC lines. (C) KRT14 and p63 expression in hESC/iPSC-derived keratinocytes suggest that they would be comparable with cells in basal layer of epidermis. NHK, normal human keratinocytes.

Supplementary Data are available online at www.liebertpub.com/scd). However, in iKCL004, but not in iKCL011, trichohyalin (TCHH) retained a DNA methylation signature characteristic of somatic cell origin (Fig. 4B). TCHH, an intermediate filament-associated protein, similar to involucrin, is transiently expressed during late stages of skin development, as well as in development of the inner root sheath of the hair follicle [29–32].

TCHH expression pattern in HEE resembles neonatal, not adult, skin

Next, we investigated whether the difference in methylation signature of TCHH in iKCL004 had an effect on expression of EDC genes and hESC/iPSC differentiation into keratinocytes. We compared the transcriptome of EDC genes in undifferentiated hESC/iPSC (T0), hESC/iPSC-derived keratinocytes (T3), and NHK from skin biopsies of two donors. The T0 and T3 transcriptomes were analyzed as biological replicates from three independent sets of differentiation experiments. Heatmaps showing relative expression for EDC genes and the principal component analysis clearly demonstrate that (i) hESC/iPSC-derived keratinocytes at T3 cluster together with NHK; and that (ii) no significant difference was observed between iKCL004 and iKCL011 (Fig. 5A, B). hESC/iPSC-derived keratinocytes, in general, express keratin 14 (KRT14) and TP63 (p63) (Fig. 5C), which are markers of keratinocytes in the stratum basale of human epidermis. No TCHH expression at protein level was detected in any of the samples (data not shown), which suggested that aberrant DNA methylation of TCHH gene did not affect hESC/iPSC differentiation into keratinocytes and did not induce its expression in KRT14⁺/p63⁺ cells.

TCHH is not normally seen in interfollicular adult skin, but is found only in the stratum granulosum (SG) of neonatal skin [30]. To examine TCHH expression in a neonatal model of SG cells, we generated multilayered HEEs from the hESC/iPSC-derived keratinocytes and primary human foreskin keratinocytes following our high-to-low humidity protocol [10,21]. All resulting HEEs, regardless of the keratinocytes' origin, had a similar TCHH expression, which resembled the pattern seen in the neonatal skin (Fig. 6). The data suggest that a difference between two iPSC clones in DNA methylation of TCHH gene did not have an obvious effect at TCHH expression pattern in three-dimensional (3D) HEEs.

Discussion

Exome sequencing of both iPSC lines and parental HFFs BJ confirmed that the reprogramming used to generate these cells did not lead to increased number of mutations. Moreover, both lines were genetically identical to their parental line. Karyotype of both iPSC lines was stable over the extended periods of in vitro culture under feeder-free conditions. Cytogenetic resilience of iKCL004 and iKCL011 lines might arise from particular aspects of cell culture, as well as reflect a genetic predisposition to chromosomal stability/instability, commonly noted with hESC lines [8,9,33,34]. hESC is derived from embryos created by in vitro fertilization, often from donors that have problems with fertility. Prevalence of balanced structural chromosomal abnormalities in patients seeking assisted conception is higher than in general population and it is reasonable to speculate that undiagnosed genetic mutations might be underlying cause of infertility, and consequently, genomic instability in the hESC lines.

Although ideally, hiPSC should be functionally and molecularly indistinguishable from hESC, often that is not the case [35]. These differences may manifest only upon differentiation, not in the pluripotent state; miRNA profile [36] and residual DNA methylation [28] are likely to influence differentiation propensity. For example, upon induction of keratinocyte differentiation with bone morphogenic protein 4 and all-trans retinoic acid, none of the three lines used in the study (iKCL004, iKCL011, and KCL034) responded equally [10]. Over time, by end of the differentiation, these differences were diminished, probably through selection and purification by selective attachment to collagen IV-coated surface as described [10]. Indeed, the miRNA profile of both iPSC lines, iKCL004 and iKCL011, was not significantly different from control hESC KCL034 line (Fig. 2). Interestingly, epigenome-wide methylation (>480,000 CpG sites) analysis using Infinium Human Methylation 450 BeadChip Kit found more differences between iKCL004 and iKCL011 (7%) than between either iKCL004 and KCL034 (5%) or iKCL011 and KCL034 (4%) (Fig. 3) [10].

