## SUPPLEMENTAL METHODS

### MethoCult Colony-Forming Unit (CFU) Assay

CFU assay was performed using the murine MethoCult GF M3434 (Stem Cell Techologies) or human MethoCult H4034 Optimum (Stem Cell Techologies) kit. Procedure was performed per manufacturer's instructions.

### Tissue Histology and Cell Imaging

A small portion of the human fetal lung was fixed in 4% paraformaldehyde for 2 hours, cryoprotected in 30% sucrose and embedded in Optimum Cutting Temperature embedding medium. 10-12µm tissue sections were made using a cryostat. Sections were rinsed with PBS prior to permeabilizing and blocking in a solution of 0.4% Triton X-100 and 10% NDS. A list of antibodies used for immunofluorescence can be found under Supplemental Table 1.

Imaging of cell cultures were performed using a Keyence BZ-X700 fluorescence microscope. Cytospins were stained using the Hema 3 Stat Pack (Fisher) and imaged using a Nikon Eclipse NiE.

### Flow cytometry

All staining and washing steps were done using a solution of 1% BSA in PBS, 5mM EDTA. Single cell suspensions were stained with the appropriate master mix of antibodies on ice for 30 minutes. A table of antibodies used for analysis can be found under Supplemental Table 1. Fc-Receptor block treatment was performed using an anti-mouse CD16/32 antibody (BioLegend, Clone: 93) or anti-human Fc receptor block (biolegend, Human TruStain FcX) for 10 minutes prior to the addition of master mix. True-Stain Monocyte Blocker (BioLegend) and Brilliant Stain Buffer Plus (BD Biosciences) were included in master mix to minimize non-specific staining. Single stain controls were processed using cells and/or beads (Invitrogen, UltraComp eBeads Plus). Flow cytometry was performed using either a BD LSRII, a Stratedigm S1000EXI or a 5-laser Cytek Aurora spectral flow cytometer. Analysis was performed using FlowJo_v10.8 (FlowJo, LLC) software.

### Single cell RNA sequencing

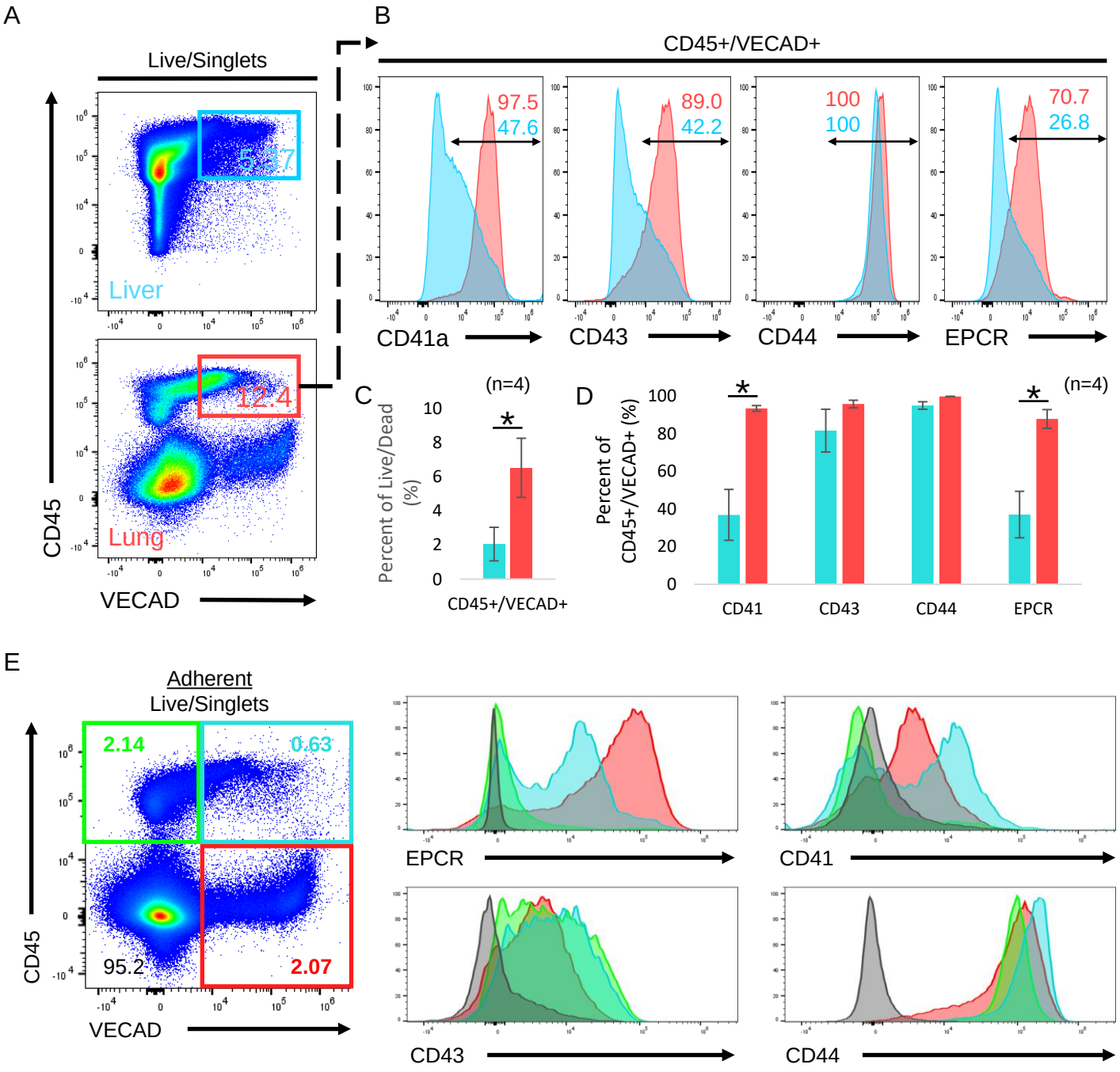
Adherent cells from days 4, 5 and 6 of explant cultures were released from Matrigel coated wells using Accutase. Collected cells were stained for VE-Cadherin-APC-Cy7 and resuspended in sort buffer (2% BSA in PBS) containing Calcein Blue AM (1:1000). Using a Beckman MoFlo Coulter Astrios, live-singlets and VECAD+ events were sorted into sort buffer (2% BSA in PBS with 5mM EDTA). Cells counts and viability after sorting were confirmed by hemocytometer using trypan blue. Suspension cells were left unsorted as VECAD expression is lost gradually as hematopoietic cells differentiate post-EHT emergence. VECAD+ sorted adherent cells and unsorted suspension cells were single cell captured using the 10X Genomics Chromium platform and prepared using the Single Cell 3' v3 kit. Library preparation and sequencing was done at the Boston University Microarray and Sequencing Resource (BUMSR) Core using the Illumina NextSeq 2000 instrument. GEO accession: GSE197290. Token: wvqjuouqdtmrtuh.

