



HAL
open science

Influence of CPM-dependent sorting on the multi-omics profile of hepatocyte-like cells matured in microscale biochips

Mathieu Danoy, Stéphane Poulain, Benedikt Scheidecker, Rachid Jellali, Yannick Tauran, Marjorie Leduc, Johanna Bruce, Françoise Gilard, Bertrand Gakiere, Hiroshi Arakawa, et al.

► To cite this version:

Mathieu Danoy, Stéphane Poulain, Benedikt Scheidecker, Rachid Jellali, Yannick Tauran, et al.. Influence of CPM-dependent sorting on the multi-omics profile of hepatocyte-like cells matured in microscale biochips. *Biochemical Engineering Journal*, 2022, 181, pp.108408. 10.1016/j.bej.2022.108408 . hal-03667767

HAL Id: hal-03667767

<https://u-paris.hal.science/hal-03667767v1>

Submitted on 11 Oct 2022

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.

Influence of CPM-dependent sorting on the multi-omics profile of hepatocyte-like cells matured in microscale biochips

Danoy Mathieu^{1,2}, Poulain Stephane³, Scheidecker Benedikt^{1,2}, Jellali Rachid⁴, Tauran Yannick^{1,5}, Leduc Marjorie⁶, Bruce Johanna⁶, Gilard Françoise⁷, Gakiere Bertrand⁷, Arakawa Hiroshi⁸, Kato Yukio⁸, Kim Soo Hyeon³, Kido Taketomo⁹, Miyajima Atsushi⁹, Sakai Yasuyuki², Leclerc Eric¹

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

³ Biomedical Microsystems Laboratory, Institute of Industrial Science, University of Tokyo; 4-6-1 Komaba; Meguro-ku; Tokyo, 153-8505, Japan.

⁴ Université de Technologie de Compiègne, CNRS, Biomechanics Bioengineering, Centre de recherche Royallieu-CS 60319, 60203 Compiègne Cedex, France

⁵ Univ Lyon, Université Claude Bernard Lyon 1, Laboratoire des Multimatériaux et Interfaces, UMR CNRS 5615, F-69622 Villeurbanne, France

⁶ Plateforme protéomique 3P5, Université de Paris, Institut Cochin, INSERM, CNRS, F-75014, France

⁷ Plateforme Métabolisme Métabolome, Institute of Plant Sciences Paris-Saclay (IPS2), CNRS, INRA, Univ. Paris-Sud, Univ. Evry, Univ. Paris-Diderot, Univ. Paris Saclay, Bâtiment 630 Rue Noetzlin, 91192 Gif-sur-Yvette cedex, France

⁸ Laboratory of Molecular Pharmacokinetics, Faculty of Pharmacy, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Kakuma-machi, Kanazawa City, Ishikawa, 920-1192, Japan.

⁹ Laboratory of Stem Cell Therapy, Institute for Quantitative Biosciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan.

* 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 modeling in disease via advanced *in vitro* physiological tissues remains a challenging issue, yet crucial for the future development of relevant tools for drug screening. In that regard, advanced technics such as human induced Pluripotent Stem Cells (hiPSCs) and organ-on-chip are promising technologies for the relevant reproduction of the *in vivo* micro physiology in *in vitro* culture conditions. In the present work, the maturation of carboxypeptidase M positive (CPM+) Hepatocyte-Like-Cells (HLCs) in microfluidic culture conditions was investigated. hiPSCs were differentiated in culture dishes until sorting and further amplified until seeding in biochips. After 2 weeks of maturation in biochips, evaluation of the cell metabolism was performed, and samples were collected for characterization via RNA sequencing, proteomics and metabolomics. The omics profile of biochips loaded with CPM+ HLCs was then compared with the one of previously cultured biochips loaded with unsorted HLCs. CPM+ HLCs presented advanced liver characteristics in terms of a higher activity of important Transcriptions Factors (TFs) such as HNF1, HNF4A, and CEBP/A, of the upregulation of steroids/corticoids-related nuclear receptors (FXR, NR3C2, NR4A1 genes), of phase I, and phase II metabolism genes (CYP1A1, CYP1B1, CYP2C18, CYP27A1, several UGT, SULT and GST genes). Further differences in the cellular reorganization, in the lipids and steroids metabolisms, in the OXPHOS respiration and in the expression of TGF β signaling were also observed. The present study introduces a novel protocol using advances hiPSCs differentiation and sorting methods as well as the organ-on-chip technology to highlight important *in vitro* liver regeneration and hepatic maturation processes.

Keywords: hepatocyte like cells, induced pluripotent stem cells, carboxypeptidase M, organ on chip, liver, nanocage, transcriptomics, proteomics, metabolomics

1. Introduction

The liver plays a central role in the metabolism of xenobiotics and in their detoxification while also being involved in regulation of lipids, proteins, and amino acids, in glucose homeostasis, and in bile synthesis (Kalra et al., 2018). In that regard, efficient *in vitro* models are highly sought for by industrials for drug testing purposes and by clinicians for the understanding of new mechanisms. Especially, microfluidics devices and organ on chip technologies are expected to be able to enhance the precision of prediction of drug safety and efficacy during preclinical tests by reproducing accurately the physiological environment of the targeted organ (Low et al., 2021). By doing so, the development of therapeutical solutions would not only be more accurate but also more cost-efficient which is a recurring issue for the industry (Esch et al., 2015). However, numerous challenges in the field remain as compatibility with manufacturing processes and standardization is often low for these models (Ma et al., 2021) and will need to be solved once satisfying models have been developed. Focusing of the reproduction of the physiology of the liver, organ on chip must go through the reproduction of its microenvironment including the canals of Hering and the liver acinus. Of this, hepatic zonation-like patterns (Allen et al., 2005; Kang et al., 2018), maintenance of primary hepatic cells for drug metabolism applications (Chao et al., 2009), and liver cocultures (Novik et al., 2010) have been proposed and have shown significant improvement of functionality of conventional culture methods. However, the combination of these models with new cutting-edge technologies in cell biology such as stem cells is still under investigation.

In that regard, hiPSCs-derived cells have been the topic of intensive research for the past few years as they are potentially allowed to differentiate into any cell type while being free from the constraint of limited cell number seen in primary cells harvested from donors. hiPSCs have also been combined with the organ on chip technology to form liver tissues, composed of Hepatocyte-Like Cells (HLCs), and which could be used for the improvement of the knowledge of liver homeostasis and liver regeneration processes as well as for the formation of mature *in vitro* liver tissue (Giobbe et al., 2015, Schepers et al., 2016; Ramme et al., 2019; Danoy et al., 2019, 2021). While the differentiation and maturation of hiPSCs-derived HLCs is expected to be enhanced by the culture in microfluidic devices either due to the presence of a physiological flow (Xie et al., 2021) or to the culture in confined environment

(Michielin et al., 2020), those cells still suffer from dominant patterns of primitiveness and are yet to lead to practical applications. A possible culprit for the unsatisfying maturation state of hiPSCs-derived cultures could be the low yield resulting from the differentiation process. A low yield in the differentiation would lead to the formation of a heterogeneous tissue which could have an uncontrolled effect on the behavior of HLCs. While positive effects have been observed (Danoy et al., 2019), with the formation of an endothelial-like tissue, due to the effect of the shear stress from the microenvironment, control over the culture is needed to improve the reliability of those models for drug screening applications.

A possible strategy to increase control over the culture is cell sorting according to specific markers. Carboxypeptidase M (CPM) has been reported as a bipotential markers which led to the formation of highly functional, proliferating hepatocytes (Kido et al., 2015). CPM+ cells are selected at the hepatic progenitor differentiation stage and can be amplified several times in culture dishes. The cells were then detached and inoculated in microfluidic biochips for maturation. Comparison of the function and of the transcriptome, proteome and metabolome profiles of those cells with previously published protocols (Unsorted HLCs matured in biochips, data published in Danoy et al., 2021) was then performed and the influence of sorting for CPM positive populations was investigated. The analysis has led to a clear separation between the two conditions and improvements of the maturation in the CPM+ HLCs model could be confirmed by both the functional analysis and the multiomics analysis. The dataset presented in this study is of value toward the improvement of the culture of HLCs within microfluidic environment and should lead toward efficient liver on chip models for pharmaceutical applications.

