Developmental Cell, Volume *51*

Supplemental Information

Self-Organization of Mouse Stem Cells into an Extended Potential Blastoid

Berna Sozen, Andy L. Cox, Joachim De Jonghe, Min Bao, Florian Hollfelder, David M. Glover, and Magdalena Zernicka-Goetz

Supplemental Figure Legends

Figure S1. Generation of blastoids. Related to Figure 1. A. Images of a single microwell from an AggreWell containing an EPSC-aggregate and single TSCs at 0h, and the emergence of a blastocyst-like structure over 96h. Bar=200μm. Image on the right shows a representative blastocyst-like structure built from EPSCs that express nuclear PDGFRa-H2B-GFP and TSCs that express ubiquitous eGFP. Bar=20μm. **B.** Low magnification transmitted-light image of EPS-blastoids generated in a typical experiment after 96h. **C.** Frequency of cystic structure formation under normoxia (20% O2) versus hypoxia (5% O2) from ESC (ES-blastoid, SB-Lif) or EPSC (EPS-blastoid, SB-EP) cells. After 96h of co-culture, all aggregates from an AggreWell, containing 1200 microwells were collected and scored over 3 separate experiments per group. Cystic structures that had a single layer of TSC-derived epithelium containing an acentric EPSC or ESC-derived inner compartment were considered as synthetic blastocyst (SB). Two-sided Student's t test, P< 0.001. Error bars, SEM. **D.** EPS-blastoid built from EPSCs that express nuclear PDGFRa-H2B-GFP and TSCs that express ubiquitous eGFP at 96h showing all three blastocyst lineages. Nanog, grey. Note PE-like cells align with blastocoel cavity. **E.** Montage panel for 59 Z-sections of a representative EPS-blastoid generated with EPSCs (expressing membrane CAG:GFP) and TSCs (expressing ubiquitous eGFP). Please note that the outside TE-like epithelium expresses ubiquitous-GFP throughout. Membrane CAG:GFP signal is confined to inner compartment and co-expresses Nanog (magenta). **F.** A representative EPS-blastoid at 100h generated with dual reporter EPSCs (membrane mTmG-tomato, nuclear PDGFRa-H2B-GFP) and TS:GFP (ubiquitous eGFP). Structure stained with Sox17 to confirm PE identity. Please note that no contribution from EPSCs is observed in the TE-like layer. **G.** EPSC-only aggregation experiment (formed under the same sequential culture conditions and duration as for EPS-blastoid experiments). No cavitation or CDX2 positive cell specification is observed. **H.** Left: Quantitation of PE positive cells per structure. Early blastocyst (EB, E3.5, *n*=11), late blastocyst (LB, E4.75, *n*=7) and EPS-blastoid (SB, 96h, *n*=20). Error bars, SEM. Right: Quantification of cell numbers in EPSC/Epi, TSC/TE, PE/PE-like layers in early blastocyst (EB, E3.5, *n*=13), late blastocyst (LB, E4.75, *n*=7) and EPS-blastoid (SB, 96h, *n*=27). Error bars, SEM.

Figure S2. Transcriptome analyses of lineages specified in synthetic blastocysts. **Related to Figure 2. A.** Upper illustration depicts cell collection for single-cell RNA-sequencing (sc-RNAseq) from ESC- (generated in media containing Serum/Lif, SB-Lif) or EPSC-derived (generated under EP conditions, SB-EP) blastocysts. UMAP dimensional reduction illustrates three clearly defined clusters in stem cell-derived blastocysts representing the three blastocyst lineages. **B. Top:** Dot-plot showing proportion and scaled expression levels of lineage fate determining genes within Epi-like (Epi-L), TE-like (TE-L) or PE-like (PE-L) cell clusters of ES- (SB-Lif) or EPS-blastoids (SB-EP) following sc-RNAseq. **Bottom:** UMAP dimensional reduction plots selected lineage specific-genes within Epi-like, TE-like or PE-like cell clusters of EPS-blastoids following sc-RNAseq. **C.** Pie charts showing percentage of cells with PE-like identity that presumably become specified from ESC- or EPSC-derived ICM-like compartment. *n*= 52/804 in ESCs; 123/1210 EPSCs. Calculated from number of PE-like cells relative to EPI-like cells in the PE-cluster from sc-RNAseq data. **D**. Volcano plots illustrating the DEGs (absolute value of log2FC higher than 1 and a Bonferroni corrected p-value lower than 0.005) for each lineage between the downsampled SB-EP and natural embryo E4.5. Downsampling of the synthetic structures was carried out to account for discrepancies in sample sizes between the samples (PE-L = 68 cells, Epi-L = 32 cells, TE-L = 38 cells). **E.** STRING

v10 (Szklarczyk et al., 2015) Biological Process Gene Ontology (GO) analysis of DEGs (absolute value of log2FC higher than 0.7 and a Bonferroni corrected p-value lower than 0.01) illustrating pathway enrichment and depletion between SB-EP and E4.5. The 8 enriched pathways in SB-EP were plotted alongside the 42 pathways with the lowest False Discovery Rated (FDR) **F.** Phylogenetic tree analysis between the lineages of the downsampled SB-EP, downsampled SB-Lif and E4.5 blastocyst on the dataset's variable features, illustrating lineage hierarchy conservation throughout the synthetic structures. **G.** UMAP dimensional reduction plot of the cells from the synthetic and natural embryos before random downsampling according to the colour-coded represented lineages, identified according to Klf2, Gata6 and Gata3 lineage marker expression.