## Bioinformatic analysis

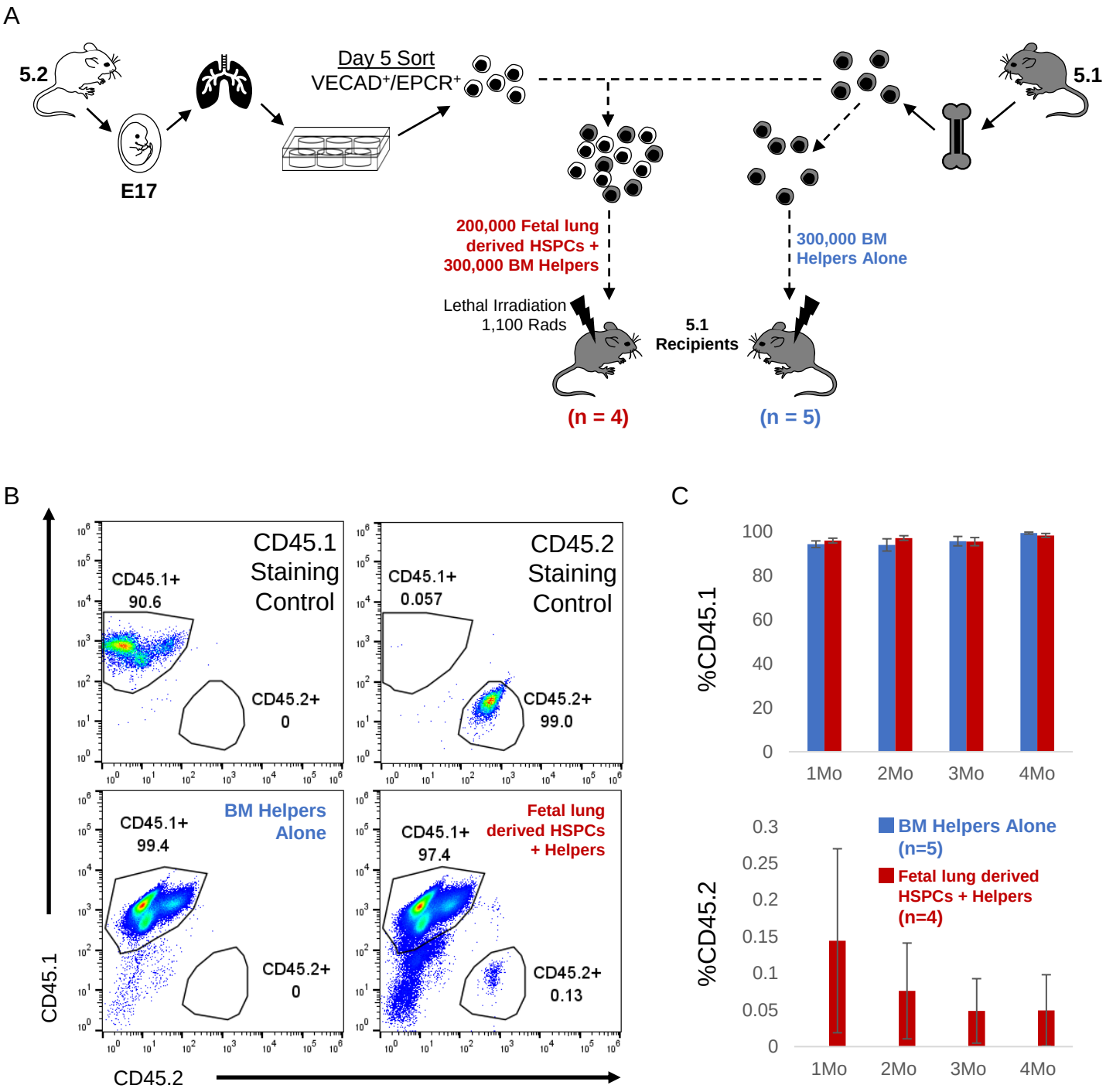
Reads were demultiplexed and aligned to the mouse genome assembly (GRCm38, Ensembl) with the STARsolo pipeline73. Further analyses were done using Seurat v. 3.1.474. After inspection of the quality control metrics, cells with more than 12% of mitochondrial content or less than 800 detected genes were excluded for downstream analyses. We normalized and scaled the UMI counts using the regularized negative binomial regression (SCTransform)75. Following the standard procedure in Seurat's pipeline, we performed linear dimensionality reduction (PCA), and used the top 20 principal components to compute both the UMAP76 and the clusters (Louvain method77) which were computed at a range of resolutions from 1.5 to 0.05 (more to fewer clusters). Cell cycle scores and classifications were done using the Seurat method78. The cut-offs for independent filtering79 prior to differential expression testing required genes: a) being detected in at least 10% of the cells of either population and b) having a natural log fold change of at least 0.25 between populations. The tests were performed using Seurat's wrapper for the MAST framework80. For a comparison on the performance of methods for single-cell differential expression see Soneson & Robinson et al.81. GEO accession: GSE197290, Token: wvqjuouqdtmrtuh.

## Statistical analysis

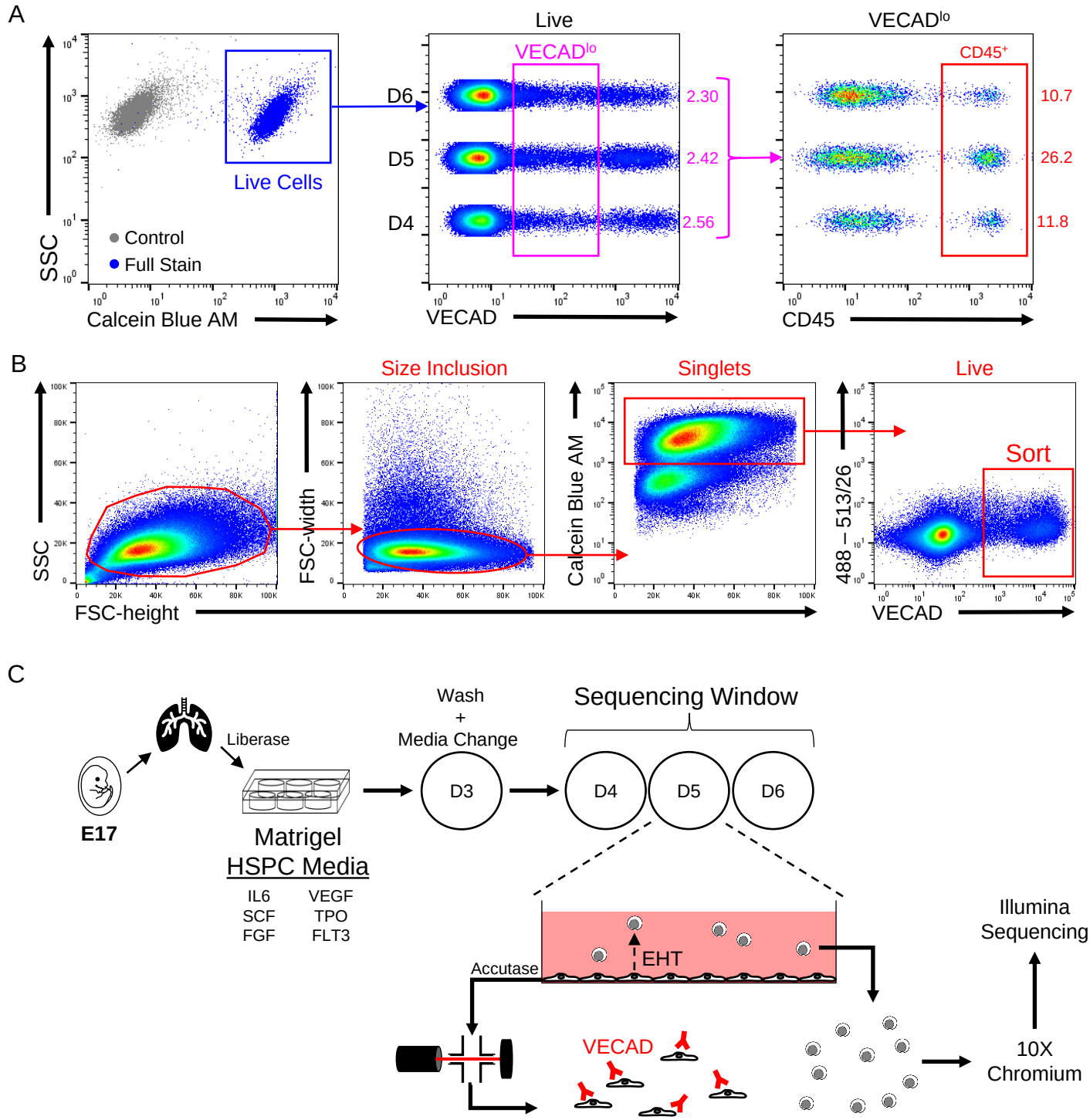
Significance for Enrichr based pathway enrichment was determined by Fisher exact test82,83. Significance for pairwise scRNA-Seq gene expression comparisons was determined using the MAST framework80. Significance for flow cytometry assays performed in biological triplicate was determined by student paired t-test.




