High-parametric protein maps reveal the spatial organization in early-developing human lung

Sanem Sariyar[1,2], Alexandros Sountoulidis[1,3], Jan Niklas Hansen[1,2,4], Sergio Marco Salas[1,5], Mariya Mardamshina[1,2], Anna Martinez Casals[1,2,4], Frederic Ballllosera Navarro[1,2,4], Zaneta Andrusivova[1,6], Xiaofei Li[7], Paulo Czarnewski[1,6], Joakim Lundeberg[1,6], Sten Linnarsson[8], Mats Nilsson[1,5], Erik Sundström[7], Christos Samakovlis[1,3,9], Emma Lundberg[1,2,4,10,*,#], Burcu Ayoglu[1,2,*,#]

[1]Science for Life Laboratory, Solna, Sweden
[2]Department of Protein Science, KTH - Royal Institute of Technology, Stockholm, Sweden
[3]Department of Molecular Biosciences, Wenner-Gren Institute, Stockholm University, Stockholm, Sweden
[4]Department of Bioengineering, Stanford University, Stanford, CA, USA
[5]Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden.
[6]Department of Gene Technology, KTH - Royal Institute of Technology, Stockholm, Sweden
[7]Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden
[8]Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden
[9]Molecular Pneumology, Cardiopulmonary Institute, Justus Liebig University, Giessen, Germany
[10]Department of Pathology, Stanford University, Stanford, CA, USA

* Equal contribution
# Correspondence should be addressed to: burcu.ayoglu@scilifelab.se or emmalu@stanford.edu

ABSTRACT

The respiratory system, encompassing the lungs, trachea, and vasculature, is essential for terrestrial life. Although recent research has illuminated aspects of lung development, such as cell lineage origins and their molecular drivers, much of our knowledge is still based on animal models, or is deduced from transcriptome analyses. In this study, conducted within the Human Developmental Cell Atlas (HDCA) initiative, we describe the spatiotemporal organization of lung during the first trimester of human gestation *in situ* and at protein level. We used high-parametric tissue imaging on human lung samples, aged 6 to 13 post-conception weeks, using a 30-plex antibody panel. Our approach yielded over 2 million individual lung cells across five developmental timepoints, with an in-depth analysis of nearly 1 million cells. We present a spatially resolved cell type composition of the developing human lung, with a particular emphasis on their proliferative states, spatial arrangement traits, and their temporal evolution throughout lung development. We also offer new insights into the emerging patterns of immune cells during lung development. To the best of our knowledge, this study is the most extensive protein-level examination of the developing human lung. The generated dataset is a valuable resource for further research into the developmental roots of human respiratory health and disease.

KEYWORDS

Human lung development, Human Developmental Cell Atlas, spatial proteomics, single-cell proteomics, proliferation

INTRODUCTION

Studies on the developing human lung may not only provide a better understanding of how the organ is formed, but is also key to developing prevention and treatment strategies for respiratory diseases[1,2]. Various cell types form and interact over time to reach the mature state of adult lung. In recent years, single-cell transcriptomics studies have identified close to 60 different cell types in the adult human lung[3,4]. Despite limited availability of human embryonic tissues, recent initiatives like the Human Developmental Cell Atlas (HDCA)[5], have facilitated single-cell or spatial transcriptomic characterization of over 80 different cell types and states in the developing human lung[6-11]. These studies have shed light on the dynamic processes involved in lung development and have provided important insights into the spatial organization of cells in the early-developing human lung.

To our knowledge, there are currently no reports providing a comprehensive overview of the single-cell resolved spatial organization of the main cell types in the developing human lung at the protein level. As proteins are the effectors of biological processes, knowledge of their distribution and relative abundance within different cell types and tissue regions can reveal functional insights that may not be evident from spatial transcriptomics data alone, or from single-cell transcriptomics data of dissociated cells, where valuable spatial information is *per se* lost.

In the present study, we undertook a spatially resolved protein analysis of the developing human lung during the first trimester of gestation. We developed a panel of oligo-barcoded antibodies specifically validated to visualize cell types in the developing human lung, and used high-parametric microscopy for simultaneous analysis of 30 proteins in sections of human prenatal lung tissue. The 30-plex antibody panel was used to generate protein data in more than 2 million single cells in the human lung at five distinct time points ranging between 6 to 13 post-conception weeks (pcw) in order to identify the spatial organization of and between main cell types, their degree of proliferation, and how these change during the first trimester of gestation.

MATERIALS AND METHODS

Prenatal tissue samples and ethics

The lung samples originated from 5-13 post-conception week (pcw) old embryos/fetuses obtained after elective medical abortions with oral and written consent given by the pregnant women. The age of each embryo/fetus was determined using clinical ultrasound, actual crown-rump length and other anatomical landmarks. The use of fetal material was approved by the Swedish National Board of Health and Welfare. The analysis using this material was approved by the Swedish Ethical Review Authority (2018/769-31).