Figure S3. Molecular features of PE and DE derivatives. **Related to Figure 2. A.** UMAP dimensional reduction shows PE cluster in stem cell-derived blastocysts and E3.5, E4.5, and DE cluster in E7.5 (Nowotschin et al., 2019). **B.** PE (*Gata6, Sox7, Rhox5*) and DE (*Cer1, Nnat, Apela*) marker expression on the dimensional reduction map. **C.** Wilcoxon rank sum test results between the synthetic PE from the extended pluripotency protocol and DE cells from the natural embyo at E7.5. p_val stands for p-value. Log2FC stands for log2- foldchange (positive values mean upregulation in PE from the extended pluripotency protocol and negative values signify downregulation).

Figure S4. Molecular features of endoderm derivatives in ES- and EPS-blastoids. **Related to Figure 2. A.** Violin graphs showing the distribution of expression levels of PE-specific genes within PE-like cell cluster of ES- (SB-Lif) or EPS-blastoids (SB-EP). **B.** UMAP clustering graphs comparing cell populations positive for *Gata6, Gata4, Pdgfra* or *Nanog,* within the PE-like cell cluster from either ESC- or EPS-derived blastoids. **C.** Dot**-**plot showing proportion and scaled expression levels of PE fate determining genes within PE-like cell cluster of ES- or EPS-blastoids in each batch following sc-RNAseq.

Figure S5. Self-organisation of blastoids into post-implantation structures *in vitro*. Related to Figure 5. A. Frequency of EPS-blastoids developed using Matrigel or 2D IVC methods. On average 11.8% of structures self-organised into a cylindrical morphology comprising aligned EPSC/TSC compartments. 160 structures scored over 8 separate experiments. 4 experiment excluded because of a total developmental failure. Lower panel shows a representative post-implantation-like structure formed by the self-organisation of EPS-blastoids after 24h *in vitro.* PDGFRa-H2B-GFP endogeneous signal defines the PE cell-derived VE-like layer. TSC-derived ExE-like compartment stained for Tfap2c (red), EPSC-derived Epi-like compartment stained for Oct4 (yellow). *n*=5 structures, 3 separate experiments. **B.** EPS-blastoids built from CAG:GFP EPSCs and wild-type TSCs transferred into IVC that showed post-implantation-like morphology transition within 20h. *n*=24 structures, 11 separate experiments. **C.** The vast majority of ESC-derived blastoids died shortly after transfer into IVC (91.7%, image on the left). 8.3% of the ES-blastoids formed aggregates with non-polarised ESC-TSC compartments (image on the right); 120 structures scored over 6 separate experiments. 5 experiments excluded because of total developmental failure. Graph on the right shows frequency of EP-blastoids developed using Matrigel or 2D IVC methods. Error bars, SEM. **D.** Individual frames from time-lapse recordings of development of EPS-blastoids built from CAG:GFP EPSCs and wild-type TSCs transferred into IVC showing post-implantation-like morphology transition over 47h. White dashed line outlines the EPSC-derived ICM-like compartment. Please note

that CAG promoter-driven transgenes become silenced in extraembryonic lineages in this particular line (Rhee et al. 2006; Griswold et al. 2011, Abe and Fujimora, 2013, Bedzhov and Zernicka-Goetz, 2014) Time-lapse images captured at 30 min intervals. Representative series from 3 separate time-lapse movies of 3 structures. Bar=20μm. Please note that the structure has attached to the imaging dish (ibidi) and spread on the surface, and that by the final time points the structure is growing upwards, with the EPS-derived EPI (as marked by OCT4 and CAG-GFP) positioned on top (panel on the right). **E.** The formation of PE cell-derived VE-like layer in IVC. The PDGFRa-H2B-GFP endogeneous signal defines the VE-like layer.

Figure S6. Decidualisation induced by implanting EPS-blastoids *in vivo*. Related to Figure 6. **A.** Immunohistochemical staining of decidua at 7.5 d.p.c and EPS-blastoid (synthetic blasocyst; SB)-induced decidua 4 days post-transfer revealing Ptgs2. Distinguishable PDZ and SDZ reveal sustainable decidua patterning. *n*=2 NB-induced; n=2 EPS-blastoid-induced decidua**. B.** Immunohistochemical staining of the decidua at 5.5 d.p.c and EPS-blastoid-induced decidua 3 or 4 days post-transfer revealing Ki67. Normal tissue proliferation was observed in all cases. *n=* 2 NB-induced; n=2 SB-induced deciduae**. C.** Immunohistoflourescence staining of the core trophoblast-stem-cell self-renewal markers showing expression restricted to ExE lineage of implanted conceptus. Decidua at 5.5 d.p.c. Cdx2, green; Eomes, white. **D.** EPS-blastoid-induced decidua with GFP (green), SOX2 (red), EOMES (white) positive signals in implanting conceptus. *n*=5 SB-induced decidua. **E.** EPS-blastoid-induced decidua with GFP (green) positive signals in implanting conceptus. Note LE is broken down as implantation progresses. Asterisks mark GFP-positive signal from the implanting conceptus, yellow arrows indicate broken epithelium. *n*=3 SB-induced decidua. **F.** *Upper panel:* Decidua at 5.5 d.p.c. showing LE before closure and disappearance at implantation site. Serial section of same decidua stained for Ptgs2 defining implantation site (right). e, embryo. *n*=3 NB-induced decidua. *Lower panel:* EPS-blastoid-induced decidua dissected 3 days after transfer showing LE before closure and disappearance at implantation site. Serial section of same decidua stained for Ptgs2 defining implantation site (right). Yellow asterisk marks implantation site. *n*=6 SB-induced decidua.

Supplementary Table 1. Oligonucleotide sequences used to barcode the final library. Related to STAR methods, droplet microfluidics scRNA-seq and library preparation

Supplementary Table 2- Sequencing and data pre-processing parameters. Related to STAR methods, library sequencing, data pre-processing RNA-sequencing and mapping of reads.

