Liver-specification of human iPSC-derived endothelial cells transplanted into mouse liver

Kiryu K. Yap, Jan Schröder, Yi-Wen Gerrand, Aleksandar Dobric, Anne M. Kong, Adrian M. Fox, Brett Knowles, Simon W. Banting, Andrew G. Elefanty, Eduoard G. Stanley, George C. Yeoh, Glen P. Lockwood, Victoria C. Cogger, Wayne A. Morrison, Jose M. Polo, Geraldine M. Mitchell

Table of contents

Supplementary figures & legends

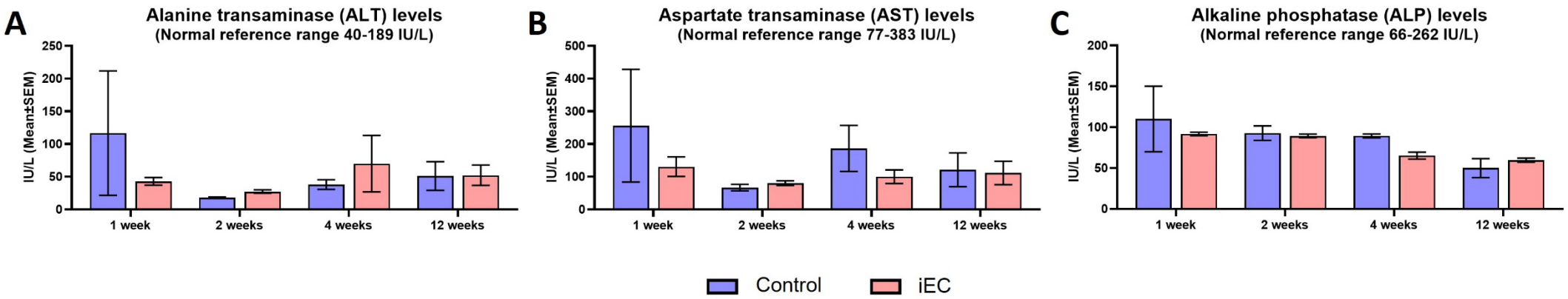


Fig. S1. Liver function tests comparing iEC-transplanted and sham control FRG mice at 1, 2, 4 and 12 weeks. (A) Alanine transaminase (ALT) levels. **(B)** Aspartate transaminase (AST) levels. **(C)** Alkaline phosphatase (ALP) levels). None of the liver enzymes were elevated across all groups and time-points. No significant difference was found between all time-points and between control and iEC groups, when analysed with two-way ANOVA and Bonferroni post-hoc analysis. N=3-5 per group.

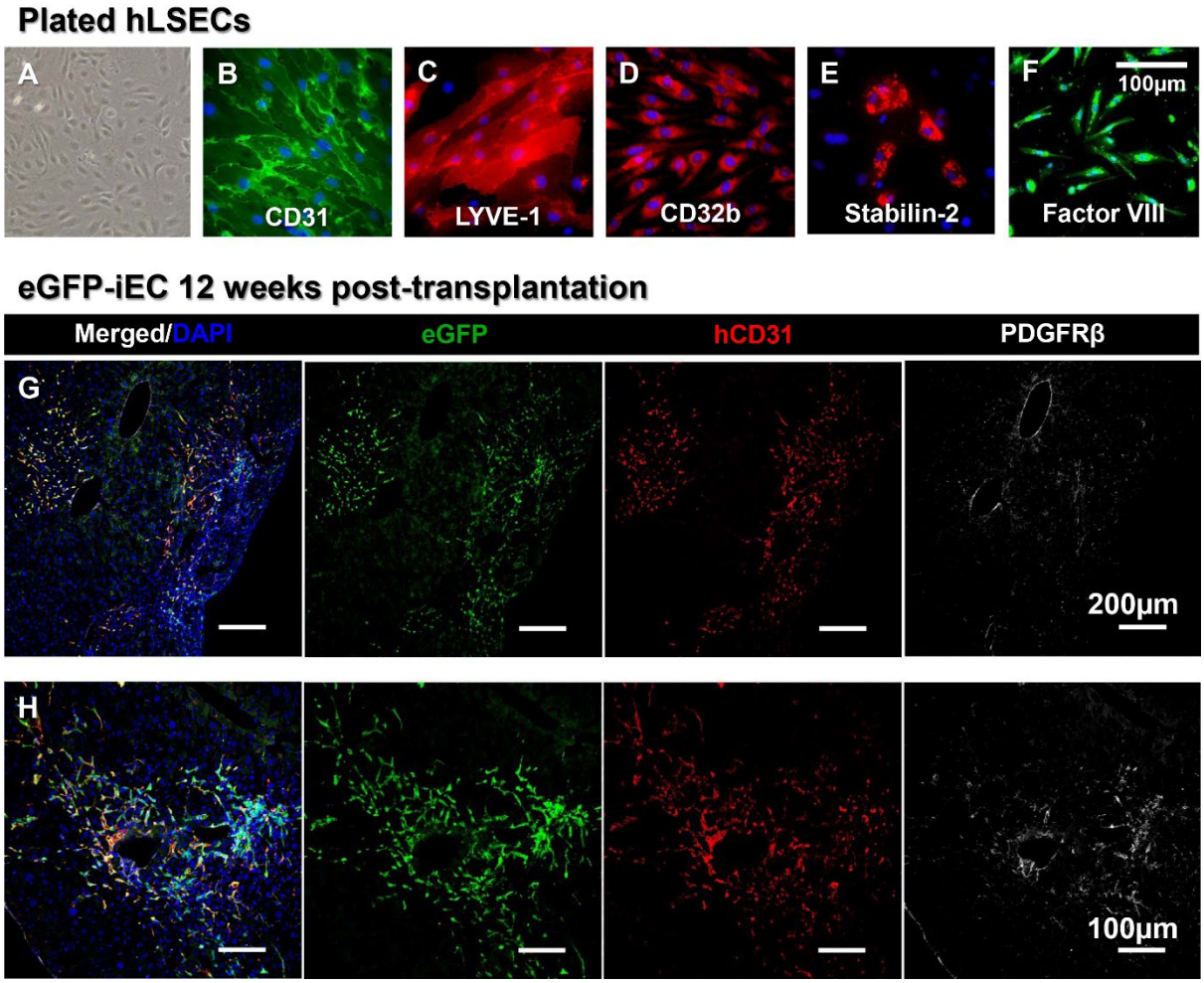
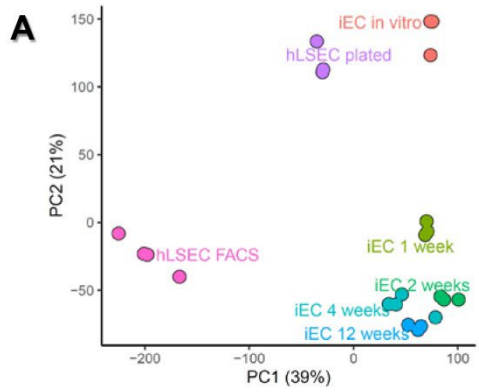
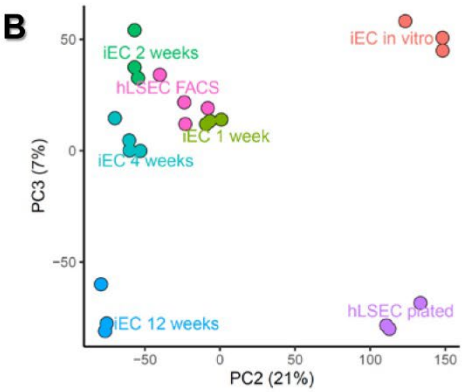
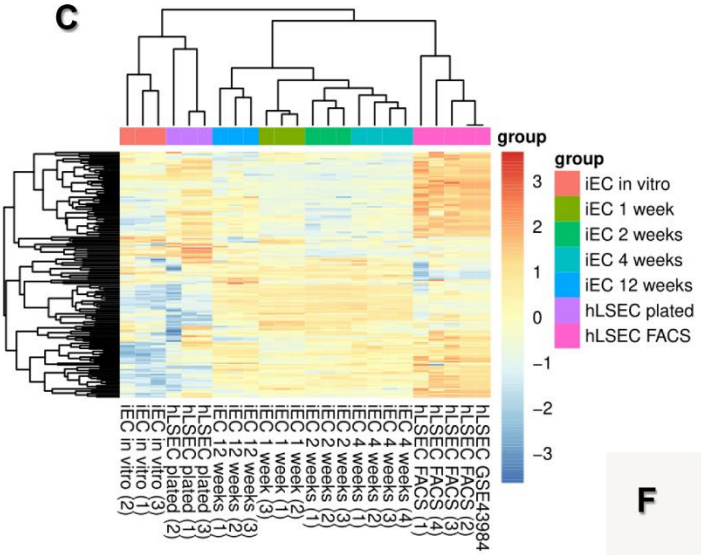
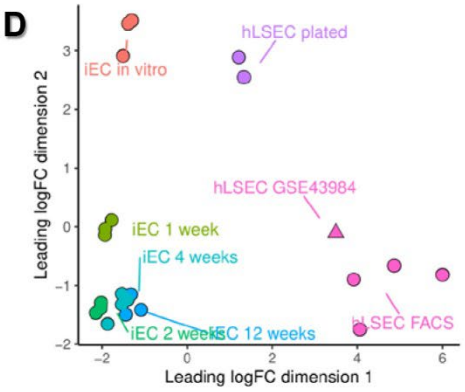