2. Material and methods

2.1. hiPSCs cell source

In this study, the hiPSCs used were obtained from the stem cell bank of the Institute of Medical Science of the University of Tokyo (TKDN4-M clones, established by Takayama et al., 2010, using a retrovirus harbouring four reprogramming factors, OCT3/4, SOX2, KLF4, and c-MYC). Cells were proliferated on vitronectin (Thermo) coated Tissue Culture Polystyrene (TCPS) dishes in Essential-8 (Gibco) culture

medium (Chen *et al.*, 2011). For proliferation, cells were seeded at a 10,000 cells/cm² density and were subcultured at 80% confluency.

2.2. Differentiation protocol

Before differentiation, hiPSCs were seeded and maintained on Matrigel (20mg/mL, Corning) coated TCPS dishes as a standard culture condition to induce differentiation into HLCs (Si-Tayeb *et al.*, 2010). Cells were seeded at a 10,000 cells/cm² density and the differentiation was initiated at 30% confluency. The protocol consisted of 3 differentiation steps before sorting, detailed in Table S1, Supp. File 1. For sorting, cells were detached with 0,25% Trypsin (Gibco), washed with DMEM containing 10% Fetal Bovine Serum (FBS), and prepared as a single cell suspension with a 70 µm cell strainer. Cells were then incubated in FcR Blocking Reagent (Miltenyi Biotec) diluted in 0,3% BSA in PBS as recommend by the manufacturer. Cells were then sequentially incubated with the 1st antibody (Anti-hCPM, M233-3, MBL), the 2nd antibody (Anti-mouse IgG Micro Beads, 130-048-402, Miltenyi Biotec) and washed after each step. Sorting was then performed using a MidiMACS (Miltenyi Biotec) and LS columns (Miltenyi Biotec). After sorting, the CPM+ population was seeded at a 20,000 cells/cm² density on laminin (Sigma) coated culture dishes and kept in the medium and culture conditions detailed in Table S1, Supp. File 1 to allow for cell proliferation (Kido *et al.*, 2015) The cells were subcultured every 6 to 7 days in the previously detailed conditions and proliferated for a total of 20 days. The resulting cells were then detached using 0,25% Trypsin (Gibco) and seeded in Matrigel (20mg/mL, Corning) coated biochips. The experiment was repeated 3 independent times and each run consisted of 6 biochips.

2.3. Maturation in microfluidic biochips

Maturation of the CPM+ cells was led in polydimethylsiloxane (PDMS) biochips. The design and preparation of those biochips was fully detailed in previously published studies (Danoy *et al.*, 2021). The general design of the microfluidic biochip is presented in Fig. S1 in Supp. File 2. Seeding was performed at a 9 cm²/biochip ratio in a 40 µL cell suspension directly in the tubing via a 200 µL pipette tip. Uniform seeding of the cells was performed within the device by adjustment with 1mL

syringes plugged at the inlet and at the outlet. Culture medium was then replaced after an adhesion step of 6 hours and every subsequent 12 hours for the first 48 hours. The perfusion circuit was composed of a bubble traps connected to the biochips in series with a peristaltic pump. PFTE tubing was selected to limit the adsorption of chemicals and growth factors. The total volume of the circuit was of 2mL. Perfusion was then initiated at a 10 μ L/min flow rate, identically to previously published studies (Danoy et al., 2021) and within the range of shear stress values observed in the culture of hepatic cells. After 12 days of perfusion, culture was halted, and samples were collected for further analysis. Normalization of all samples was performed by counting the number of cells at the end of the maturation step. For each experiment, controls in conventional culture dishes were also performed.

2.4. Albumin measurements

Quantification of the levels of Albumin was performed as a potent marker for tissue maturity in hiPSCs-derived HLCs (Ang et al., 2018) via ELISA sandwich assays using anti-Human Albumin IgG (Bethyl, Japan, capture antibody) and anti-Human Albumin IgG coupled with peroxidase (Bethyl, Japan, detection antibody). Revelation of peroxidase was done via a H₂O₂/OPD mixture, and the plate was read at 490 nm using iMark Microplate reader (Bio-Rad).

2.5. RT-qPCR

Isolation and purification of total RNA from samples was performed with Trizol™ Reagent (Life Technologies) following the manufacturer's instructions. Sample concentration and quality was assessed with a BioSpec-nano (Shimadzu Scientific Instruments). The ReverTra Ace qPCR RT kit (TOYOBO) was used to perform reverse transcription of each RNA sample into cDNA with 1 μ g of total RNA. The THUNDERBIRD SYBR qPCR Mix (TOYOBO) was used according to the manufacturer's recommendations to perform real-time quantitative PCR with a StepOnePlus Real-Time PCR system (Applied Biosystems). Samples were issued

from 6 biochips in total evenly distributed between 3 independent experiments. The primer sequences used for ALB, AFP, CYP3A4, HNF4A, and CD144 are given in [Table S2 \(Supp. File 1\)](#). ACTB (β -Actin) was used as the reference gene and undifferentiated hiPS cells as the reference sample for the normalization of gene expression data. Total Liver RNA (5 Donor Pool, R1234149-P, Single Donor Pool, R1234149-50, BioChain Institute, Inc.) samples were used for comparison.

2.6. Immunohistochemistry

Samples for immunohistochemistry were fixed in 4% paraformaldehyde overnight at 4°C, washed several times with PBS and stored in PBS at 4°C until staining. Before staining, samples were permeabilized with 0,1% Triton X-100 (PlusOne) for 15min and blocked for 2h with a gelatin buffer. The primary antibodies used in this study were goat anti-albumin (A80-129A, Bethyl) and rabbit anti-CYP3A4 (ab135813, abcam). The following secondary antibodies were used: donkey anti-goat Alexa Fluor 488 (ab150129, abcam) and donkey anti-rabbit Alexa Fluor 568 (A10042, Thermofisher). All antibodies were diluted in the range recommended by the manufacturers. Incubation of primary antibodies was done overnight at 4°C and followed by several washing steps. Incubation of secondary antibodies was done for 3h at room temperature in the dark. After incubation, samples were further washed with PBS and incubated with DAPI (342-07431, Dojindo) following the manufacturer's recommendations. Images were taken using a confocal microscope (PowerIX70, Olympus). Negative controls were performed by performing all the previously described steps except for the incubation with the secondary antibody.

2.7. Periodic Acid Schiff staining

Staining was done on samples previously fixed in 4% paraformaldehyde overnight at 4°C and then washed with PBS for storage. A PAS staining kit (Muto Chemicals) was used following the manufacturer's recommendations.

2.8. Drug screening

Metabolism of drugs in samples was assessed via a cocktail-substrate approach (Kozakai *et al.*, 2014). The drug cocktail (500 µL, Components in **Table S3 in Supp. File 1**) was incubated for 1 hour in biochips removed from the perfusion loop. After sampling, 20 µL of culture medium was mixed with 40 µL of water and 140 µL of methanol containing 100 nM of imipramine as an internal standard. Solutions were then vortexed and centrifuged at 21,500 ×g for 5 min at 4°C. The resulting supernatant was then collected and used for the LC-MS/MS analysis. Metabolites were determined using a LCMS8050 triple quadrupole mass spectrometer (Shimadzu) coupled with an LC-30A system (Shimadzu).

