Frontiers in Bioscience Landmark

Front. Biosci. (Landmark Ed) **2022**; 27(8): 239
https://doi.org/10.31083/j.fbl2708239

Original Research

Universal Markers for hiPSCs Residue Detection

Hui Shi[1], Tiejun Feng[2], Rui Wang[2], Lida Wu[2,*], Yuchun Gu[1,2,*]

[1]Molecular Pharmacology Laboratory, Institute of Molecular Medicine, Peking University, 100871 Beijing, China
[2]Allife Medical Science and Technology Co., Ltd. Economic and Technological Development Zone, 100176 Beijing, China
*Correspondence: wldpaper@pku.edu.cn (Lida Wu); ycgu@pku.edu.cn (Yuchun Gu)

Academic Editor: Graham Pawelec
Submitted: 27 May 2022 Revised: 8 July 2022 Accepted: 25 July 2022 Published: 11 August 2022

Abstract

Background: Residual undifferentiated induced pluripotent stem cells (iPSCs) detection is essential for both Embryonic Stem Cells (ESCs) and iPSCs application in final cell therapy products. However, specific differentiated cells require specific genes for residual detection; identifying the suitable marker is costly and time-consuming. Thus, a universal marker for iPSCs residue detection for all three germline cells would greatly benefit PSC-derived cellular therapies. **Methods**: Next-generation sequencing (NGS) was performed on total RNAs isolated from the iPSC cell lines and embryonic stem cells (H9), the top 30 expressed genes were selected as candidates. By analysis expression fold change comparing iPSC cells to the differentiated cells, seven genes were highly expressed in iPSCs but showed minimal background expression in differentiated cells. Tissue expression pattern of the candidate genes were explored in the Genotype-Tissue Expression (GTEx) project database, candidate genes were narrowed down to two genes. Spike-in experiments were performed to determine the detection limit and correlation with the number of iPSCs and gene expression by ddPCR. **Results**: By next-generation sequencing (NGS), we identified two marker genes (ESRG and ZSCAN10) suitable for universal undifferentiated iPSC detection. Both ESRG and ZSCAN10 are highly expressed in iPSCs. ZSCAN10 is slightly expressed in the testis, pituitary, and cerebellum; ESRG is highly expressed in the vagina and scarcely expressed in the other tissues. Furthermore, the ddPCR method with a probe and primers for ESRG and ZSCAN10 detected a trace of undifferentiated hiPSCs to a spiked level of 0.0001%. **Conclusions**: These results suggest that targeting ESRG/ZSCAN10 transcripts is highly sensitive, quantitative, and could be broadly applied to quality control of almost all iPSC-derived cell therapy products.

Keywords: human pluripotent stem cell; cell therapy; ddPCR; *ESRG*; *ZSCAN10*

1. Introduction

Human induced pluripotent stem cells (hiPSCs) can indefinitely self-renew and differentiate into almost any specialized cell type [1,2]. Therapies that have been based on PSC have shown great promises in patient-specific cell therapy, potentially providing regenerative medicine therapies for many life-threatening diseases. Thus, an increasing number of cell therapies are in clinical development and have shown good preliminary clinical efficacies [3–5]. However, one particular issue associated with PSC-based cell therapies is the presence of residual undifferentiated hiPSCs in the drug products, which can lead to tumorigenesis [6,7]. Therefore, it is essential to establish highly sensitive assays to detect residual undifferentiated hiPSCs.

Currently, several strategies, such as flow cytometry [8,9], quantitative real-time polymerase chain reaction (qRT-PCR) [8,10,11], digital droplet PCR [12], miRNA targets [13] and highly efficient culture system [14], have been implemented to identify unwanted undifferentiated PSC residues and allow safe transplantation. Most of these methods were based on detecting undifferentiated cell marker expression, except for the culture system platform. For instance, Lin-28 Homolog A (*Lin28A*), although it is widely clinics for detecting undifferentiated cells in retinal pigment epithelial (RPE) cells derived from hiPSCs, similar efficacies in hepatic differentiation could not be replicated [11]. Also, despite *OCT4* and *TDGF1* being useful detection markers, they are mainly specific to hiPSC-derived neural stem cells [15]. Thus, an important limitation in current research is that the markers identified/developed are specific to certain cell types and cannot be used in other cell lineages. Further, continuous research for identifying suitable markers is costly and time-consuming. Therefore, developing a universal marker for hiPSCs residue detection for all three germinal layers cells would be of great significance for iPSC-derived cell therapy products.

This study focused on identifying marker genes suitable for hiPSCs residue detection for any cell type derived from hiPSCs. Using next-generation sequencing analysis on the top 30 hiPSC-expressed genes, 7 potential candidate genes were identified. Then, the expression pattern was explored using the Genotype-Tissue Expression (GTEx) project database to narrow these 7 genes to the 2 most potential gene candidates (*ESRG* and *ZSCAN10*). The droplet digital PCR (ddPCR) method with corresponding primer-probe pairs was then used and showed a trace of undifferentiated hiPSCs, which spiked to a level of 0.0001%. This result suggests that targeting *ESRG*/*ZSCAN10* transcripts could be highly sensitive and quantitative and could

be broadly applied to assess the quality of hiPSC-based cell replacement products.