**Supplemental Figure 1. Assessment of EHT and pre-HSC markers on fetal liver and fetal lung explants.** (A) Comparison of EHT and pre-HSC markers between fetal liver and fetal lung explants. (B) Comparison of EHT and pre-HSC markers between populations differentiated by VECAD and CD45 expression. *P < 0.05.



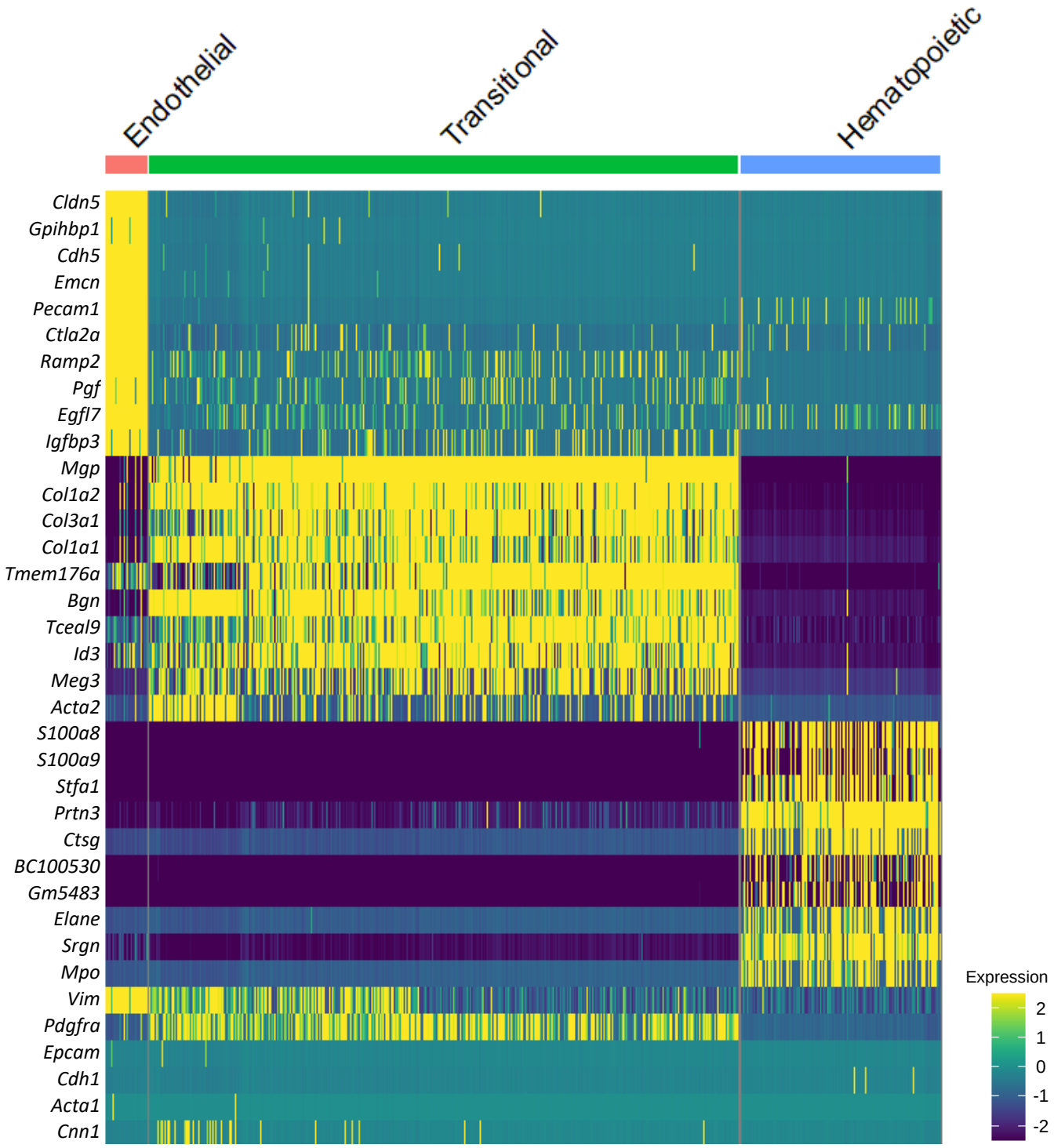

**Supplemental Figure 2. Competitive transplantation assay of fetal lung explant derived HSPCs.** (A) Schematic of transplantation assay. In brief, fetal lung explants from CD45.2 mice were cultured as described in the methods section. On day 5, all cells (adherent and suspension) were collected and stained for VECAD and EPCR. Live/VECAD$^+$/EPCR$^+$ cells were sorted using a MoFlo Astrios cell sorter. 200,000 Live/VECAD$^+$/EPCR$^+$/CD45.2$^+$ cells were combined with 300,000 bulk BM cells isolated from CD45.1 mice. (B) Representative flow cytometric assessment of CD45.1 versus CD45.2 cells in peripheral blood for evidence of engraftment. (C) Degree of engraftment assessed monthly for 4 months.