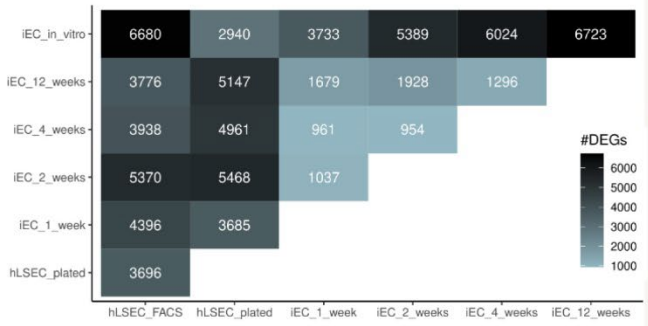
Fig. S2. Immunofluorescence characterisation of plated hLSECs and eGFP-iECs transplanted into FRG mouse livers at 12 weeks. (A) hLSECs plated in fibronectin coated monolayer culture demonstrate typical endothelial cobble-stone morphology, and are largely **(B)** CD31+, **(C)** LYVE-1+, **(D)** CD32b+, occasionally **(E)** Stabilin-2+, and also express **(F)** human factor VIII. **(G, H)** iEC transplantations into FRG mouse livers were performed using a second hiPSC line to confirm reproducibility. An eGFP reporter hiPSC line was used for lineage tracing. Similar to results found with TdTom-iECs, at 12 weeks iECs robustly repopulated the mouse liver vasculature, with a large majority of eGFP+/hCD31+/PDGFRβ- endothelial cells expanding along the sinusoids, and a small proportion of eGFP+/hCD31-/PDGFRβ+ stromal cells forming small clusters not associated with sinusoids. Scale bars, 100µm **(A, B, C, D, E, F, H)**, 200µm **(G)**.


A
iEC in vitro
hLSEC plated
hLSEC FACS
iEC 1 week
iEC 2 weeks
iEC 4 weeks
iEC 12 weeks
PC2 (21%)
PC1 (39%)


B
iEC 2 weeks
iEC in vitro
hLSEC FACS
iEC 1 week
iEC 4 weeks
iEC 12 weeks
hLSEC plated
PC3 (7%)
PC2 (21%)


C
group
iEC in vitro
iEC 1 week
iEC 2 weeks
iEC 4 weeks
iEC 12 weeks
hLSEC plated
hLSEC FACS


D
iEC in vitro
hLSEC plated
hLSEC GSE43984
iEC 1 week
iEC 4 weeks
iEC 2 weeks
iEC 12 weeks
hLSEC FACS
Leading logFC dimension 2
Leading logFC dimension 1


E
iEC_in_vitro
iEC_12_weeks
iEC_4_weeks
iEC_2_weeks
iEC_1_week
hLSEC_plated
hLSEC_FACS
hLSEC_plated
iEC_1_week
iEC_2_weeks
iEC_4_weeks
iEC_12_weeks
#DEGs

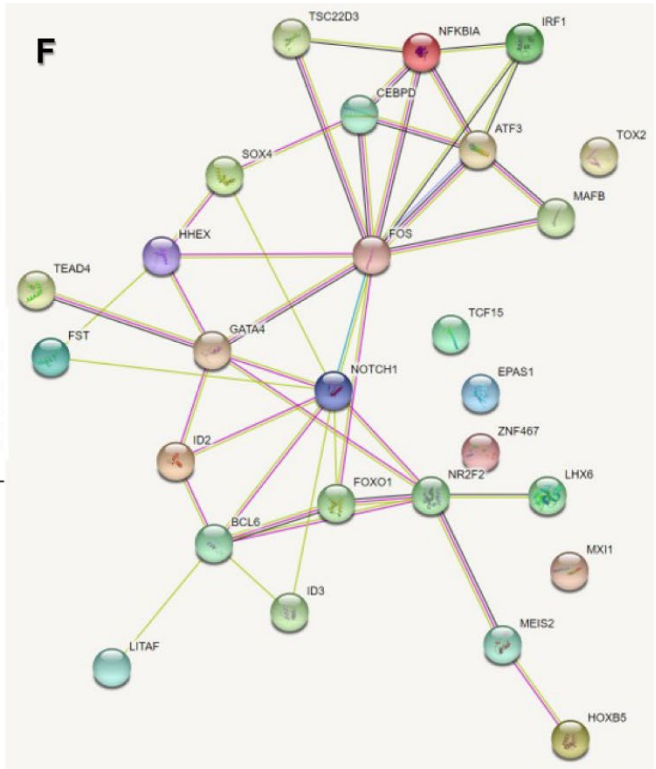

F
TSC22D3
NFKBIA
IRF1
CEBPD
ATF3
TOX2
SOX4
MAFB
HHEX
FOS
TEAD4
TCF15
FST
GATA4
NOTCH1
EPAS1
ZNF467
ID2
NR2F2
LHX6
FOXO1
BCL6
MXI1
ID3
MEIS2
LITAF
HOXB5

