



HAL
open science

Transcriptome profiling of hiPSCs-derived LSECs with nanoCAGE

Mathieu Danoy, Stéphane Poulain, Yuta Kouji, Yannick Tauran, Benedikt Scheidecker, Taketomo Kido, Atsushi Miyajima, Yasuyuki Sakai, Charles Plessy, Eric Leclerc

► **To cite this version:**

Mathieu Danoy, Stéphane Poulain, Yuta Kouji, Yannick Tauran, Benedikt Scheidecker, et al.. Transcriptome profiling of hiPSCs-derived LSECs with nanoCAGE. *Molecular Omics*, 2020, 10.1039/c9mo00135b . hal-02438467

HAL Id: hal-02438467

<https://hal.science/hal-02438467>

Submitted on 5 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Morpho

Transcriptome profiling of hiPSCs-derived LSECs with nanoCAGE

Mathieu Danoy¹⁺, Stéphane Poulain²⁺, Yuta Kou³, **Yannick Tauran^{1,4}**,
Benedikt Scheidecker⁵, Taketomo Kido³, Atsushi Miyajima³, Yasuyuki
Sakai⁵, Charles Plessy², Eric Leclerc^{1*}

¹ CNRS UMI 2820; Laboratory for Integrated Micro Mechatronic Systems, Institute of Industrial Science, University of Tokyo; 4-6-1 Komaba; Meguro-ku; Tokyo, 153-8505, Japan.

² RIKEN Center for Integrative Medical Sciences, Division of Genomic Medicine, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan.

³ Laboratory of Stem Cell Therapy, Institute for Quantitative Biosciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo, 113-0032, Japan

⁴ Univ. Lyon, Université Claude Bernard Lyon 1, Laboratoire des Multimatériaux et Interfaces, UMR CNRS 5615, F-69622 Villeurbanne, France

⁵ Department of Chemical System Engineering, graduate school of Engineering, the University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan

+ authors with equivalent contributions

* corresponding author: Eric Leclerc

CNRS UMI 2820; Laboratory for Integrated Micro Mechatronic Systems,
Institute of Industrial Science, University of Tokyo; 4-6-1 Komaba; Meguro-ku;
Tokyo, 153-8505, Japan

Abstract

Liver Sinusoidal Endothelial Cells (LSECs) are an important component of the liver as they compose the microvasculature which allows the supply of oxygen, blood, and nutrients. However, maintenance of those cells *in-vitro* remains challenging as they tend to rapidly lose some of their characteristics such as fenestration or as their immortalized counterparts present poor characteristics. In this work, human induced pluripotent stem cells (hiPSCs) have been differentiated toward an LSECs phenotype. After differentiation, the RNA quantification allowed to demonstrate high expression of specific vascular markers (CD31, CD144, and STAB2). Immunostaining performed on the cells were found to be positive for both Stabilin-1 and Stabilin-2. Whole transcriptome analysis performed with the nanoCAGE method further confirmed the overall vascular commitment of the cells. The gene expression profile revealed the upregulation of the APLN, LYVE1, VWF, ESAM and ANGPT2 genes while the VEGFA appeared downregulated. Analysis of promoter motif activities highlighted several transcription factors (TFs) of interest in LSECs (IRF2, ERG, MEIS2, SPI1, IRF7, WRNIP1, HIC2, NFIX_NFIB, BATF, and PATZ1). Based on this investigation, we compiled the regulatory network involving the relevant TFs, their target genes as well as their related signaling pathways. The proposed hiPSCs-derived LSECs model and its regulatory network were then confirmed by comparing the experimental data

to primary human LSECs reference datasets. Thus, the presented model appears as a promising tool to generate more complex *in-vitro* liver multi-cellular tissues.

Keywords: induced pluripotent stem cells, iPSCs, liver sinusoidal endothelial cells, LSECs, differentiation, transcriptomic, nanoCAGE

1. Introduction

The liver is composed of three networks (lymphatic, biliary and, vascular) with specific functions. The lymphatic network, composed of lymphatic cells, white blood cells, and lymphocytes, serves as a route for part of the filtered plasma and is central in the immune system. The biliary network, composed of biliary cells, is responsible for the retrieval of the bile produced by hepatocytes. Finally, the microvasculature is a more complex structure composed of several cell types and is responsible for the transport of blood, for the transfer of substrates to the liver parenchyma and for the regulation of the lipoprotein traffic ^[1]. In this structure, the crosstalk between hepatocytes, forming most of the liver mass, the endothelial cells, forming the vascular walls, and the pericytes (or hepatic stellate cells, HSCs), acting as a scaffold for the latter, has been extensively studied ^[1]. In details, the coculture of hepatocytes and endothelial cells has been shown to be of great importance to the maintenance of the hepatocyte's functionality *in-vitro* ^[2]. In those cultures, cell/cell contact, as well as secretion factors, have been shown to be central to the enhancement of the functionality and to the stability of the coculture.