2. Materials and Methods

2.1 Reprogramming hiPSC Lines

The cells were then transduced using the Episomal hiPSC Reprogramming Vectors (Life Technologies), followed by culturing on matrigel-coated dishes using StemSpan™ SFEM/StemSpan™ CD34+ Expansion Supplement (Stem Cell Technologies). The half medium was changed with Essential 8 medium (Life Technologies) each day from day 1 to day 7 until small colonies were formed. From the 7th day, Essential 8 was used as the main median and was changed daily for 2–3 weeks. When colonies developed, they were individually retrieved (for the first 8 passages), mechanically dissociated, and used for expansion into individual hiPSC lines. Every 5 days, the hiPSCs were passaged with 0.5 mM EDTA (Sigma). Three colonies were chosen for further experiments based on molecular testing and morphology.

2.2 Cell Culture

Matrigel plates (Corning) were used for hiPSCs and H9 cell culture. The medium used was the Essential 8 medium supplemented with penicillin (50 U/mL) and streptomycin (50 mg/mL) (Life Technologies). When the cell confluency reached 70–80%, EDTA solution (Sigma) was used for undifferentiated colonies passage.

2.3 Embryoid Body (EB) Formation and Spontaneous Differentiation

Here, 0.5 mM EDTA was used to detach hiPSCs, suspended in the Essential 8 medium containing 10 μM of Y-27632 (Abcam), seeded into low attachment dishes (Corning), then incubated at 37 °C with 5% CO_2 for 24 h. Following EB formation, they were cultured in an E6 medium (Life Technologies), which was changed every 3 days. Following an incubation period of 10–14 days, spontaneous differentiated EBs were obtained and used for further experiments.

2.4 Differentiation of Cardiomyocytes (CMs) from hiPSCs

The differentiation of cardiomyocytes from hiPSCs was induced, as reported by Cao *et al*. [16] but with some adjustements. In short, the hiPSCs were incubated for 5 min in 0.5 mM EDTA and PBS, then transferred into Corning low attachment dishes containing Essential 8 medium for 2 days before cardiomyocyte induction. Induction was performed as follows: On day 0, the EBs were treated using a cocktail of 3 μM CHIR99021 (Stemgent), 25 ng/mL BMP4, 50 μg/mL L-ascorbic acid (Sigma), 400 μM 1-thioglycerol, DMEM/F12 combined with B27-vitamin A (Life Technologies) for 2 days; on day 2, the initial medium was replaced with fresh medium; day 3: treatment with 2 μM C59 (Abcam), 10 ng/mL BMP4 (Peprotech) in RPMI combined with B27-insulin for 2 days; on day 5, the medium was replaced with RPMI combined with B27 (Life Technologies) and replaced every 3 days. After induction, the differentiated cells were retrieved and stored for further analysis.

2.5 Hepatocyte-Like Cell Differentiation of hiPSCs

Hepatocyte-like cell differentiation was derived from hiPSCs, as previously reported [17]. Briefly, the hPSCs were seeded for hepatocyte differentiation using LN-521 (BioLamina) for 24 h in Essential 8 medium. On day 1, the medium was changed to fresh RPMI 1640 medium (Life Technologies), to which 50 ng/mL of Wnt 3a (Peprotech), 100 ng/mL of Activin A (Peprotech), 2% of B27 (Life Technologies) and 1% of penicillin or streptomycin (Life Technologies) were added. The medium was replaced with fresh medium on a daily basis for 5 days. On day 6, the medium was switched to a cocktail of KSR/DMSO combined with 0.5% GlutaMAX (Life Technologies), 1% DMSO (Sigma), 1% non-essential amino acids (Life Technologies), 1% penicillin/streptomycin, 80% knockout DMEM (Life Technologies), 20% knockout serum replacement (Life Technologies) and 0.1 mM beta-mercaptoethanol (Life Technologies); which were replaced every day for 5 days. Then, the cocktail medium consisting of HepatoZYME basal (Life Technologies) combined with 20 ng/mL OSM (Peprotech), 10 ng/mL HGF (Peprotech), 10 μM hydrocortisone 21-hemisuccinate sodium salt (Sigma), 1% GlutaMAX and 1% penicillin/streptomycin was used and replaced with fresh ones every 2 days for 7–10 days. When the cells reached standard differentiation, they were retrieved and stored for further use.

2.6 Neural Stem Cell Differentiation of hiPSCs

The hiPSCs were plated onto Geltrex-coated six-well plates after splitting into cell clumps at a density of 2–2.5 $\times$ 10^4 cells per cm^2. After a day, the Gibco PSC Neural Induction Medium (Life Technologies) containing a Neurobasal medium and Gibco PSC neural induction supplement was used. From days 0–4, medium was changed every other day and changed to a daily basis as from day 4 because the cells reached full confluence quickly. On day 7, the primitive NSCs were detached with Accutase (Life Technologies) transferred onto Geltrex-coated dishes, at a density of 0.5–1 $\times$ 10^5 cells per cm^2, and maintained in an NSC expansion medium consisting of a neural induction supplement and 50% of Neurobasal and Advanced DMEM/F12 medium. The NSC expansion medium was changed on a daily basis until they reached confluence on day 5 since the NSC plating.