Fig. S3. Further bulk RNAseq analysis. (A) Principal component analysis (PCA) plot of bulkRNAseq samples, showing the first principal component (PC1) and second principal component (PC2). **(B)** PCA plot of bulkRNAseq samples showing PC2 and PC3. **(C)** Clustered heatmap (double dendrogram) of bulk RNAseq samples based on their whole transcriptome. Note the clustering of a publicly available dataset (hLSEC GSE43984) with samples in the hLSEC FACS group, and clustering of the *in vitro* samples together (iEC *in vitro* and hLSEC plated) and the *ex vivo* samples together (iEC 1, 2, 4, 12 weeks, hLSEC FACS and hLSEC GSE43984). **(D)** Multi-dimensional scaling plot showing dimensions 1 and 2, incorporating all samples of the bulk RNAseq dataset together with the publicly available dataset for hLSEC (GSE43984). Note the clustering of hLSEC GSE43984 with hLSEC FACS. **(E)** Differential gene expression matrix showing the number of differentially expressed genes (DEGs) between different samples in the bulk RNAseq dataset. **(F)** Protein-protein interaction network from STRING analysis of the 27 transcription factors predicted to drive LSEC specification from the Mogrify webtool, showing that *NOTCH1*, *GATA4*, and *FOS* are central factors in the network.

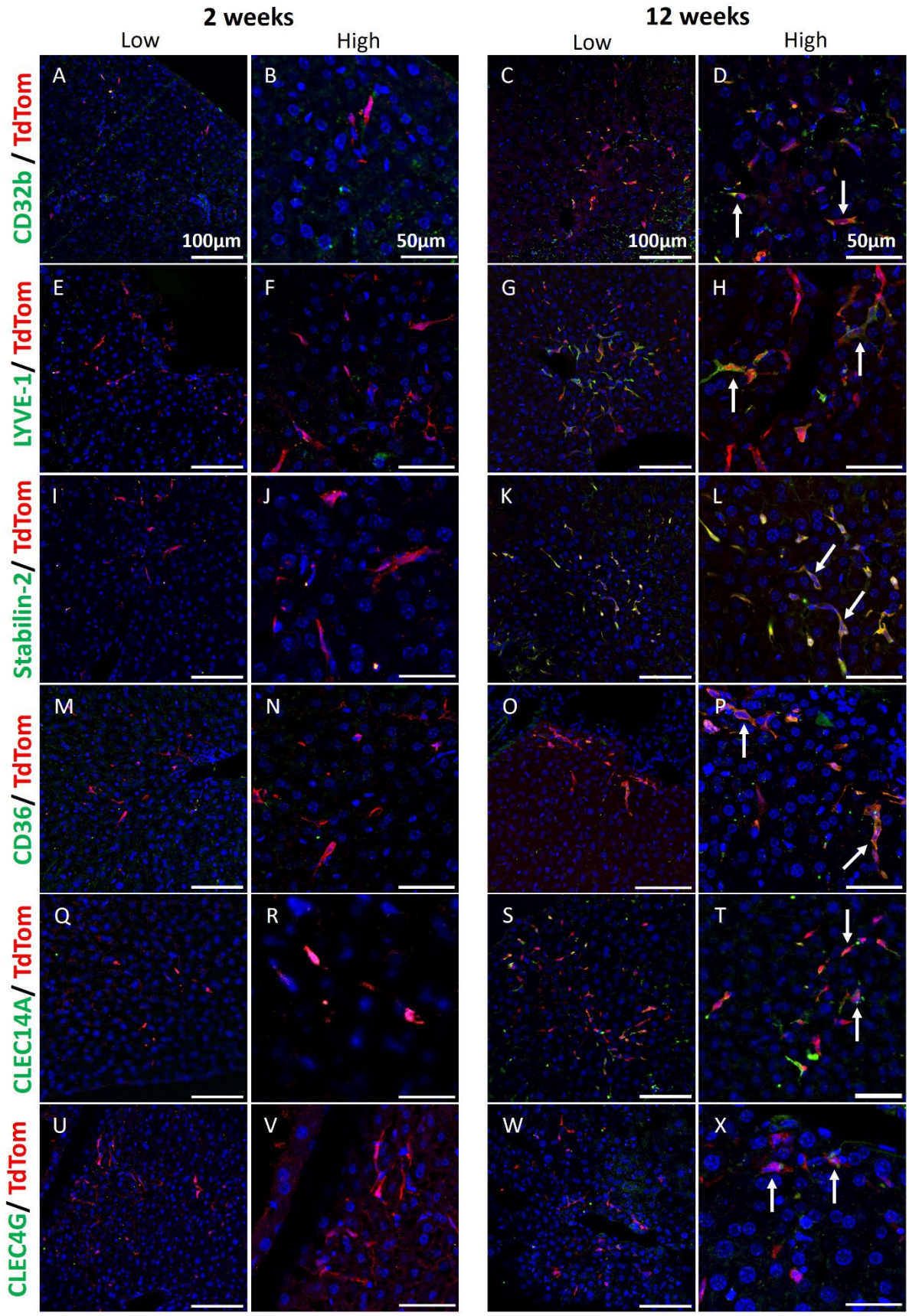


Fig. S4. The expression of LSEC markers in TdTom+ iECs at 2 and 12 weeks.

(A-D) CD32b, **(E-H)** LYVE-1, **(I-L)** Stabilin-2, **(M-P)** CD36, **(Q-T)** CLEC14A, **(U-X)** CLEC4G (all green). At 2 weeks, none of these markers were detected in TdTom+ iECs (red). At 12 weeks, all markers were seen in TdTom+ iECs (white arrows) although this expression was heterogeneous. CLEC4G was expressed in only a few cells at very low levels. Scale bars 100µm **(A ,C, E, G, I, K, M, O, Q, S, U, W)**, 50µm **(B, D, F, H, J, L, N, P, R, T, V, X)**.

