Electronic Supplementary Material (ESI) for Molecular Omics.
This journal is © The Royal Society of Chemistry 2022

Supplementary File 2

1. Honeycomb technology

A silicon mold was fabricated using SU-8-based negative photolithography (MicroChem, Newton, MA, USA) on a silicon wafer to make honeycomb microwells (diameter: 126 μm; depth: 148 μm) and the surface was coated with CHF3, as previously described.[1] PDMS prepolymer and curing agent (Silpot 184, Dow Corning Toray, Tokyo, Japan) were mixed at 10:1 (w/w), poured onto the silicon mold and cured at 70℃ for at least 2 hours. 15 mm-diameter rounds were cut from the PDMS microwell to fit the 24-well tissue culture plate.

2. RT-qPCR assays

At the end of the experiments, 500 μl of Trizol™ Reagent (Life Technologies) were added in each biochip/well and cells were lysed by pipetting. The Trizol solutions containing cell lysates were stored at -80℃ until use. Total RNAs were extracted and purified from the Trizol solutions using a hybrid protocol combining Trizol and the RNeasy Mini Kit (QIAGEN 74104) following the manufacturer's instructions. Briefly, 100 μL of chloroform were added to Trizol solution, vortexed and incubated 5 min at room temperature (RT). After centrifugation (15 min at 120000 g), the upper aqueous phase was collected, mixed with 250 μL of isopropyl alcohol and incubated at RT for 10 min. The RNA precipitate forms a gel-like pellet after centrifugation at 120000 g for 15 min. After supernatant removing, 1 mL of ethanol was added to RNA pellet and the solution centrifugated 5 min at 7500 g. Finally, ethanol was removed, the RNA pellet dried (air dry for 5-10 min) and 20 μL of RNase-free water were added to the RNA pellet.

The concentrations and qualities of the RNAs extracted were assessed using a BioSpec-nano (Shimadzu Scientific Instruments). Reverse-transcription into cDNA was performed from 0.5 μg of total RNA using the ReverTra Ace qPCR RT Master Mix with gDNA Remover (TOYOBO). Real-time quantitative PCR was then performed with the THUNDERBIRD SYBR qPCR Mix (TOYOBO) according to the manufacturer's protocol and a StepOnePlus Real-Time PCR system (Applied Biosystems). Primer sequences of genes are shown in Table S1. β-Actin

was used as the reference gene and hiPSCs in stage 1 of maturation (iPSC-ST1) were used as the reference sample for the normalization of gene expression data.

Table S1. Primers used in RTqPCR analysis.

Gene*	Sequences
PDX1	f_ CTTGGAAACCAACAACTATTCAC r_ ATTAAGCATTTCCCACAAACA
NKX2.2	f_ ATGTAAACGTTCTGACAACT r_ TTCCATATTTGAGAAATGTTTGC
NKX6.1	f_ TCAACAGCTGCGTGATTTTC r_ CCAAGAAGAAGCAGGACTCG
NGN3	f_ TTGCGCCGGTAGAAAGGATGAC r_ TCAGTGCCAACTCGCTCTTAGG
INS	f_ CATCAGAAGAGGCCATCAAG r_ TCTTGGGTGTGTAGAAGAAGC
GCG	f_ CAGAAGAGGTCGCCATTGTT r_ TGGCTAGCAGGTGATGTTGT
SLC30A8	f_ GAGCGCCTGCTGTATCCTG r_ TGCACAAAAGCAGCTCTGAC
MAFA	f_ GTCAGCAAGGAGGAGGTGATC r_ TCACCAACTTCTCGTATTTCTCCT
GLUT2	f_ TGGGCTGAGGAAGAGACTGT r_ CCCATCAAGAGAGCTCCAACT
GCK	f_ CACTGCTGAGATGCTCTTCGAC r_ CCACGACATTGTTCCCTTCTG
SST	f_ CCCAGACTCCGTCAGTTTCTG r_ TCATTCTCCGTCTGGTTGGGT
PCSK1	f_ TGATCCCACAAACGAGAACAAAC r_ TGTGATTATTTGCTTGCATGGCA
UCN3	f_ GGCCTCCCCCACAAGTTCT r_ TCTCTTTGCCCTCCTCCTCC
PTF1A	f_ ATAGAAAACGAACCACCATTTGAGT r_ CAGGACGTTTTCTGGCCAGA
B_actin	f_ CCTCATGAAGATCCTCACCGA r_ TTGCCAATGGTGATGACCTGG

***PDX1*: pancreatic and duodenal homeobox 1 ; *NKX2.2* : NK2 homeobox 2 ; *NKS6.1* : NK6 homeobox 1 ; *NGN3* : neurogenin 3 ; *INS* : insulin ; *GCG* : glucagon ; *SLC30A8* : solute carrier family 30 member 8 ; *MAFA* : MAF bZIP transcription factor A ; *GLUT2* : solute carrier family 2 member 2 ; *GCK* : glucokinase ; *SST* : somatostatin ; *PCSK1* : proprotein convertase subtilisin/kexin type 1 ; *UCN3* : urocortin 3 ; *PTF1A* : pancreas associated transcription factor 1a.