2.7 RNA Extraction and cDNA Preparation

Total RNA was isolated from the cells (hiPSCs, H9 and EBs) with Trizol (Life Technologies) following the manufacturer's protocol, which was then purified using the column cleanup (Qiagen). cDNA was synthesized

by reverse transcription following the manufacturer's instructions using TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix (Transgene).

2.8 ddPCR

The BioRad QX200 Droplet Digital PCR instrument with corresponding EvaGreen ddPCR SuperMix (Bio-Rad) was used. After preparing the ddPCR reaction mixtures, the QX200 droplet generator (Bio-Rad) was used to produce droplets using the following conditions: 10 min at 95 °C, followed by 40 cycles of denaturation for 30 s at 95 °C, 60 s of annealing/extension at 60 °C followed by 1 min at 72 °C, 5 min at 4 °C and 5 min at 90 °C, then maintained at 12 °C. Following the PCR reaction, the data were assessed using QuantaSoft (Bio-Rad). The corresponding fluorescence amplitude threshold of each gene was determined manually by comparing their distribution in distilled water.

2.9 Immunofluorescence Staining

Cells were fixed with 4% paraformaldehyde (Sigma-Aldrich) for 15 min at room temperature, permeabilized with 0.1% Triton-X (Sigma-Aldrich) for 10 min at room temperature, blocked in 5% BSA (Sigma-Aldrich) 1h at room temperature, followed by immunocytochemistry staining with the Embryonic Stem Cell Marker Panel (ab109884, abcam), Mouse monoclonal [1F11] to cTnT (ab10214, abcam), Nestin Polyclonal antibody (19483-1-AP, proteintech), AFP Polyclonal antibody (14550-1-AP, proteintech), Albumin Polyclonal antibody (16475-1-AP, proteintech) overnight at 4 °C in 1% BSA. Cells were washed two times, for 10 min, and then incubated for 1 h at room temperature in the dark with secondary antibodies Alexa Fluor 488 goat anti-mouse IgG, AlexaFluor 488 goat anti-rabbit IgG (H+L) or Alexa Fluor 594 goat anti-rabbit IgG (H+L) (all Life Technologies). Cells were washed two times, for 10 min, nuclei were stained with DAPI (Sigma-Aldrich) for 5 min , and washed three times, for 10 min. then The cells were imaged with ImageXpress Pico (MolecularDevices).

2.10 mRNA Sequencing by Illumina HiSeq

Total RNA was obtained from the samples with TRIzol (Invitrogen, Carlsbad, CA, USA). RNA integrity was determined on the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). RNA quality and purity were determined using 1% agarose gel and the NanoDrop equipment (Thermo Fisher Scientific Inc., Waltham, MA, USA). NGS library were prepared using 1 μg of total RNA at RIN value >6.5. The Poly(A) mRNA Magnetic Isolation Module was used for mRNA isolation. The First-Strand Synthesis Reaction Buffer was used for mRNA fragmentation. Random Primers were used for mRNA priming. To synthesize the first-strand of cDNA, ProtoScript II Reverse Transcriptase was used, and for the second-strand cDNA, the Second Strand Synthesis Enzyme Mix was used. The End Prep Enzyme Mix was then used to repair the opposite ends of the purified double-stranded cDNA and add a dA-tail, in one reaction. Next, T-A ligation was performed to both ends to add adaptors. An appropriate size of the adaptor-ligated DNA was chosen with Beads, and fragments with a size of approximately ~420 bp (approximate insert size, 300 bp) were recaptured. Next, amplification with 13 cycles of PCR using P5 and P7 primers was performed for each sample. The P7 primers, which contained a six-base index, were used for multiplexing. After this, cleaning of the PCR products were performed with beads, Qsep100 was used for validation (Biopticm, Taiwan, China) and a Fluorometer (Qubit 3.0; Invitrogen) was used for quantification. Libraries having different indexes were multiplexed and loaded onto the Illumina HiSeq instrument (Illumina, San Diego, CA, USA). The configuration was set at 2 $\times$ 150 bp paired-end configuration for sequencing. The HiSeq Control Software (HCS) + OLB + GAPipeline-1.6 (Illumina) on the HiSeq instrument was used for image analysis and base calling. Lastly, GENEWIZ (Suzhou, China) was used to analyze the sequences.

2.11 Data Analyses

Upstream analyses of RNA-seq data were performed using a series of software, including fastp (version 0.14.1, https://github.com/OpenGene/fastp) was used to do the quality control of raw reads, star (version 2.7.10, https://github.com/alexdobin/STAR) was used to do the reads mapping to hg19, htseq (version 0.11.3, https://github.com/simon-anders/htseq) was used to do the reads counting, countToFPKM (version 1.0.0, https://github.com/AAlhendi1707/countToFPKM) was used to do the fpkm normalizing. Downstream analyses of RNA-seq data were performed using R (version 3.6.0, https://www.r-project.org/). pheatmap (version 1.0.12, https://github.com/raivokolde/pheatmap) was used to generate the heatmap plot. The description of the heatmap is: the color is drawn by the fpkm value,expressing the difference between samples, with the color bar depicting the extents of the expression value accordingly.