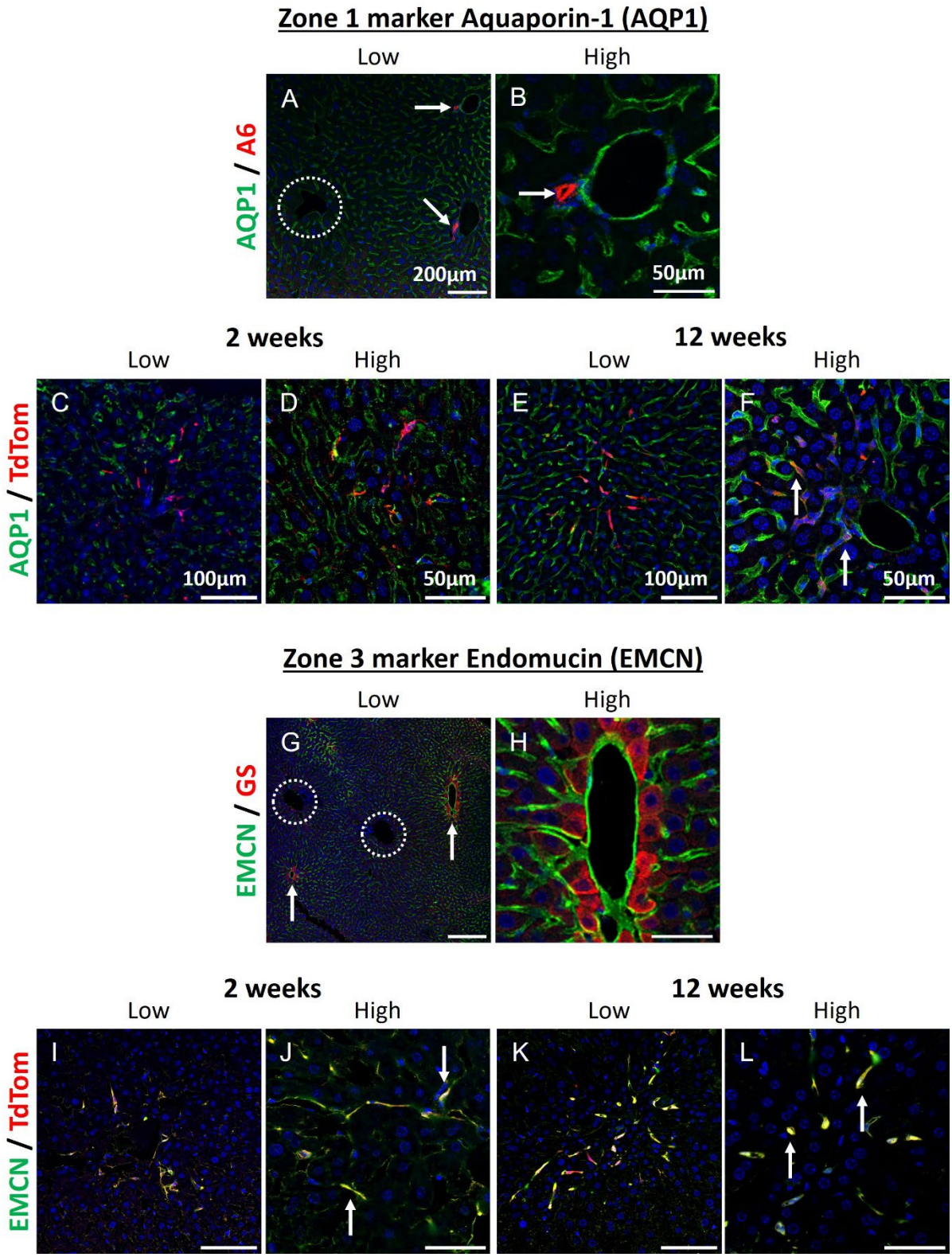


Fig. S5. Expression of zonated LSEC markers in TdTom+ iECs at 2 weeks and 12 weeks.
(A,B) Aquaporin-1 (AQP1) (green) was expressed in native periportal Zone 1 LSECs marked by the presence of A6+ bile ducts (red, arrows) in the mouse liver. AQP1 expression was low in perivenous Zone 3 (white dotted circle), confirming that it is a Zone 1 marker in the liver. AQP1 expression was negative in TdTom+ iECs (red) at 2 weeks **(C,D)**, but present at 12 weeks (arrows) **(E,F)**. **(G,H)** Endomucin (EMCN) (green) was expressed in perivenous Zone 3 LSECs in close proximity with glutamine synthetase (GS)+ hepatocytes (red, arrows) in the mouse liver. EMCN was not expressed in periportal Zone 1 (white dotted circles), confirming that it is a Zone 3 marker in the liver. EMCN was widely expressed in TdTom+ iECs at both 2 weeks **(I,J)** and 12 weeks **(K,L)** (arrows). Scale bars 200µm **(A,G)**, 100µm **(C, E, I, K,)**, 50µm **(B, D, F, H, J, L)**.

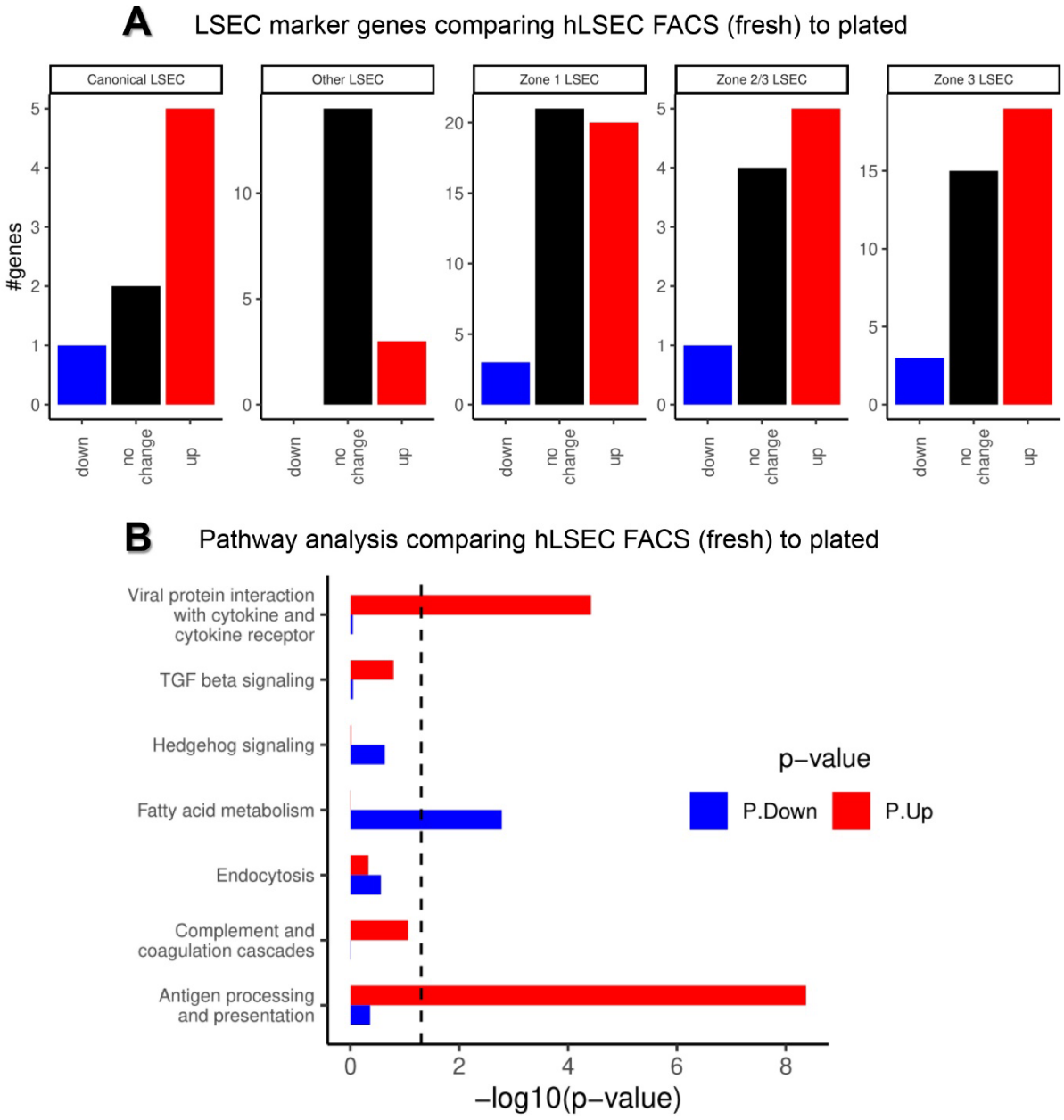


Fig. S6. Comparing the bulk transcriptome of freshly isolated hLSEC (hLSEC FACS) and monolayer cultured hLSEC (hLSEC plated). (A) Comparison across the five curated gene groups (canonical LSEC, other LSEC, zone 1 LSEC, zone 2/3 LSEC and zone 3 LSEC) indicates that fresh and plated hLSECs share many genes across all 5 groups. However, fresh hLSECs also express many more genes across all 5 groups compared to plated LSECs, particularly genes in the zone 1 and zone 3 LSEC groups. **(B)** Comparing the enrichment of key LSEC pathways indicates that hLSEC FACS are enriched in viral protein interaction with cytokine and cytokine receptor, TGF beta signaling, complement and coagulation cascades, and antigen processing and presentation pathways. hLSEC plated are enriched in hedgehog signaling, fatty acid metabolism, and endocytosis pathways.