Supplementary Table 3- Commands for UMAP dimensional reduction plotting and cluster identification. Related to STAR methods, batch-effect correction and UMAP dimensional reduction.

Figure S1

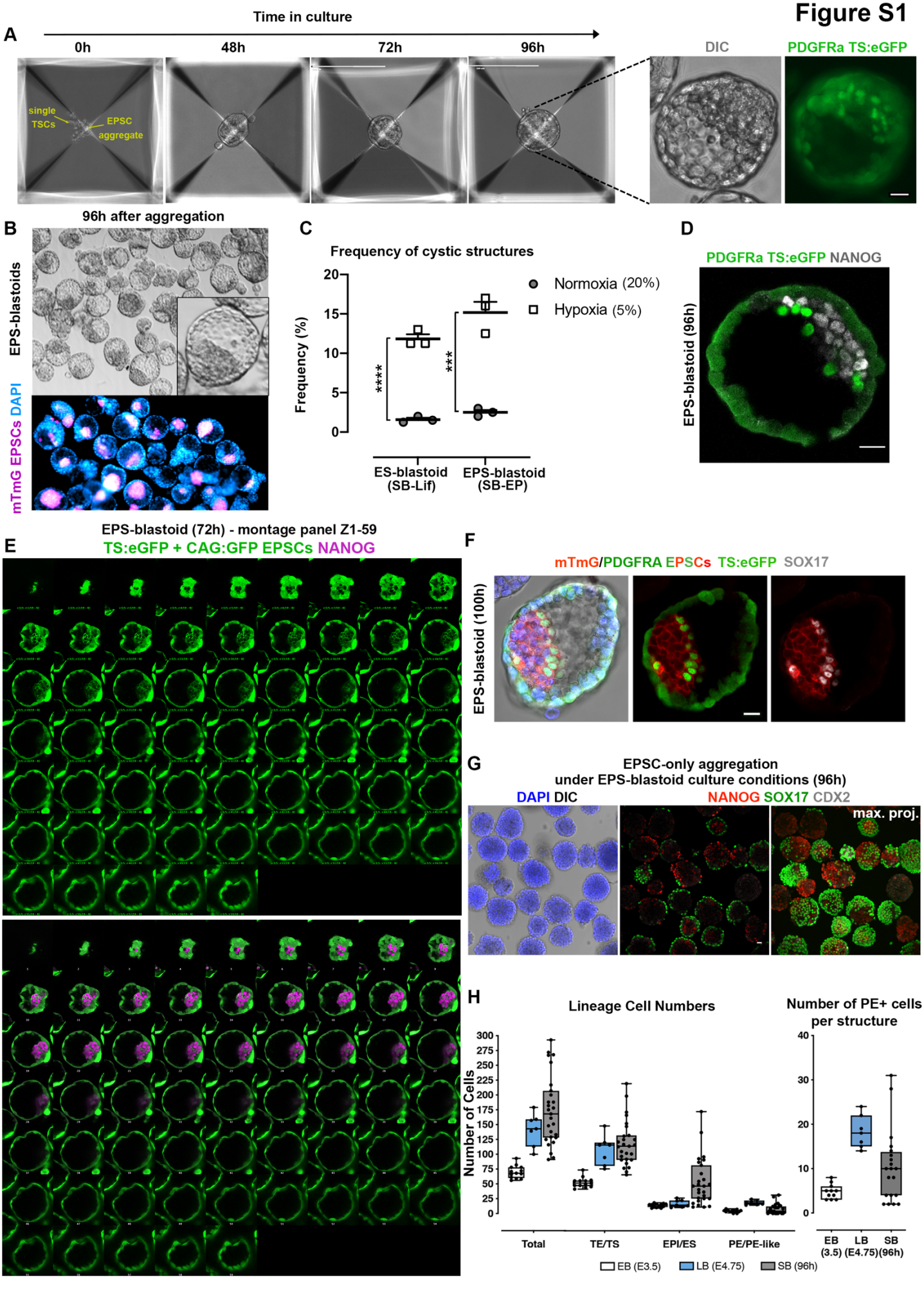

A
Time in culture
0h
48h
72h
96h
single TSCs
EPSC aggregate
DIC
PDGFRa TS:eGFP
B
96h after aggregation
EPS-blastoids
mTmG EPSCs DAPI
C
Frequency of cystic structures
Frequency (%)
Normoxia (20%)
Hypoxia (5%)
****
***
ES-blastoid (SB-Lif)
EPS-blastoid (SB-EP)
D
PDGFRa TS:eGFP NANOG
EPS-blastoid (96h)
E
EPS-blastoid (72h) - montage panel Z1-59
TS:eGFP + CAG:GFP EPSCs NANOG
F
mTmG/PDGFRA EPSCs TS:eGFP SOX17
EPS-blastoid (100h)
G
EPSC-only aggregation
under EPS-blastoid culture conditions (96h)
DAPI DIC
NANOG SOX17 CDX2
max. proj.
H
Lineage Cell Numbers
Number of Cells
Total
TE/TS
EPI/ES
PE/PE-like
Number of PE+ cells per structure
EB (3.5)
LB (E4.75)
SB (96h)
EB (E3.5)
LB (E4.75)
SB (96h)