Tissue treatment for spatial analyses

The left lung from each donor was snap-frozen in cryo-matrix for spatial analyses. The sectioning of 10 µm-thick tissues was performed with a cryostat (Leica CM3050S) on slides (Epredia, J1800AMNZ) as previously described[7]. Sections were dried on a silica gel at 37ºC for 15 minutes and then stored at -80ºC.

Immunostaining for high-parametric imaging

Tissues were put on Drierite beads (Alfa Aesar, K13Z026) directly after obtaining the samples from a -80°C freezer, according to the manufacturer's protocols (Akoya Biosciences, Marlborough, MA). Tissues were first incubated in acetone (Sigma, 650501) for 10 minutes, then in hydration buffer (Akoya Biosciences, 7000008) for 2 minutes, and fixed in 1.6% paraformaldehyde (Thermo Scientific, 043368.9M) diluted in hydration buffer for 10 minutes. Tissues were then incubated in a staining buffer (Akoya Biosciences, 7000008) for 30 minutes. After equilibration, tissues were incubated with blocking buffer (Akoya Biosciences, 7000008) and primary antibody cocktail (dilutions indicated in ***Supplementary Table 1***) for 3 hours inside a humidity chamber and at room temperature (RT). The samples were washed with a staining buffer for 2 minutes, and then fixed in 1.6% PFA diluted in a storage buffer (Akoya Biosciences, 7000008) for 10 minutes. The samples were then washed with phosphate-buffered saline (PBS), incubated in cold methanol (Sigma, 322415) (4°C) for 5 minutes, and washed again with PBS.

Finally, the samples were fixed in a fixative reagent (Akoya Biosciences, 7000008) and washed again with PBS before being stored in storage buffer at 4°C.

Indirect immunofluorescence

Samples were equilibrated at RT for 15 minutes and fixed with 4% paraformaldehyde (Thermo Scientific, 043368.9M) for 15 minutes. After fixation, samples were washed 3x in PBS for 15 minutes. Antibodies, diluted in PBS containing 0.3% Triton (Sigma, T8787), were added (200 µl /slide) and samples were incubated overnight at 4°C. The samples were blocked for 30 minutes in 1xTBS (Medicago, 097500100) containing 0.5% TSA blocking reagent (Perkin Elmer, FP1020) and Hoechst (Invitrogen, H3570). Secondary antibodies were diluted in a blocking buffer and incubated for 90 minutes at RT. After incubation, slides were washed 3x with TBS-Tween for 15 minutes. Coverslips (VWR, 631-0147) were mounted on the slides using Fluoroshield mounting medium (Invitrogen, 00495802).

Antibody conjugation to build multiplexed panels

The high-parametric antibody panel consisted of both a selection of pre-conjugated antibodies (n=23, Akoya Biosciences) and antibodies available through other commercial sources that were conjugated in house (n=49, ***Supplementary Table 1***). The latter were free from additives or preservatives that might potentially affect the conjugation efficiency. For conjugation, the following protocol was used (Akoya Biosciences): Filters with 50 kDa molecular weight cut off were first blocked using 500 µl of filter blocking solution (Akoya Biosciences, 7000009). For each conjugation, 50 µg of antibody was used and if the volume was less than 100 µl, PBS was added to reach 100 µl. Antibodies were transferred on blocked filters and centrifuged at 12,000 x g for 8 minutes. The reduction mixture (Akoya Biosciences, 7000009) was prepared, put on the filters containing antibodies, and incubated for 30 minutes at RT. The reduction solution was removed by centrifugation (12,000 x g) and exchanged with the conjugation solution (Akoya Biosciences, 7000009). Each oligonucleotide barcode was rehydrated in 10 µl nuclease-free water (Life Technologies, AM9937) and 210 µl of conjugation solution was added. The corresponding barcodes were added on top of the reduced primary antibodies and incubated for 2 hours at RT. After centrifugation (12,000 x g), purification solution (Akoya Biosciences, 7000009) was added,

and antibodies were purified in several steps of centrifugation. Then, 100 µl of antibody storage buffer (Akoya Biosciences, 7000009) was added on top of antibodies and centrifuged (3000 x g) for 2 minutes.

Antibody conjugation was validated by manual incubation with PhenoCycler reporters: After staining with the conjugated antibodies, the tissue was incubated in a screening buffer containing 10x PhenoCycler buffer (Akoya Biosciences, 7000001), nuclease-free water and DMSO. Meanwhile, a reporter stock solution was prepared including screening buffer, assay reagent (Akoya Biosciences, 7000002) and nuclear stain (Akoya Biosciences, 7000003). Here, 2.5 µl from each reporter was mixed with the reporter stock solution and the end volume was 100 µl for each slide. Slides were incubated with the reporters for 5 minutes in the dark and then washed with the screening buffer several times before mounting on a slide using Fluoroshield mounting medium (Invitrogen, 00495802).

Automated collection of high-parametric immunofluorescence data

Reporters were diluted in reporter stock solution composed of nuclease-free water (Life Technologies, AM9937), 10x PhenoCycler buffer, assay reagent, and nuclear stain. Diluted reporters were put into the corresponding wells according to the experimental plan. Fully automated imaging was performed with the PhenoCycler Fusion system (Akoya Biosciences) ***(Supplementary Figure 1)***. Acquired images were processed by the PhenoCycler processor (Akoya Biosciences) and ".qptiff"-image files were obtained as output.