3. Results

3.1 Marker Genes Highly Expressed in hiPSCs

Marker genes were selected for detecting residual undifferentiated cells using the experimental scheme illustrated in Fig. 1A. Three hiPSC cell lines were constructed from PBMCs for further experiments based on molecular testing and morphology (**Supplementary Fig. 1A**). To identify candidate genes with high expression in hiPSCs but minimally expressed in differentiated cells, next-generation sequencing (NGS) was performed on the total RNAs isolated from the hiPSC cell lines and embryonic stem cells (H9), a panel of 250 stem cells markers was selected as a candidates pool [18], and a heat map was plotted (Fig. 1B), following which the top 30 expressed genes were picked.

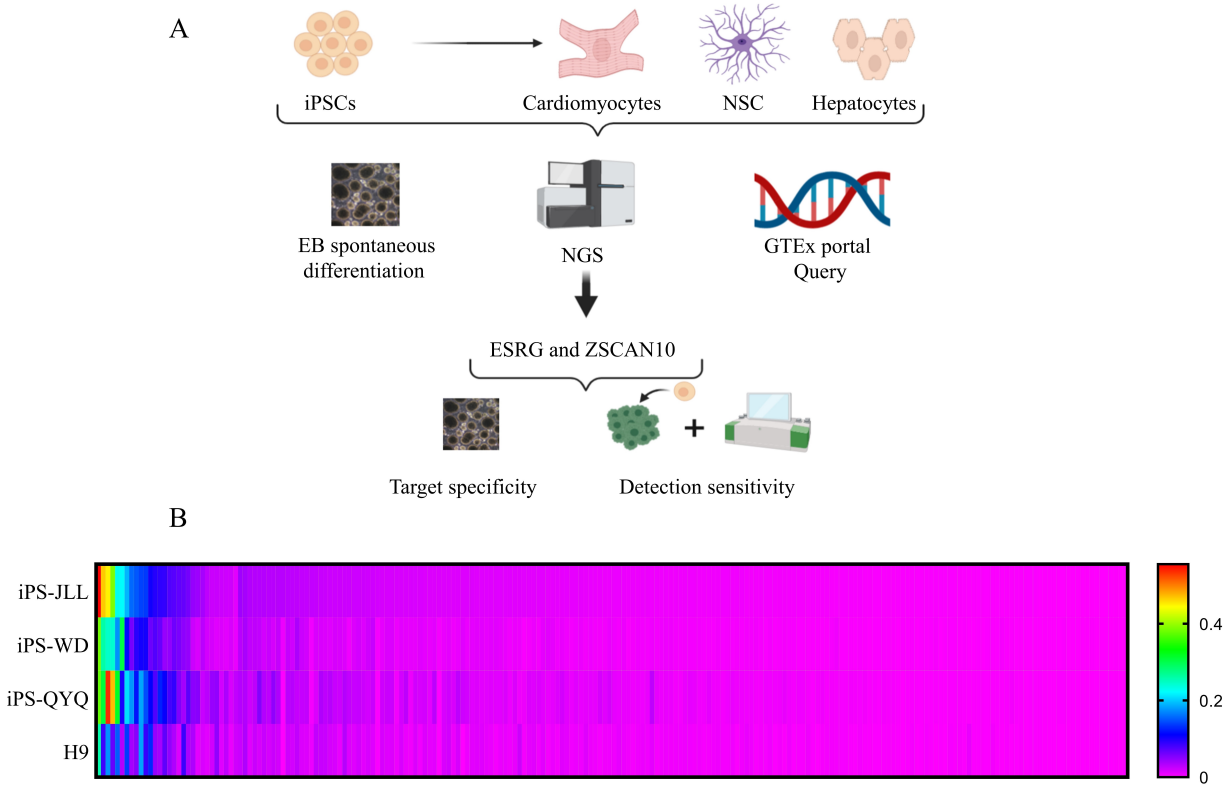

Fig. 1. Experimental scheme and stem cell genes analysis. (A) Experimental scheme for universal hiPSCs residue detection markers identification. Differentiate hiPSC into cardiomyocytes, neural stem cells and hepatocytes, then perform whole-exome sequencing on the hiPSCs and differentiated cells to analyze gene expression. Genes that were highly expressed in hiPSCs and were simultaneously low or not expressed in differentiated cells were selected as the candidate genes. NGS sequencing on spontaneous differentiated EB and GEtex query showed that the candidate genes with low expression in various tissues were the target genes. Finally, analyze the sensitivity of the target genes by digital PCR. (B) Expression heat map of a panel of 250 stem cell genes; sorting gene expression levels from high to bottom.

3.2 Marker Genes Slightly Expressed in hiPSC-Derived Cells

hiPSCs were differentiated into neural stem cells (ectoderm), hepatocytes-like cells (endoderm) and cardiomyocytes (mesoderm) to identify genes specifically expressed in differentiated cells. hiPSCs derived neural stem cells were identify by the expresion of *NESTIN* and *SOX2* (**Supplementary Fig. 2A**), hiPSCs derived hepatocytes-like cells were identify by the expresion of *AFP* and *ALB* (**Supplementary Fig. 3A**), hiPSCs derived cardiomyocytes were identify by the expresion of *cTnT* (**Supplementary Fig. 4A**). Next-generation sequencing (NGS) was also performed on total RNAs isolated from the hiPSC differentiated cells following differentiation. Next, the mRNA seq data on the top 30 expressed genes was explored (Fig. 2A). Several genes, such as *POU5F1*, *ESRG*, *DNMT3B*, *GAL*, *L1TD1*, *TDGF1, NSCAN10* and *DPPA4*, were highly expressed in hiPSCs and had minimal expression in differentiated cells.

