Oxygen control in bioreactor drives high yield production of functional hiPSC-like hepatocytes for advanced liver disease modelling

Pedro Vicente, Joana I. Almeida, Inês E. Crespo, Nikolaus Virgolini, Inês A. Isidro, Maria Eréndira Calleja-Cervantes, Juan R. Rodriguez-Madoz, Felipe Prosper, Paula M. Alves, Margarida Serra

Oxygen control in bioreactor drives high yield production of functional hiPSC-like hepatocytes for advanced liver disease modelling

Pedro Vicente, Joana I. Almeida, Inês E. Crespo, Nikolaus Virgolini, Inês A. Isidro, Maria Eréndira Calleja-Cervantes, Juan R. Rodriguez-Madoz, Felipe Prosper, Paula M. Alves, Margarida Serra

Table of Contents

Supplementary Table and Figures

Table S1 - Primers used for RT-qPCR analysis.

Gene	Forward primer (5'-3') sequence	Reverse primer (5'-3') sequence
GAPDH1	AATGAAGGGGTCATTGATGG	AAGGTGAAGGTCGGAGTCAA
A1AT	CACCCACGATATCATCACCA	CCCCATTGCTGAAGACCTTA
AGXT	GAGATCATGGGTGGCCTTG	GTCACGCGGTCCACATTCT
AFP	TTTGGGCTGCTCGCTATGAC	TTGCTGCCTTTGTTTGGAAGC
CYP3A4	AAGTCGCCTCGAAGATACACA	AAGGAGAGAACACTGCTCGTG
ALB	ACACAAGCCCAAGGCAACAA	TATCGTCAGCCTTGCAGCAC
HNF4α	AGAGCAGGAATGGGAAGGAT	GCAGTGGCTTCAACATGAGA
CA9	TTTGCCAGAGTTGACGAGGC	TCTTCCAAGCGAGACAGCAAC

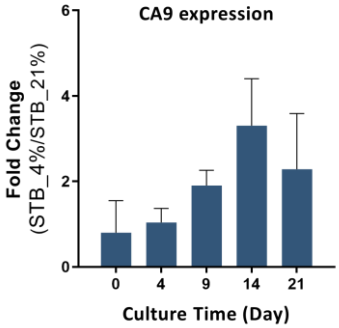

Figure S1. Higher CA9 gene expression in STB_4%O_2. Fold change of CA9 gene expression in STB_4%O_2 normalized for CA9 expression in STB_21%O_2 condition. Gene expression was analyzed by qRT-PCR quantified using the $2^{-\Delta\Delta CT}$ method. CA9, carbonic anhydrase 9.

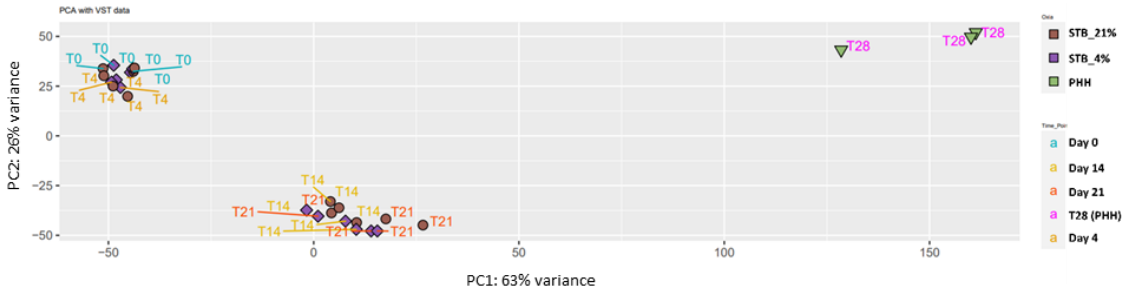

Figure S2. RNA-seq and principal component analysis along hepatic differentiation in STB. Principal components analysis of HLC from STB_21%O_2 (brown), STB_4%O_2 (purple) and PHH (green) cell subsets. The percentage of variance explained by PC1 and PC2 is depicted. T0, day 0; T4, day 4; T14, day 14; T21, day 21 and T28, PHH.