Chromatography was performed on a CAPCELL PAK C18 MG III (ID 2.0 mm×50 mm; Osaka Soda Co. Ltd., Osaka, Japan) at 50°C by means of step-gradient elution (flow rate, 0,4 mL/min) as follows: 0 to 0,5 min, 95% A/ 5% B; 0.5 to 3.0 min, 95% A/ 5% B to 20% A/ 80% B; 3.0 to 4.0 min, 20% A/ 80% B; 4.0 to 4.1 min, 20% A/ 80% B to 95% A/ 5% B; 4.1 to 5.5 min, 95% A/ 5% B; (A, water containing 0.1% formic acid; B, acetonitrile containing 0,1% formic acid). The mass number detected for each of the compounds are listed in **Table S3 in Supp. File 1**. Analytical standard curves for each metabolite were made using commercially available products. Analysis of the data was made using the Lab solutions software (version 5.89, Shimadzu). Inducibility of the tissue for CYP3A4 was performed by incubation with culture medium supplemented with 5 µM of rifampicin and 2 µM of 3-Methylcholanthrene (3-MC).

2.9. Statistical analysis

Statistical analysis for the dataset issued from the Albumin measurements, the RT-qPCR, and the drug screening was done by ANOVA to evaluate differences between the groups. When the null hypothesis was rejected, a posthoc Tukey HSD test was performed and differences with $P < 0,05$ (*), $P < 0,01$ (**), and $P < 0,001$ (***), were highlighted and considered statistically significant.

2.10. Transcriptomic analysis

Transcriptome profiling was performed via nanoCAGE as previously described (Poulain *et al.* 2017). Total RNAs were extracted from samples using

Trizol™ Reagent (Life Technologies) and PureLink RNA mini kit (ThermoFisher) following the manufacturer's instructions. Reverse-transcription was done with 250 ng of each sample using pseudo-random primers to reduce the amount of sequencing reads mapping to ribosomal DNA (Arnaud et al., 2016). The cDNA samples obtained were tagged by specific index sequences, amplified by PCR and then multiplexed to produce a nanoCAGE library that was finally sequenced paired-end (9.47 pM + 10% PhiX) on MiSeq system with the Reagent Kit v2 50-cycles (Illumina).

FASTQ files generated by the sequencer were processed for CAGEscan analysis with the MOIRAI pipeline OP-WORKFLOW-CAGEscan-short-reads-v2.0 (Plessy et al., 2010; Kratz et al., 2014; Hasegawa et al., 2014). Briefly, sequencing reads were demultiplexed and trimmed to 50 nt with the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Then, reads originating from rRNA or oligo-artifacts were removed with TagDust (v2.33) (Lassmann, 2015) and the remaining reads were aligned on the human genome (hg19 and hg38 assemblies) with BWA (v0.7.5a) (Li and Durbin, 2009). Non-proper pairs of reads and PCR duplicates were filtered out with samtools (v0.1.19) (Li et al., 2009), and properly paired reads belonging to the same molecule were finally clustered with the tools "pairedBamToBed12" version 1.1 (<https://github.com/charles-plessy/pairedBamToBed12>). The CAGER package (Haberle et al., 2015) was used with custom R scripts to create gene expression tables that were uploaded on the iDEP server (Ge et al., 2018) (<http://bioinformatics.sdstate.edu/idep/>) for Differential Gene Expression (DGE) with the DESeq2 method (Love et al., 2014) and Pathway Gene Set Enrichment Analysis (PGSEA). Motif Activity Response Analysis (MARA) was done via the ISMARA webserver (Balwierz, et al., 2014) (<https://ismara.unibas.ch/mara/>). Sequence data that support the findings of this study can be found in the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>) with the series accession code GSE179713.

2.11. Proteomic analysis

Protein samples were reduced and alkylated with 10mM TCEP, 50 mM chloroacetamide, 2% SDS and 100 mM Tris/HCl pH8.5 for 5 min at 95°C. The concentration of proteins in each sample was measured using SDS-PAGE and

Imagelab software (Biorad). For each sample, 30 µg of proteins were digested overnight at 37°C with 1µg trypsin (Promega) using S-TRAP as previously described (Ludwig et al., 2018). Peptides were separated in 5 fractions using strong cation exchange (SCX) resin (Ishihama et al., 2006). In details, the peptides were loaded into pipette-tip columns (made by stacking six layers of a 3M Empore cation extraction disk into a 200 µL micropipette tip) and column conditioning was done with acetonitrile (ACN). Samples acidified with 0,1% Trifluoroacetic acid (TFA) were loaded in the column and further washed with 0,1% TFA. Peptides were finally successively eluted using 20% ACN, 0,05% formic acid and ammonium acetate at 75mM, 125mM, 200mM and 300mM. The 5th fraction was eluted in 1.4% NH₄OH, 80% ACN.

After speed-vacuum drying, fractions were solubilized in 10 µL of 0,1% TFA, 10% ACN. A U3000 RSLC nanoflow-HPLC system coupled to an Orbitrap fusion MS analysis (Thermo Fisher Scientific) were used for Liquid chromatography and mass spectrometry analyses. 1µl of each sample were concentrated and washed on a C₁₈ reverse-phase precolumn (3µm particle size, 100 Å pore size, 75 µm inner diameter, 2 cm length, Thermo Fischer Scientific), and then separated using a C₁₈ reverse-phase analytical column (2 µm particle size, 100 Å pore size, 75 µm inner diameter, 25 cm length, Thermo Fischer Scientific). A 3h gradient starting from 99% of solvent A (0,1% formic acid) to 55% of solvent B (80% ACN and 0,085% formic acid) was used. The Fusion mass spectrometer acquired data throughout the elution process and operated in a data-dependent scheme with full MS scans acquired with the Orbitrap, followed by as many MS/MS ion trap HCD spectra 3 seconds which could fit (data-dependent acquisition with top speed mode: 3s cycle) using the following settings for full MS: automatic gain control (AGC) target value: 1e6, maximum ion injection time (MIIT): 60 ms, resolution: 6.10e4, m/z range 350–1500. For HCD MS/MS: Quadrupole filtering, Normalised Collision Energy: 30. Ion trap rapid detection: isolation width: 1.6 Th, minimum signal threshold: 5000, AGC: 1.10e5, MIIT: 60 ms, resolution: 3.10e4. Peptides with undefined charge state or charge state of 1 or over 7 were excluded from fragmentation, a dynamic exclusion time was set at 30 s.

Maxquant v1.6.15.0 (Cox et al., 2014) was used to analyze the spectrometry data. The database used was a concatenation of human sequences from the Uniprot-Swissprot database (Uniprot, release 2020-03) and an incremented list of

contaminants. The enzyme specificity was trypsin. The precursor mass tolerance and the fragment mass tolerance were set at 4,5 and 20 ppm respectively. Carbamidomethylation of cysteins was set as constant modification and acetylation of protein N-terminus and oxidation of methionines were set as variable modification. Second peptide search was allowed, and minimal length of peptides was set at 7 amino acids. The maximum for the False discovery rate (FDR) was kept at 1% on both peptides and proteins. Both unique and razor peptides were used for label-free protein quantification (LFQ). At least 2 ratio counts were required for LFQ. All experiments were analyzed simultaneously with the “match between runs” option with a match time window of 0,7 min and an alignment time window of 20 min.

2.12. Metabolomic analysis

The GC-MS technology was used for the extraction of the metabolomic profile. Sample preparation and metabolites extraction were performed according to previously published protocols (Danoy et al., 2021). 250 µL of culture medium were added to 500 µL of a frozen solution (-20°C) of water:acetonitrile:isopropanol (2:3:3) containing 4 mg/L of adonitol, 2.75 mg/L of α-aminobutyric acid solution (αABA), and placed in an Eppendorf thermomixer for 10 min at 4°C with the shaking speed set at 1500 rpm. Two centrifugation steps at 14000 rpm for 10 min were performed to remove insoluble materials. Three aliquots of each extract (200 µL) were dried for 4 h at 35 °C in a speed-vac and stored at -80°C until analysis.