Differently methylated CpG sites are in general randomly scattered across the genome. Within one gene, DNA methylation often is significantly different on only one or two CpG sites (Supplementary Table S1; Supplementary Figs. S1–S3).

FIG. 6. The expression pattern of TCHH in human epidermal equivalents is similar to human neonatal skin. Localization in the granular layer corresponds to expression of other proteins involved in cornification such as LCE2B.

One of the few genes in which DNA methylation was significantly different on most of its CpG sites (12 out of 21) was *TCHH* (Supplementary Table S1; Fig. 4). *TCHH* is normally expressed in the inner root sheet and medulla of hair follicles and transiently in the SG and corneum of the epidermis during late stages of fetal skin development [30]. It is also present in isolated cells of the SG of some regions of epidermis (Supplementary Fig. S4) [37]. While in normal skin, *TCHH* expression correlates with expression of filaggrin, this correlation is lost in psoriatic skin [38]. By forming multiple complex cross-links with itself and other structural proteins, including epiplakin, involucrin, and SPRRs, *TCHH* is important for the mechanical strength of the hair follicle and epithelial tissues such as hard palate or filiform papillae of the tongue [39]. *TCHH* is also thought to be a potential major autoantigen in alopecia areata [40]. CpG sites within *TCHH* were mostly hypomethylated in HFFs and hypermethylated in the hESC KCL034 line (Fig. 4). *TCHH* methylation pattern suggested that iKCL004 retained methylation signature of the parental fibroblasts, whereas iKCL011 matched the hESC line. We found, however, no differences in *TCHH* expression pattern in HEE (Fig. 6) suggesting that DNA methylation of the *TCHH* gene in establishing an iPSC population may not play an essential role.

The data presented in this study highlight our limited understanding of the epigenetic heterogeneity. It is generally considered that faithful genetic and epigenetic reprogramming is critical to obtain fully naive “ground state” of pluripotency.” Indeed, DNA methylation defects have been linked to multiple human diseases [41]. However, as there are significant differences among the genomes of healthy human individuals that do not affect normal fetal and postnatal growth, it is quite plausible that the same rule can also be applied for the epigenome—DNA methylation status may not be equally important for regulation of all genes. Differentiation and tissue formation may, in some cases, attenuate variations in iPSC methylome. These findings should enhance future research using these 3D cultures, since iPSC methylome differences produced by the transformation and culture techniques do not result in phenotypic variation.

Acknowledgments

This work was supported, in part, by grants from the Medical Research Council UK (G0801061 to D.I.), the NIH (R21 ARO61583 and R01 AR051930), the Research Service of the Department of Veterans Affairs (to T.M.M.), and the Dystrophic Epidermolysis Bullosa Research Association (DebRA) (to J.A.M.). A.P. was a DebRA-supported PhD student. We thank Prof Caroline Ogilvie from the Genetics Center at Guy's Hospital for a critical reading of the article.

Author Disclosure Statement

No competing financial interests exist.