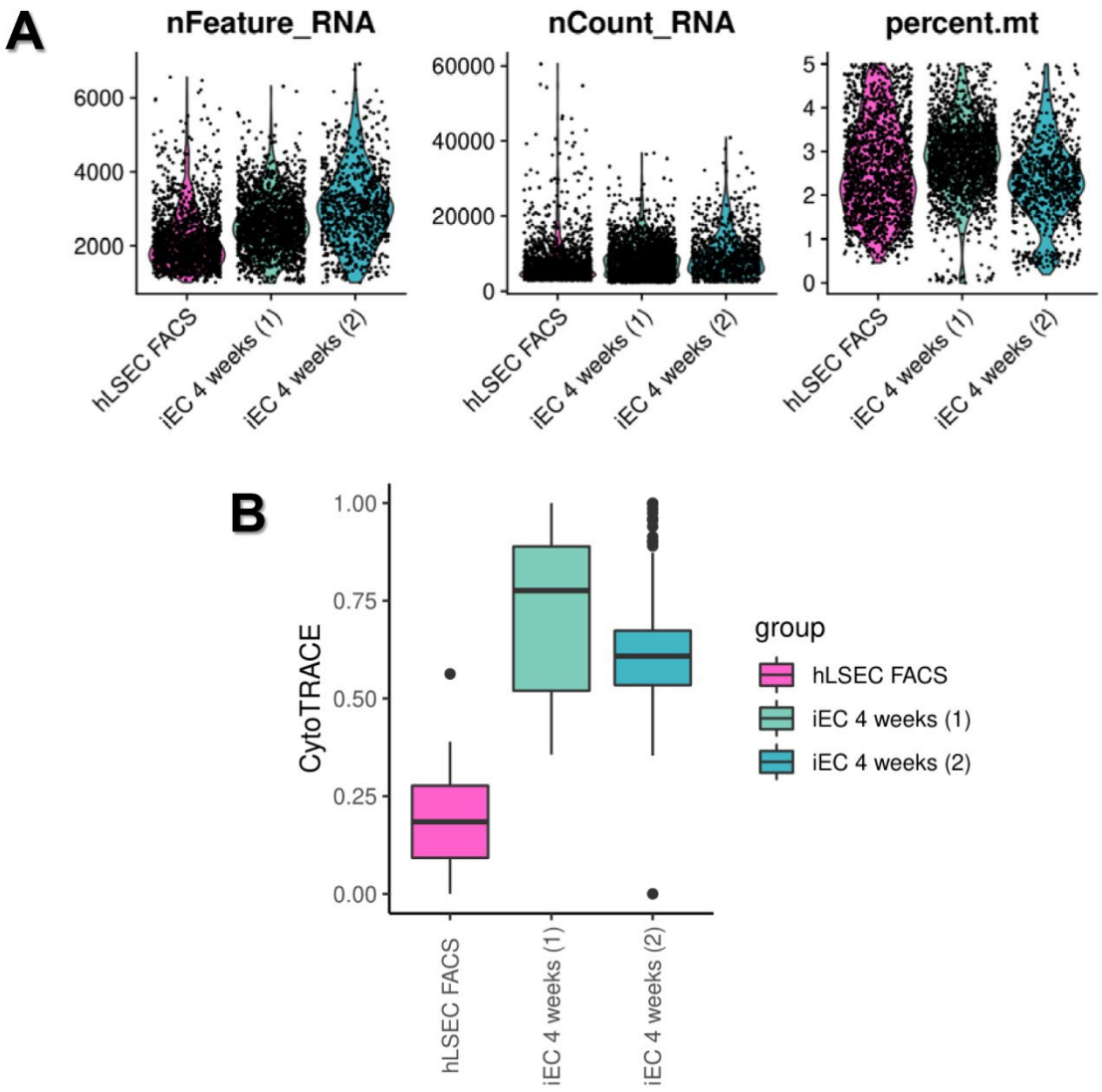


Fig. S7. Quality control of scRNAseq data and assessment of differentiation status. (A) Quality control violin plots indicating the number of genes per sample (nFeature_RNA), the number of unique molecular identifiers (UMIs) (nCount_RNA) and the percentage of UMI mapping to mitochondrial genes (percent.mt). The plots depict cells derived from a total of 4452 cells across 3 samples. **(B)** CytoTRACE computational analysis of differentiation status of cells within each scRNAseq sample demonstrates that hLSEC FACS contain the most differentiated cells, and both iEC 4 week samples (1 and 2) contain less differentiated cells.