Analysis of high-parametric imaging data

Image processing and segmentation

The ".qptiff" files were opened in FIJI[12] and were converted into separate .tiff images for each marker channel. These single-channel images were then uploaded to our in-house developed software PIPEX (https://github.com/CellProfiling/pipex.git), performing segmentation and various downstream single-cell data analysis. DAPI and EPCAM marker images were preprocessed (normalization, background subtraction and otsu thresholding) to get the most

significant intensity thresholds. Then, the IF Stardist ML model[13] was executed over the preprocessed DAPI marker to get the pure and expanded nuclei shapes (nuclei masks). A watershed algorithm was executed over the preprocessed EPCAM marker to get the membrane boundary shape information (membrane mask). The nuclei mask expansions (not the pure nuclei shapes) were cut to fit the membrane boundaries calculated in the membrane mask. In the locations where no/incorrect nuclei were detected by Stardist, the membrane shapes themselves were kept as segmented cells. These were considered as final segmented cells. A sanity filtering was applied using cell size and a combination of DAPI/EPCAM intensities to discard segmented cells that were too small and/or inconsistent with the overall intensity levels. With this final combined mask, .csv files were generated for all segmented cells, including mean intensities and binarizations for all the original marker images (including original DAPI and EPCAM images) ***(Supplementary Figure 2)***. Next, Perseus[14] was used for each sample to exclude the areas presumably not belonging to the lung tissue, or showing sectioning or staining artefacts **(*Supplementary Figure 3*)**. The output after this data clean-up was used for the downstream quantitative analysis. A manual background subtraction was performed prior to the clustering step. For each marker, images were inspected using FIJI, where background values were measured in regions of interest (ROIs) where no signal was observed and where the background signal appeared relatively high (***Supplementary Table 2***). Background values defined for each marker were subtracted from the values in the segmented dataset.

Clustering and annotation of cell populations

Segmented and background-subtracted values were used to generate annotated data matrices using Anndata[15]. Data was then normalized using Arcsinh normalization[16]. To identify the cell type populations present *in situ*, Scanpy's[17] PCA (sc.tl.pca(adata)), find neighbors (sc.pp.neighbors(adata,n_neighbors= 10)), UMAP (sc.tl.umap(adata,min_dist= 0.1, spread= 0.3, negative_sample_rate= 4)) and Leiden clustering (sc.tl.leiden(adata,resolution=1)) functions were used. After the first round of clustering, clusters were investigated and clusters with artefacts were removed (***Supplementary Figure 4***).

Datasets of each sample after artefact removal were merged using ComBat (sc.pp.combat(adata, key='age'))[18]. The same parameters as for the initial clustering were used for clustering of the

merged data. Density plots (sc.pl.embedding_density(adata, groupby='age') were used to investigate the efficiency of data merging (***Supplementary Figure 5A***). The merged data was visualized in a UMAP colored for each week (***Supplementary Figure 5A)*** and clusters of the merged dataset were annotated for each cell type in accordance with the dot plots (sc.pl.dotplot (adata, markers, groupby= 'leiden', dendrogram=True, size_title=None) (***Supplementary Figure 5B***). Based on the cell type annotations, the percentage of each cell type was calculated.

For further analysis of individual weeks, clustering was performed using the previously mentioned parameters. Cell type annotations were performed based on the dot plots (***Supplementary Figure 6)***. After annotations, clusters belonging to the same cell type were merged (***Supplementary Figure 7***). Some cell types were annotated further for positivity for certain markers such as Vim+ (Vimentin positive) mesenchymal or Vim+DCN+ (Vimentin and Decorin positive) mesenchymal. For calculating and comparing the main cell type percentages in each week, clusters with these detailed annotations were merged back with their main cell type cluster, *e.g.* Vim+ mesenchymal, Vim+ DCN+ mesenchymal, CD90+ mesenchymal, and Vim+ CD56+ mesenchymal clusters were combined with mesenchymal cluster for a comparison across different weeks. Similarly, immature airway smooth muscle cells, adventitial fibroblasts, airway fibroblasts, chondroblasts, and neural cell clusters were combined with mesenchymal cell cluster, and myeloid and lymphoid immune cells were combined as immune cluster for comparison across different weeks. Correlation between the expression patterns of markers was calculated using Pearson correlation (***Supplementary Figure 8***).

In order to do a comparison between cell type percentages obtained in our current protein and previous transcriptomics datasets, single-cell RNA sequencing data from Sountoulidis *et al.*[7] was used and cells were grouped into similar cell type annotations (***Figure 2C***).