Still, the lack of complex *in-vitro* models including mature human hepatocytes remains one of the major bottlenecks in the pharmaceutical industry to provide

relatively low-cost toxicological data and predictive information on the effects of xenobiotics ^[3]. In that regard, promising alternatives such as the differentiation and maturation of hiPSCs have been developed. As such, several investigations have led to the differentiation of hepatocyte-like cells (HLCs) from hiPSCs ^[4-6]. In parallel, hiPSCs have also been used to provide a reliable source for other hepatic cells such as HSCs ^[7] and LSECs ^[8].

Especially, the produced LSECs have been shown to be positive to LSECs-specific markers such as Stabilin-2. In this work, we further extend the characterization of these cells by performing whole transcriptome analysis with nanoCAGE ^[9,10]. The analysis allowed to provide additional information regarding the gene expression patterns and promoter activities involved during the differentiation process of LSECs. In addition, the analysis was validated by comparing the gene expression profiles of the LSECs derived from hiPSCs with reference FANTOM5 CAGE datasets obtained from primary human LSECs samples and with various human endothelial cell models ^[11,12].

2. Materials and methods

2.1 iPS culture and differentiation protocol

The hiPSCs used in this study (454E2) were provided by the cell bank of the Riken BioResource Research Center (RIKEN BRC). 454E2 hiPSCs were cultured and differentiated following the previously published protocols for LSECs differentiation ^[8]. After differentiation, the cells were harvested and stored in liquid nitrogen for storage. When required, the needed number of cells was seeded on

plates (1.5×10^4 cells/cm²) coated with 20ug/mL fibronectin (Life Technologies) for 1 hour at 37C, for further culture and proliferation for 9 days. Experiments have been performed 3 independent times and 3 samples for each analysis were collected for each independent experiment.

2.2 Primary culture for morphological comparison

Hepatic Sinusoidal Endothelial Cells were purchased from ScienCell Research Laboratories and cultured as recommended by the provider for morphological comparison.

2.3 CAGE transcriptome profiling

NanoCAGE sequencing libraries were prepared following the previously described protocol ^[10]. 4 undifferentiated hiPSCs and 4 hiPSCs-derived LSECs samples issued from 4 independent experiments were used for this analysis. Total RNA extraction and library preparation steps were performed as previously detailed ^[13]. The pool of nanoCAGE samples was sequenced paired-end (9.13 pM + 10% PhiX) on MiSeq system with the Reagent Kit v3 150-cycles (Illumina).

2.4 CAGE data processing

NanoCAGE sequencing data were processed for CAGEScan analysis as previously detailed ^[14]. Gene expression matrices were created with custom R scripts using the CAGEr package ^[15], then uploaded on the iDEP server (version 8.1) ^[16]

(<http://ge-lab.org:3838/idep/>) for differential gene expression (DGE) and Pathway Gene Set Enrichment Analysis (PGSEA). In iDEP the EdgeR:log2(CPM=c) method was used to transform the read count for clustering. The DGE analysis was conducted with DESeq2 ^[17], applying a false discovery rate (FDR) cutoff of 0,1 (corresponding to an adjusted p_value of 0,1) and a setting the minimum fold change to 5. The PGSEA was performed with a pathway significant cutoff set at 0,2. The ISMARA webserver ^[18] (<https://ismara.unibas.ch/mara/>) was used for Motif Activity Response Analysis (MARA) and the Ingenuity Pathway Analysis software (Qiagen) was used for pathway analysis. Datasets were deposited to Zenodo (FASTQ files DOI: 10.5281/zenodo.2572390 and MOIRAI output DOI: 10.5281/zenodo.2572394).

2.5 RT-qPCR

RT-qPCR was performed following the protocol previously described ^[14]. Primer sequences of STAB2, CD31, VEGF, and CD144 are shown in **Supp. File 1**. The expression of each gene was normalized with the reference gene ACTB (β -Actin) and then plotted in folds of the expression found in undifferentiated hiPSCs.

2.6 Immunostaining assays

Samples were stained following the protocol previously described ^[14]. The antibodies used were: CD144 (rabbit, LS-B2138, LifeSpan Biosciences), Stabilin-2 (rabbit, ab121893, abcam), Stabilin-1 (mouse, H00023166-M05, Abnova), **CD31-PE (mouse, 555446, BD Bioscience)**, anti-rabbit Alexa Fluor 568 (donkey, A10042,

ThermoFisher), anti-mouse Alexa Fluor 647 (donkey, ab150107, abcam) and anti-mouse (goat, A21424, Life Technologies). All fluorescent images were taken using a confocal microscope (PowerIX70, Olympus).