The GC-MS analysis performed in this study was done as previously described (Fiehn 2006; Fiehn et al., 2008). Before analysis, samples were thawed and dried for 1,5h at 35 °C. 10 µL of 20 mg/mL methoxyamine in pyridine were then added before the reaction was performed at 30 °C for 90min while shaking. 90 µL of N-methyl-N-trimethylsilyl-trifluoroacetamide (MSTFA, Regis Technologies) were then added as the reaction was continued for 30 min at 37 °C. Samples were then cooled and 100 µL was transferred to an Agilent vial for injection. 4h after derivatization, 1 µL of the samples was injected in splitless mode on an Agilent 7890B gas chromatograph coupled to an Agilent 5977A mass spectrometer. The column used was a Rxi-5SilMS from Restek (30 m with 10 m Integra-Guard column). An injection in split mode with a ratio of 1:30 was systematically performed for saturated compound quantification. Temperature of the oven was set at 60 °C for 1 min and

ramped up to 325 °C for 10 min at a 10 °C/min speed. The constant flow of helium was set at 1,1 mL/min. Temperatures were set as follows: injector: 250 °C, transfer line: 290 °C, source: 230 °C, and quadrupole 150 °C. The quadrupole mass spectrometer was switched on after a 5,90 min solvent delay time, scanning from 50-600 u. Samples were randomized and fatty acid methyl ester mix (C8, C9, C10, C12, C14, C16, C18, C20, C22, C24, C26, C28, C30) was injected in the middle of the queue for external RI calibration.

AMDIS (<http://chemdata.nist.gov/mass-spc/amdis/>) was used for the analysis of the raw Agilent datafiles. Metabolite identification was done via the Agilent Fiehn GC/MS Metabolomics RTL Library (version of June 2008). Masshunter Quantitative Analysis (Agilent) in splitless and split 30 modes was used to determine peak areas. Due to occasional errors in the automated peak integration, integration was verified manually for each compound and peak areas were normalized to ribitol. All metabolite contents are expressed in arbitrary units (semi-quantitative determination). The XLSTAT.2016 software (Addinsoft) and MetaboAnalyst 4.0 (Chong *et al.*, 2018) was used to perform the metabolomic multivariate data analysis. Maximum separation between groups and exploration of the variables that contributed to this separation was done by applying Supervised projections to latent structures discriminant analysis (PLS-DA). The significant differences in the presence of metabolites were highlighted using variable importance in the projection value (VIP > 1) and Student's t-test (P < 0.05).

2.13. Data from the unsorted HLCs differentiation protocol

The dataset used for comparison in this study was produced for and published in Danoy *et al.*, 2021. The protocol for the differentiation of unsorted HLCs from hiPSCs is reminded in [Table S4 in Supp. File 1](#).

3. Results

3.1. Functional characterization of CPM+ HLCs

In this work, the CPM+ sorted HLCs matured and described in conventional culture dishes by Kido *et al.*, 2015 were seeded in PDMS microfluidic

biochips. After sorting on day 2, the cells exhibited a typical morphology of liver progenitor cells (LPCs) and evolved toward an hepatoblast-like morphology after maturation and subcultures as described (Fig. 1A). After 12 days of maturation in microfluidic biochip, the CPM+ HLCs tissue presented several distinct morphologies (Fig. 1B). Especially, zones in which the cells presented a morphology typical of fetal hepatocytes, smaller in size than mature hepatocytes but with a large nucleus could be observed. The remaining tissue appeared as a mix of fibroblastic tissue, particularly in areas close to the channel walls. Comparatively, in the data previously obtained with unsorted HLCs (Fig. 1C), the tissue appeared denser and largely conjunctive. HLCs could be found particularly in the center of the channels, seemed less numerous, and appeared larger than in the CPM+ HLCs biochips. In culture dishes controls, the presence of cells with a morphology typical of fetal hepatocytes could also be confirmed (Fig. S2 in Supp. File 2). The cells appeared relatively small and densely packed. Cell aggregates and fibroblastic tissues could also be observed in certain areas.

In regard to Albumin production, a gradual increase was observed during the maturation in biochips, reaching $20 \mu\text{g}/\text{mL}/10^6$ cells at the end of the culture. Compared to the data previously obtained with unsorted HLCs, a significant increase of one order of magnitude was observed in CPM+ HLCs (Fig. 1D). In culture dishes controls, the same tendency in the increase of the production of albumin was observed (Fig. S3A in Supp. File 2). While albumin was found to be lower in control culture dishes than in biochips ($12,7 \mu\text{g}/\text{mL}/10^6$ cells, Fig. S3B in Supp. File 2), no statistical difference was observed. Concerning mRNA levels, CPM+ HLCs did not show any statistical difference with unsorted HLCs when cultured in biochips (Fig. 1E). In both cases, levels of ALB, CYP3A4, and HNF4A remained lower than those obtained with Liver Total RNA samples. No statistical difference in the levels of AFP were detected between all groups. In addition, CD144 levels were found to be lower in both CPM+ HLCs and unsorted HLCs in biochips than in the Liver Total RNA samples, indicating the absence of endothelial cells. Finally, when compared to the culture in conventional dishes as a control, CPM+ HLCs exhibited lower levels of AFP and ALB, and higher levels of CD144 (Fig. S3C in Supp. File 2). No significant difference was observed in the levels of CYP3A4 and HNF4A.

Immunohistochemistry on CPM+ HLCs and unsorted HLCs revealed expression of both Albumin and CYP3A4 in each condition (Fig. 2A) while staining for

CYP3A4 was found to be significantly higher in the former. Those markers could also be observed in CPM+ HLCs cultured in conventional dishes as a control but the intensity of the staining for CYP3A4 was found to be lower than in CPM+ HLCs cultured in biochips (Fig. S4 in Supp. File 2). Glycogen storage was also confirmed in CPM+ HLCs and unsorted HLCs cultured in biochips via PAS staining (Fig. 2B) and also found to be more functional in the former. Glycogen storage was also confirmed in CPM+ HLCs cultured in conventional dishes as a control but appeared weaker than in cells cultured in biochips (Fig. S4 in Supp. File 2). Finally, CYP450 activity could be confirmed in both CPM+ HLCs and unsorted HLCs cultured in biochips (Fig. 2C). Notably, higher production of paracetamol (via metabolism of phenacetin by CYP1A2), of hydroxymephenytoin (via metabolism of S-mephenytoin by CYP2C19), and of hydroxybufuralol (via metabolism of bufuralol by CYP2D6) were found in the former. No significant difference was observed in the production of 1-hydroxymidazolam (via metabolism of midazolam by CYP3A4) and of hydroxybupropion (via metabolism of bupropion by CYP2B6). Also, no activity for CYP2A6, CYP2C8, and CYP2C9 were not detected (data not shown).

3.2. Transcriptomic analysis

Comparison of CPM+ HLCs matured in biochips with previously differentiated unsorted HLCs, with undifferentiated hiPSCs and with Primary Human Hepatocytes (PHHs) was performed via nanoCAGE transcriptome analysis. Motif Activity Response Analysis (MARA) was performed on sequencing data and highlighted 176 transcription factors (TFs) that were significantly modulated (z -value of TF motif activity above 1) between the 4 conditions. Additional multivariate analysis on the TF motif activity matrix allowed to extract a selection of those motifs as shown in the heatmap presented in Fig. 3A (complete data available in Supp. file 3). The analysis confirmed that differentiation toward a mature hepatocyte's phenotype was more advanced in CPM+ HLCs matured in biochips than in unsorted HLCs as shown by higher levels of activity for the HNF1A_HNF1B, CEBPA, HNF4A, ESRA_ESR2 motifs. However, comparison with PHHs highlighted that the activity levels of those TF motifs were yet to attain those of mature hepatocytes in CPM+ HLCs matured in biochips. In the comparison with unsorted HLCs matured in biochips, CPM+ HLCs showed an increase in the activity of motifs related to EBF1, ID4_TFC4_SNAI2,

WT1_MTF1_ZBTB7B, TAL1, IRF2_STAT2, CEBPE_CEBPD, MEIS2, HIC2, and MAF_NRL. Similarly, this comparison showed a decrease in the activity of motifs related to SP3, MECP2, EGR2_EGR3, ZNF711_TFAP2A_TFAP, GMEB2, YY1_YY2, and MYBL2, BPTF in the latter. In both cases, the levels observed in CPM+ HLCs matured in biochips were found to be getting closer to the ones observed in PHHs. In parallel, higher activity of motifs related to WRNIP1, BATF, TEAD3_TEAD1, FOSL2_SMARCC1, RELA, XBP1, FOSL1, and SRF was found in both CPM+ HLCs and unsorted HLCs when compared to PHHs.