Analysis of cell type adjacency

Spatial localization patterns of annotated cell type clusters were analyzed using spatial neighbors function (sq.gr.spatial_neighbors(adata, n_neighs=10, coord_type="generic")) in Squidpy[19], which creates a graph from spatial coordinates based on the number of neighbors selected, which was 10 for this analysis. To visualize the connectivity between cells, xy-scatter plots were used

(***Figure 3A***). To quantitatively analyze the cell type adjacency patterns, enrichment analysis was performed (sq.gr.nhood_enrichment(adata_new, cluster_key="annotation")), and enrichment Z-scores were plotted in heatmaps (***Figure 3B).***

Identification of spatial domains

UTAG package[20] was used for analysis of the tissue topography by unsupervised identification of spatial cellular domains. To carry out the unsupervised domain analysis in weeks 6 and 8.5, originally selected areas (***Supplementary Figure 3A***) could be used fully, but due to the large tissue size in the 11-, 12-, and 13-week-old lung samples, only a portion of the sample was used (***Supplementary Figure 9A***). First, Anndata object was created for the selected areas, including the annotation matrices based on normalized data, clusters, and annotations for each cluster. UTAG domains were calculated by combining both adjacency and expression matrices. Their dot products were clustered using Leiden clustering with the function utag(adata, slide_key="roi", max_dist=12, normalization_mode="l1_norm", apply_clustering=True, clustering_method = 'leiden', resolutions = [0.05, 0.1, 0.3]). For the clustering resolution, a value of 0.3 was selected for all weeks except for week 12, where a value of 0.5 was chosen. Identified spatial domains were investigated and domains with similar marker expressions and spatial distributions were combined. Identified domains were plotted in xy-scatter plots.

For quantitative analysis of the spatial distribution of the identified domains, Ripley's L was calculated using Squidpy's squidpy.gr.ripley function with default settings (***Figure 4C***). A model of physical proximity between the identified spatial domains was created by using the "Draw network" function with default settings (***Figure 4B***). Finally, their annotations were done using bar plots showing the cell type composition in each spatial domain (***Supplementary Figure 9B)***.

Analysis of proliferation patterns

Using the background-subtracted data frame for all weeks, a Ki67 threshold corresponding to the 13th-15th percentile was selected in each week. All the Ki67+ cells were then integrated into the Anndata object and were annotated as 'proliferating". Proliferating cells were counted and proliferating fractions were calculated for each week (***Figure 5B-C***). To compare the degree of proliferation across different cell types in different weeks, broadly annotated cell types were used

(***Supplementary Figure 10A***). Spatial localization patterns of proliferating and non-proliferating cells were analyzed using sq.gr.nhood_enrichment(adata, cluster_key="proliferation") function and Z-score outputs were visualized in heatmaps (***Figure 5D***).

The degree of proliferation was compared further between SOX2^{high} airway regions, larger SOX2^{high} airway regions and SOX9^{high} airway regions (***Figure 5E)***. For this, .h5ad files were uploaded to TissUUmaps[21] and airway regions were selected on images for each of these three categories (***Supplementary Figure 10B, Supplementary Table 3***). Here, only 11-, 12- and 13-week-old samples were used, as the overall size of airway-like structures in earlier weeks did not allow to distinguish well further between these three structures. In these three categories of selected regions, the numbers of proliferating cells were calculated. Cell number counts for proliferating and non-proliferating cells were then used to perform Fisher's exact test to calculate the statistical significance of differences in proliferation patterns in these regions (***Figure 5F, Supplementary Figure 10C***).

Subclustering of immune cells

For each week, the immune cell cluster was subclustered using the following markers: CD3, CD4, CD44, CD45, CD56, CD68, CD163, HLA-DR, Ki67 and MRC1. For clustering, Scanpy's PCA (sc.tl.pca(adata)), find neighbors (sc.pp.neighbors(adata,n_neighbors= 10)), UMAP (sc.tl.umap(adata,min_dist= 0.1, spread= 0.3, negative_sample_rate= 4)) and Leiden clustering (sc.tl.leiden(adata,resolution=0.2)) was used. For week 8.5 and 12, a resolution parameter of 0.3 was used. The resulting clusters were annotated by inspecting their expression profiles in dot plots (***Supplementary Figure 11***) and clusters with the same annotations were merged (***Figure 6A***). In weeks 11,12, and 13, B and NK cell annotations were integrated to the analysis with the help of manual selection based on the marker expression and co-expression of CD19 and CD45 for B cells, and CD56 and CD45 for NK cells. The percentages of the cell types were calculated based on the cell counts (***Figure 6B***). Based on Ki67 expression, proliferating immune cell subtypes were identified and integrated to the immune cell Anndata object. The fraction of proliferating cell types were calculated based on the cell counts (***Figure 6D***). Spatial proximity between immune cell subtypes were calculated using sq.gr.spatial_neighbors with n_neighs set to 10 and sq.gr.nhood_enrichment functions and plotted with sq.pl.spatial_scatter function by setting the

connectivity key to ‘spatial connectivities’. A model of physical proximity between cell types was created by using the “Draw network” function with default settings and as an input ‘count’ from sq.gr.nhood_enrichment was used (***Figure 6E***).