2.7 Western Blots

Isolation and solubilization of proteins after TRIzol™ extraction of RNA from cells has been based on a previously published protocol ^[19]. Shortly, as cells were lysed with TRIzol™ Reagent, after extracting RNA from the aqueous phase with chloroform (200 µL / mL of Trizol™), the proteins from the phenol red phase were precipitated by isopropanol. After washing two times with ethanol, the protein pellet was dried and then resuspended in a solubilization solution (100 mM Tris pH 8, 5 % SDS, 140 mM NaCl, 20 mM EDTA supplemented with phosphatase and protease inhibitors). Total proteins were then finally solubilized after incubation at 50°C for 10 hours.

The protein concentration was then determined (BCA protein assay Kit, Pierce®, Thermo scientific, IL, USA), identical protein amount of cell extracts was separated by NuPAGE (4-12% precast gels) (Life technologies) and electrotransferred to polyvinylidene difluoride (PVDF) membranes (Merck Millipore). Membranes were probed with either HRP-conjugated primary antibodies or primary antibodies, followed by an incubation with the corresponding HRP secondary antibodies as indicated below. Protein-antibody complexes were then visualized by chemiluminescence with the SuperSignal® West Dura Extended Duration Substrate (Thermo scientific), using a LAS-3000 imaging system (Fujifilm). The following antibodies were used for the experiment: Stabilin-2 (rabbit, PA5-76192,

ThermoFisher) combined with anti-rabbit HRP-conjugated antibody (donkey, 711-035-152, Jackson ImmunoResearch Laboratories), CD144 HRP-conjugated (rabbit, SPC-1290D-HRP, Funakoshi) and Beta-Actin HRP conjugated (mouse, ab20272, abcam).

3. Results

3.1 Morphology and **expression** for typical LSECs markers

After differentiation and maturation, the hiPSCs-derived LSECs displayed a typical LSECs behavior. Upon seeding, the cells exhibited an elongated shape (Fig. 1A). After 3 days of culture, the cells reached confluence (Fig. 1B) and started to proliferate while displaying a more elongated profile after 6 (Fig. 1C) and 9 days of culture (Fig. 1D). After 9 days of culture, the LSECs exhibited a morphology very distinct from undifferentiated hiPSCs (Fig. 1E) and comparable to the one observed in primary human LSECs (Fig. 1F). Immunostaining assays confirmed that after 9 days of culture, the LSECs were still expressing CD144, Stabilin-1, CD31 and more interestingly, Stabilin-2, confirming the liver specificity of the differentiated endothelial cells (Fig. 2A-G). In addition, Western Blots have confirmed the expression of Stabilin-2 and CD144 (Fig. 2H).

3.2 Gene expression of the hiPSCs-derived LSECs

RT-qPCR performed on key control genes (Fig. 3) allowed to confirm the maintenance of the LSECs phenotype during proliferation. The mRNA levels of CD31, CD144 and STAB2 were all found upregulated in LSECs derived cells when

compared to undifferentiated iPSC. In addition, VEGF mRNA levels appeared downregulated in the LSECs cultures.

3.3 Transcriptomic analysis of the hiPSCs-derived LSECs

The methodology of the present work in terms of transcriptomic has been illustrated in Supp. File 1. Firstly, we used the transcriptomic data obtained from nanoCAGE sequencing of biological quadruplicates representative of four independent experiments to compare the gene expression profiles of hiPSCs-derived LSECs (4 samples) and undifferentiated hiPSCs (4 samples plus 3 technical replicates). Therefore, expression tables were created and analyzed online with iDEP (version 8.1). The DGE analysis conducted with DESeq2 ^[17], applying a false discovery rate cutoff of 0.1 and a minimum fold change of 5, contributed to clearly separate both culture conditions and led to extract respectively a list of 1482 genes up-regulated in LSECs and of 1921 genes over-expressed in hiPSCs. The complete list of differentially expressed genes is provided in **Supp. File 2** while the corresponding heatmap is presented in **Fig. 4A**. Among the top 400 genes discriminating the two cell types, the LSECs exhibited an upregulation of typical endothelial genes such as APLN (Apelin), LYVE1, PECAM1, VWF (Von Willebrand Factor), CDH5 (Endothelial Cadherin), ESAM, ESM1, STAB1, ANGPT2 (Angiopoietin 2), ICAM1, ICAM2, PEAR1, as well as an upregulation of TFs such as Wnt2B, ERG, MEIS2, EPAS1, PPARG, PLAGL1 and MAF. In addition, other genes such as WLS, FZD4, BMP4, EHD2, EHD3, MECOM and NFATC2 were also found to be upregulated in hiPSCs-derived LSECs while downregulation of the VEGF-A, the HMMR, and the VLDLR genes was observed.