3. Immunostaining

After transfer to an untreated TCPS 24-well plate, the spheroids were washed with phosphate buffer saline solution (PBS) and fixed in paraformaldehyde 4% at 4°C overnight. In order to perform the immunohistochemistry (IHC) staining in a 3D structure, the spheroids were permeabilized with 1% Triton X100 in PBS for 3 hours at 4°C and washed 3 times with PBS for 30 min. Then, the spheroids were blocked with a gelatin buffer for 24 hours at 4°C. Primary antibodies (Table S2) were incubated for 24 hours at 4°C in a BSA/PBS solution. After washing with PBS, secondary antibodies (Table S2) were further incubated overnight in a BSA/PBS solution at 4°C in the dark. Finally, the nuclei were stained with DAPI (342-07431, Dojindo) at 1/1000 for 30 min at room temperature (RT) in the dark. All the incubations and washing steps were carried out using a shaker. Observations were made using an Olympus IX-81 confocal laser-scanning microscope.

Table S2. Primary and secondary antibodies used for islets immunostaining*.

Immunostaining	Primary antibody	Secondary antibody
Insulin	*Mouse anti-insulin (ab6995)* ***Abcam***	*Goat anti-mouse Alexa Fluor® 594 (ab150116)* ***Abcam***
Glucagon	*Sheep anti-glucagon (ab 36232)* ***Abcam***	*Donkey anti-sheep Alexa Fluor 647 (ab 150179)* ***Abcam***
MAFA	*Rabbit anti-MAFA (ab26405)* ***Abcam***	*Goat Anti-rabbit Alexa Fluor® 680 (A-21109)* ***ThermoFisher***
PDX1	*Goat anti-PDX1 (ab347383)* ***Abcam***	*Donkey Anti-goat Alexa Fluor® 488 (ab150129)* ***Abcam***
GCK	*Mouse anti-GCK (sc-17819)* ***Santa Cruz***	*Donkey anti-mouse Alexa Fluor® 647 (ab150107)* ***Abcam***

* All antibodies were diluted in the range recommended by the manufacturers.

4. CAGE transcriptome profiling

NanoCAGE libraries were generated and sequenced as previously described by Poulain et al.2017.[2] Total RNAs were extracted from cell samples conserved in Trizol™ Reagent (Life

Technologies) and the PureLink RNA mini kit (ThermoFisher) following the manufacturer's instructions. Fifty nanograms of each total RNA were retrotranscribed to produce cDNA samples individually tagged by specific barcode sequences. Samples were subsequently PCR-amplified and multiplexed to produce a nanoCAGE library that was sequenced paired-end (9.1 pM + 10% PhiX) on a MiSeq system with the reagent kit v3 150-cycles (Illumina).

NanoCAGE sequencing data (FASTQ files) were subsequently processed for CAGEscan analysis with the cagescan-pipeline workflow (https://gitlab.com/mcfrith/cagescan-pipeline and https://github.com/oist/plessy_CAGEscan_Nextflow).[3] Briefly, Tagdust v2.33 was used to demutiplex the sequencing reads and filter out reads originating from ribosomal RNA or oligo-artifacts.[4] Remaining reads were then aligned on the human genome (hg19 and hg38) with LAST (https://gitlab.com/mcfrith/last).[5] Resulting .BED files were processed with custom R scripts using the CAGEr package[6] to produce gene expression tables that were uploaded online for differential gene expression (DGE) and pathway gene set enrichment (PGSEA) analysis with iDEP version 0.92 (http://bioinformatics.sdstate.edu/idep92/).[7] The EdgeR:log2(CPM+c) option was used to transform read counts for clustering. The DGE analysis was conducted with limma-voom,[8] applying a false discovery rate (FDR) cutoff of 0.2 and setting the minimum fold change to 1.1. Gene Set Enrichment Analysis (GSEA) and Pathway Gene Set Enrichment Analysis PGSEA was performed with a pathway significance cutoff set at 0.2. The ISMARA webserver (https://ismara.unibas.ch/mara/) was used for Motif Activity Response Analysis (MARA).[9] MetaboAnalyst 4.0 was used to perform multivariate data analysis with the transcription factor motif activity matrix extracted from ISMARA.[10]