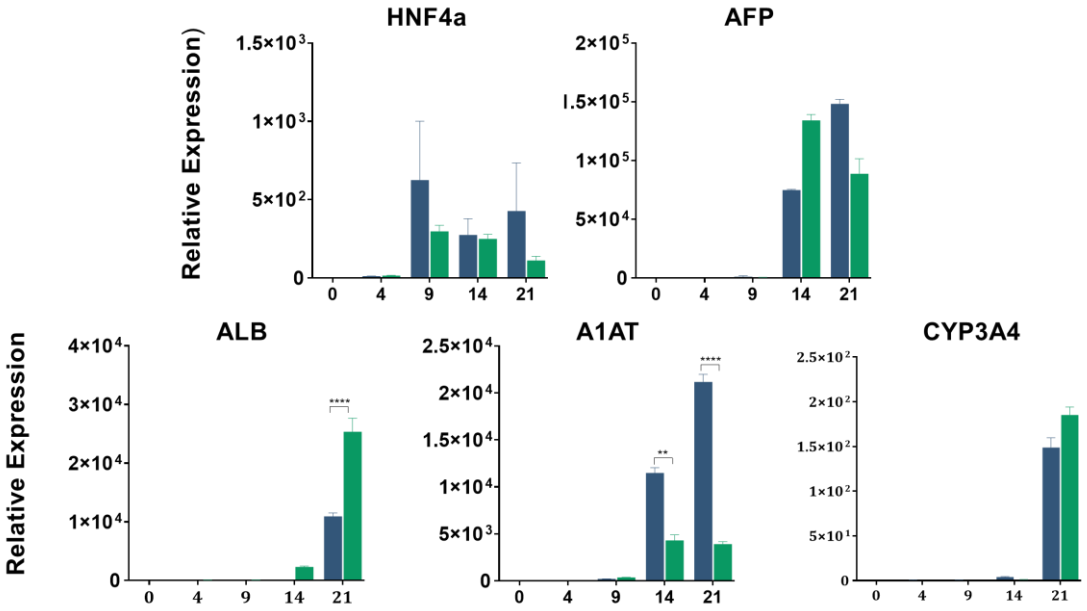

Figure S3. Gene expression of the two different hiPSC lines, PH1 (blue) and ChiPSC18 (green) along hepatic differentiation. Relative expression of hepatic specific markers, alpha fetoprotein (AFP), hepatocyte nuclear factor 4α (HNF4a), albumin (ALB), α-1 antitrypsin (A1AT), and cytochrome P450 subunit 3A4 (CYP3A4) by qRT-PCR quantified using the $2^{-\Delta\Delta CT}$ method relatively to day 0 of differentiation and normalized for GAPDH housekeeping gene.

PH1.HLC	GLC	GLN	LAC	NH3
Metabolic flux (nmol/(10^6.cell.day))	-5.6 ± 0.70	0.9 ± 0.14	8.66.5 ± 1.27	0.4 ± 0.06
CHiPSC18.HLC	**GLC**	**GLN**	**LAC**	**NH3**
Metabolic flux (nmol/(10^6.cell.day))	-5.5 ± 0.88	0.6 ± 0.12	6.5 ± 1.20	0.4 ± 0.06

Figure S4. Metabolic fluxes of GLC, GLN, GLU, LAC and NH3 metabolites.

Supplementary Methods

Primary cultures of human hepatocytes

Primary human hepatocytes (PHH) from an adult male donor were provided by Dr. Bodo Beck (Institute of Human Genetics, University of Cologne Medical Center) in the scope of the ERA-NET E-Rare 3 research program. Briefly, PHH were thawed in Hepatocyte recovery medium (CHRM, Gibco®) according to the manufacturer's instructions and seeded onto 24-well collagen-coated culture plates (BD) at a density of 0.4×10^6 cells/well. PHH were cultured for 3 days in Williams' E (WE) medium (Sigma-Aldrich) supplemented with 10% v/v fetal bovine serum (FBS) and Hepatocyte Maintenance Supplements (Gibco) in a humidified atmosphere of 5% CO_2 at 37°C.

Cell viability. For cell viability assessment two methods were used: (i) the enzyme substrate fluorescein diacetate (Sigma-Aldrich) and TO-PRO®-3 dye (0.5 μM; Thermo Fisher Scientific) and (ii) the trypan blue (Thermo Fischer Scientific) exclusion method, as previously described[1,2]. For the first method, direct staining of the aggregates was performed followed by observation at the inverted fluorescence microscope (DMI6000, Leica, Wetzlar, Germany), as described elsewhere[2]. Representative images were taken using a digital camera (Leica DFC 360 FX). For the latter method, aggregates were centrifuged at 100 × g for 1 min, resuspended in DPBS and centrifuged at 100 × g for 1 min. Samples were then dissociated through incubation with TrypLE Select (Gibco Life Technologies) for 5-10 min at 37◦C with agitation using a thermomixer (Eppendorf). Viable cells were quantified by Trypan Blue exclusion, as described elsewhere[2] Data were processed and visualized using GraphPad Prism.