To identify highly selective markers for undifferentiated hiPSCs, expression fold change analysis on the RNA sequencing data comparing hiPSC cells to the differentiated cells was performed. An expression fold change >20 was set as the criteria for candidate genes. Our findings showed the expression fold change of *ZSCAN10*, *ESRG*, *PTPRZ1*, *GAL*, *TDGF1*, *ZIC2*, *POU5F1*, *SOX2*, *L1TD1* and *DNMT3B* was >20 when differentiating hiPSCs to hepatocyte-like cells (Fig. 2B). Comparing hiPSCs with neural stem cells, *TDGF1*, *ESRG*, *ZSCAN10*, *GAL*, *L1TD1*, *EPCAM*, *DNMT3B*, *PIM2* and *POU5F1* met the selection criteria (Fig. 2C). *ESRG*, *ZSCAN10*, *ZIC2*, *DPPA4*, *SOX2*, *DNMT3B*, *PTPRZ1*, *POU5F1*, *THY1*, *GAL*, *LIN28A*, *L1TD1*, *TDGF1*, *TERF1* and *PIM2* were selected as the candidates when differentiating hiPSCs to cardiomyocytes (Fig. 2D). Seven overlapping genes were found between the three sets of analyses (Fig. 2E), among which the expression fold change of four genes (*ZSCAN10*, *ESRG*, *GAL* and TDGF1) was >50. Therefore, these seven genes were selected as candidates because they were highly expressed and abundant in hiPSC.

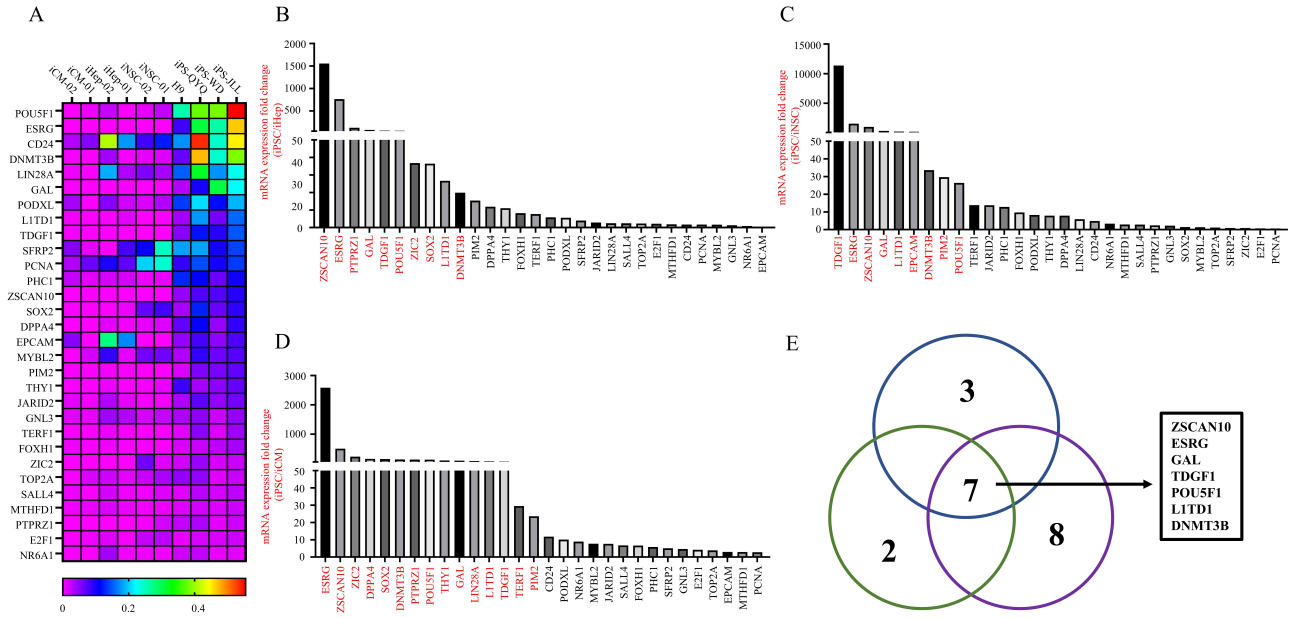

Fig. 2. hiPSCs residue detection candidate genes screening. (A) Expression heat map of 30 stem cell genes of the hiPSCs, H9 and hiPSC-derived hepatocyte-like cells (iHep), cardiomyocytes (iCM), and NSC (iNSC). The gene expression levels were sorted from high to bottom. (B) mRNA expression fold change between hiPSCs and hepatocyte-like cells. The genes marked in red are those whose expression fold change exceeded 20. (C) mRNA expression fold change between hiPSCs and NSC. The genes marked in red are those whose expression fold change exceeded 20. (D) mRNA expression fold change between hiPSCs and cardiomyocytes. The genes marked in red are those whose expression fold change exceeded 20. (E) Screening strategy for identifying candidate genes.