5. Metabolomic analysis

The collected culture medium (250 µL) was completed with 500 µL of 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 shaking at 1500 rpm. Insoluble material was removed by two centrifugations steps at

14000 rpm for 10 min. Three aliquots of each extract (100 µL) were dried for 4 h at 35 °C in a speed-vac and stored at -80°C until analysis.

For GC-MS injection, samples were taken out of -80°C, warmed for 15 min and dried again for 1.5 h at 35 °C before adding 10 µL of 20 mg/mL methoxyamine in pyridine and the reaction was performed for 90 min at 30 °C with shaking. Then, 90 µL of N-methyl-N-trimethylsilyl-trifluoroacetamide (MSTFA, Regis Technologies) was added and the reaction continued for 30 min at 37 °C. After cooling, 100 µL was transferred to an Agilent vial for injection. Four hours after derivatization, 1 µL of sample was injected in splitless mode on an Agilent 7890B gas chromatograph coupled to an Agilent 5977A quadrupole mass spectrometer. The column was a Rxi-5SilMS from Restek (30 m with 10 m Integra-Guard column - ref 13623-127). An injection in split mode with a ratio of 1:30 was systematically performed for saturated compound quantification. The oven temperature ramp was 60 °C for 1 min then 10 °C/min to 325 °C for 10 min. Helium constant flow was 1.1 mL/min. Temperatures were the following: 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.

Raw Agilent datafiles were analyzed with AMDIS (http://chemdata.nist.gov/mass-spc/amdis/). The Agilent Fiehn GC/MS Metabolomics RTL Library (version June 2008) was employed for metabolite identifications. Peak areas were determined with the Masshunter Quantitative Analysis (Agilent) in splitless and split 30 modes. Because automated peak integration was occasionally erroneous, integration was verified manually for each compound and peak areas were normalized to ribitol. Metabolite contents are expressed in arbitrary units (semi-quantitative determination).

Specific references associated to the protocols:

1 M. Shinohara, H. Kimura, K. Montagne, K. Komori, T. Fujii and Y. Sakai, *Biotechnol. Prog.*, **2014**, 30, 178-187.

2 S. Poulain, S. Kato, O. Arnaud, J. E. Morlighem, M. Suzuki, C. Plessy and M. Harbers, in *Promoter Associated RNA. Methods in Molecular Biology*, ed. S. Napoli, Humana Press, New York, 2020, NanoCAGE: A method for the analysis of coding and noncoding 5′-capped transcriptomes, 57-109.

3 C. Plessy, N. Bertin, H. Takahashi, R. Simone, M. Salimullah, T. Lassmann, M. Vitezic, J. Severin, S. Olivarius, D. Lazarevic, N. Hornig, V. Orlando, I. Bell, H. Gao, J. Dumais, P. Kapranov, H. Wang, C. A. Davis, T. R. Gingeras, J. Kawai, C. O. Daub, Y. Hayashizaki, S. Gustincich and P. Carninci, *Nat. Methods*, 2010, **7**, 528-534.

4 T. Lassmann, *BMC Bioinformatics*, 2015, **16**, 24.

5 S. M. Kiełbasa, R. Wan, K. Sato, P. Horton and M. C. Frith, *Genome Res.*, 2011, **21**, 487-493.

6 V. Haberle, A. R. Forrest, Y. Hayashizaki, P. Carninci and B. Lenhard, *Nucleic Acids Res.*, 2015, **43**, e51.

7 S. X. Ge, E. W. Son and R. Yao, *BMC Bioinformatics*, 2018, **19**, 534.

8 C. W. Law, Y. Chen, W. Shi and g. k. Smyth, *Genome Biol.*, 2014, **15**, R29.

9 P. J. Balwierz, M. Pachkov, P. Arnold, A. J. Gruber, M. Zavolan and E. van Nimwegen, *Genome Res.*, 2014, **24**, 869-884.

10 Z. Pang, J. Chong, G. Zhou, GD. A. de Lima Morais, L. Chang, M. Barrette, C. Gauthier, P. E. Jacques, S. Li and J. Xia, *Nucleic Acids Res.,* 2021, **49**, W388-W396.