Figure S2

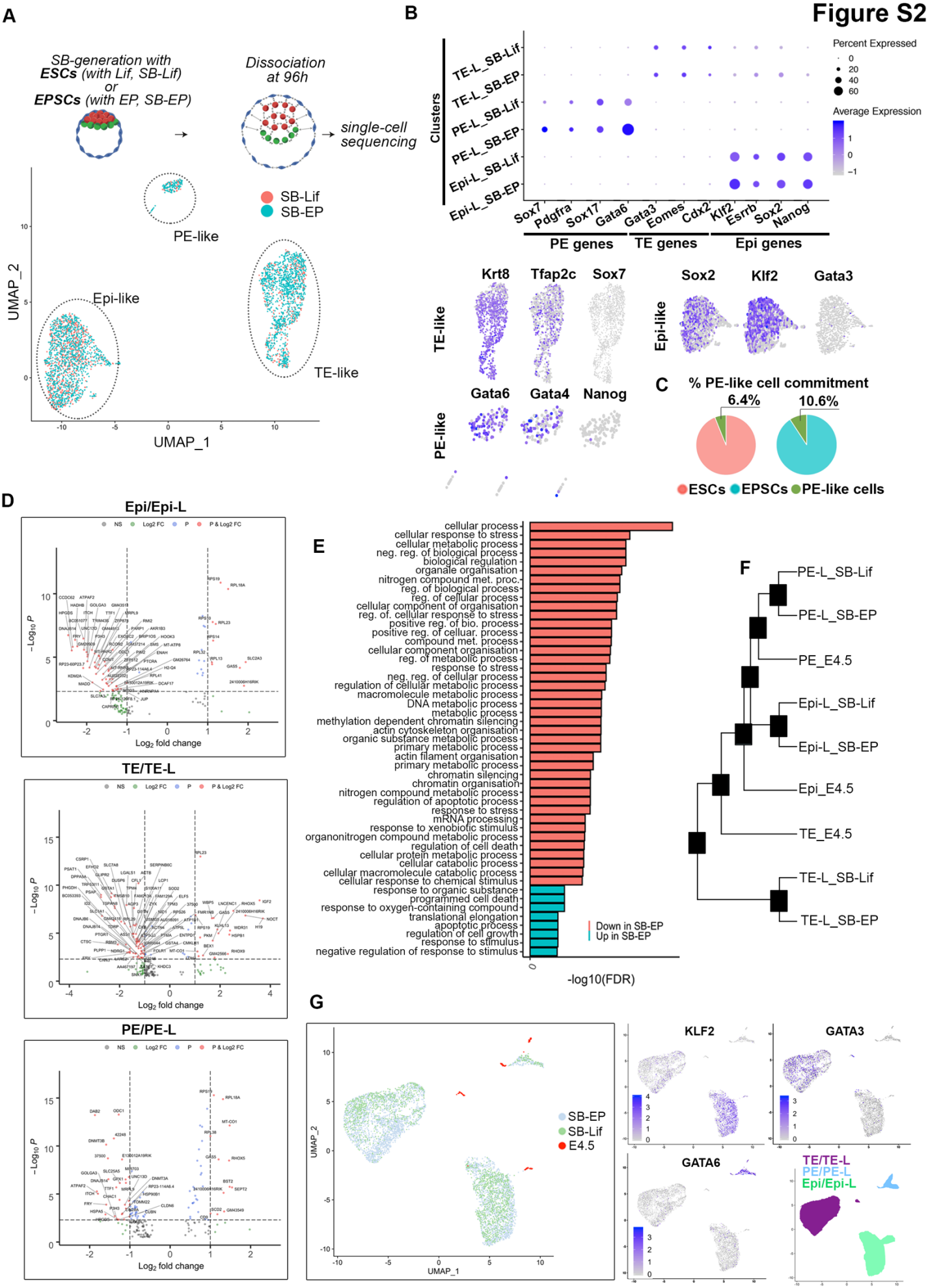
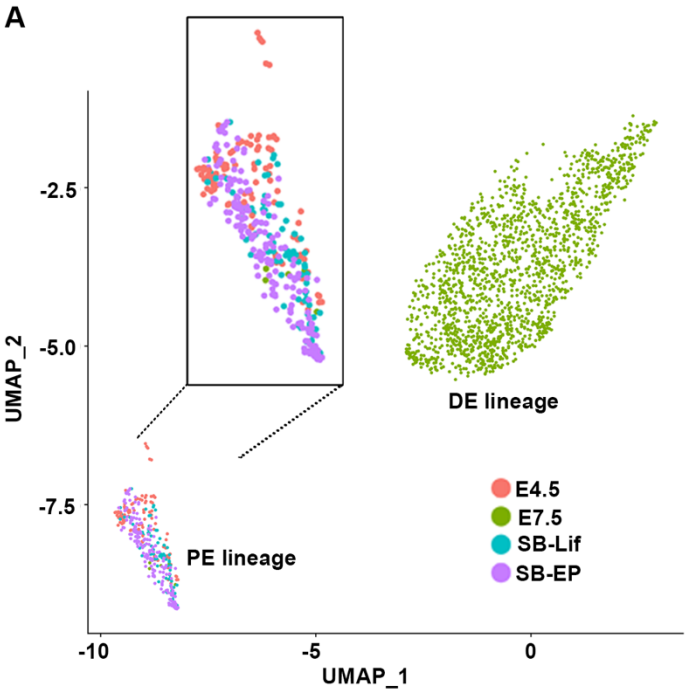
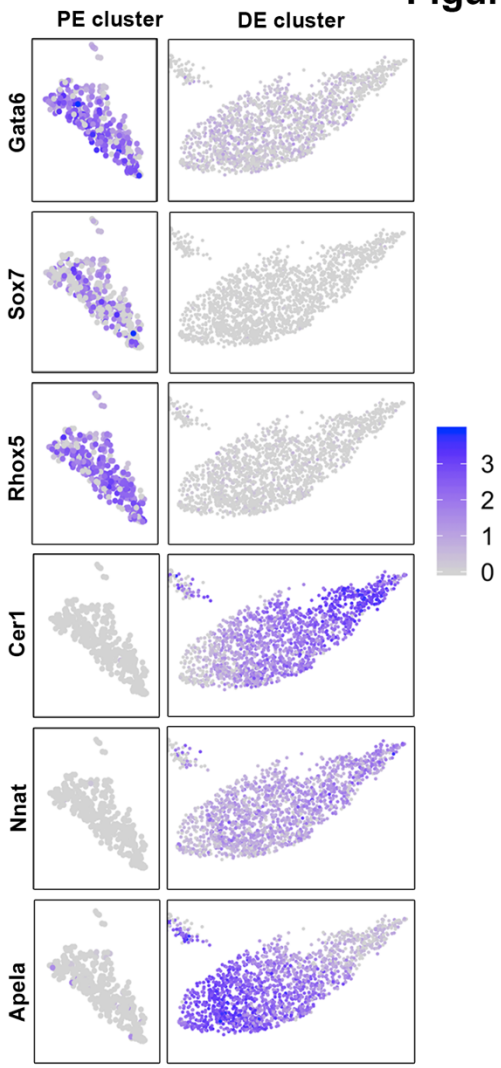