Gene Ontology (GO) analysis was performed with the list of differentially expressed genes using several pathway databases. In the GO_biological process, the top pathways over-represented in hiPSCs-derived LSECs were found to be characterized by annotations such as “Angiogenesis”, “Circulatory system development”, “Blood vessel morphogenesis and development”, and “Tube development”. In parallel, hiPSCs were defined by annotations such as “Cell development” and “Cell morphogenesis”. In the GO_cellular component, hiPSCs were characterized by “Ribosomes” and “Chromosomes” while the hiPSCs-derived LSECs were defined by “Extracellular region” and “Cell membrane”. The complete list of the different GO_biological processes, GO_cellular components, GO_molecular functions and KEGG metabolic pathways annotations extracted from the analysis are given in [Supp. File 3](#).

Further, PGSEA was performed with a false discovery rate of 0.2 using the GO and KEGG reference gene sets. In hiPSCs-derived LSECs, “Angiogenesis”, “Cardio-vascular system development”, and “Blood vessel development” were found to be among the top GO_biological process pathways underlined by the analysis. Analysis in the same conditions with the KEGG metabolic pathways dataset revealed terms such as “Fluid shear stress and atherosclerosis” and “NFKB signaling” for hiPSCs-derived LSECs while hiPSCs were characterized by “Cell cycle” and “DNA replication”. In the analysis performed with the GO_molecular function dataset, the hiPSCs-derived LSECs were defined by “MAPK binding” and “TGFB binding” while the hiPSCs were characterized by “RNA and DNA processing”. Heatmaps showing the top 30 GO_biological process and GO_molecular function annotations discriminating the two cell types are shown in [Fig. 4](#).

Finally, the genes differentially expressed between the hiPSCs and the hiPSCs-derived LSECs conditions were examined with the Ingenuity Pathway Analysis (IPA) software. Among the top 5 canonical pathways extracted by IPA, “Hepatic fibrosis and stellate cells activation” (p_value = 0,0005) and “Actin cytoskeleton reorganization” (p_value = 0,002) were highlighted. Potential upstream regulators such as COL18A1, ERG, IFNA2, TGFB1, and CLOCK were also suggested. In addition, COL18A1 and VEGF-A were found to be the main gene network regulators displayed and the main physiological system and development pathway highlighted was the “Cardio-vascular system”. A complete summary of the analysis is given in [Supp. File 4](#).

3.4 Motif activity response analysis of the hiPSCs-derived LSECs

An analysis of nanoCAGE sequencing data using ISMARA was performed to complete the investigation of the potential TFs involved in the differentiation of hiPSCs-derived LSECs. The analysis led to the extraction of IRF2, ERG, MEIS2, SPI1, IRF7, WRNIP1, HIC2, NFIX_NFIB, BATF and PATZ1 as the top 10 TF motifs with high activity in hiPSCs-derived LSECs when compared to hiPSCs ([Table 1](#)). In addition, MYCN, CEBPZ, ETV5, E2F7, ZEB1, MTA3, TAF1, KLF16_SP2, PBX3, and E2F5 were highlighted as the top 10 TF motifs with high activity in hiPSCs when compared to hiPSCs-derived LSECs ([Table 1](#)).

Based on the information provided by ISMARA, an interactive regulatory network involving the main TFs with an upregulated motif activity in hiPSCs-derived LSECs was built ([Fig. 5](#)). The described motifs were found to be related to several typical endothelial targets such as STAB1, LYVE1, EPAS1, ESAM, CDH5, VWF or

APLN. Those targets were also found to be differentially expressed in both cell types. ISMARA allowed relating the TFs HIC2 and WRNIP1 to their involvement in the “Endothelial differentiation” GO_biological process. The analysis allowed to extract TFs such as ARTF6 and NFKB as well as specific metabolic pathways (ERK/MAPK, Interferon gamma) or biological processes (Extra Cellular Matrix (ECM)) in KEGG and in the GO_biological process. Details on the top 10 up and down-regulated TFs are given in [Supp. File 5](#) and [Supp. File 6](#).

3.5 Comparison with FANTOM5 primary LSECs data

To confirm the validity of the model, the analysis previously performed was repeated on hiPSCs and primary human LSECs samples from publicly available CAGE data provided by the FANTOM5 consortium project (<http://fantom.gsc.riken.jp/5/>). The ISMARA analysis of these samples revealed that ERG, SPI1, IRF1, PPARG, REL, KLF3, KLF4, NFATC3, ZNF711 and ELF2 were the top 10 TFs with an upregulated motif activity in primary human LSECs when compared to hiPSCs from the FANTOM5 database ([Table 2](#)). Among those TFs, the motif activity of SPI1, IRF2_IRF1, PATZ1_KLF4, and ERG appeared consistent with the results obtained in the analysis performed on hiPSCs-derived LSECs ([Fig. 6](#)). In addition, REL, KLF3, NFATC2_NFATC3, ZNF711, PPARG, and ELF2 were also found to be overactivated in hiPSCs-derived LSECs but with lower z_values.