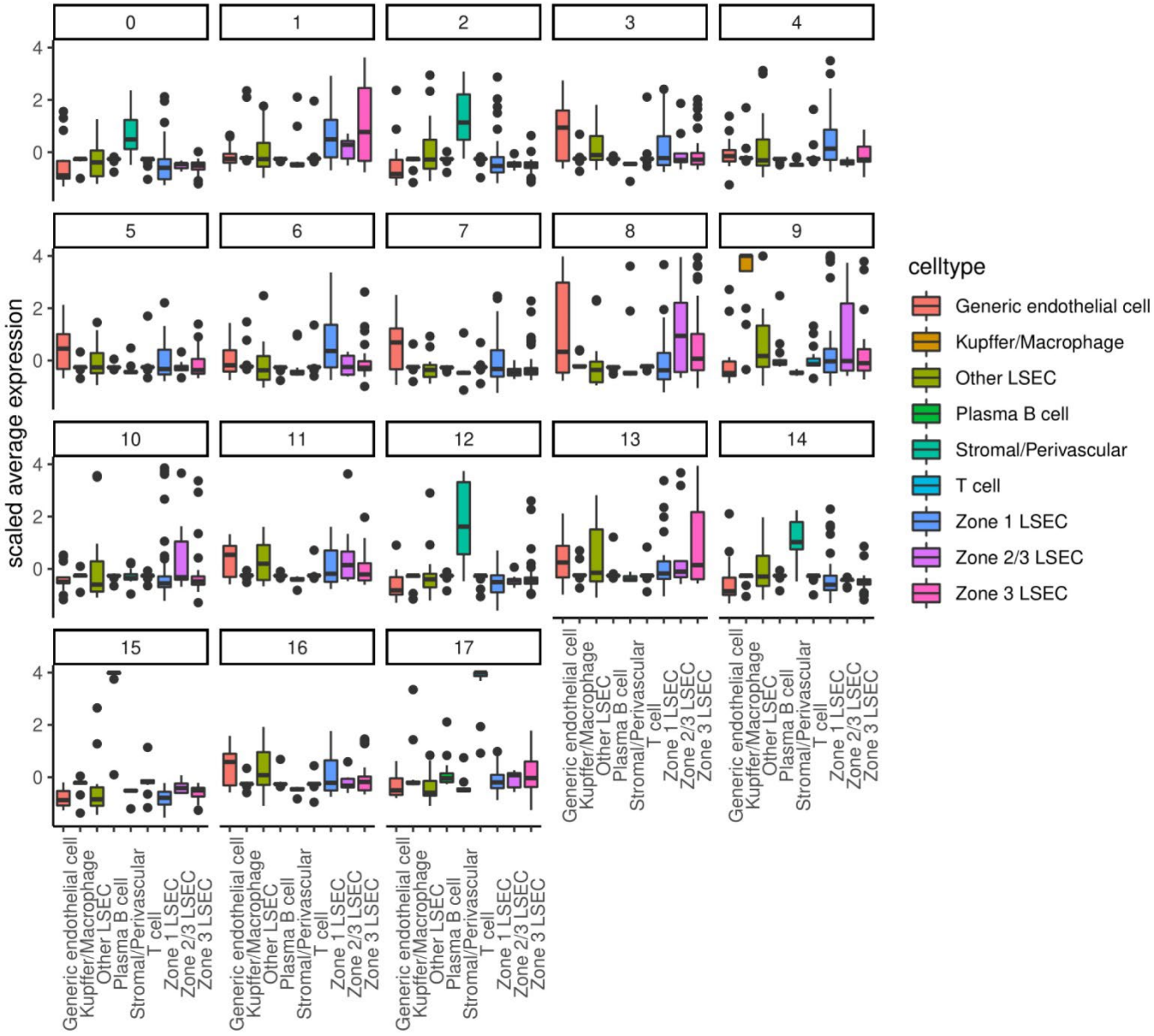


Fig. S8. Decision plots used to assign identity to each cell subpopulation in the scRNAseq dataset. The average expression of genes associated with 9 different cell types found in the liver (generic endothelial cell, Kupffer/macrophage, other LSEC, plasma B cell, stromal/perivascular, T cell, zone 1 LSEC, zone 2/3 LSEC, and zone 3 LSEC) for each of the 18 subpopulations identified through automated clustering is shown.

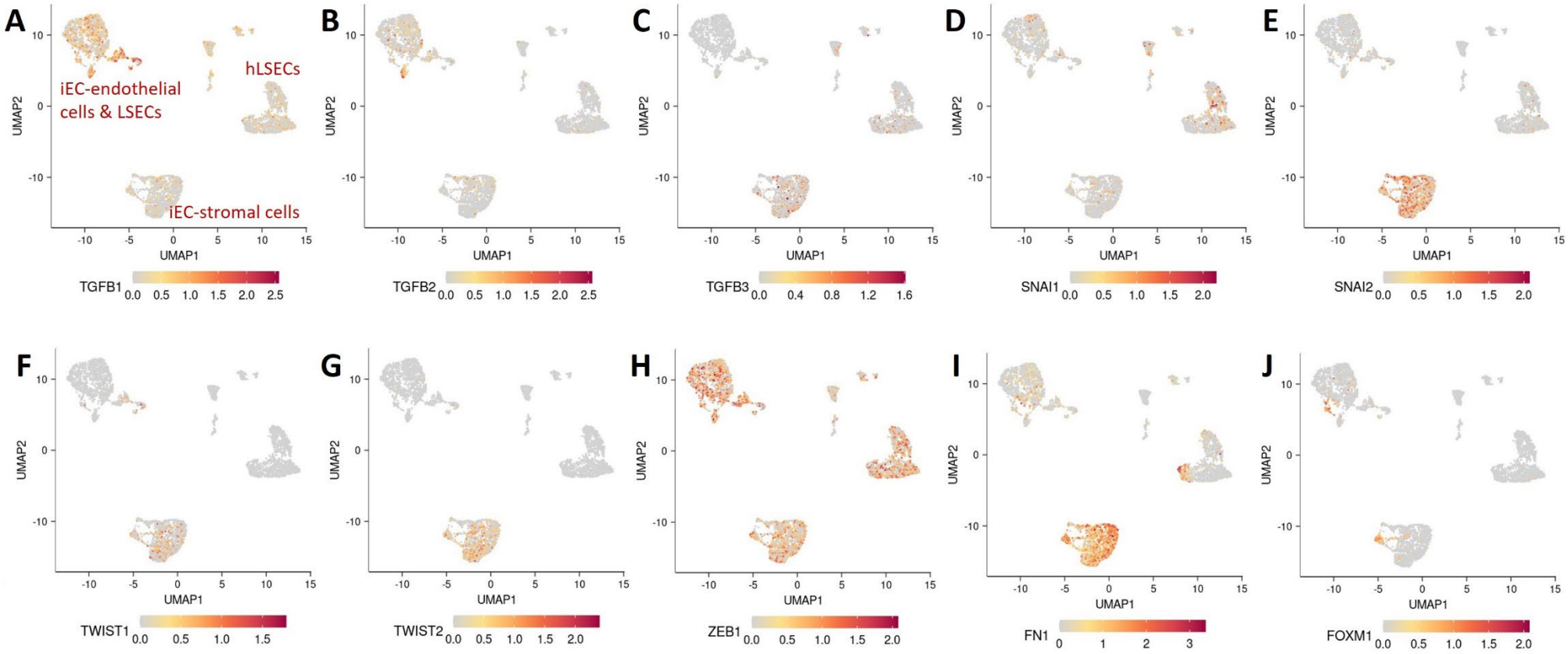


Fig. S9. ScRNAseq UMAP expression plots of genes associated with endothelial-mesenchymal transition. (A) *TGFB1,* **(B)** *TGFB2,* **(C)** *TGFB3,* **(D)** *SNAI1,* **(E)** *SNAI2,* **(F)** *TWIST1,* **(G)** *TWIST2,* **(H)** *ZEB1,* **(I)** *FN1,* **(J)** *FOXM1.* ScRNAseq included cells from primary hLSECs, and iECs post-transplantation at 4 weeks which formed two major populations, consisting of endothelial cells/LSECs and stromal cells. These major populations are annotated in **(A)** in red and apply to all UMAP expression plots.