A
SB-generation with
ESCs (with Lif, SB-Lif)
or
EPSCs (with EP, SB-EP)
Dissociation
at 96h
single-cell
sequencing
SB-Lif
SB-EP
PE-like
Epi-like
TE-like
UMAP_2
UMAP_1
B
Clusters
TE-L_SB-Lif
TE-L_SB-EP
PE-L_SB-Lif
PE-L_SB-EP
Epi-L_SB-Lif
Epi-L_SB-EP
Sox7
Pdgfra
Sox17
Gata6
Gata3
Eomes
Cdx2
Klf2
Esrrb
Sox2
Nanog
PE genes
TE genes
Epi genes
Percent Expressed
0
20
40
60
Average Expression
1
0
−1
TE-like
Krt8
Tfap2c
Sox7
Epi-like
Sox2
Klf2
Gata3
PE-like
Gata6
Gata4
Nanog
C
% PE-like cell commitment
6.4%
10.6%
ESCs
EPSCs
PE-like cells
D
Epi/Epi-L
TE/TE-L
PE/PE-L
NS
Log2 FC
P
P & Log2 FC
−Log10 P
Log2 fold change
E
cellular process
cellular response to stress
cellular metabolic process
neg. reg. of biological process
biological regulation
organelle organisation
nitrogen compound met. proc.
reg. of biological process
reg. of cellular process
cellular component of organisation
reg. of. cellular response to stress
positive reg. of bio. process
positive reg. of cellular. process
compound met. process
cellular component organisation
reg. of metabolic process
response to stress
neg. reg. of cellular process
regulation of cellular metabolic process
macromolecule metabolic process
DNA metabolic process
metabolic process
methylation dependent chromatin silencing
actin cytoskeleton organisation
organic substance metabolic process
primary metabolic process
actin filament organisation
primary metabolic process
chromatin silencing
chromatin organisation
nitrogen compound metabolic process
regulation of apoptotic process
response to stress
mRNA processing
response to xenobiotic stimulus
organonitrogen compound metabolic process
regulation of cell death
cellular protein metabolic process
cellular catabolic process
response to organic substance
cellular macromolecule catabolic process
cellular response to chemical stimulus
programmed cell death
response to oxygen-containing compound
translational elongation
apoptotic process
regulation of cell growth
response to stimulus
negative regulation of response to stimulus
Down in SB-EP
Up in SB-EP
0
-log10(FDR)
F
PE-L_SB-Lif
PE-L_SB-EP
PE_E4.5
Epi-L_SB-Lif
Epi-L_SB-EP
Epi_E4.5
TE_E4.5
TE-L_SB-Lif
TE-L_SB-EP
G
SB-EP
SB-Lif
E4.5
UMAP_2
UMAP_1
KLF2
GATA3
GATA6
TE/TE-L
PE/PE-L
Epi/Epi-L