Furthermore, TFC4, MAZ, ZEB1, RFX3, REST, PBX3, MTA3, SP2, ETV1, and AHR were highlighted as the top 10 TFs in hiPSCs from the FANTOM5 database. Among those targets, ZEB1, PBX3, KL16_SP2, and MTA3 were found to follow a similar tendency in our previous analysis with hiPSCs and hiPSCs-derived LSECs

(Fig. 6). The MAZ, ETV1, AHR and REST TFs also presented a motif overactivity in our comparison between hiPSCs and hiPSCs-derived LSECs but with lower z_values than those observed with the FANTOM5 database's samples.

Eventually, as a validation of the hiPSCs-derived LSECs model, a comparison with the primary LSECs from the FANTOM5 database was made using ISMARA. TFs extracted from the analysis are presented in Supp. File 7. The results obtained confirmed the common patterns for TFs such as ERG, RELA, SPI1, IRF2_IRF1. However, different profiles regarding CREB1, BHLHE40 and PITX1 were observed as those TFs were found to be over-expressed in the primary LSECs from the FANTOM5 database.

4. Discussion

In this work, the transcriptomics analysis of hiPSCs-derived LSECs obtained following the previously described method ^[8] was performed. Consistently with their previous findings, the expression of Stabilin-2 in the cells was confirmed. Expressions of Stabilin-1 and CD144 was also confirmed at both the gene and protein levels. In the literature, the differentiation process of LSECs was related to genes such as Wnt2, FDZ4, FDZ5, FDZ9, WLS, VEGFR and NRP2, to several TFs (GATA4, LMO3, MAF and TCFEC) and to endocytosis-related molecules (LYVE1, EHD3, STAB1, STAB2) ^[20]. Consistently, our hiPSCs-derived LSECs presented an upregulation of those markers when compared to hiPSCs. In addition, the presence of several typical endothelial markers such as ASM and APLN and TFs such as EPAS1 was confirmed, assessing of the quality of the model of hiPSCs-derived LSECs.

With the help of the data generated from ISMARA processing of the nanoCAGE transcriptomic analysis, a regulatory network inferred from the direct comparison of hiPSCs and hiPSCs-derived LSECs could be proposed. This network includes some critical TFs related to the endothelial differentiation and to the vascular development. Among these, ERG was reported to be involved in the stability of the vascular tissue ^[21] and HIC2 to be involved in the maturation of the vascular system ^[22]. The presented network also included PATZ1 which was reported to prevent endothelial cells from diabetic damages ^[23] and senescence ^[24]. MEIS2 was reported to be involved in the hemogenic endothelial specification ^[25] as well as to be a potential key marker of LSECs (together with C-MAF and GATA4 ^[26]). In the analysis performed in the presented work, MEIS2 could be connected to the VEGF/VEGFR signaling and to the hyaluronan reactome. This is consistent with STAB2 being reported to be a hyaluronic acid receptor through the NFKB regulation ^[27], and NFKB interacting with SPI1, which linked it to MEIS2 in the network presented in **Fig. 5**. In addition, WRNIP1, also part of the network, was possibly linked to the VEGF/VEGFR signaling, to the hyaluronan reactome, to the “Endothelial differentiation” GO_biological process and to several endothelial target genes such as EPAS1 and ESAM. To validate the results obtained, a similar motif activity response analysis using CAGE transcriptomics data obtained from hiPSCs and human primary LSECs samples, made publicly available by the FANTOM5 consortium was performed. In both cases (hiPSCs vs human primary LSECs from the FANTOM5 consortium and hiPSCs vs hiPSCs-derived LSECs), the analysis led to close TF expression patterns and motif activity profiles. For instance, ERG, SPI1, IRF1_IRF2, PATZ1_KLF4 were extracted as highly active TF binding motifs in both the hiPSCs-derived LSECs and

primary human LSECs. Those results emphasize the relative proximity of the hiPSCs-derived LSECs to their primary counterparts.

Interestingly, NFIB was extracted in our comparison between hiPSCs and hiPSCs-derived LSECs. NFIB, while being reported as an important factor in brain and lung development ^[28], has also been shown to be a regulator of the VWF expression ^[29]. Using bioinformatics tools in our analysis, several genes (LYVE1, VWF) and TFs such as NRP2 and MAF ^[20,26] could be linked to this gene. A link with NOTCH4 could also be established, this gene being reported to be essential against the malformation of the hepatic vascularization ^[30] and for the protection against liver metastasis ^[31], opening the application of the presented cells to different *in-vitro* models.