References

1. Takahashi K and S Yamanaka. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126:663–676.
2. Sugiura M, Y Kasama, R Araki, Y Hoki, M Sunayama, M Uda, M Nakamura, S Ando and M Abe. (2014). Induced pluripotent stem cell generation-associated point mutations arise during the initial stages of the conversion of these cells. *Stem Cell Rep* 2:52–63.
3. Young MA, DE Larson, CW Sun, DR George, L Ding, CA Miller, L Lin, KM Pawlik, K Chen, et al. (2012). Background mutations in parental cells account for most of the genetic heterogeneity of induced pluripotent stem cells. *Cell Stem Cell* 10:570–582.
4. Kim K, R Zhao, A Doi, K Ng, J Unternaehrer, P Cahan, H Huo, YH Loh, MJ Aryee, et al. (2011). Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat Biotechnol* 29:1117–1119.
5. Ohi Y, H Qin, C Hong, L Blouin, JM Polo, T Guo, Z Qi, SL Downey, PD Manos, et al. (2011). Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPSC cells. *Nat Cell Biol* 13:541–549.
6. Lister R, M Pelizzola, YS Kida, RD Hawkins, JR Nery, G Hon, J Antosiewicz-Bourget, R O'Malley, et al. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* 471:68–73.
7. Polo JM, S Liu, ME Figueroa, W Kulalart, S Eminli, KY Tan, E Apostolou, M Stadtfeld, Y Li, et al. (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat Biotechnol* 28:848–855.
8. Stephenson E, CM Ogilvie, H Patel, G Cornwell, L Jacquet, N Kadeva, P Braude and D Illic. (2010). Safety paradigm: genetic evaluation of therapeutic grade human embryonic stem cells. *J R Soc Interface* 7:S677–S688.
9. Liang G and Y Zhang. (2013). Genetic and epigenetic variations in iPSCs: potential causes and implications for application. *Cell Stem Cell* 13:149–159.
10. Petrova A, A Celli, L Jacquet, D Dafou, D Crumrine, M Hupe, M Arno, C Hobbs, A Cvorovic, et al. (2014). 3D in vitro model of a functional epidermal permeability barrier from human embryonic stem cells and induced pluripotent stem cells. *Stem Cell Rep* 2:675–689.
11. Jackson B, CM Tilli, MJ Hardman, AA Avilion, MC MacLeod, GS Ashcroft and C Byrne. (2005). Late cornified envelope family in differentiating epithelia—response to calcium and ultraviolet irradiation. *J Invest Dermatol* 124:1062–1070.
12. Contzler R, B Favre, M Huber and D Hohl. (2005). Cornulin, a new member of the “fused gene” family, is expressed during epidermal differentiation. *J Invest Dermatol* 124:990–997.
13. Marshall D, MJ Hardman, KM Nield and C Byrne. (2001). Differentially expressed late constituents of the epidermal cornified envelope. *Proc Natl Acad Sci U S A* 98:13031–13036.
14. South AP, A Cabral, JH Ives, CH James, G Mirza, I Marenholz, D Mischke, C Backendorf, J Ragoussis and D Nizetic. (1999). Human epidermal differentiation complex in a single 2.5 Mbp long continuum of overlapping DNA cloned in bacteria integrating physical and transcript maps. *J Invest Dermatol* 112:910–918.
15. Lioumi M, CA Ferguson, PT Sharpe, T Freeman, I Marenholz, D Mischke, C Heizmann and J Ragoussis. (1999). Isolation and characterization of human and mouse ZIRTL, a member of the IRT1 family of transporters, mapping within the epidermal differentiation complex. *Genomics* 62:272–280.
16. Zhao XP and JT Elder. (2002). Evidence for local control of gene expression in the epidermal differentiation complex. *Exp Dermatol* 11:406–412.