Figure S3

A



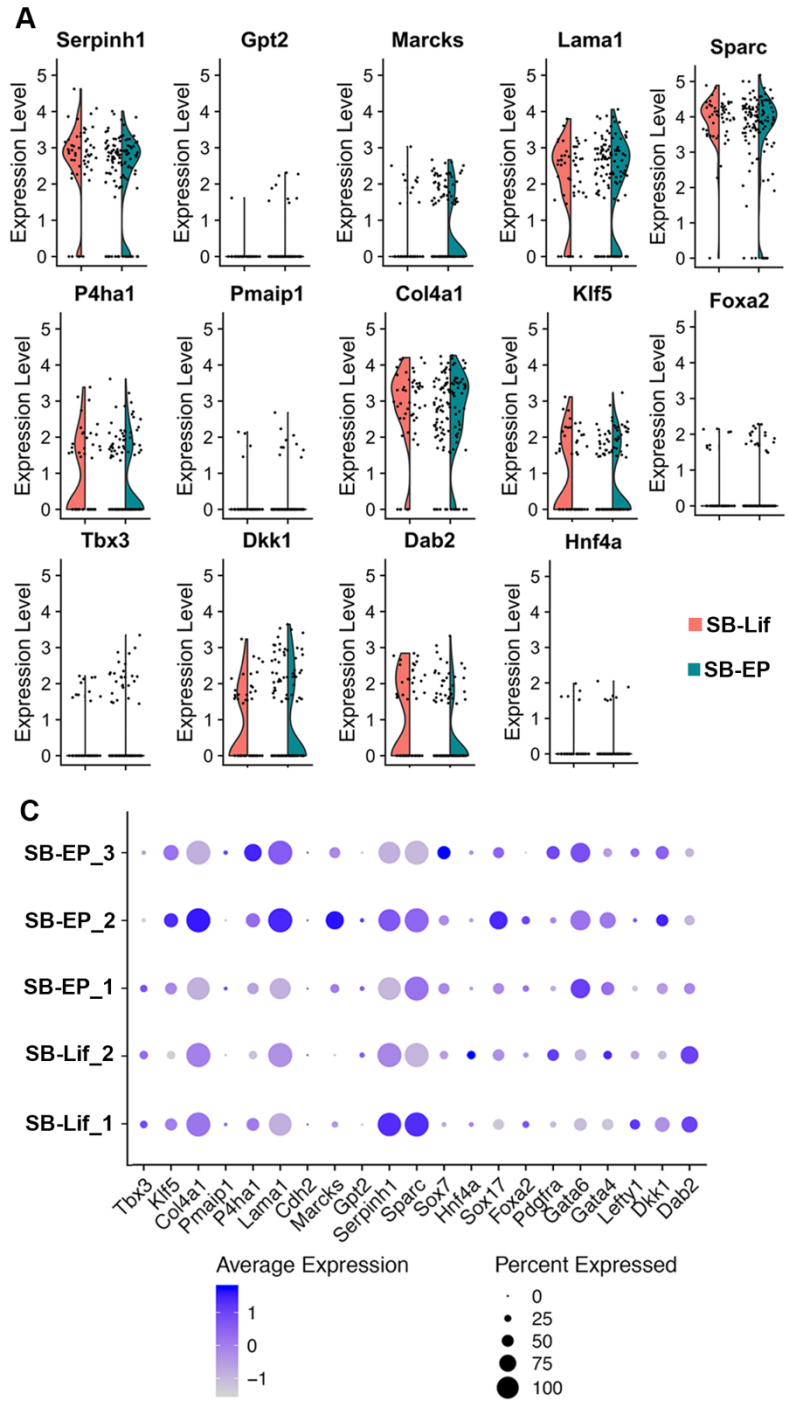
B

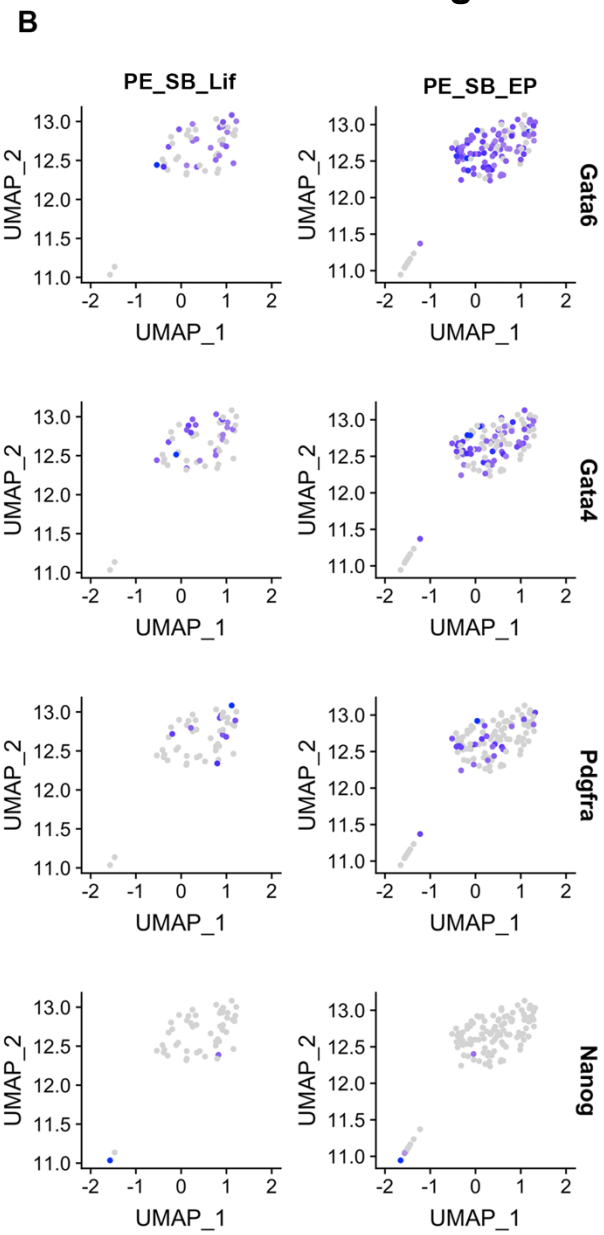


C

Gene	p_val	avg_logFC
Sox7	1.18E-117	1.456435069
Gata6	1.62E-48	1.978008538
Rhox5	2.50E-102	2.359566947
Nnat	8.73E-86	-0.86926672
Apela	1.84E-88	-2.10569594
Cer1	1.04E-88	-1.694802598

Figure S4


A
Serpinh1
Gpt2
Marcks
Lama1
Sparc
P4ha1
Pmaip1
Col4a1
Klf5
Foxa2
Tbx3
Dkk1
Dab2
Hnf4a
Expression Level
SB-Lif
SB-EP


B
PE_SB_Lif
PE_SB_EP
Gata6
Gata4
Pdgfra
Nanog
UMAP_1
UMAP_2

C
SB-EP_3
SB-EP_2
SB-EP_1
SB-Lif_2
SB-Lif_1
Tbx3
Klf5
Col4a1
Pmaip1
P4ha1
Lama1
Cdh2
Marcks
Gpt2
Serpinh1
Sparc
Sox7
Hnf4a
Sox17
Foxa2
Pdgfra
Gata6
Gata4
Lefty1
Dkk1
Dab2
Average Expression
Percent Expressed