Comparative analyses of immune cells encircling vasculature

Groups of immune cells encircling arteries were observed in 11-, 12- and 13-week-old samples and comparative analyses were performed for these cells. First, .qptiff files for these weeks were loaded to Qupath and with the help from endothelial cell marker CD144 and smooth muscle marker ACTA2, arterial structures in each tissue section were identified on the images. Immune cells encircling these arterial structures were selected manually based on the CD45 marker expression. These manually selected immune cells were collected in a .json file and cell ID information was gathered. Gathered cell IDs were added to the Anndata object and tagged as “arterial”. To analyze the composition of cell types that were within a 50 μm radius around the selected immune cells, an in-house script based on calculation of Euclidean distance between the centers of two cells was used (***Supplementary Figure 12A***). For differential marker analysis, the sc.tl.rank_genes_groups(adata, 'leiden', method='t-test',corr_method='benjamini-hochberg') function was used, where T-test was applied, the p-value was corrected using Benjamini-Hochberg correction. In the output, markers were ranked based on Z-scores (***Figure 7C***).

RESULTS

Development of a multiplexed antibody panel to characterize the developing human lung microenvironment

To comprehensively characterize the cellular heterogeneity and spatial organization of developing human lung tissue, we initially selected 73 antibodies that could potentially be used in a multiplexed antibody panel. We tested each antibody in three consecutive validation steps: (1) indirect immunofluorescence staining (IF); (2) validation of the oligo barcode-conjugated antibody; (3) high-parametric imaging (***Figure 1A***). We excluded antibodies that did not yield a sufficient signal to noise ratio (SNR) and/or showed unspecific or no staining. After these stringent validation rounds, 30 of the 73 antibodies were selected to be included in the final antibody panel and their SNRs were optimized further by titrating the antibody dilution factors (***Figure 1B, Supplementary Table 1***). Among the antibodies that were excluded, 43% were excluded after initial IF screening due to non-specific staining patterns, 19% were excluded after barcode-conjugation, and 38% were excluded after evaluation of staining performance in optimization runs as part of the multiplexed panel. These results highlight the increased complexity of multiplexed antibody assays and emphasize the importance of validation efforts.

The panel design was based on cell type markers proposed in recent single-cell transcriptomic studies of human adult[3,22] and developing lung[7] (***Figure 1C, Table 1***). We included general epithelial markers, such as EPCAM (Epithelial cell adhesion molecule), PanCK (Pan Cytokeratin), CDH-1 (Epithelial cadherin, E-Cadherin), TTF1 (Transcription termination factor 1, also known as Nkx2.1), as well as SOX2 and SOX9 (SRY-box transcription factor 2 and 9) in the panel to map the proximal and distal epithelial cells[23]. For endothelial cells, we incorporated CD31, CD34, CD123, CD144 and CLDN5 (Claudin 5) as general endothelial markers, PDPN (Podoplanin) as a lymphatic endothelial marker and PRX (Periaxin) as a capillary marker. We added markers such as VIM (Vimentin), DCN (Decorin), and COL1A1 (Collagen Type I) to map the mesenchymal cells, and ACTA2 to map smooth muscle cells. In addition to the general immune markers of CD45 and HLA-DR (Human leukocyte antigen DR isotype), we included other markers to further delineate immune cells, such as CD3 and CD4 for T cells and CD68, CD163 and MRC1 (Mannose

receptor C-type 1) for macrophages, and CD19 for B cells. We also included Ki67 to identify proliferative cell states (***Figure 1C***).

Spatially resolved cell type composition of the developing human lung

Using the validated multiplexed antibody panel, we performed tissue imaging on 6-, 8.5-, 11-, 12-, and 13-post-conception-week-old human lung tissue sections ***(Figure 1D-F, Supplementary Figure 13)***. We segmented images using DAPI as a nuclear marker and EPCAM as a membrane marker, and selected regions for downstream analysis. For each single cell, we determined the intensity for each protein marker and merged the resulting single-cell resolved protein expression data for all weeks (***Supplementary Figure 5A***). Next, we annotated the identified cell type clusters finely (***Supplementary Figure 5B)*** and combined clusters belonging to the same major cell type for a broader annotation (***Supplementary Figure 14A***). As shown in ***Supplementary Figure 14B***, the expression of markers such as SOX2, SOX9, TTF1, EPCAM, and E-cadherin (CDH1) was high in epithelial cell clusters. Also, CD44 was expressed in both SOX2^{high} (proximal) and SOX9^{high} (distal) epithelial cells, with relatively higher expression in the latter cell cluster. In endothelial and lymphatic endothelial cell clusters, relatively higher levels of CD31, CD144, and CD90 were observed. Lymphatic endothelial cells exhibited lower levels of CD123 than endothelial cells, but higher levels of PDPN. As summarized in ***Supplementary Figure 14C***, the most abundant cell type in the developing human lung was mesenchymal, accounting for 31.4% of the cell population, followed by endothelial cells (27.4%), SOX9^{high} epithelial cells (17.9%), SOX2^{high} epithelial cells (7.6%), immune cells (5%), airway smooth muscle cells (4.7%) and chondroblasts (0.6%).