17. Zhao XP and JT Elder. (1997). Positional cloning of novel skin-specific genes from the human epidermal differentiation complex. *Genomics* 45:250–258.
18. Marenholz I, A Volz, A Ziegler, A Davies, I Ragoussis, BP Korge and D Mischke. (1996). Genetic analysis of the epidermal differentiation complex (EDC) on human chromosome 1q21: chromosomal orientation, new markers, and a 6-Mb YAC contig. *Genomics* 37:295–302.
19. Mischke D, BP Korge, I Marenholz, A Volz and A Ziegler. (1996). Genes encoding structural proteins of epidermal cornification and S100 calcium-binding proteins form a gene complex (“epidermal differentiation complex”) on human chromosome 1q21. *J Invest Dermatol* 106:989–992.
20. Devito L, A Petrova, C Miere, S Codognotto, N Blakely, A Lovatt, C Ogilvie, Y Khalaf and D Ilic. (2014). Cost-effective master cell bank validation of multiple clinical-grade human pluripotent stem cell lines from a single donor. *Stem Cell Transl Med* 3:1116–1124.
21. Sun R, A Celli, D Crumrine, M Hupe, LC Adame, SD Pennypacker, K Park, Y Uchida, KR Feingold, et al. (2015). Lowered humidity produces human epidermal equivalents with enhanced barrier properties. *Tissue Eng Part C Methods* 21:15–22.
22. Meyer SU, S Kaiser, C Wagner, C Thirion and MW Pfaffl. (2012). Profound effect of profiling platform and normalization strategy on detection of differentially expressed microRNAs—a comparative study. *PLoS One* 7:e38946.
23. Ilic D, M Mao-Qiang, D Crumrine, G Dolganov, N Larocque, P Xu, M Demerjian, BE Brown, ST Lim, et al. (2007). Focal adhesion kinase controls pH-dependent epidermal barrier homeostasis by regulating actin-directed Na⁺/H⁺ exchanger 1 plasma membrane localization. *Am J Pathol* 170:2055–2067.
24. Stephenson E, L Jacquet, C Miere, V Wood, N Kadeva, G Cornwell, S Codognotto, Y Dajani, P Braude and D Ilic. (2012). Derivation and propagation of human embryonic stem cell lines from frozen embryos in an animal product-free environment. *Nat Protoc* 7:1366–1381.
25. Warren L, PD Manos, T Ahfeldt, YH Loh, H Li, F Lau, W Ebina, PK Mandal, ZD Smith, et al. (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* 7:618–630.
26. Vitaloni M, J Pulecio, J Bilic, B Kuebler, L Laricchia-Robbio and JC Izpisua Belmonte. (2014). MicroRNAs contribute to induced pluripotent stem cell somatic donor memory. *J Biol Chem* 289:2084–2098.
27. Wang B, P Howel, S Bruheim, J Ju, LB Owen, O Fodstad and Y Xi. (2011). Systematic evaluation of three microRNA profiling platforms: microarray, beads array, and quantitative real-time PCR array. *PLoS One* 6:e17167.
28. Kim K, A Doi, B Wen, K Ng, R Zhao, P Cahan, J Kim, MJ Aryee, H Ji, et al. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature* 467:285–290.
29. Steinert PM and LN Marekov. (1999). Multiple roles for trichohyalin in the inner root sheath. *Exp Dermatol* 8:331–332.
30. Lee SC, JB Lee, JP Kook, JJ Seo, KI Nam, SS Park and YP Kim. (1999). Expression of differentiation markers during fetal skin development in humans: immunohistochemical studies on the precursor proteins forming the cornified cell envelope. *J Invest Dermatol* 112:882–886.
31. Lee SC, M Wang, OW McBride, JE O’Keefe, IG Kim and PM Steinert. (1993). Human trichohyalin gene is clustered with the genes for other epidermal structural proteins and calcium-binding proteins at chromosomal locus 1q21. *J Invest Dermatol* 100:65–68.
32. Mlitz V, B Strasser, K Jaeger, M Hermann, M Ghannadan, M Buchberger, L Alibardi, E Tschachler and L Eckhart. (2014). Trichohyalin-like proteins have evolutionarily conserved roles in the morphogenesis of skin appendages. *J Invest Dermatol* 134:2685–2692.
33. Hanson C and G Caisander. (2005). Human embryonic stem cells and chromosome stability. *APMIS* 113:751–755.
34. Catalina P, R Montes, G Ligerio, L Sanchez, T de la Cueva, C Bueno, PE Leone and P Menendez. (2008). Human ESCs predisposition to karyotypic instability: is a matter of culture adaptation or differential vulnerability among hESC lines due to inherent properties? *Mol Cancer* 7:76.
35. Chin MH, MJ Mason, W Xie, S Volinia, M Singer, C Peterson, G Ambartsumyan, O Aimiwu, L Richter, et al. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5:111–123.
36. Neveu P, MJ Kye, S Qi, DE Buchholz, DO Clegg, M Sahin, IH Park, KS Kim, GQ Daley, et al. (2010). MicroRNA profiling reveals two distinct p53-related human pluripotent stem cell states. *Cell Stem Cell* 7:671–681.
37. O’Keefe EJ, EH Hamilton, SC Lee and P Steinert. (1993). Trichohyalin: a structural protein of hair, tongue, nail, and epidermis. *J Invest Dermatol* 101:65S–71S.
38. Ishida-Yamamoto A, Y Hashimoto, M Manabe, WM O’Guin WM, BA Dale and H Iizuka. (1997). Distinctive expression of filaggrin and trichohyalin during various pathways of epithelial differentiation. *Br J Dermatol* 137:9–16.
39. Steinert PM, DA Parry and LN Marekov. (2003). Trichohyalin mechanically strengthens the hair follicle: multiple cross-bridging roles in the inner root sheath. *J Biol Chem* 278:41409–41419.
40. Leung MC, CW Sutton, DA Fenton and DJ Tobin. (2010). Trichohyalin is a potential major autoantigen in human alopecia areata. *J Proteome Res* 9:5153–5163.
41. Robertson KD. (2005). DNA methylation and human disease. *Nat Rev Genet* 6:597–610.

Address correspondence to:

Dusko Ilic
Assisted Conception Unit
Stem Cell Laboratory
11th Floor, Tower Wing, Guy’s Hospital
London SE1 9RT
United Kingdom

E-mail: dusko.ilic@kcl.ac.uk

Received for publication June 3, 2016

Accepted after revision July 26, 2016

Prepublished on Liebert Instant Online July 26, 2016