Stabilin-2 has been shown to be a crucial LSECs marker ^[32] and its involvement in the regulation of VWF has been demonstrated ^[33]. Additionally, STAB2 has been shown to regulate by calcineurin/NFAT signaling in myoblasts ^[34]. The STAB2 gene is located on chromosome 12 and nanoCAGE data has revealed that RCOR1, MTA3, FOXP3, HMX3, EMX2, ESR2, ZNF652, ZNF384, HOXC8, TGFI1, TGFI2, POU3F2, PKNOX1, MAFB, PAX5, ZEB1, MAF, and MEIS2 are among the proteins that could bind to the STAB2 gene promoter and could be involved in its regulation. When comparing hiPSCs with hiPSCs-derived LSECs, PAX5, ZEB1, and MEIS2 were extracted by ISMARA and iDEP post-processing. Among those genes, ZNF652 was found to be related to STAB2 in both the comparison between hiPSCs and hiPSCs-derived LSECs and in the comparison between hiPSCs and primary human LSECs from the FANTOM5 database. When comparing hiPSCs-derived LSECs to hiPSCs, the top 3 targets extracted using ISMARA for ZNF652 genes, were TM4SF18, APLN (apelin) and ESM1 (endothelial

cell-specific molecule 1). Other targets included the like of AGGF1 (angiogenic factor with G patch and FHA domains 1). In addition, ZNF652, *via* the FLI1 TF, is linked to HIC2, WRNIP1, and SPI1 that were among the top regulated TFs in the analysis performed. Finally, ZNF652, *via* the FLI1 TF, is linked to HIC2, WRNIP1 and SPI1 TFs that were among the top regulated TFs in the same analysis. FLI1 was notably linked to the NFATC2_NFATC3 TFs in ISMARA (NFAT calcineurin-dependent signaling) and has been shown to be an important gene in the maintenance of endothelial cells differentiation ^[35]. Those results further emphasize the central role of Stabilin-2 and its different linked genes in the expression of and in the differentiation toward the LSECs phenotype. **Further analysis of the hiPSCs-derived LSECs including important phenomenon such as fenestration or a direct comparison with primary human LSECs will provide further insights on the presented work and should still be worked on in the near future as the present dataset did not allow such comparison.**

Conclusion

In this work, an in-depth analysis of the differentiation of hiPSCs toward an LSECs phenotype was performed. The different assays allowed to assess the phenotype of the LSECs after differentiation and maturation. In addition, mRNA sequencing based on the nanoCAGE technology allowed to characterize in detail the cells obtained. Typical endothelial TFs such as ERG were identified, and the subsequent targets and pathways were studied. Finally, by comparing the obtained data to primary human samples, we were able to propose a regulatory network describing the gene interactions observed in hiPSCs-derived LSECs. Those results

further improve our understanding of the model and will be of use for its further integration in multi-cellular models for efficient *in-vitro* modeling of the liver using the hiPSCs technology.

Figure Legend

Figure 1: Cell morphologies of hiPSCs-derived LSECs observed after 1 day of culture (A), after 3 days of culture (B), after 6 days of culture (C) and after 9 days of culture (D). Cell morphologies of undifferentiated hiPSCs (E) and primary human LSECs (F) for comparison. Scale bar = 200 μm .

Figure 2: Immunostainings after 9 days of culture. Staining for DAPI (blue, A) with corresponding immunostainings for Stabilin-2 (red, B), and Stabilin-1 (magenta, C). Staining for DAPI (blue, D) with corresponding immunostaining for CD144 (green, E). Staining for DAPI (blue, F) with corresponding immunostaining for CD31 (red, G). Total protein extracts in duplicate (1 and 2) from hiPSCs and hiPSCs-derived LSECs analyzed by Western Blots for Stabilin-2, CD-144 and Beta-Actin (H). Full images of Western Blots are given in Supp. File 1. Scale bar = 200 μm .

Figure 3: Gene expression levels (Logarithmic scale) of STAB2, CD31, VEGF, and CD144 measured by RT-qPCR after 9 days of culture.

Figure 4: Heat map clustering of the hiPSCs vs hiPSCs-derived LSECs comparison (A). GO_biological process (B) and GO_molecular function (C) annotations discriminated in the comparison between hiPSCs and hiPSCs-derived LSECs.

Figure 5: Interaction network connecting the different TF motifs highlighted by MARA analysis and their target pathways.

Figure 6: Motifs and comparison of the Z value of the motif activity of the TFs identified by ISMARA in the comparison between hiPSCs and hiPSCs-derived LSECs (PS) and in the comparison between hiPSCs and human primary LSECs from the FANTOM5 database (FT). TFs with a higher motif activity in LSECs are presented in the top while TFs with a higher motif activity in hiPSCs are presented in the bottom of the graph.