Figure S5

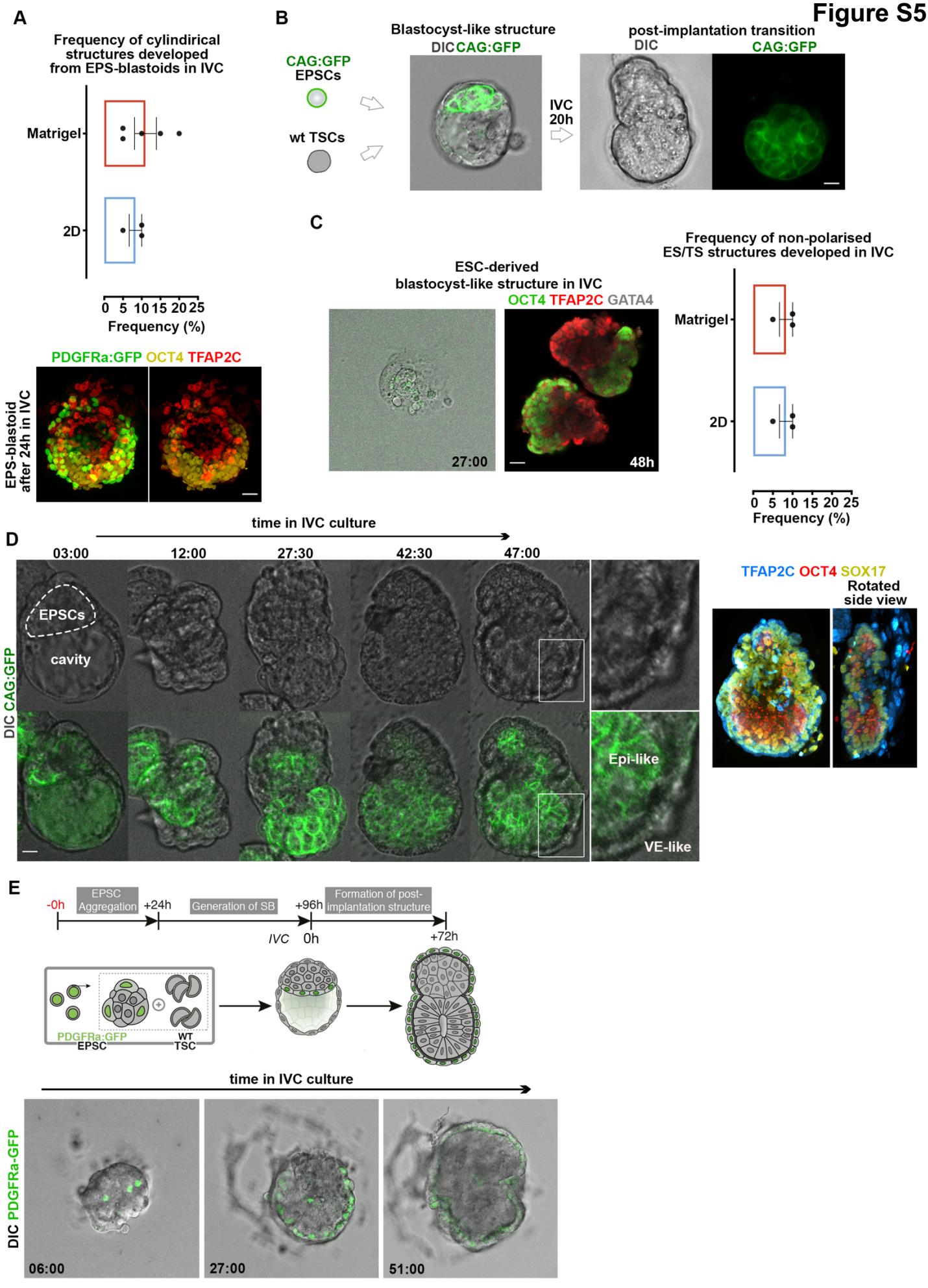
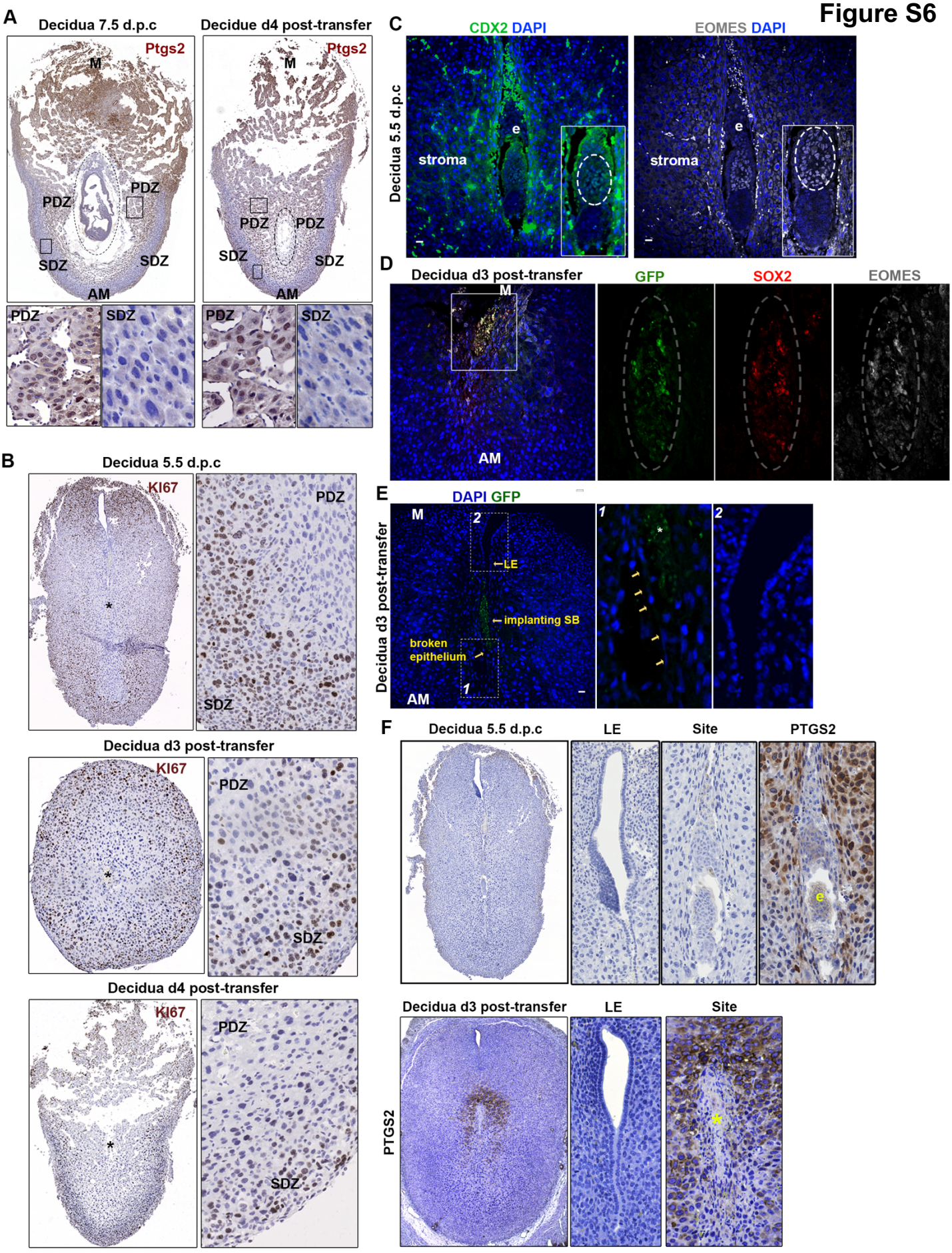