Next, we clustered the data individually for each week to monitor how the composition and spatial localization of the major cell types change in time and space over the course of early lung development (***Supplementary Figure 6-7***). The cell type annotation for each cell was visualized in scatterplots to assess and confirm microanatomic consistency (***Figure 2A***). Endothelial structures had already formed at week 6, while continuous endothelial structures appeared from week 8.5 onwards (***Figure 2D***). Lymphatic endothelial structures were not observed at week 6, but were present at all later weeks (***Figure 2E***, ***Supplementary Figure 15***). Immune cells were

observed starting from week 6 (***Figure 2A***). The expression patterns of each marker across the annotated cell types (***Figure 2B)*** as well as correlations between the markers (***Supplementary Figure 8)*** were in agreement with previous reports[6,7].

Mesenchymal cells were the most abundant cell type in each week ***(Figure 2C)*** and the highest fraction of them was observed at week 6 (57.7% of the total cell population). Endothelial cells were the second most abundant cell type, with an increasing trend during development (8.5% at week 6 and 20-30% at weeks 8.5, 11 and 13). Endothelial cells were followed by SOX9^{high} epithelial cells, SOX2^{high} epithelial cells, airway smooth muscle cells and immune cells. Immune cells reached their highest abundance in week 11 and 12, constituting 7.2% and 7.3% of the total cell population, respectively. SOX2^{high} epithelial cells reached their highest abundance in week 12 and 13 (9.4% and 9.6%, respectively). Similarly, SOX9^{high} epithelial cells reached their highest abundance in week 12 (15.9%). Lymphatic endothelial cells, which were not identified in week 6, were present in week 8.5, peaked in week 11 and their abundance decreased to 1.4% and 1% in weeks 12 and 13, respectively (***Figure 2E***, ***Supplementary Figure 15)***.

Our findings present a cell type distribution that aligns with the patterns identified in the single-cell transcriptomics data reported by Sountoulidis *et al.* 2023[7], where mesenchymal cells emerged as the most abundant cell type, while lymphatic endothelial cells were the least abundant cell type across all the weeks (***Figure 2C***). Notably, our spatial protein dataset revealed higher proportions of epithelial, endothelial and lymphatic endothelial cells when compared with the single-cell transcriptomics data. These differences can be attributed to tissue dissociation bias inherent in the single-cell transcriptomics approach, leading to an underrepresentation of these cell types. Consequently, this underscores the benefit of spatial omics methodologies in offering a less biased representation of the cell type composition within the tissue.

Cell type adjacency patterns in early human lung development

Next, we aimed to characterize the cell type adjacency patterns in the developing human lung as the communication between neighboring cells orchestrates the development of the organ. Analysis

of the enrichment of cellular adjacency patterns revealed that in all weeks, cells were predominantly, and in several cases exclusively, neighboring with cells of their own type (***Figure 3A***). SOX9^{high} epithelial cells, SOX2^{high} epithelial cells, airway smooth muscle cells, vascular smooth muscle cells (from week 8.5 onwards), and chondroblasts (from week 11 onwards) were the cell types with most homotypic adjacency patterns (***Figure 3B***). In contrast, mesenchymal cells, in particular Ki67+ proliferating mesenchymal cells, neighbored more frequently with cells of other types rather than their own, indicating the dispersed pattern of the mesenchyme throughout the tissue space and the absence of centralized niches of mesenchymal proliferation. In general, endothelial-epithelial cells, and mesenchymal-epithelial cells were the cell type pairs sharing adjacency to the least degree.

While the most frequently adjacent cell type pairs varied across different weeks, underscoring the dynamic character of the lung tissue development during the first trimester of gestation, airway smooth muscle-epithelial cells were the cell type pair with enriched adjacency from week 8.5 onwards (***Figure 3B***). Here, airway smooth muscle cells were neighboring more frequently with SOX2^{high} proximal epithelial cells than with SOX9^{high} distal epithelial cells, in particular at week 8.5, indicating their potential role in driving branching morphogenesis[10]. In summary, the analysis of the adjacency patterns within and between the cell types present in the developing human lung provided an in-depth, protein-level overview of the complex architecture of the developing organ.

Evolution of spatial domains during early human lung development

Tissue function depends on multi-cellular units, that we here refer to as “domains”, which consist of higher-order interactions between one or more cell types. In addition to the analysis of enrichment of adjacency patterns of the cell types we annotated, we also performed an unsupervised analysis of the tissue topography by inferring the presence of distinct spatial domains. These domains, which might contain one or more of the annotated cell types, were identified based solely on the spatially aggregated expression values in the tissue, i.e. without use of the cell type annotations[20]. This analysis revealed on average nine different spatial domains in developing human lungs, ranging between six to eleven at different developmental weeks (***Figure 4A, Supplementary Figure 9***).