Direct comparison between CPM+ HLCs and unsorted HLCs matured in biochips allowed to extract another set of TF motifs with a differential activity between those two conditions (list of the top 50 motifs shown in [Fig. 3B](#)). Among those, HNF1, HNF4, FOXA3_FOXC2, and EBF1 were found to have a higher activity in CPM+ HLCs. Within the list of top 10 TFs with a differential activity between the two conditions, ZNF711_TFAP2A_TFAP, TAF1, NRF1, MECP2, YY1_YY2, NFATC2_NFATC3, and ZBTB33_CHD2 were found to have a higher activity in unsorted HLCs and EBF1, ID4_TFC4_SNAI2, and MEIS2 were found to have a higher activity in CPM+ HLCs. Involvement of those TFs in regard to hepatic differentiation and to other important liver processes are detailed in [Table 1](#).

Comparison of the transcriptome data obtained for unsorted HLCs and CPM+ HLCs matured in biochips was further performed using iDEP9.1. The Differential Gene Expression (DGE) analysis led to the extraction of 2394 genes differentially expressed between the two conditions with an FDR cutoff of 0,1 and a minimum fold change of 2 ([Fig. 4A](#), the full list of differentially expressed genes is provided in [Supp. File 4](#)). Among those, overexpression of CYP1A1, CYP1B1, CYP2C18, CYP27A1, UGT2A3, 2B15, 2B4, 2B7, SULT1A1, 1B1, 1C2, 1E1, NR1H4 (FXR, bile acid receptor), NR3C2 (aldosterone receptor), and NR4A1 was found in CPM+ HLCs. In addition, Pathway Gene Set Enrichment Analysis (PGSEA) led to the identification of differentially regulated pathways and gene ontologies between the two culture conditions. Results obtained using the KEGG metabolic pathways, the GO_molecular function, and the GO_biological process gene sets are respectively shown on the heatmaps presented in [Fig. 4B, 4C and 4D](#). These results revealed an upregulation of the expression of genes related to drug metabolism, CYP450, glutathione metabolism, oxidative phosphorylation, steroid hormone biosynthesis, and to pentose phosphate-glucuronate pathway in CPM+ HLCs. Besides, mRNA expression levels

of genes related to Hippo signaling, TGF β signaling, and to extra cellular matrix (ECM) processing were found upregulated in unsorted HLCs.

3.3. Proteomic analysis

Comparison of the proteome in unsorted HLCs and CPM+ HLCs matured in biochips was performed. 861 proteins were found to be differentially expressed in the two conditions with a *p*-value below 0,05 among the 4553 proteins detected (Supp. File 5). The heatmap made from the top 70 proteins differentially expressed (with the lowest *p*-value) between the two conditions is given in Fig. 5A. Pathway analysis with iDEP9.1 on the complete list of proteins highlighted overactivation of lipids, carbohydrates, aerobic, energy, and mitochondrial processes in CPM+ HLCs. Similarly, processes related to the vascular development and to the cytoskeleton were highlighted in unsorted HLCs (Fig. 5B).

Uploading of the extracted proteins within the KEGG database (Supp. File 6) led to the identification of a first group of pathways related to the modulation of the cellular reorganization via the focal adhesion (18 hits), the regulation of actin cytoskeleton (16 hits), and the ECM receptor (10 hits). This translated, in CPM+ HLCs, into the low expression of different types of collagens (COL1A2, COL3A1, COL6A2 COL4A2), vimentin, vitronectin, alpha actinin 1, filamin A, and filamin B, and into the high expression of profilin-1, profilin-2, and laminin. Important changes in the metabolism of lipids, cholesterol, and steroids were also detected via the PPAR signaling, the metabolism of fatty acid, of sphingolipid, of glycerophospholipid, and of glycerolipid, and in the biosynthesis of steroids (14, 18, 9, 5, 5, and 5 hits respectively). This translated, in CPM+ HLCs, into the high expression of *ACOX1*, 3, *ACSL1*, *ACSL3*, *ACSL5*, *CPT1A*, *CPT2*, *APOA1*, *APOL2*, *PLPT*, *SCP2*, *HMGCS2*, *ACACA*, *ACADVL*, *ACAT1*, *HADHA*, *HADHB*, *HADH*, *HSD17B4*, *AGK*, *GPD1*, *GPD2*, *AKR1B10*, *AKR1B1*, *SORT1*, *NCEH1*, and *DHCR24*, and into low expression of *EHHADH*, *NPC1*, *ABCA1*, *DHCR7*, and *APOE*. Finally, significant differences were also found in the metabolism of carbohydrates via glycolysis/gluconeogenesis, the metabolism of pyruvate and galactose, via pentose phosphate, and the metabolism of fructose (15, 11, 10, and 8 hits respectively). This translated, in CPM+ HLCs, into the high expression of *G6PD*, *GPI*, *PFKL*, *PFKM*, *PGM2*, *LDHB*, *HK1*, *PKM*, *ACSS2*, *PCK2*, and *PDHA1* and into the low expression of *PGAM1* and *SORD*. Notably,

modulation of the TCA pathways (13 hits) and of the oxidative phosphorylation (21 hits) were also found and translated, in CPM+ HLCs, into high expression of *SDHA*, *SDHB*, *PDHA1*, *PCK2*, *IDH3A*, *IDH3G*, *SUCLA2*, *SUCLG2*, *FH*, *OGDH*, *ACO1*, and *ACO9* and high levels of 2 ATPases, of 10 NADH ubiquinone oxidoreductases, 1 cytochrome C oxidase, and 5 ubiquinol cytochrome C reductases respectively. In agreement with the high expression of OXPHOS proteins in CPM+ HLCs, high levels of glutathione related proteins (*GPX1*, *GPX4*, *GSTM2*, and *GSTM3*) were also found.

In regard to inflammatory pathways, no significant difference was observed between CPM+ HLCs and unsorted HLCs matured in biochips. Especially, no significant difference in proteins related to inflammatory cytokines such as *TNF α* , *TGF β* , and interleukins were detected. However, modulation of NF κ B pathways (6 hits) was found and translated in CPM+ HLCs, into high expression of *RELA*, *BCL2L13*, and *BCL2L1* and lower expression of *IKBIP* and *TGFBI*. In relation to those, modulation of PI3K-AKT, MAPK (20 and 13 hits respectively) were found as potential controlling pathways. In that regard, low levels of GNB1, high levels of PRKCA, RELA, PPP2CA, YWHAE, mTOR, and NRAS were also observed. Finally, in relation with the functionality of the tissue, no significant difference was found in regard to CYP450 enzymes and to liver transcription factors such as *HNF1*, *HNF3*, *HNF4*, *PXR*, and *CAR*. However, higher levels for few proteins related to phase II drug metabolism such as 2 GST and 3 UGT was detected in CPM+ HLCs while higher levels of *SULT2A1* was detected in unsorted HLCs.

3.4. Metabolomic analysis

As the composition for the maturation culture media of CPM+ HLCs and unsorted HLCs differed, comparison between media used in culture and basal medium was performed in each case to extract the metabolomic profile.