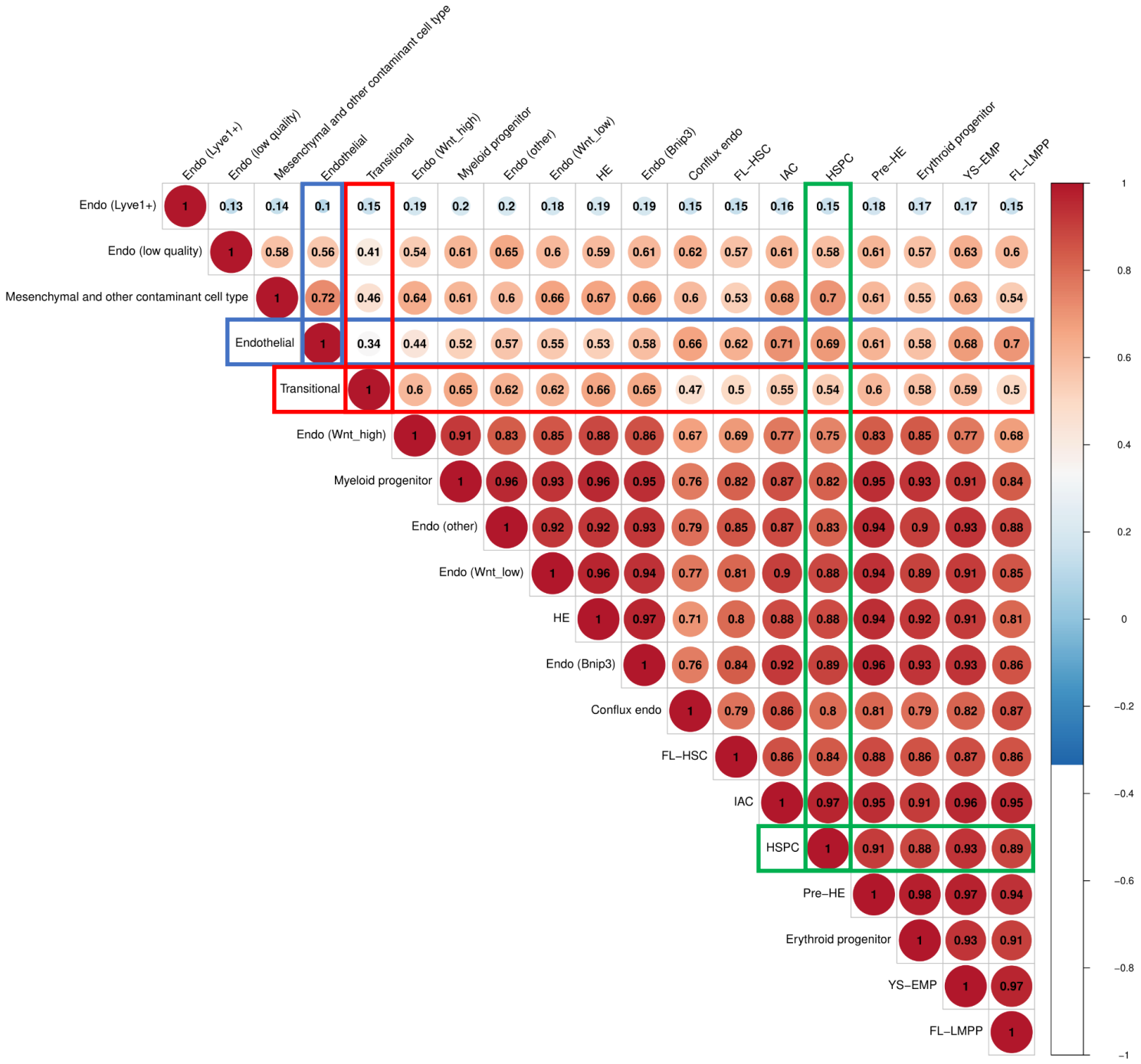

D

| Metric | Day 4 Suspension | Day 4 Adherent | Day 5 Suspension | Day 5 Adherent | Day 6 Suspension | Day 6 Adherent |
|---|---|---|---|---|---|---|
| Estimated Number of Cells | 2693 | 532 | 2497 | 1099 | 1398 | 1915 |
| Mean Reads per Cell | 16139 | 68754 | 17906 | 34019 | 39497 | 21106 |
| Median Genes per Cell | 2863 | 5681 | 2994 | 4098 | 3883 | 3162 |

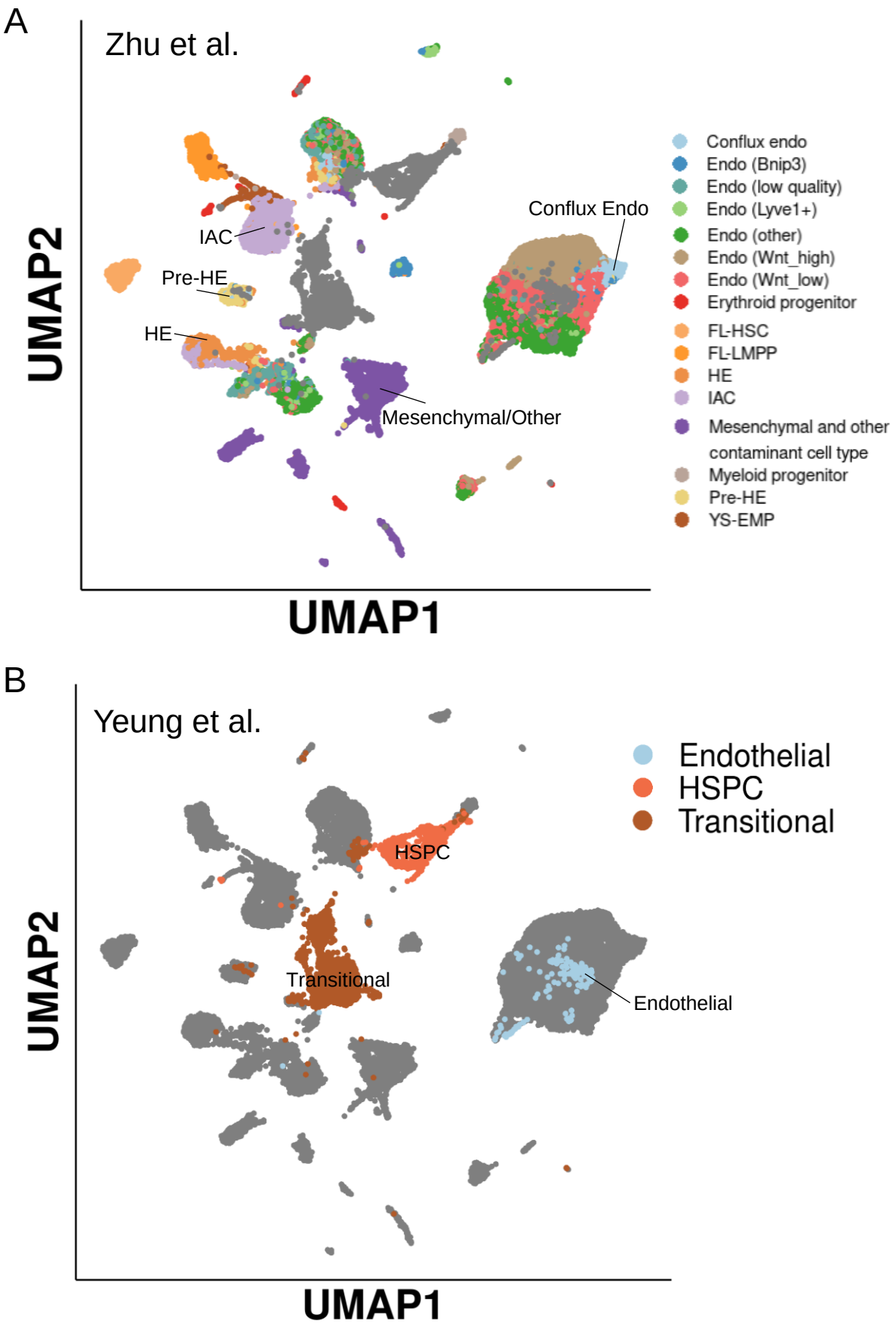
**Supplemental Figure 3. Optimization of scRNA-Seq protocol and analysis.** (A) Comparison of the fraction of VECAD$^{lo}$ cells that are CD45$^{+}$. (B) Sorting strategy for scRNA-Seq of adherent cells from fetal lung explants. (C) Schematic of scRNA-Seq procedure. (D) Table showing baseline metrics of scRNA-Seq analysis.