3.3 Tissue Expression Pattern of the Candidate Genes

To determine whether the candidate marker genes could be used for identifying residual undifferentiated hiPSCs in any cell type in the body, Multi-Gene Query was conducted through The GTEx database (UTL: https://gtexportal.org/home/) (Fig. 3A). Five genes (*ZSCAN10*, *ESRG*, *FOXH1*, *LIN28A*, and *DPPA4*) which were abundantly expressed in hiPSCs but minimally expressed in other tissues (Fig. 3A, in red) were selected as suitable for universal residual hiPSCs detection. We also generated the tissue expression pattern of several genes (Fig. 3B). A slight expression of *ZSCAN10* was observed in the testis, pituitary and cerebellum, while the expression of *ESRG* was high in the vagina but low in the other tissues. The other candidate genes were abundantly expressed in many tissues (**Supplementary Fig. 5A**). Combining the differentiation data above (Fig. 2E), *ZSCAN10* and *ESRG* were selected as the ideal markers for universal residual hiPSCs detection.

A spontaneous embryonic body (EB) differentiation was performed to validate the selection of *ZSCAN10* and *ESRG*, as EB contains all three germ layers. For spontaneous differentiation, self-aggregated EBs were formed and transferred to Corning low attachment dishes. The medium was switched to E6 medium for 14 days. Following differentiation, NGS was also performed on total RNAs isolated from the EBs. Next, the mRNA seq data on the top 30 expressed genes was assessed (Fig. 3C). *ZSCAN10* and *ESRG* were found highly expressed in hiPSCs but scarcely expressed in differentiated EBs, with an expression fold change of 146 and 574, respectively (Fig. 3D). These findings showed that *ZSCAN10* and *ESRG* were suitable genes for universal residual hiPSCs detection.

3.4 Detection Sensitivity of Candidate Marker Gene

Here Spike-in experiments were performed. A serial of 2, 20, 200, 2000 and 20,000 hiPSCs was mixed with 293T cells to get 2 million cells; then, the RNA was isolated for ddPCR. The copy number of *ESRG* and *ZSCAN10* gene expression correlated with hiPSCs numbers (Fig. 4A,B). Two hiPSCs in a background of 2×10^6 cells were detected with at least 0.0001% sensitivity. However, *ESRG* performed better with a higher copy number (Fig. 4A), possibly because of a higher expression level of *ESRG* in hiPSCs (Fig. 2D), which was consistent with the RNA seq data.

4. Discussion

This study identified two target genes for residual undifferentiated hiPSCs detection in a broad range of cells derived from hiPSCs. The ddPCR assay efficiently detected as little as 0.0001% undifferentiated hiPSCs spiked in Hek293T cells, with *ESRG* as a target gene. Further, the copy number of the *ESRG* mRNA correlated with hiPSC numbers, enabling us to quantitatively establish the number of residual undifferentiated hiPSCs.