A
Frequency of cylindirical structures developed from EPS-blastoids in IVC
Matrigel
2D
0 5 10 15 20 25
Frequency (%)
PDGFRa:GFP OCT4 TFAP2C
EPS-blastoid after 24h in IVC
B
Blastocyst-like structure
DIC CAG:GFP
CAG:GFP EPSCs
wt TSCs
IVC 20h
post-implantation transition
DIC
CAG:GFP
C
ESC-derived blastocyst-like structure in IVC
OCT4 TFAP2C GATA4
27:00
48h
Frequency of non-polarised ES/TS structures developed in IVC
Matrigel
2D
0 5 10 15 20 25
Frequency (%)
D
time in IVC culture
03:00
12:00
27:30
42:30
47:00
EPSCs
cavity
DIC CAG:GFP
Epi-like
VE-like
TFAP2C OCT4 SOX17
Rotated side view
E
-0h
EPSC Aggregation
+24h
Generation of SB
+96h
Formation of post-implantation structure
IVC 0h
+72h
PDGFRa:GFP
EPSC
WT
TSC
time in IVC culture
DIC PDGFRa-GFP
06:00
27:00
51:00

Figure S6


A
Decidua 7.5 d.p.c
Decidue d4 post-transfer
Ptgs2
M
PDZ
SDZ
AM
B
Decidua 5.5 d.p.c
KI67
Decidua d3 post-transfer
Decidua d4 post-transfer
C
CDX2 DAPI
EOMES DAPI
Decidua 5.5 d.p.c
stroma
e
D
Decidua d3 post-transfer
GFP
SOX2
EOMES
E
DAPI GFP
LE
implanting SB
broken epithelium
F
Decidua 5.5 d.p.c
LE
Site
PTGS2
Decidua d3 post-transfer
Decidua d4 post-transfer

Supplementary Table 1- oligonucleotide sequences used to barcode the final library

SB LIF 1_P5	AATGATACGGCGACCACCGAGATCTACACCCTTCGCATCGTCGGCAGCGTC
SB LIF 2_P5	AATGATACGGCGACCACCGAGATCTACACGCCGTCGATCGTCGGCAGCGTC
SB EP_1_P5	AATGATACGGCGACCACCGAGATCTACACACACGATCTCGTCGGCAGCGTC
SB EP_2_P5	AATGATACGGCGACCACCGAGATCTACACGATATCCATCGTCGGCAGCGTC
SB EP_3_P5	AATGATACGGCGACCACCGAGATCTACACCCAACATTTCGTCGGCAGCGTC
E4.5 Blastocysts	AATGATACGGCGACCACCGAGATCTACACTTCGCTGATCGTCGGCAGCGTC
PCR_barcode_P7	CAAGCAGAAGACGGCATACGAGATGGGTGTCGGGTGCAG

Supplementary Table 2- Sequencing and data pre-processing parameters

2a Sequencing parameters:

Platform: Nextseq 75bp High output
RE-HYB No
Read type Paired-end
Cycles read1 61
Cycles read2 14
Indexing DualIndex
Cycles index 1 8
Cycles index 2 8

2b Parameters of the yaml indrop file:

Pl umi_quantification_arguments:
m : 10
u : 1
d : 600
split-ambigs: False
min_non_polyA: 15
output_unaligned_reads_to_other_fastq: False
low_complexity_mask: False
bowtie_arguments:
m : 200 n:1
l : 12
e : 120

Supplementary Table 3- Commands for UMAP dimensional reduction plotting and cluster identification

```
RunUMAP(reduction = "harmony", dims = 1:20) %>%
FindNeighbors(reduction = "harmony", dims = 1:20) %>%
FindClusters(resolution = 0.5) %>%     identity()
```