In the comparison of media used in presence of CPM+ HLCs and basal medium 26 metabolites were found to be differentially expressed between the two conditions (p_value below 0,05, [Supp. File 7](#), PLDS-DA and model accuracy in [Fig. S5 Supp. File 2](#)). The top 25 metabolites differentially expressed between the two conditions are illustrated in the heatmap given in [Fig. 6A](#). Among those, lower levels of linoleic acid, L-cystine, L-tryptophan, L-serine, L-leucine. DL-isoleucine, putrescine, and citramalic acid were detected in media used in presence of CPM+ HLCs while

higher levels of lyxose, citric acid, threonic acid, beta hydroxyisovaleric acid, 2 ketoisocaproic acid, L valine, azelaic acid, isocitric acid, xylulose, pyruvic acid, and L-kynerunine were detected in media used in presence of CPM+ HLCs. Pathway analysis on the metabolites differentially expressed between the two conditions revealed enrichment of the valine, leucine, and isoleucine degradation, of the citric acid cycle, of the transfer of acetyl groups into mitochondria, of the Warburg effect and of ammonia recycling as the top 5 biological processes (Fig. 6B).

In the comparison of media used in presence of unsorted HLCs and basal medium 23 metabolites were found to be differentially expressed between the two conditions (p -value below 0,05, Supp. File 7). The top 25 metabolites differentially expressed between the two conditions are illustrated in the heatmap given in Fig. 6C. Among those, lower levels of oxalic and azaleic acids were detected in media used in presence of unsorted HLCs while higher levels of aspartic acid, glycerol, xylulose, glyceric acid, lactic acid, xylulose, capric acid, hexanoic acid, palmitoleic acid, lauric acid, D mannose, alpha ketoglutaric acid, fumaric acid, threonic acid, 4 trans hydroxy proline, D threitol, and 2-ketoisocaproic acid were detected in media used in presence of unsorted HLCs. Pathway analysis on the metabolites differentially expressed between the two conditions revealed enrichment of the beta oxidation, of the urea cycle, of the Warburg effect, of the phosphatidylethanolamine and phosphatidylcholine biosynthesis as the top 5 biological process (Fig. 6D).

4. Discussion

Previously published strategies in our group (Danoy et al., 2019, 2021) were based on the conviction that cellular heterogeneity, caused by the differentiation process, would lead to an improved maturation due to the presence of non-parenchymal cells sharing the same ascendance. However, in both monoculture and in coculture with non-parenchymal cells, primitive patterns were still found to be dominant and further differentiation did not lead to any specific improvement. In that regard, CPM+ was previously reported as a remarkable bipotential marker and characteristic of highly functional HLCs populations (Kido et al., 2015). In the present investigation, confirmation of the compatibility of CPM+ HLCs with the organ on chip technology was obtained and comparison with unsorted hiPSCs-derived HLCs populations matured in microfluidic biochips was performed. Functional

characterization of CPM+ HLCs confirmed their improved performance over unsorted HLCs. This could be seen in terms of albumin production but also in terms of CYP450 activity which was qualitatively confirmed by immunofluorescence imaging. Especially, albumin levels ($30\mu\text{g}/10^6$ cells), were found to be comparable to the ones observed in adult primary human hepatocytes ($25\mu\text{g}/10^6$ cells), and much higher than those observed in fetal hepatocytes or cell line such as HepG2 ($7.5\mu\text{g}/10^6$ cells and $6\mu\text{g}/10^6$ cells respectively, Rowe et al., 2013, Nishikawa et al, 2017). In addition, improvement of the metabolism of carbohydrates was also detected as highlighted by the results of the proteomic analysis, the metabolomic analysis, and qualitatively by the PAS staining. While the presented protocol has shown improvement in the function of the tissue, it is important to note its pros and cons. Indeed, as cells can be proliferated and passaged, the CPM+ HLCs protocol presents an alternative to conventional differentiation protocols as it allows for the production of substantial number of cells after sorting. Though, as selection of cells at the LPC stage leads to an initial loss in terms of cell number, it also triggers the need for an extensive cellular proliferation which causes the overall differentiation protocol to be much longer than conventional protocols. For that reason, the CPM+ HLCs differentiation protocol described here is in total, 54 days long, while the previously used unsorted HLCs protocol was only 34 days long. Nevertheless, maturation in microfluidic biochips was led in both cases for a total of 14 days, including 12 days in perfusion and cells were, in both cases, loaded into the biochips at a state that is defined as progenitor (Kido et al., 2015, Danoy et al, 2019).

The signature extracted for each condition by the multiomics analysis further confirmed the improved maturation in CPM+ HLCs. This was illustrated by higher motif activity levels of important liver transcription factors such as the HNF1 and HNF4, by higher activation of important hepatic nuclear receptors, and by higher expression of phase I and phase II metabolism genes. Improved maturation could also be related to the upregulation of genes and proteins involved in the oxidative phosphorylation. In addition, overexpression of proteins related to the TCA pathway as well as low levels of TCA intermediates and substrates (citric acid, isocitric acid) were also detected. Although cycles between OXPHOS respiration and glycolysis respiration were reported during liver regeneration and hepatic maturation (Hopkinson et al., 2017), the OXPHOS signature is an important marker of the

transition between hepatic specification and hepatic maturation, further confirming the latter (Hopkinson et al., 2017).

Signaling related to lipids and bile acids were also found to be clearly distinct in CPM+ HLCs and unsorted HLCs matured in biochips in all multiomics dataset. Particularly, this was found in the form of the higher expression of proteins involved in fatty acid beta oxidation, steroids biosynthesis, and lipids transport in CPM+ HLCs and of the lipid production (lauric acid, capric acid, palmitoleic acid) only observed in unsorted HLCs. Usually, profiles of transient lipid accumulations coupled with temporal steatosis and endothelium reticulum stress are associated to liver regeneration (Rudnick et al., 2012). Indeed, lipids are component of the cells membrane that are requested for stages of hepatocyte proliferation, and restoration of a normal and healthy lipid metabolism is found as regenerative processes are concluded (Rudnick et al., 2012; Della Fazia et al., 2018). In the healthy liver, this process is reported to lead to the activation of XBP1, CHOP (DDIT3), and ATF4 and to the recovery of the liver mass (Della Fazia et al., 2018). In the presented dataset, intense lipids production was shown in unsorted HLCs as well as a higher motif activity of the transcription factors XBP1 and DDIT3 (Supp. File 3). The tissue was also found to be dense and exhibiting cell proliferation in terms of morphologies, ECM production, and pathways related to connective tissue development and ECM assembly (Fig. 5D), highlighting differences between CPM+ HLCs and unsorted HLCs in terms of liver regeneration and maturation. The maturation of unsorted HLCs in biochips was previously found to be linked with a complex modulation of the TGF β activation (Danoy et al., 2021). Especially, this modulation would be consistent with the responses observed in unsorted HLCs such as cell proliferation and tissue reorganization. However, modulation of TGF β appeared lower in CPM+ HLCs at the transcriptome level while no difference could be observed in the proteomic analysis which indicates a possible regulation mainly done at the transcriptional level. As TGF β has been reported to be sequentially activated during processes related to liver regeneration and during the hepatic maturation (Apte et al., 2009; Michalopoulos, 2010), it would be further consistent with the differences observed between unsorted HLCs and CPM+ HLCs in regard to liver regeneration and maturation.

Finally, morphologies within the tissues in unsorted HLCs and CPM+ HLCs matured in biochips were found to be strikingly different. In unsorted HLCs, the tissue

was characterized as being conjunctive which was confirmed by the transcriptomic (via the GO_molecular function heatmap, through ECM structures and collagen binding) and the proteomic analysis with a higher production of ECM related proteins (collagen, vimentin, vitronectin, and filamin). In addition, as cells are unsorted after the differentiation process, the cell population is expected to be more heterogeneous than in CPM+ HLCs matured in biochips. It can be suggested that mechanical stimulation and that the local microenvironment could have contributed to trigger the differentiation of several subpopulations which couldn't be clearly identified. In that regard, while complex endothelial structures could not be observed in unsorted HLCs, the proteomic analysis clearly extracted pathways related to the vascular development and to the cardiovascular development in those cells (Danoy et al., 2021) which was not the case in CPM+ HLCs. As the endothelium is known to be supporting of the liver function (Bale et al., 2015) and to be highly linked to complex angiogenesis processes during liver regeneration (Poisson et al., 2017), it is highly expected that coculture with LSECs would be needed to further improve the function of CPM+ HLCs matured in biochips.