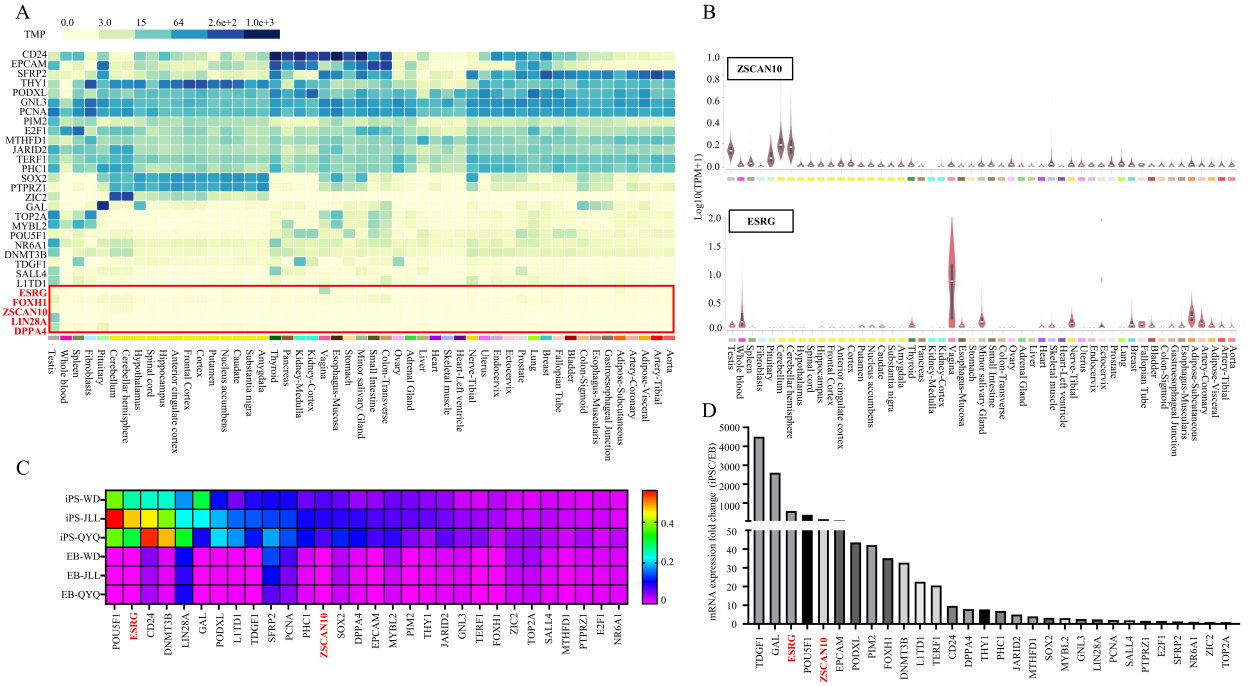

Fig. 3. Tissue expression pattern of the candidate genes. (A) Gene expression level in various tissues. Data were from the Multi-Gene Query through The Genotype-Tissue Expression (GTEx) project database. (B) *ESRG* and ZSCAN 10 expression level in various tissues. Data were from the Multi-Gene Query through The Genotype-Tissue Expression (GTEx) project database. (C) Expression heat map of 30 stem cell genes of the spontaneous differentiated EBs; Gene expression levels were sorted from high to bottom. (D) mRNA expression fold change between hiPSCs and spontaneous differentiated EBs. The genes marked in red are those whose expression fold change exceeded 20.

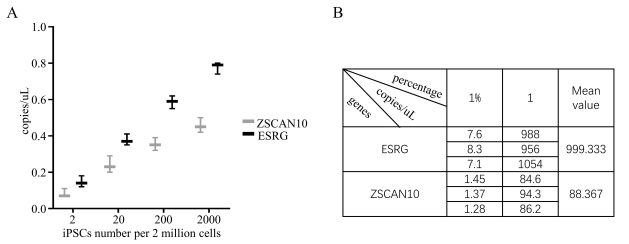

Fig. 4. Detection sensitivity via spike-in experiments. (A) mRNA copy numbers of *ESRG* and *ZSCAN10* in samples containing 2000, 200, 20 and 2 hiPSC samples in 2 million 293T cells. N = 3. (B) mRNA copy numbers of *ESRG* and *ZSCAN10* in 2 million hiPSCs samples and samples containing 20,000 hiPSCs in 2 million 293T cells. N = 3.

In previous studies, several strategies were implemented to identify residual undifferentiated hiPSCs from hiPSC-derived cell products. Most of the methods aimed to detect the expressions of pluripotent gene markers. As a result, the levels of targeted genes had to be low enough in the normal somatic cells of interest. Currently, several methods, such as flow cytometry, Elisa, immunofluorescence detection, cell culture system, qPCR and ddPCR, are being used in clinical practice. Since hiPSCs in minimal amounts could lead to teratoma, the sensitivity of the detection method should be very high. Based on the current consensus, the sensitivity should not exceed 0.001%. Considering that the sensitivity of other methods does not conform to the current consensus or cannot be quantitative, we chose the ddPCR to verify our target genes. The results showed that with the primers designed for *ESRG* and *ZSCAN10*, we could detect one hiPSCs residue in 1 million cells, demonstrating a sensitivity of 0.0001%, which is ten times more sensitive than the current consensus.

Screening marker genes to detect hiPSC residues has always been an obstacle for stem cell clinical research and drug development. The most used marker in research is Lin28A, which is used for residual detection in hiPSC-derived RPE cells. So far, many hiPSC-derived cell products have been developed, many of which, such as cardiomyocytes, pancreatic islet cells, NK cells and neural stem cells, were also used in clinical research [19]. However, there are no detailed reports about the detection methods, marker genes and hiPSCs residues detection standards. Further, there are currently no unified marker genes for different hiPSC-derived cell products to detect hiPSC residues. Thus, in this study, we aimed to identify potentially suitable genes that could be used to detect hiPSC residues in all cell types derived from hiPSCs.

5 Conclusions

The universal hiPSCs residue detection markers' criteria are high expression in hiPSCs, with low or no expression in somatic cells. To assess whether a gene could meet these criteria, the most intuitive method would be to detect its expression level in all somatic cells, which is quite impractical. Therefore, we used The Genotype-Tissue Expression (GTEx) project database to query the expression of target genes in as many tissues as possible and screened for genes that were minimally expressed in other tissues but highly expressed in hiPSCs. One way to test the stemness of hiPSCs in *in vitro* settings is through spontaneous EB differentiation, which contains cells of three germ layers and is a cell cluster with a large number of cell types. Therefore, we tested the gene expression of Spontaneously differentiated EB using NGS to validate the identified target genes. Taken together, based on the methodologies and findings of this study, the proposed target genes, *ESRG* and *ZSCAN10*, demonstrated promising clinical reliability as universal markers for hiPSCs residue detection.

Institutional Review Board Statement

None.

Data Availability Statement

Datasets used could be made provided upon reasonable request to the corresponding author.

Author Contributions

LW and YG conceived and designed the experiments. HS and TF performed all qPCR and ddPCR experiments. HS, TF, and RW differentiated the hiPSCs to neural stem cells, hepatocytes and cardiomyocytes. LW did all the NGS data analysis. LW—literature assessment and primary manuscript writing. YG—manuscript revision.