**Supplemental Figure 4. Heatmap of DGE analysis of fetal lung scRNA-Seq.** Top 10 differentially expressed genes for each cluster presented as a heat map. Common markers for fibroblasts (*Vim*, *Pdgfra*), epithelial cells (*Epcam*, *Cdh1*), and smooth muscle cells (*Acta1*, *Cnn1*) are also provided. Full enrichment analysis can be found in Supplemental_Data.

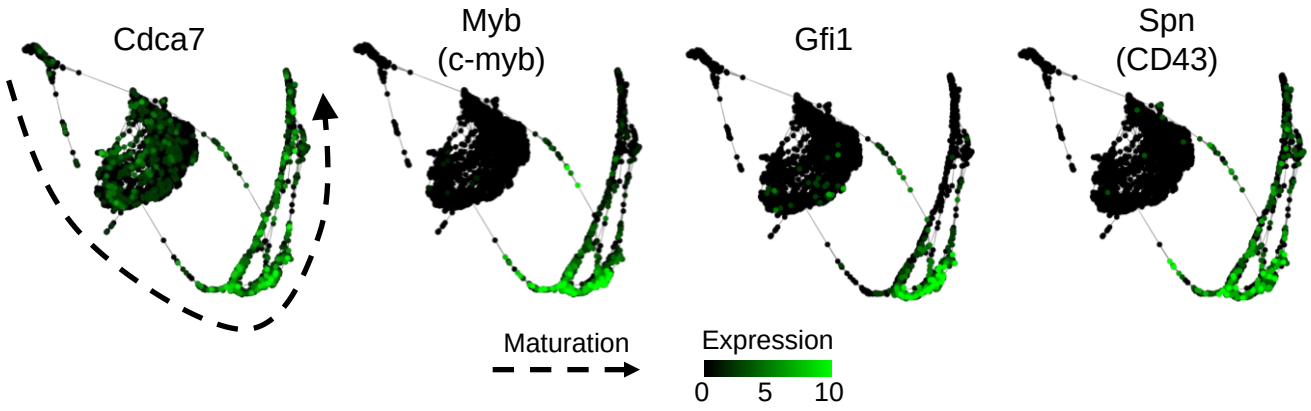



**Supplemental Figure 5. Pearson Correlation Plot Comparing Clusters Identified by Yeung et al. Versus Zhu et al.** scRNA-Seq datasets from Yeung et al. and Zhu et al.[35] were harmonized and analyzed by Pearson correlation analysis to assess for similarities in gene expression profile. Endothelial, Transitional, and HSPC clusters as defined by Yeung et al. are highlighted with blue, red, and green boxes, respectively, to assist in interpretation.




**Supplemental Figure 6. UMAP Projection of Harmonized Datasets from Yeung et al. and Zhu et al.** scRNA-Seq datasets from Yeung et al. and Zhu et al.[35] were harmonized and plotted by UMAP projection. (A) Cluster identities as defined by Zhu et al. are shown, and select clusters are highlighted to provide further clarity. (B) Cluster identities as defined here by Yeung et al. are shown.