5. Conclusion

In the present manuscript, a new protocol combining the differentiation and maturation of CPM+ HLCs with the organ on chip technology has been proposed. While the tissue still exhibited immature feature such as the high expression of AFP, significant improvement was observed in terms of function when compared to unsorted HLCs culture in similar conditions. The analysis of the transcriptomic, proteomic and metabolomic profiles contributed to clearly separate both conditions and further confirmed the improvement maturation in CPM+ HLCs. As the sorted population was logically found to be more homogeneous, it can be expected that coculture with non-parenchymal cell would provide significant improvement toward the establishment of an hiPSCs-based liver on a chip.

Acknowledgments

This work was supported by the French Investissements d'Avenir in the framework of the iLite project (ANR-16-RHUS-0005). The molds used for the fabrication of the PDMS biochips were made by the LAAS in the frame of RENATECH support. Stéphane Poulain was supported by the JSPS Grant-in-aid for Scientific Research (S) 16H06328. The authors thank the FANTOM5 consortium for generously allowing the use of their PHHs RNA samples and the RIKEN Center for Integrative Medical Sciences, Yokohama, Japan, for providing access to their computer clusters for transcriptomic data processing. We thank Charles Plessy from OIST, Japan, for fruitful discussion regarding the nanoCAGE analysis. We acknowledge the 3P5 Proteom'IC platform (Université de Paris, Institut Cochin, INSERM, U1016, CNRS, UMR8104, F-75014 PARIS, France) and François Guillonnet for their precious advice on the proteomic analysis and Cédric Broussard for monitoring the project. The Orbitrap Fusion mass spectrometer was acquired with funds from the FEDER through the « Operational Programme for Competitiveness Factors and employment 2007-2013 », and from the « Canceropole Ile de France». The authors also thank the JSPS C2C Program.

Conflict of interest statement

The authors declare no conflict of interest.

Ethical issue

No ethical approval was required for the work presented in this study.

Data availability statement

All data are provided in supplementary files. Additional information can be provided by directly contacting the authors

Figure captions

Fig.1: Morphologies observed in CPM+ HLCs during the amplification phase (A) and after maturation in biochips for a total of 14 days (B). Morphologies observed in unsorted HLCs after maturation in biochips for a total of 14 days (C). Production of albumin during the maturation of CPM+ HLCs in biochips and comparison with

unsorted HLCs on the last day of maturation (D). mRNA levels measured by RT-qPCR in unsorted HLCs and CPM+ HLCs matured in biochips and in Liver Total RNA samples (E). The unsorted HLCs dataset was produced for and published in [Danoy et al., 2021](#).

Fig.2: Immunohistochemistry for Albumin and CYP3A4 performed on unsorted and CPM+ HLCs matured in biochips (A). Expression of glycogen storage obtained by PAS staining in unsorted and CPM+ HLCs matured in biochips (B). Metabolites production via CYP450 activity measured by the cocktail-based approach in unsorted and CPM+ HLCs matured in biochips (C). The unsorted HLCs dataset was produced for and published in [Danoy et al., 2021](#).

Fig.3: Heatmap of the motifs activity matrix for important transcription factors extracted by ISMARA processing using the metaboanalist multivariate analysis (A). Comparison between CPM+ HLCs and unsorted HLCs (B).

Fig.4: Analysis of the transcriptomic dataset using iDEP9.1. Heatmap of differentially expressed genes discriminating the two culture conditions (DESeq2 processing with FDR cutoff = 0,1 and min fold change = 2) (A). Heatmaps of the KEGG (B), GO_molecular function (C), and GO_biological process (D) gene sets extracted from the pathway analysis via the PGSEA formalism.

Fig.5: Heatmap of the top 70 discriminated proteins in the proteomic dataset in the comparison between CPM+ HLCs and unsorted HLCs matured in biochips (A). Heatmap of the GO_Biological process extracted from the pathway analysis *via* the PGSEA formalism by injecting the proteins differentially expressed (*p_value* below 0,05) in iDEP9.1 (B).

Fig.6: Heatmap of the top 25 discriminated metabolites in the metabolomic dataset in the comparison of media used in presence of CPM+ HLCs and basal medium (A). Metabolites Set Enriched pathway using the top metabolites differentially expressed in CPM+ HLCs matured in biochips (B). Heatmap of the top 25 discriminated metabolites in the metabolomic dataset in the comparison of media used in presence

of unsorted HLCs and basal medium (C). Metabolites Set Enriched pathway using the top metabolites differentially expressed in unsorted HLCs matured in biochips (D).

Table 1: Top 10 TFs motifs extracted from ISMARA in the comparison between unsorted HLCs and CPM+ HLCs matured in biochips. PCC stands for the motif with the strongest Pearson correlation coefficient. Details are provided in Supp. File 4

References

Allen, J. W., Khetani, S. R., & Bhatia, S. N. (2005). In vitro zonation and toxicity in a hepatocyte bioreactor. *Toxicological sciences*, *84*(1), 110-119.

Ang, L. T., Tan, A. K. Y., Autio, M. I., Goh, S. H., Choo, S. H., Lee, K. L., ... & Lim, B. (2018). A roadmap for human liver differentiation from pluripotent stem cells. *Cell reports*, *22*(8), 2190-2205.

Apte, U., Gkretsi, V., Bowen, W. C., Mars, W. M., Luo, J. H., Donthamsetty, S., ... & Michalopoulos, G. K. (2009). Enhanced liver regeneration following changes induced by hepatocyte-specific genetic ablation of integrin-linked kinase. *Hepatology*, *50*(3), 844-851.

Arnaud, O., Kato, S., Poulain, S., & Plessy, C. (2016). Targeted reduction of highly abundant transcripts using pseudo-random primers. *Biotechniques*, *60*(4), 169-174.

Bale, S. S., Golberg, I., Jindal, R., McCarty, W. J., Luitje, M., Hegde, M., ... & Yarmush, M. L. (2015). Long-term coculture strategies for primary hepatocytes and liver sinusoidal endothelial cells. *Tissue Engineering Part C: Methods*, *21*(4), 413-422.

Balwierz, P. J., Pachkov, M., Arnold, P., Gruber, A. J., Zavolan, M., & van Nimwegen, E. (2014). ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs. *Genome research*, *24*(5), 869-884.

Chao, P., Maguire, T., Novik, E., Cheng, K. C., & Yarmush, M. L. (2009). Evaluation of a microfluidic based cell culture platform with primary human hepatocytes for the prediction of hepatic clearance in human. *Biochemical pharmacology*, *78*(6), 625-632.

Chen, G., Gulbranson, D. R., Hou, Z., Bolin, J. M., Ruotti, V., Probasco, M. D., ... & Thomson, J. A. (2011). Chemically defined conditions for human iPSC derivation and culture. *Nature methods*, *8*(5), 424-429.

Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., ... & Xia, J. (2018). MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic acids research*, *46*(W1), W486-W494.

Cox, J., Hein, M. Y., Lubner, C. A., Paron, I., Nagaraj, N., & Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & cellular proteomics*, 13(9), 2513-2526.

Danoy, M., Bernier, M. L., Kimura, K., Poulain, S., Kato, S., Mori, D., ... & Leclerc, E. (2019). Optimized protocol for the hepatic differentiation of induced pluripotent stem cells in a fluidic microenvironment. *Biotechnology and bioengineering*, 116(7), 1762-1776.

Danoy, M., Tauran, Y., Poulain, S., Jellali, R., Bruce, J., Leduc, M., ... & Leclerc, E. (2021). Multi-omics analysis of hiPSCs-derived HLCs matured on-chip revealed patterns typical of liver regeneration. *Biotechnology and Bioengineering*.