Models of physical proximity between these identified spatial domains revealed that the domains representative of the mesenchyme and endothelium were the ones with closest proximity to each other across all weeks (***Figure 4B***). From week 8.5 onwards, the networks between the different domains represented the high-order tissue assembly from the airway epithelium to the smooth muscle surrounding the airway epithelium, and to the connective mesenchyme and endothelium. This suggests that the foundations of a mature lung tissue topography are initiated during early fetal development[24]. Week 6 contained an additional mesothelium domain along this main anatomical axis of tissue assembly (***Figure 4B***). The proximity between SOX2^{high} and SOX9^{high} epithelium domains became gradually more frequent in particular in week 12 and 13, indicating the presence of airways containing both proximal and distal ends in these weeks. In all weeks, smooth muscle domains were in closer proximity more with SOX2^{high} proximal epithelium domains than with SOX9^{high} distal epithelium domains, illustrating again the microanatomy of developing airways where smooth muscle cells surround only SOX2^{high} proximal epithelium[10].

The least frequent proximity patterns were generally observed between domains representative of vascular smooth muscle cells and chondroblasts, or between immune and epithelium domains, and between endothelium and SOX2^{high} epithelium domains. In later weeks, the domains having the least frequent proximity with other domains were chondroblasts and lymphatic endothelial cells (***Figure 4B***). This may suggest a significance of the topological segregation of these domains for the formation of the organ.

When the distribution patterns of the identified domains were investigated, the spatial domain representative of mesenchyme exhibited the most congregated spatial pattern compared to the other domains (***Figure 4C***). From week 8.5 onwards, the domain representing endothelium was the second most congregated domain. In week 11, in addition to the mesenchyme and endothelium domains, the SOX2^{high} epithelium domain also exhibited a congregated pattern (***Figure 4C***). The spatially most dispersed domain was representing immune cells (***Figure 4C***). In addition to the immune domain, from week 8.5 onwards, lymphatic endothelium domain; from week 11 onwards, neuronal domain; at week 12, Ki67+ mesenchyme domain; and at week 13, chondroblast domain exhibited relatively more dispersed spatial patterns (***Figure 4C***). In summary, our data suggests that distinct spatial domains start emerging in the developing human lung as early as in week 6 and they start to form the blueprint for a mature lung tissue organization.

Proliferation patterns during early human lung development

We next explored the developing lung's outgrowth patterns by studying the degree of proliferation in different cell types at different stages. To achieve this, we analyzed Ki67 expression of the individual cell types as a proxy for their proliferation[25] (***Figure 5A***). We observed a gradual decrease in proliferation over time, with 36% of all cells being proliferative in week 6, as compared to 14% at week 13 (***Figure 5B***). Although proliferating cells were observed among all cell types, the majority of proliferating cells consisted of mesenchymal and endothelial cells (***Figure 5C***).

In earlier weeks, SOX9^{high} epithelial cells contained the highest fraction of proliferating cells (60-70%) compared to other cell types, followed by SOX2^{high} epithelial cells. In later weeks, endothelial cells were the other most proliferating cell type whereas airway smooth muscle cells and lymphatic endothelial cells contained the lowest fraction of proliferating cells (***Figure 5B***). In earlier weeks, there was a wide difference in the degree of proliferation across the different cell types, *e.g.* 73% in SOX9^{high} epithelial cells versus 15% in airway smooth muscle cells at week 6. However, in later weeks the difference in the degree of proliferation was smaller across the different cell types, *e.g.* 17% in endothelial cells versus 7% in airway smooth muscle cells at week 13 (***Figure 5B***). This suggests that the lung development is mainly focused on forming and expanding airways in the early weeks, and the endothelial and mesenchymal expansion, including smooth muscle, temporally follows the epithelial outgrowth.

Comparison of the enrichment of cellular adjacency among different proliferating cell types and their non-proliferating counterparts revealed that non-proliferating cells had generally more homotypic adjacency patterns compared to proliferating cells (***Figure 5D***). The most homotypic adjacency pattern was identified among proliferating SOX9^{high} epithelial cells, which could indicate formation of a proliferation niche at the tip of the outgrowing epithelium, whereas the least homotypic adjacency patterns were identified among proliferating immune cells, proliferating lymphatic endothelial cells and proliferating vascular smooth muscle cells indicating that these proliferating cells neighbored more frequently with cells of other types rather than their own (***Figure 5D***).

Although proliferating cells were predominantly neighboring with each other and with their non-proliferating counterparts, we observed enriched adjacency patterns between proliferating cells of different cell types (*e.g.* proliferating endothelial and mesenchymal cells), and between proliferating cells and non-proliferating cells of different cell types (*e.g.* proliferating vascular smooth muscle cells and endothelial cells). Throughout all weeks, the least frequent adjacency patterns were observed between mesenchymal and proliferating epithelial cells (***Figure 5D***).