A



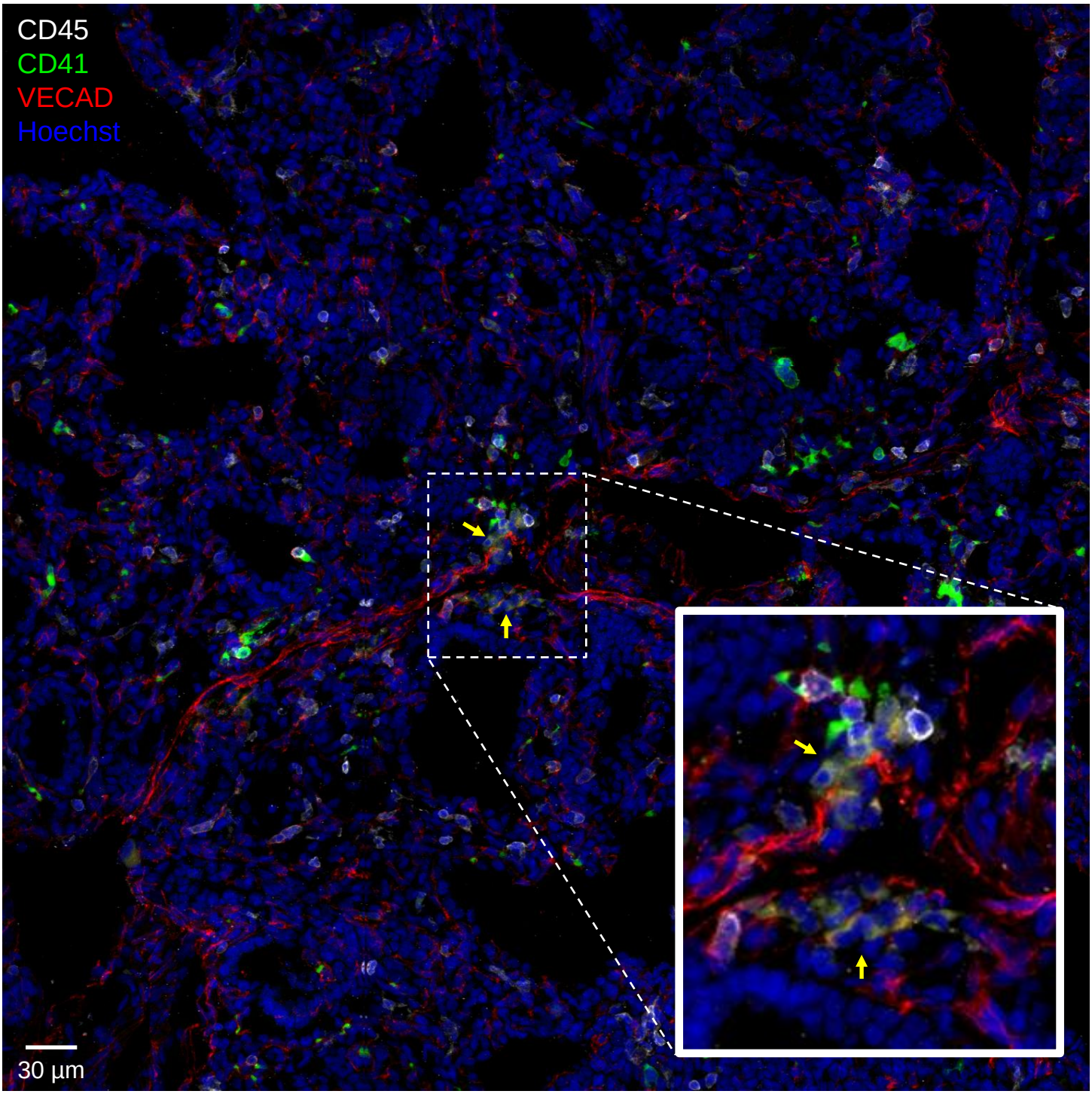

B

| Cluster | Index | Name | p-value |
|---|---|---|---|
| Endothelial | 1 | regulation of angiogenesis (GO:0045765) | 3.55E-08 |
| | 2 | apical junction assembly (GO:0043297) | 2.14E-06 |
| | 3 | bicellular tight junction assembly (GO:0070830) | 2.58E-05 |
| | 4 | blood vessel morphogenesis (GO:0048514) | 2.58E-05 |
| | 5 | tight junction assembly (GO:0120192) | 2.58E-05 |
| | 6 | glomerulus vasculature development (GO:0072012) | 2.58E-05 |
| | 7 | cell-cell junction assembly (GO:0007043) | 4.79E-05 |
| | 8 | maintenance of blood-brain barrier (GO:0035633) | 5.71E-05 |
| | 9 | extracellular structure organization (GO:0043062) | 5.71E-05 |
| | 10 | external encapsulating structure organization (GO:0045229) | 5.71E-05 |
| Transitional | 1 | extracellular matrix organization (GO:0030198) | 3.99E-05 |
| | 2 | skin morphogenesis (GO:0043589) | 0.000173 |
| | 3 | muscle contraction (GO:0006936) | 0.000173 |
| | 4 | extracellular structure organization (GO:0043062) | 0.000173 |
| | 5 | external encapsulating structure organization (GO:0045229) | 0.000173 |
| | 6 | skin development (GO:0043588) | 0.000222 |
| | 7 | supramolecular fiber organization (GO:0097435) | 2.69E-04 |
| | 8 | collagen fibril organization (GO:0030199) | 2.82E-04 |
| | 9 | regulation of membrane protein ectodomain proteolysis (GO:0051043) | 1.95E-03 |
| | 10 | negative regulation of dendritic cell differentiation (GO:2001199) | 4.07E-03 |
| Hematopoietic | 1 | neutrophil degranulation (GO:0043312) | 2.92E-13 |
| | 2 | neutrophil activation involved in immune response (GO:0002283) | 2.92E-13 |
| | 3 | neutrophil mediated immunity (GO:0002446) | 2.92E-13 |
| | 4 | defense response to bacterium (GO:0042742) | 2.61E-09 |
| | 5 | leukocyte aggregation (GO:0070486) | 0.000104 |
| | 6 | defense response to Gram-negative bacterium (GO:0050829) | 0.000112 |
| | 7 | neutrophil migration (GO:1990266) | 1.25E-04 |
| | 8 | innate immune response (GO:0045087) | 7.64E-04 |
| | 9 | regulation of phagocytosis (GO:0050764) | 8.85E-04 |
| | 10 | myeloid cell activation involved in immune response (GO:0002275) | 8.85E-04 |

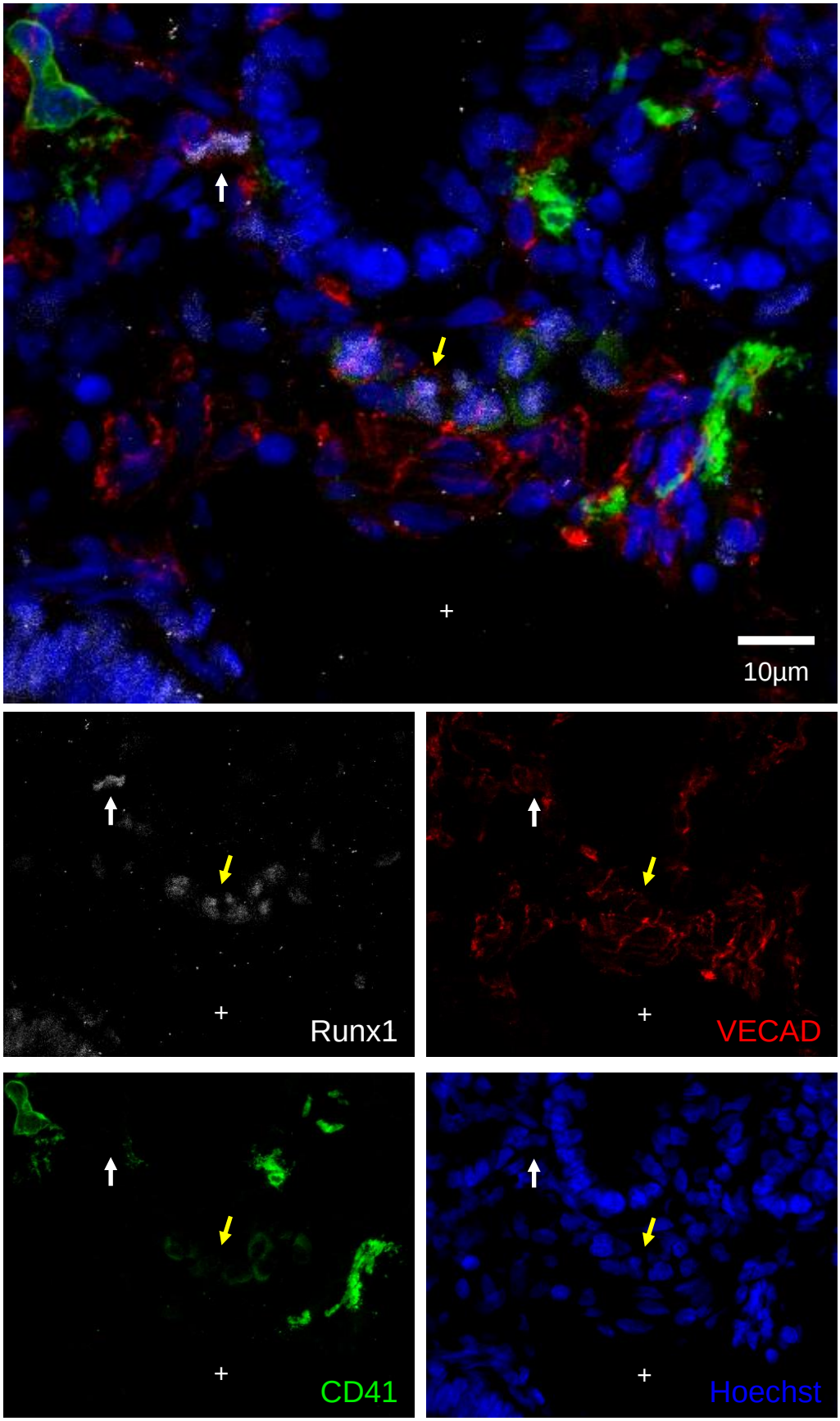
**Supplemental Figure 7. SPRING and DGE analysis of fetal lung scRNA-Seq.** (A) SPRING plots showing the expression of key hematopoietic commitment markers. The dotted arrow represents the trajectory of maturation. (B) Extended list of Gene Ontology biological processes enrichment analysis. Full enrichment analysis can be found in Supplemental_Data.