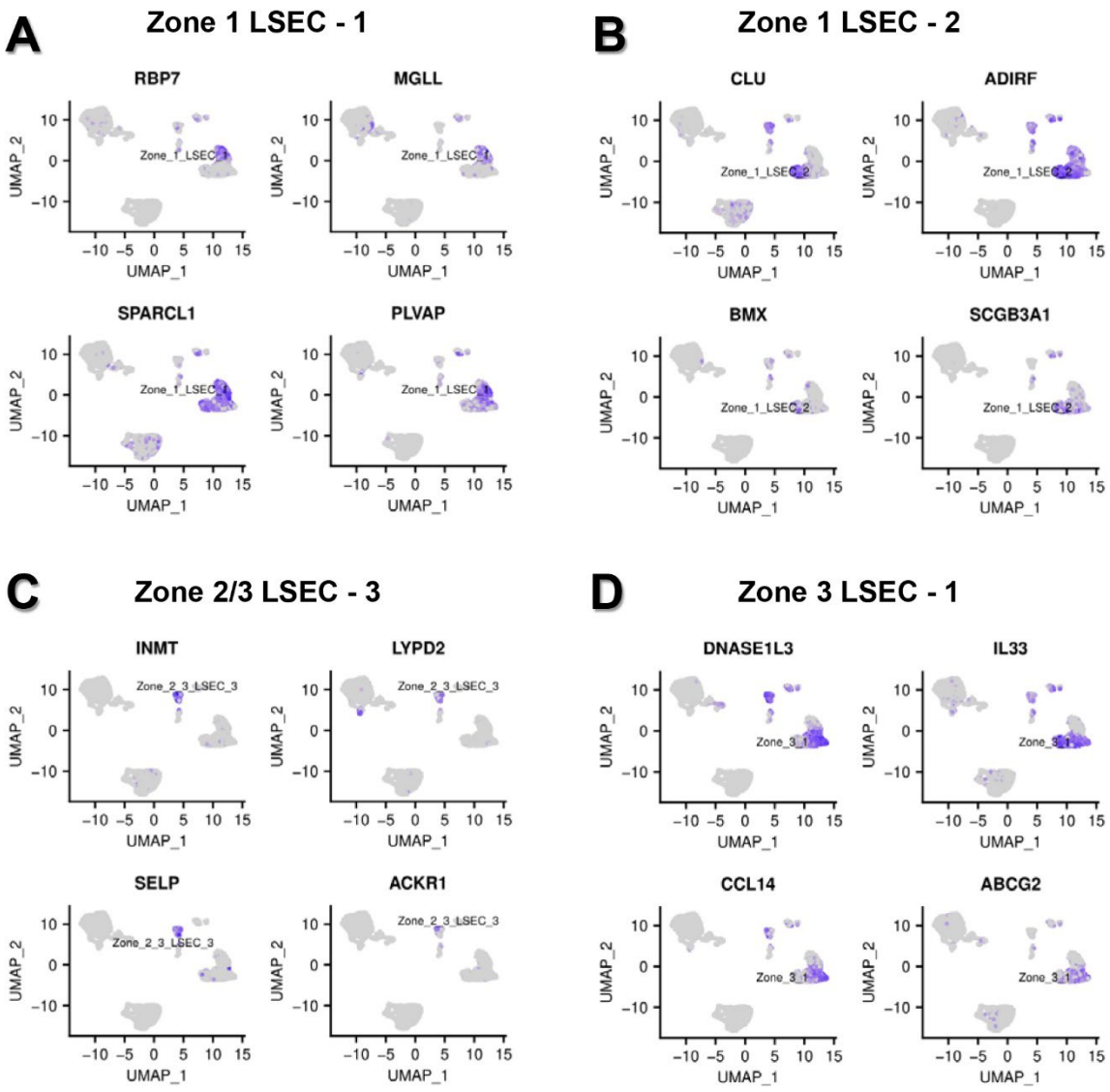


Fig. S10. Expression plots of top genes associated with LSEC subpopulations in the hLSEC scRNAseq sample. (A) Expression plots of the top 4 genes expressed in the Zone 1 LSEC – 1 subpopulation. **(B)** Expression plots of the top 4 genes expressed in the Zone 1 LSEC – 2 subpopulation. **(C)** Expression plots of the top 4 genes expressed in the Zone 2/3 LSEC – 3 subpopulation. **(D)** Expression plots of the top 4 genes expressed in the Zone 3 LSEC – 1 subpopulation.

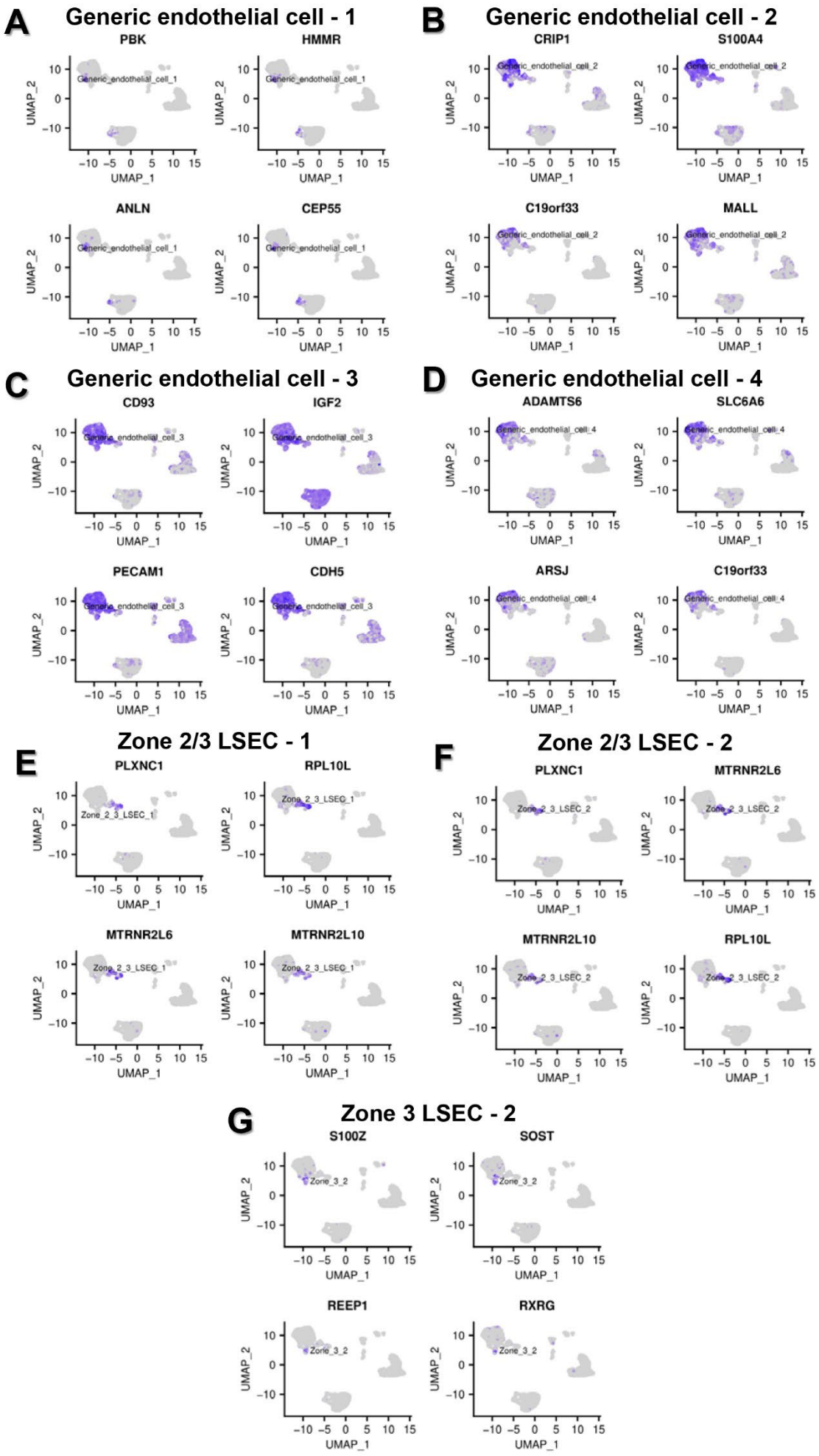

A
Generic endothelial cell - 1
PBK
HMMR
ANLN
CEP55
B
Generic endothelial cell - 2
CRIP1
S100A4
C19orf33
MALL
C
Generic endothelial cell - 3
CD93
IGF2
PECAM1
CDH5
D
Generic endothelial cell - 4
ADAMTS6
SLC6A6
ARSJ
C19orf33
E
Zone 2/3 LSEC - 1
PLXNC1
RPL10L
MTRNR2L6
MTRNR2L10
F
Zone 2/3 LSEC - 2
PLXNC1
MTRNR2L6
MTRNR2L10
RPL10L
G
Zone 3 LSEC - 2
S100Z
SOST
REEP1
RXRG
UMAP_1
UMAP_2