Aggregate size and concentration. Aggregates sampled from the bioreactor were distributed in 96-well plates (100 μL/well) and counted using an inverted microscope (CKX31, Olympus). Images of cell aggregates were collected (Leica DFC 360 FX) and analysed in Image J open-source software (Rasband, WS, ImageJ, U. S. National Institutes of Health, Bethesda, MD, USA, http://imagej.nih.gov/ij/, 1997–2012) for estimation of aggregate size, as described elsewhere[1]

Bulk RNA-seq analysis

RNA was isolated using RNeasy Mini Kit (Qiagen, 74104) according to the manufacturer's instructions. Roughly 100ng of total RNA were used for the transcriptomic interrogation of samples using Illumina's Stranded Total RNA Prep Ligation with Ribo-Zero Plus according to the manufacturer's instructions. Briefly, cytoplasmic and mitochondrial rRNAs as well as beta globin transcripts were depleted from the samples. The remaining RNA was fragmented and reverse-transcribed. A second strand cDNA synthesis step removed the RNA template while incorporating dUTP in place of dTTP in order to preserve strand specificity. Next, double-stranded cDNA was A-tailed, then ligated to Illumina anchors bearing T-overhangs. PCR-amplification of the library allowed the barcoding of the samples with 10bp dual indexes and the completion of Illumina sequences for cluster generation. Libraries were quantified with Qubit dsDNA HS Assay Kit and their profile was examined using Agilent's HS D1000 ScreenTape Assay. Sequencing was carried out in an Illumina NextSeq500 using single-end, dual-index sequencing (Rd1: 71 cycles; i7: 10 cycles; i5: 10 cycles) at a depth of 30 million reads per sample. Samples were demultiplexed using Illumina bcl2fastqsoftware (version 1.2.4) and quality assessed using FastQC. Next, Cutadapt[3] was used to trim Illumina adapter sequences, polyA-tails and reads deemed too short (less than 20 base pairs). Remaining reads were subsequently aligned to the human reference genome (GRCh38) using STAR (version 2.6.1)[4] with default settings. The number of reads aligned to each gene was determined with featureCounts, from the Subread package[5].

Downstream data analysis was performed using the statistical software R. Genes with a lower expression of 5 counts per million (cpm) in more than 20% of samples in a group were removed from the analysis. DESeq2[6] was employed to perform differential expression analysis. Genes with a fold change of at least ± 2 and an adjusted p-value <0.05 were considered significantly differentially expressed, unless stated otherwise. For visualization purposes and to remove unwanted variability, normalized and variance stabilized data was further processed using the RemoveBatchEffect function from the Limma package[7]. To identify pathway and gene ontology

terms associated to the differential gene list caused by hypoxic conditions, the R package clusterProfiler[8] was employed. Batch effect associated with library preparation was removed by introducing "library" variable into the model formula (~library+condition) for DESeq2 analysis, and for gene ontology categories TCGABiolinks[9] was employed.

Metabolic profiling

Culture supernatants were collected after centrifugation (300g, 5 min) at specific time points (day 0,1, 3, 4, 9, 14 and day 21) representing relevant time points in a hepatic differentiation process. Supernatants were analyzed in Cedex Bio Analyzer (ROCHE) to calculate the concentration of four metabolites of interest (Glucose – GLC; Glutamine – GLN; Lactate – LAC; Ammonia – NH3;). Specific metabolic fluxes were calculated according to the general mass balance equation: $q = (\Delta C/\Delta t - D \times (Cin - Cout))/XV$ average, where q is the specific synthesis rate, $\Delta C/\Delta t$ is change rate in the supernatant, D is the dilution rate, Cin and Cout are the inlet and outlet concentrations, and XV average is the average of viable cell concentrations during the time period Δt.

Supplementary References

1. Abecasis, B. *et al.* Expansion of 3D human induced pluripotent stem cell aggregates in bioreactors: Bioprocess intensification and scaling-up approaches. *J Biotechnol* **246**, 81–93 (2017).

2. Serra, M. *et al.* Improving expansion of pluripotent human embryonic stem cells in perfused bioreactors through oxygen control. *J Biotechnol* **148**, 208–215 (2010).

3. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**, 10–12 (2011).

4. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

5. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

6. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

7. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**, Article3 (2004).

8. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

9. Colaprico, A. *et al.* TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* **44**, e71 (2016).