Ethics Approval and Consent to Participate

Peripheral blood mononuclear cells (PBMCs) for hiPSC induction were isolated from adult peripheral blood that with informed written consent and approval by the Institute of Ethics Committee Review Board in Beijing Luhe Hospital (ethical approval no. 2021-LHYW-043-02).

Acknowledgment

Not applicable.

Funding

This work is supported by grants held by YG (973 Project No. 2013CB531206).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/j.fbl2708239.

References

[1] Shi Y, Inoue H, Wu JC, Yamanaka S. Induced pluripotent stem cell technology: a decade of progress. Nature Reviews Drug Discovery. 2017; 16: 115–130.

[2] Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, *et al*. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell. 2007; 131: 861–872.

[3] Taylor C, Peacock S, Chaudhry A, Bradley JA, Bolton E. Generating an iPSC Bank for HLA-Matched Tissue Transplantation Based on Known Donor and Recipient HLA Types. Cell Stem Cell. 2012; 11: 147–152.

[4] Schweitzer JS, Song B, Herrington TM, Park T, Lee N, Ko S, *et al*. Personalized iPSC-Derived Dopamine Progenitor Cells for Parkinson's Disease. New England Journal of Medicine. 2020; 382: 1926–1932.

[5] Selvaraj S, Kyba M, Perlingeiro RCR. Pluripotent Stem Cell-Based Therapeutics for Muscular Dystrophies. Trends in Molecular Medicine. 2019; 25: 803–816.

[6] Trounson A, McDonald C. Stem Cell Therapies in Clinical Trials: Progress and Challenges. Cell Stem Cell. 2015; 17: 11–22.

[7] Andrews PW, Ben-David U, Benvenisty N, Coffey P, Eggan K, Knowles BB, *et al*. Assessing the Safety of Human Pluripotent Stem Cells and their Derivatives for Clinical Applications. Stem Cell Reports. 2017; 9: 1–4.

[8] Kuroda T, Yasuda S, Kusakawa S, Hirata N, Kanda Y, Suzuki K, *et al*. Highly sensitive in vitro methods for detection of residual undifferentiated cells in retinal pigment epithelial cells derived from human iPS cells. PLoS ONE. 2012; 7: e37342.

[9] Han J, Qian X, Wu Q, Jha R, Duan J, Yang Z, *et al*. Novel surface-enhanced Raman scattering-based assays for ultrasensitive detection of human pluripotent stem cells. Biomaterials. 2016; 105: 66–76.

[10] Kuroda T, Yasuda S, Sato Y. In Vitro Detection of Residual Undifferentiated Cells in Retinal Pigment Epithelial Cells Derived from Human Induced Pluripotent Stem Cells. Methods in Molecular Biology. 2014; 1210: 183–192.

[11] Sekine K, Tsuzuki S, Yasui R, Kobayashi T, Ikeda K, Hamada Y, *et al*. Robust detection of undifferentiated iPSC among differentiated cells. Scientific Reports. 2020; 10: 10293.

[12] Kuroda T, Yasuda S, Matsuyama S, Tano K, Kusakawa S, Sawa Y, *et al*. Highly sensitive droplet digital PCR method for detection of residual undifferentiated cells in cardiomyocytes derived from human pluripotent stem cells. Regenerative Therapy. 2015; 2: 17–23.

[13] Parr CJC, Katayama S, Miki K, Kuang Y, Yoshida Y, Morizane A, *et al*. MicroRNA-302 switch to identify and eliminate undifferentiated human pluripotent stem cells. Scientific Reports. 2016; 6: 32532.

[14] Tano K, Yasuda S, Kuroda T, Saito H, Umezawa A, Sato Y. A novel in vitro method for detecting undifferentiated human pluripotent stem cells as impurities in cell therapy products using a highly efficient culture system. PLoS ONE. 2014; 9: e110496.

[15] Artyuhov AS, Dashinimaev EB, Mescheryakova NV, Ashikhmina AA, Vorotelyak EA, Vasiliev AV. Detection of small numbers of iPSCs in different heterogeneous cell mixtures with highly sensitive droplet digital PCR. Molecular Biology Reports. 2019; 46: 6675–6683.

[16] Cao N, Liang H, Huang J, Wang J, Chen Y, Chen Z, *et al*. Highly

efficient induction and long-term maintenance of multipotent cardiovascular progenitors from human pluripotent stem cells under defined conditions. Cell Research. 2013; 23: 1119–1132.
[17] Wang Y, Alhaque S, Cameron K, Meseguer-Ripolles J, Lucendo-Villarin B, Rashidi H, *et al*. Defined and Scalable Generation of Hepatocyte-like Cells from Human Pluripotent Stem Cells. Journal of Visualized Experiments. 2017; 55355.
[18] Maguire CT, Demarest BL, Hill JT, Palmer JD, Brothman AR, Yost HJ, *et al*. Genome-wide analysis reveals the unique stem cell identity of human amniocytes. PLoS ONE. 2013; 8: e53372.
[19] Yamanaka S. Pluripotent Stem Cell-Based Cell Therapy—Promise and Challenges. Cell Stem Cell. 2020; 27: 523–531.