Fig. S11. Expression plots of top genes associated with endothelial subpopulations in the iEC 4 week scRNAseq samples. (A) Expression plots of the top 4 genes expressed in the generic endothelial cell – 1 subpopulation. **(B)** Expression plots of the top 4 genes expressed in the generic endothelial cell – 2 subpopulation. **(C)** Expression plots of the top 4 genes expressed in the generic endothelial cell – 3 subpopulation. **(D)** Expression plots of the top 4 genes expressed in the generic endothelial cell – 4 subpopulation. **(E)** Expression plots of the top 4 genes expressed in the Zone 2/3 LSEC – 1 subpopulation. **(F)** Expression plots of the top 4 genes expressed in the Zone 2/3 LSEC – 2 subpopulation.
(G) Expression plots of the top 4 genes expressed in the Zone 3 LSEC – 2 subpopulation.

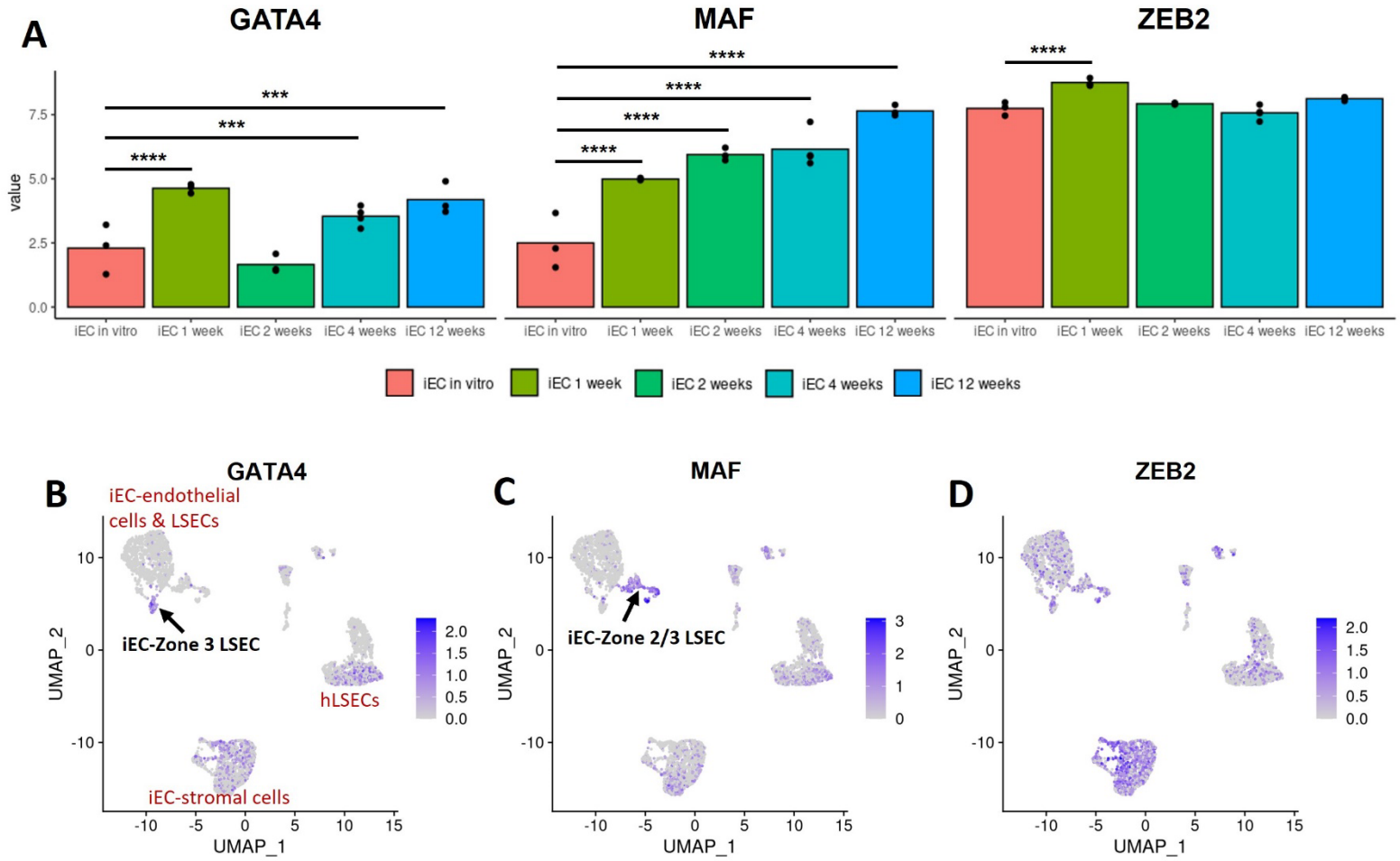


Fig. S12. Bulk and single RNAseq analysis of *GATA4*, *MAF, ZEB2*. (A)

BulkRNAseq analysis comparing iEC *in vitro* to iECs post-transplantation at 1 week, 2 weeks, 4 weeks and 12 weeks. *** $p \leq 0.001$, **** $p \leq 0.0001$, N=3-4 samples per group. **(B)** ScRNAseq UMAP expression plot of *GATA4*, which is enriched in the iEC-derived Zone 3 LSEC cluster. **(C)** ScRNAseq UMAP expression plot of *MAF*, which is enriched in the iEC-derived Zone 2/3 LSEC cluster. **(D)** ScRNAseq UMAP expression plot of *ZEB2,* which is widely expressed across all clusters especially iEC-derived stromal cells.

Supplementary table legends (tables provided in separate excel files)

Table S1. Curated list of genes used to identify LSECs (canonical LSEC markers), zonal supopulations of LSECs, stromal/perivascular cells and other non-parenchymal cells in the liver.

Table S2. KEGG pathway analysis of bulk RNAseq samples.

Table S3. Top genes associated with each cluster in scRNAseq analysis.

Table S4. DEGs between zone 1 and zone 3 hLSEC scRNAseq clusters, and pathway analysis of zone 1 and zone 3 subpopulations.

Table S5. Top 100 DEGs between the bulk transcriptome of *ex vivo* iEC 1 and 12 weeks, and *ex vivo* hLSEC FACS and *in vitro* iEC.

Table S6. Marker genes and transcription factors associated with LSEC specification.