Reference

- [1] Sørensen, K. K., Simon-Santamaria, J., McCuskey, R. S., & Smedsrød, B.. Liver sinusoidal endothelial cells. *Comp Physiol.* 2011; 5(4), 1751-1774.
- [2] Bale, S. S., Golberg, I., Jindal, R., McCarty, W. J., Luitje, M., Hegde, M., et al. Long-term coculture strategies for primary hepatocytes and liver sinusoidal endothelial cells. *Tissue Eng Pt C-Meth.* 2014; 21(4), 413-422.
- [3] Chen, M., Bisgin, H., Tong, L., Hong, H., Fang, H., Borlak, J., et al.. Toward predictive models for drug-induced liver injury in humans: are we there yet? *Biomark med.* 2014; 8(2), 201-213.
- [4] Touboul, T., Hannan, N. R., Corbineau, S., Martinez, A., Martinet, C., Branchereau, S., et al. Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development. *Hepatology.* 2010; 51(5), 1754-1765.
- [5] Hay, D. C., Fletcher, J., Payne, C., Terrace, J. D., Gallagher, R. C., Snoeys, J., et al. Highly efficient differentiation of hESCs to functional hepatic endoderm requires ActivinA and Wnt3a signaling. *P Natl Acad Sci Usa.* 2008;105(34), 12301-12306.
- [6] Cai, J., Zhao, Y., Liu, Y., Ye, F., Song, Z., Qin, H., et al.. Directed differentiation of human embryonic stem cells into functional hepatic cells. *Hepatology.* 2007;45(5), 1229-1239.

[7] Coll, M., Perea, L., Boon, R., Leite, S. B., Vallverdu, J., Mannaerts, I., et al.. Generation of hepatic stellate cells from human pluripotent stem cells enables in vitro modeling of liver fibrosis. *Cell stem cell*. 2018;23(1), 101-113.

[8] Kouji, Y., Kido, T., Ito, T., Oyama, H., Chen, S. W., Katou, et al.. An in vitro human liver model by iPSC-derived parenchymal and non-parenchymal cells. *Stem cell rep*. 2017;9(2), 490-498.

[9] Plessy, C., Bertin, N., Takahashi, H., Simone, R., Salimullah, M., Lassmann, T., et al.. Linking promoters to functional transcripts in small samples with nanoCAGE and CAGEscan. *Nat methods*. 2010;7(7), 528.

[10] Poulain, S., Kato, S., Arnaud, O., Morlighem, J. É., Suzuki, M., Plessy, C., et al.. NanoCAGE: a method for the analysis of coding and noncoding 5'-capped transcriptomes. In *Promoter Associated RNA* (pp. 57-109). Humana Press, New York, NY 2017.

[11] Lizio, M., Harshbarger, J., Shimoji, H., Severin, J., Kasukawa, T., Sahin, et al.. Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol*. 2015; 16(1), 22.

[12] Lizio, M., Harshbarger, J., Abugessaisa, I., Noguchi, S., Kondo, A., Severin, J., et al.. Update of the FANTOM web resource: high resolution transcriptome of diverse cell types in mammals. *Nucleic acids res*. 2016; gkw995.

- [13] Bernier, M. L., Poulain, S., Tauran, Y., Danoy, M., Shinohara, M., Kimura, K., et al.. Profiling of derived-hepatocyte progenitors from induced pluripotent stem cells using nanoCAGE promoter analysis. *Biochem eng j.* 2019;142, 7-17.
- [14] Danoy, M., Bernier, M. L., Kimura, K., Poulain, S., Kato, S., Mori, D., et al.. Optimized protocol for the hepatic differentiation of induced pluripotent stem cells in a fluidic microenvironment. *Biotechnol bioeng.* 2019.
- [15] Haberle, V., Forrest, A. R., Hayashizaki, Y., Carninci, P., & Lenhard, B.. CAGEr: precise TSS data retrieval and high-resolution promoterome mining for integrative analyses. *Nucleic acids res.* 2015;43(8), e51-e51.
- [16] Ge, S. X.. iDEP: An integrated web application for differential expression and pathway analysis. *bioRxiv.* 2017;148411.
- [17] Love, M., Anders, S., & Huber, W.. Differential analysis of count data—the DESeq2 package. *Genome Biol.* 2014;15(550), 10-1186.
- [18] Balwierz, P. J., Pachkov, M., Arnold, P., Gruber, A. J., Zavolan, M., & van Nimwegen, E.. ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs. *Genome res.* 2014;24(5), 869-884.
- [19] Kopec, A. M., Rivera, P. D., Lacagnina, M. J., Hanamsagar, R., & Bilbo, S. D. Optimized solubilization of TRIzol-precipitated protein permits Western blotting analysis to maximize data available from brain tissue." *J neurosci meth.* 2017;280: 64-76.

