#### Supplementary experimental material

#### **Materials and Methods**

## Maintenance and differentiation of EpiSC in culture

Four EpiSC lines were used: Mixl1-early EpiSC (PS4) from the epiblast of E6.0 pre-streak stage embryo and Mixl1-intermediate (CAV1) and -late (CAV2 and CAV4) EpiSC from the epiblast of E5.5 cavity-stage embryos. EpiSCs were maintained as an adherent layer on fibroblasts as previously described [1]. For differentiation of EpiSC, ROCK inhibitor (Tocris,10µM)-treated EpiSC colonies were dissociated into single cells and approximately 500 cells were seeded into each microwell on the Aggrewell<sup>™</sup> 400 plate (STEMCELL Technologies) to form embryoid bodies. Embryoid bodies were cultured in Aggrewell<sup>™</sup> 400 plate on days 0-2 and in ultra low attachment plates (Corning) on days 2-4.

To induce cell differentiation, embryoid bodies were cultured in RPMI media (RPMI, 20% Knockout serum replacement, Glutamax, Penicillin & Streptomycin and  $\beta$ -Mercaptoethanol ) supplemented with combinations of Activin A (R&D, 100ng/mL), BMP4 (R&D , 50ng/mL), Nodal (R&D, 100ng/mL) and PI3 kinase inhibitor (Ly 294002, Sigma, 10 $\mu$ M). The optimal protocols for the generation of definitive endoderm-like cells (days 0-4: Activin + Ly or days 0-4: Nodal) were used subsequently for endoderm differentiation of the EpiSCs.

To optimize the protocol for the differentiation of foregut endoderm, embryoid bodies generated from Mixl1intermediate EpiSCs were cultured for 2 days in RPMI media supplemented with Activin A and Ly. Day 2 embryoid bodies were dissociated into single cells and cultured on Matrigel<sup>™</sup> (BD Biosciences) with Activin A and Ly-supplemented RPMI media for another 2 days. On day 4 to 6, the cells were treated with RPMI media, Differential response of Epiblast stem cells to Nodal and Activin signalling: a paradigm of early endoderm development in the embryo. Kaufman-Francis *et al.*, (2014) supplemented with 4 different combinations of FGF10 (R&D, 50ng/mL), KAAD-cyclopamine (cyc, Merck Millipore, 0.25µM), BMP4 (50ng/mL) and retinoic acid (RA, Sigma, 2µM) as outlined in Supplementary figure S5a, to test the efficiency of differentiation.

From the aforementioned studies, an optimal protocol was devised for testing the ability of the EpiSC to differentiate to endoderm progenitors of pancreatic and liver cells after culturing in either Activin + Ly or Nodal on days 0-4. Embryoid bodies derived from the three Mixl1 categories were cultured in Activin A (100ng/mL) and Ly (10 $\mu$ M) or Nodal (100ng/mL)-supplemented RPMI media for 2 days. Embryoid bodies were dissociated on day 2 and cultured on Matrigel<sup>TM</sup> in the Activin + Ly or Nodal-supplemented RPMI media for another 2 days. Thereafter, cells were cultured in the same protocol: FGF10 (50ng/mL) and cyc (0.25 $\mu$ M) in RPMI media on days 4-6, and DMEM media (Knockout DMEM, 20% Knockout serum replacer, Glutamax, Penicillin & Streptomycin, B27 and  $\beta$ -Mercaptoethanol) supplemented with FGF10 (50ng/mL), RA (2 $\mu$ M) and cyc (0.25 $\mu$ M) or SB431542 (SB, 10 $\mu$ M) on days 6 to 9 (Supplementary Figure S5).

## Gene expression analysis

At specific time points of directed differentiation, cells were harvested for gene expression analysis. RNA extracted from cell samples with RNeasy Micro Kit (Qiagen) was reverse transcribed to cDNA using the RT<sub>2</sub> HT First Strand Kit (Qiagen). PCR was performed with BioMix<sub>TM</sub> Red (Bioline) using primers for foregut endoderm markers. qPCR was performed with the custom RT2 Profiler PCR Array (proprietary primer sequences, Qiagen) using RT<sub>2</sub> SYBR Green ROX FAST Mastermix (Qiagen). Relative difference in threshold cycles of each gene of interest, normalised against Day-0 EpiSC value, was calculated using the -ΔΔCT method.