Della Fazia, M. A., & Servillo, G. (2018). Foie GRAS and liver regeneration: a fat dilemma. *Cell Stress*, 2(7), 162.

Esch, E. W., Bahinski, A., & Huh, D. (2015). Organs-on-chips at the frontiers of drug discovery. *Nature reviews Drug discovery*, 14(4), 248-260.

Fiehn, O. (2006). Metabolite profiling in Arabidopsis. In *Arabidopsis protocols* (pp. 439-447). Humana Press.

Fiehn, O., Wohlgemuth, G., Scholz, M., Kind, T., Lee, D. Y., Lu, Y., ... & Nikolau, B. (2008). Quality control for plant metabolomics: reporting MSI-compliant studies. *The Plant Journal*, 53(4), 691-704.

Ge, S.X., Son, E.W. & Yao, R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* 19, 534 (2018).

Giobbe, G. G., Michielin, F., Luni, C., Giulitti, S., Martewicz, S., Dupont, S., ... & Elvassore, N. (2015). Functional differentiation of human pluripotent stem cells on a chip. *Nature methods*, 12(7), 637-640.

Haberle, V., Forrest, A. R., Hayashizaki, Y., Carninci, P., & Lenhard, B. (2015). CAGEr: precise TSS data retrieval and high-resolution promoterome mining for integrative analyses. *Nucleic acids research*, 43(8), e51-e51.

Hasegawa, A., Daub, C., Carninci, P., Hayashizaki, Y., & Lassmann, T. (2014). MOIRAI: a compact workflow system for CAGE analysis. *BMC bioinformatics*, 15(1), 1-7.

Hopkinson, B. M., Desler, C., Kalisz, M., Vestentoft, P. S., Juel Rasmussen, L., & Bisgaard, H. C. (2017). Bioenergetic changes during differentiation of human embryonic stem cells along the hepatic lineage. *Oxidative medicine and cellular longevity*, 2017.

Ishihama, Y., Rappsilber, J., & Mann, M. (2006). Modular stop and go extraction tips with stacked disks for parallel and multidimensional Peptide fractionation in proteomics. *Journal of proteome research*, 5(4), 988-994.

Kalra, A., Yetiskul, E., Wehrle, C. J., & Tuma, F. (2018). Physiology, liver.

Kang, Y. B. A., Eo, J., Mert, S., Yarmush, M. L., & Usta, O. B. (2018). Metabolic patterning on a chip: towards in vitro liver zonation of primary rat and human hepatocytes. *Scientific reports*, 8(1), 1-13.

Kido, T., Kouji, Y., Suzuki, K., Kobayashi, A., Miura, Y., Chern, E. Y., ... & Miyajima, A. (2015). CPM is a useful cell surface marker to isolate expandable bi-potential liver progenitor cells derived from human iPS cells. *Stem Cell Reports*, 5(4), 508-515.

Kozakai, K., Yamada, Y., Oshikata, M., Kawase, T., Suzuki, E., Haramaki, Y., & Taniguchi, H. (2013). Cocktail-substrate approach-based high-throughput assay for evaluation of direct and time-dependent inhibition of multiple cytochrome P450 isoforms. *Drug metabolism and pharmacokinetics*, DMPK-13.

Kratz, A., Beguin, P., Kaneko, M., Chimura, T., Suzuki, A. M., Matsunaga, A., ... & Launey, T. (2014). Digital expression profiling of the compartmentalized transcriptome of Purkinje neurons. *Genome research*, 24(8), 1396-1410.

Lassmann, T. (2015). TagDust2: a generic method to extract reads from sequencing data. *BMC bioinformatics*, 16(1), 1-8.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics*, 25(14), 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078-2079.

M. Love, S. Anders and W. Huber, Differential analysis of count data – the DESeq2 package, *Genome Biol.*, 2014, 15(550), 10–1186.

Low, L. A., Mummery, C., Berridge, B. R., Austin, C. P., & Tagle, D. A. (2021). Organs-on-chips: into the next decade. *Nature Reviews Drug Discovery*, 20(5), 345-361.

Ludwig, K. R., Schroll, M. M., & Hummon, A. B. (2018). Comparison of in-solution, FASP, and S-trap based digestion methods for bottom-up proteomic studies. *Journal of proteome research*, 17(7), 2480-2490.

Ma, C., Peng, Y., Li, H., & Chen, W. (2021). Organ-on-a-chip: a new paradigm for drug development. *Trends in pharmacological sciences*, 42(2), 119-133.

Michalopoulos, G. K. (2010). Liver regeneration after partial hepatectomy: critical analysis of mechanistic dilemmas. *The American journal of pathology*, 176(1), 2-13.

Michielin, F., Giobbe, G. G., Luni, C., Hu, Q., Maroni, I., Orford, M. R., ... & Elvassore, N. (2020). The microfluidic environment reveals a hidden role of self-organizing extracellular matrix in hepatic commitment and organoid formation of hiPSCs. *Cell Reports*, 33(9), 108453.

Nishikawa, T., Tanaka, Y., Nishikawa, M., Ogino, Y., Kusamori, K., Mizuno, N., ... & Takakura, Y. (2017). Optimization of albumin secretion and metabolic activity of cytochrome P450 1A1 of human hepatoblastoma HepG2 cells in multicellular spheroids by controlling spheroid size. *Biological and Pharmaceutical Bulletin*, 40(3), 334-338.

Novik, E., Maguire, T. J., Chao, P., Cheng, K. C., & Yarmush, M. L. (2010). A microfluidic hepatic coculture platform for cell-based drug metabolism studies. *Biochemical pharmacology*, 79(7), 1036-1044.

Plessy, C., Bertin, N., Takahashi, H., Simone, R., Salimullah, M., Lassmann, T., ... & Carninci, P. (2010). Linking promoters to functional transcripts in small samples with nanoCAGE and CAGEscan. *Nature methods*, 7(7), 528-534.

Poisson, J., Lemoine, S., Boulanger, C., Durand, F., Moreau, R., Valla, D., & Rautou, P. E. (2017). Liver sinusoidal endothelial cells: Physiology and role in liver diseases. *Journal of hepatology*, 66(1), 212-227.

Poulain, S., Kato, S., Arnaud, O., Morlighem, J. É., Suzuki, M., Plessy, C., & Harbers, M. (2017). 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.

Ramme, A. P., Koenig, L., Hasenberg, T., Schwenk, C., Magauer, C., Faust, D., ... & Dehne, E. M. (2019). Autologous induced pluripotent stem cell-derived four-organ-chip. *Future science OA*, 5(8), FSO413.

Rowe, C., Gerrard, D. T., Jenkins, R., Berry, A., Durkin, K., Sundstrom, L., ... & Hanley, N. A. (2013). Proteome - wide analyses of human hepatocytes during differentiation and dedifferentiation. *Hepatology*, 58(2), 799-809.

Rudnick, D. A., & Davidson, N. O. (2012). Functional relationships between lipid metabolism and liver regeneration. *International journal of hepatology*, 2012.

Schepers, A., Li, C., Chhabra, A., Seney, B. T., & Bhatia, S. (2016). Engineering a perfusable 3D human liver platform from iPS cells. *Lab on a Chip*, 16(14), 2644-2653.

Si - Tayeb, K., Noto, F. K., Nagaoka, M., Li, J., Battle, M. A., Duris, C., ... & Duncan, S. A. (2010). Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. *Hepatology*, 51(1), 297-305.

Takayama, N., Nishimura, S., Nakamura, S., Shimizu, T., Ohnishi, R., Endo, H., ... & Eto, K. (2010). Transient activation of c-MYC expression is critical for efficient platelet generation from human induced pluripotent stem cells. *Journal of Experimental Medicine*, 207(13), 2817-2830.

Xie, Y., Yao, J., Jin, W., Ren, L., & Li, X. (2021). Induction and Maturation of Hepatocyte-Like Cells. *Vitro: Focus on Technological Advances and Challenges. Front. Cell Dev. Biol*, 9, 765980.