- [20] Géraud, C., Schledzewski, K., Demory, A., Klein, D., Kaus, M., Peyre, F., et al.. Liver sinusoidal endothelium: a microenvironment-dependent differentiation program in rat including the novel junctional protein liver endothelial differentiation-associated protein-1. *Hepatology*. 2010;52(1), 313-326.
- [21] Shah, A. V., Birdsey, G. M., Peghaire, C., Pitulescu, M. E., Dufton, N. P., Yang, Y., et al.. The endothelial transcription factor ERG mediates Angiopoietin-1-dependent control of Notch signalling and vascular stability. *Nat com*. 2017;8, 16002.
- [22] Dykes, I. M., van Bueren, K. L., & Scambler, P. J.. HIC2 regulates isoform switching during maturation of the cardiovascular system. *J mol cell cardiol*. 2018;114, 29-37.
- [23] Cui, Y. X., Hua, Y. Z., Wang, N., Chen, X., Wang, F., Liu, J. Y., ... et al.. miR-24 suppression of POZ/BTB and AT-hook-containing zinc finger protein 1 (PATZ1) protects endothelial cell from diabetic damage. *Biochem bioph res co*. 2016;480(4), 682-689.
- [24] Cho, J. H., Kim, M. J., Kim, K. J., & Kim, J. R.. POZ/BTB and AT-hook-containing zinc finger protein 1 (PATZ1) inhibits endothelial cell senescence through a p53 dependent pathway. *Cell death differ*. 2012;19(4), 703.
- [25] Wang, M., Wang, H., Wen, Y., Chen, X., Liu, X., Gao, J., et al.. MEIS2 regulates endothelial to hematopoietic transition of human embryonic stem cells by targeting TAL1. *Stem cell res ther*. 2018;9(1), 340.

[26] de Haan, W., Dheedene, W., Boon, R., Hoekman, B., Vandervoort, P., Verfaillie, C., et al.. Expression of Liver Sinusoidal Endothelial Cell Markers can be induced in Endothelial Stem Cells by Overexpression of C-MAF, GATA4 and MEIS2. In *Biology meets technology for liver toxicity testing*, Date: 2015/12/02-2015/12/03, Location: Leuven.

[27] Pandey, M. S., & Weigel, P. H.. Hyaluronic acid receptor for endocytosis (HARE)-mediated endocytosis of hyaluronan, heparin, dermatan sulfate, and acetylated low density lipoprotein (AcLDL), but not chondroitin sulfate types A, C, D, or E, activates NF- κ B-regulated gene expression. *J Biol Chem.* 2014;289(3), 1756-1767.

[28] Steele-Perkins, G., Plachez, C., Butz, K. G., Yang, G., Bachurski, C. J., Kinsman, S. L., ... et al.. The transcription factor gene Nfib is essential for both lung maturation and brain development. *Mol cell biol.* 2005;25(2), 685-698.

[29] Nassiri, M., Liu, J., Kulak, S., Uwiera, R. R., Aird, W. C., Ballermann, B. J., et al.. Repressors NFI and NFY participate in organ-specific regulation of von Willebrand factor promoter activity in transgenic mice. *Arterioscl throm vas.* 2010;30(7), 1423-1429.

[30] Cuervo, H., Nielsen, C. M., Simonetto, D. A., Ferrell, L., Shah, V. H., et al.. Endothelial notch signaling is essential to prevent hepatic vascular malformations in mice. *Hepatology.* 2016;64(4), 1302-1316.

[31] Wohlfeil, S. A., Häfele, V., Dietsch, B., Schledzewski, K., Winkler, M., Zierow, J., et al.. Hepatic endothelial Notch activation protects against liver metastasis by regulating endothelial-tumor cell adhesion independent of angiocrine signaling. *Cancer res.* 2019 ;79(3), 598-610.

[32] Poisson, J., Lemoine, S., Boulanger, C., Durand, F., Moreau, R., Valla, D., et al.. Liver sinusoidal endothelial cells: physiology and role in liver diseases. *J hepatol.* 2017;66(1), 212-227.

[33] Swystun, L. L., Lai, J. D., Notley, C., Georgescu, I., Paine, A. S., Mewburn, J., et al.. The endothelial cell receptor stabilin-2 regulates VWF-FVIII complex half-life and immunogenicity. *J clin investi.* 2018;128(9).

[34] Park, S. Y., Yun, Y., Lim, J. S., Kim, M. J., Kim, S. Y., Kim, J. E., et al.. Stabilin-2 modulates the efficiency of myoblast fusion during myogenic differentiation and muscle regeneration. *Nat com.* 2016 ;7, 10871.

[35] Nagai, N., Ohguchi, H., Nakaki, R., Matsumura, Y., Kanki, Y., Sakai, J., et al.. Downregulation of ERG and FLI1 expression in endothelial cells triggers endothelial-to-mesenchymal transition. *PLoS genet.* 2018;14(11), e1007826.