| | Gestational Age | Fetal Age | Gender | Chromosomal Trisomy |
|---|---|---|---|---|
| Donor Sample 1 | 20 Weeks | 18 Weeks | Unknown | None Detected |
| Donor Sample 2 | 24 Weeks | 22 Weeks | Female | None Detected |
| Donor Sample 3 | 21 Weeks | 19 Weeks | Unknown | None Detected |
| Donor Sample 4 | 24 Weeks | 22 Weeks | Male | None Detected |

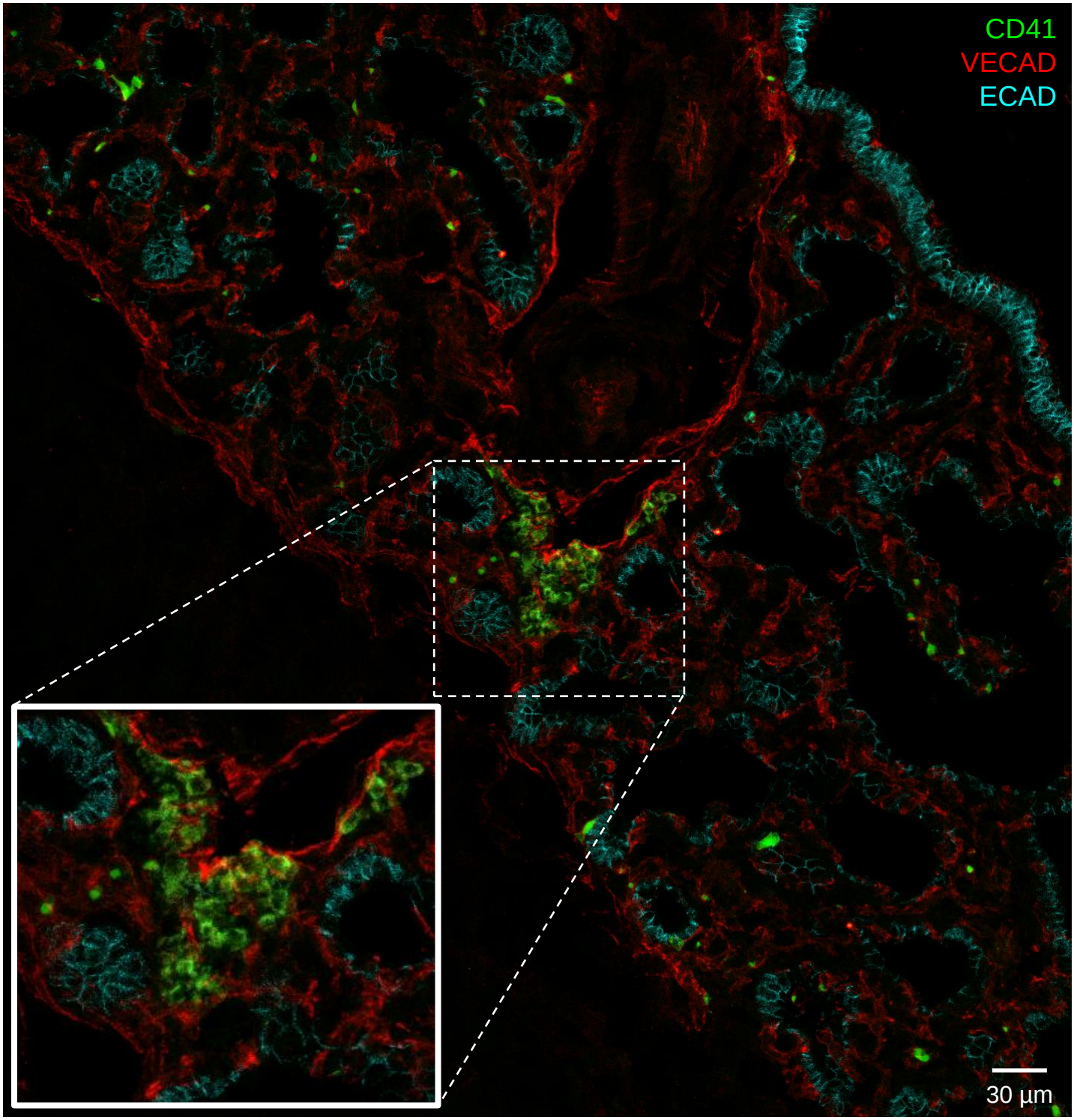
**Supplemental Table 1. Characteristics of human fetal donor tissue samples.**