To assess the differentiation of endoderm precursors in Day-4 embryoid bodies, the expression of *Mixl1* (primitive streak marker) *Sox17* and *Foxa2* (definitive endoderm markers), *Oct4* and *Nanog* (pluripotency markers) *T* and *Meox1* (mesoderm markers), *Sox7* (extraembryonic endoderm marker) and *Sox1* and *Pax6* 

(ectoderm markers) was determined by qPCR. To detect the presence of foregut endoderm cells, the expression of Capn6, Rbm47, B4galt6, Igfbp5, Rhou, 1810019J16Rik, and Cldn4 (see later section: "Identification of endoderm-expressed genes in the foregut") was assessed by PCR. Actb was used as the loading control. For cells harvested on day 6 of the differentiation protocol, the expression of Pyy, Hnf1b, Onecut1, Hnf4a, Nkx2-1, Tbx1, Sox2 (foregut markers) and Cdx2 (hindgut marker) and were analysis. For cells harvested on day 9 of the differentiation protocol, the expression of Prox1 and Pdx1 (pancreatic progenitors markers), Afp and Hhex (hepatic progenitor markers) and Krt19 and Itgb4 (cholangiocyte markers) was analysed by qPCR.

To assess the differences in the level of expression of TGFβ signalling target and pathway genes in day 4 embryoid bodies derived from *Mix1*-early, Mixl1-intermediate and Mixl1 -late EpiSC in response to Nodal or Activin A, a high throughput gene expression analysis (BioMarkTM HD System, Fluidigm) was performed using the TGFβ signalling targets RT<sub>2</sub> Profiler PCR array (Qiagen, Cat # PAMM-235Z). Data generated was analysed using RT<sub>2</sub> Profiler PCR Array Data Analysis version 3.5 (Qiagen). Gene functional annotation of specific subset of genes (Supplementary Figure S6) was performed using DAVID Bioinformatics Resources

(http://david.abcc.ncifcrf.gov/)

## Fluorescence activated cell sorting

EpiSCs were detached from the fibroblast layer by collagenase treatment and dissociated with TryPLE Select (Invitrogen). CXCR4 staining was performed using mouse CXCR4 phycoerythrin MAb, Rat IgG2B (R&D FAB 21651P) according to manufacturer's protocol. Rat IgG2B Phycoerythrin isotype control (R&D IC013P) was used for isotype control. The cells were sorted by FACS Canto and quantification of CXCR4+ cells was performed using the FACS Diva software (BD).

## Immunostaining

Definitive endoderm-like cells in Day-4 embryoid bodies were detected by immunostaining using Foxa2 and Sox17 antibodies. Cryosectioned samples were rinsed in PBS for 2x5minutes and blocked in CAS-Block (Invitrogen) for 15 minutes and stained with primary antibodies for 30 minutes at room temperature. Following rinsing with PBS for 3 x 5 minutes, the samples were incubated with secondary antibodies and DAPI (Sigma) for 30 minutes. For the detection of definitive endoderm, primary antibodies: mouse anti Sox17 (1:1000; Abcam, ab84990) and goat anti Foxa2 (1:1000; Santa Cruz, SC-6554), and secondary antibodies: Alexa Fluor 594 donkey anti-goat IgG (1:1000) were used.

# Identification of endoderm-expressed genes in the foregut

RNA samples extracted from dissected foregut endoderm and combined headfold and heart tissue (ectoderm and mesoderm) were prepared for microarray analysis using Affymetrix mouse MOE430 GeneChips. Two independent experiments were carried out (Supplementary Fig. S4a). The microarray data were normalized using the MAS5 algorithm implemented in GCOS 1.2 (Affymetrix). Genes that were called as present in the endoderm and which showed greater than 2-fold higher expression in the endoderm than the mesoderm/ectoderm samples in both experiments (Supplementary Table S2), and for which no expression data from early postimplantation was previously published were selected for further analysis by in situ hybridization. The gene expression data of this study are available at

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE36453

## In situ hybridization

Embryos were collected, fixed in 4% paraformaldehyde, dehydrated through a methanol series and stored at -20<sub>o</sub>C. Automated whole mount in situ hybridization was carried out on an InSituPro (Intavis Bioanalytical Instruments) machine, using a protocol [3] modified as described previously [2]. Antisense riboprobes for endoderm-expressed genes were generated (Ampliscribe, Epicentre) from plasmid inserts amplified with

either M13 forward and reverse primers or with specific primers designed to incorporate a T7 promoter into the reverse oligonucleotide (Supplementary Table S1).

# References

[1] Kojima Y, Kaufman-Francis K, Studdert JB, Steiner KA, Power MD, Loebel DA, et al. The transcriptional and

functional properties of mouse epiblast stem cells resemble the anterior primitive streak. Cell stem cell.

2014;14(1):107-20. doi: 10.1016/j.stem.2013.09.014

[2] Loebel DA, Studdert JB, Power M, Radziewic T, Jones V, Coultas L, Jackson Y, Rao RS, Steiner K, Fossat N,

Robb L, Tam PP. Rhou maintains the epithelial architecture and facilitates differentiation of the foregut

endoderm. Development 2011; 138:4511-22. doi: 10.1242/dev.063867.

[3] Wilkinson DG, Nieto MA. Detection of messenger RNA by in situ hybridization to tissue sections and whole mounts. Methods Enzymol. 1993;225:361-73.