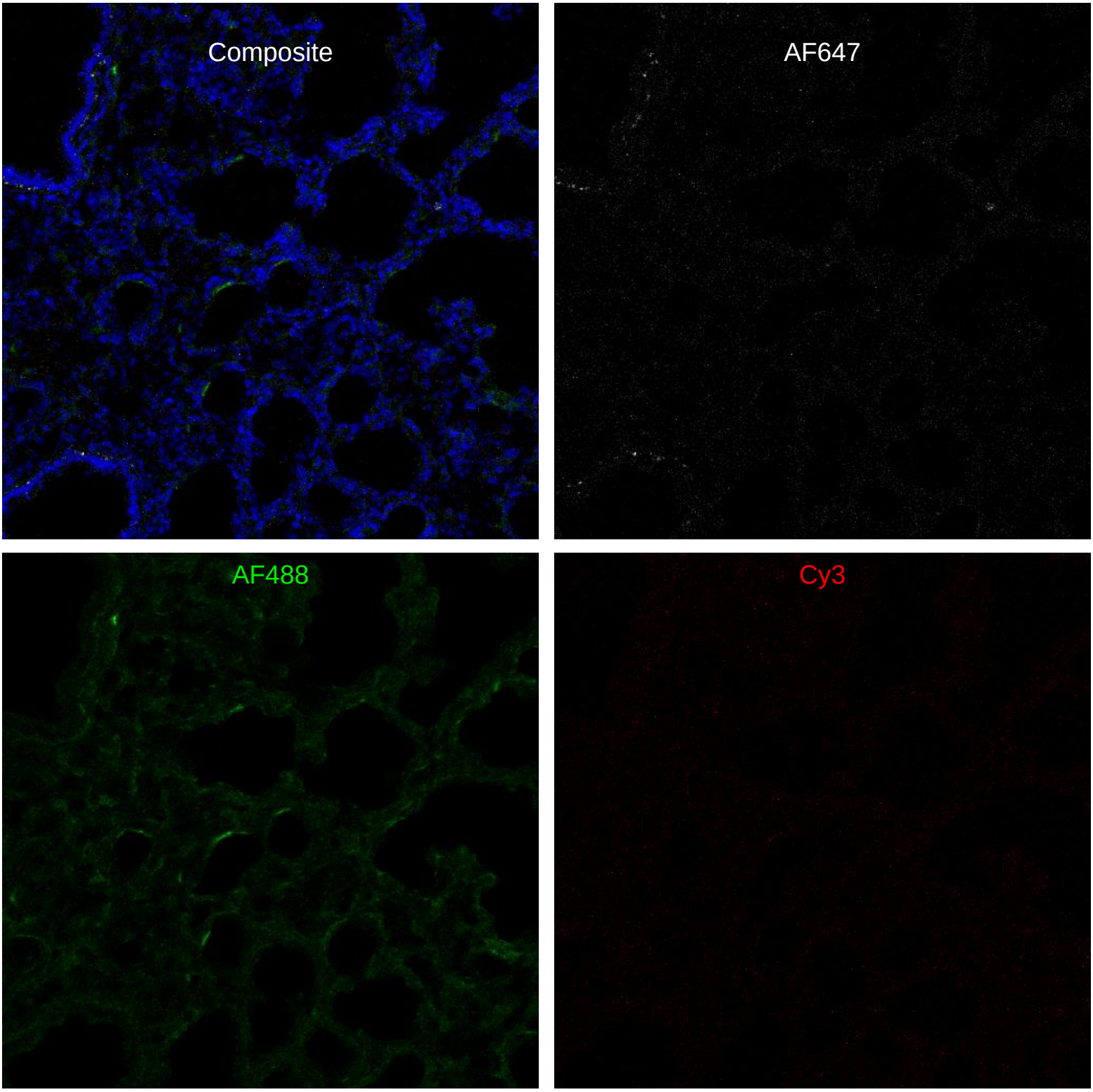

**Supplemental Figure 8. Wide field image of immunofluorescent staining of in-situ fetal lung EHT.** 12µm sections of a fixed frozen post-conception week 20 human fetal lung stained with the broad hematopoietic marker CD45, the EHT marker CD41, the endothelial marker VECAD, and the nuclear stain Hoechst. Cells co-expressing VECAD, CD41 and CD45 are highlighted with yellow arrows, demonstrating that EHT clusters are found in discrete regions.




**Supplemental Figure 9. Immunofluorescent staining shows Runx1 expression during in-situ fetal lung EHT.** 12µm sections of a fixed frozen post-conception week 20 human fetal lung stained with the EHT marker CD41, the nuclear EHT marker Runx1, the endothelial marker VECAD, and the nuclear stain Hoechst. Cells that are Runx1+ and VECAD+ are highlighted with a white arrow, while cells that are Runx1+ and CD41+ are highlighted with a yellow arrow. The lumen of a large vessel lined with VECAD$^+$ staining is highlighted with a "+".




**Supplemental Figure 10. Wide field image of immunofluorescent staining of in-situ fetal lung EHT clusters oriented towards the lung interstitium.** 12µm sections of a fixed frozen post-conception week 20 human fetal lung stained with the EHT marker CD41, the endothelial marker VECAD, and the epithelial marker epithelial cadherin (ECAD). Clusters of $CD41^{+}$ cells are closely associated with a large $VECAD^{+}$ vessel but are expanding into the interstitial space of the lung towards $ECAD^{+}$ developing lung epithelium.




**Supplemental Figure 11. Wide field image of secondary only control immunofluorescent staining of fetal lung.** Post-conception week 20 human fetal lungs were fixed in 4% paraformaldehyde, then frozen in OCT. 12µm cryosections were then blocked with a blocking solution containing bovine serum albumin, normal donkey serum, and triton x-100 for 30 minutes. Control sections were left to incubate overnight at 4$^{o}$C in a diluted blocking solution without any primary antibodies, then subsequently washed and stained with the corresponding secondary antibodies.

| Marker | Species | Conjugate | Clone | Vendor | Catalog. # | Dilution |
|---|---|---|---|---|---|---|
| CD235a | Human | PE | HIR2 | BD Biosciences | 555570 | 1:1000 |
| CD34 | Human | BV421 | 581 | BD Biosciences | 562577 | 1:100 |
| CD41 | Human | BV421 | HIP8 | Biolegend | 303730 | 1:200 |
| CD41 | Human | Unconjugated | HIP8 | BD Biosciences | BDB555465 | 1:100 |
| CD42b | Human | APC | HIP1 | BD Biosciences | 551061 | 1:200 |
| CD45 | Human | APC | 2D1 | Biolegend | 368512 | 1:200 |
| CD45 | Human | PE-Cy7 | 2D1 | Biolegend | 368532 | 1:200 |
| EPCR (CD201) | Human | APC | RCR-401 | Biolegend | 351906 | 1:100 |
| GPI80 | Human | PE | 3H9 | MBL | D087-5 | 1:50 |
| KDR | Human | PE | 89106 | BD Biosciences | 560494 | 1:200 |
| KDR | Human | APC | 89106 | BD Biosciences | 560871 | 1:200 |
| Vascular Endothelial Cadherin (VECAD/CD144) | Human | AF647 | 55-7H1 | BD Biosciences | 561567 | 1:100 |
| Vascular Endothelial Cadherin (VECAD/CD144) | Human | PE | 55-7H1 | BD Biosciences | 561714 | 1:100 |
| Vascular Endothelial Cadherin (VECAD/CD144) | Human | Unconjugated | Polyclonal | R&D Systems | AF938 | 1:100 |
| Anti-RUNX1 / AML1 + RUNX3 + RUNX2 | Human | Unconjugated | Monoclonal | Abcam | ab92336 | 1:100 |
| CD44 | Human/Murine | BV711 | IM7 | Biolegend | 103057 | 1:800 |
| CD45 | Human/Murine | Unconjugated | Polyclonal | Invitrogen | PIPA587427 | 1:100 |
| Epithelial Cadherin (E-Cadherin) | Human/Murine | Unconjugated | DECMA-1 | Santa Cruz Biotechnology | sc-59778 | 1:50 |
| CD150 | Murine | BV711 | TC15-12F12.2 | Biolegend | 115941 | 1:400 |
| CD41 | Murine | BV421 | MWReg30 | Biolegend | 133911 | 1:400 |
| CD43 | Murine | FITC | S11 | Biolegend | 143204 | 1:100 |
| CD45 | Murine | BUV395 | 30-F11 | BD Biosciences | 564279 | 1:200 |
| CD48 | Murine | BV605 | HM48-1 | Biolegend | 103441 | 1:100 |
| EPCR | Murine | APC | eBio1560 | eBioscience | 50-151-23 | 1:200 |
| Kit | Murine | BV510 | ACK2 | Biolegend | 135119 | 1:400 |
| Kit | Murine | APC | 2B8 | Biolegend | 105812 | 1:200 |
| Lineage Cocktail | Murine | FITC | Cocktail | Invitrogen | 22777072 | 1:5 |
| Sca-1 (Ly-6A/E) | Murine | PE | D7 | eBioscience | 50-110-49 | 1:200 |
| Vascular Endothelial Cadherin (VECAD/CD144) | Murine | Biotin | BV13 | Biolegend | 138008 | 1:200 |
| Streptavidin | N/A | APC-Cy7 | N/A | Biolegend | 405208 | 1:200 |
| Calcein Blue AM | Viability Dye | N/A | N/A | Invitrogen | C1429 | 1:1000 |
| Live/Dead Fixable Blue | Viability Dye | N/A | N/A | Invitrogen | L34961 | 1:800 |

**Supplemental Table 2. Antibodies used for flow cytometry and immunofluorescent microscopy.**