We investigated further the proliferation patterns in airway structures in later weeks by comparing the degree of proliferation in SOX2^{high} proximal airway regions and SOX9^{high} distal tips (***Figure 5E***). Notably, the SOX9^{high} distal tips contained a significantly higher fraction of proliferating cells compared to SOX2^{high} proximal airway regions, which are found at the epithelial stalks (***Figure 5F***). Furthermore, from week 12 onwards, the relatively smaller proximal airway regions contained a significantly higher fraction of proliferating cells compared to larger ones (***Supplementary Figure 10C***). These results show that while the bulk of proliferation occurs in the distal tips, the cells in the stalks are more quiescent. This indicates a positional hierarchy of proliferation along the proximal-distal axis of the developing epithelial network. In addition, the proliferative cells in the large proximal airways presumably contributes to their diameter expansion. Collectively, these differential proliferation patterns in the epithelium provide valuable insights into the mechanisms of airway branching and elongation during early human lung development.

Immune landscape in early human lung development

We next analyzed the immune compartment in depth by subclustering the immune cells (***Figure 6A, Supplementary Figure 11***). We observed that macrophages were the most abundant immune cell type during early lung development, supporting previously reported transcriptomics-based findings[6] (***Figure 6B***). Additionally, our data revealed that macrophages became positive for the endocytic receptor MRC1 (CD206) between week 6 and 8.5, possibly pinpointing the period during which the phenotypic heterogeneity of macrophages starts to increase (***Figure 6A***). B and NK cells were the least abundant immune cell types during early lung development. ILC & T cells emerged from 8.5 weeks onwards, and they were the second most abundant immune cell type with the exception of week 12, where dendritic cells were the second most abundant cell type (***Figure***

6B). Dendritic cells, NK cells and B cells emerged from 11 weeks onwards (***Figure 6B***) emphasizing the distinct timelines of immune cell formation during development.

The immune cell type containing the highest fraction of proliferating cells was B cells (37-62% between weeks 11-13) and while the fraction of proliferating B cells increased, the fraction of proliferating macrophages and ILC & T cells decreased gradually (***Figure 6D***). As shown in ***Figure 6E***, macrophages were more frequently in close proximity with ILC & T cells and dendritic cells, and less frequently with B cells and NK cells. These spatial proximity patterns especially between macrophages and ILC & T cells might have important implications for tissue homeostasis[26].

Starting at week 11, we observed groups of CD45+ immune cells encircling the arterial vessels in the developing human lung (***Figure 7A***). We manually selected these "arterial immune cells" in the images for weeks 11 to 13 (***Supplementary Figure 12A***) and a comparison of their neighborhood composition within a radius of 50 µm distance to the neighborhood composition of randomly selected cells of any type confirmed their distinct spatial arrangement around the vasculature (***Supplementary Figure 12B***). Comparison of their neighborhood composition to the one of randomly selected immune cells in the rest of the tissue revealed that while the closest neighbors for both arterial and randomly positioned immune cells consisted predominantly of other immune cells, mesenchymal cells, and endothelial cells, the closest neighbors of arterial immune cells were other arterial immune cells, vascular smooth muscle cells, and pericytes (***Figure 7B***). These arterial immune cells were mostly macrophages and ILC & T cells (***Figure 7D***). Differential marker analysis between the arterial immune cells and randomly selected immune cells revealed that arterial immune cells expressed relatively higher levels of CD45, CD44, CD90, and HLA-DR and relatively lower levels of CD163 and MRC1 (***Figure 7C***). The special arrangement of these immune cells around the arteries suggests a possible role in vascular growth and remodeling.

DISCUSSION

Human lung development during the first trimester is a dynamic process marked by intricate morphological changes and interactions among different cell types. Our study delved into this complexity by analyzing around 1 million single cells *in situ* in human lungs aged 6 to 13 post-conception weeks using an in-house developed 30-plex antibody panel. We characterized the composition and spatial organization of cell types within these developing human lungs and tracked the temporal evolution and establishment of stable spatial domains representing functional tissue units. Additionally, we explored proliferation patterns across different cell types and microanatomic regions of the developing lung tissue.

Lung development hinges on key processes including cell proliferation, differentiation, migration and proximal-distal patterning. Our multiplexed antibody panel was meticulously curated to target markers involved in as many of these processes as possible. Rigorous testing and validation of over 70 antibodies resulted in an optimized 30-plex assay (***Figure 1B***, ***Supplementary Table 1)***. Thus, our study not only presents a comprehensive proteomic view of early lung development but also serves as a resource documenting detailed performance of over 70 antibodies on fresh-frozen human developmental lung specimens.

The 30-plex antibody panel enabled us to localize the major types of cells present in the developing human lung. Comparison of cell type abundances derived from this protein-based spatial analysis to our previous single-cell transcriptomics analysis[7] revealed similar trends with notable differences in specific cell types such as endothelial cells, especially in week 12. These discrepancies may originate from a variety of factors, including tissue dissociation protocols causing a cell composition bias in single-cell transcriptomics data[27] or segmentation inaccuracies in our multiplexed protein imaging data. Crucially, these comparative analyses emphasize the distinct advantage of spatial omics techniques in capturing a relatively more faithful representation of cell type distribution within tissues.

One of the unique advantages of single-cell resolved high-parametric immunofluorescence is the possibility to explore the microenvironmental cell type configuration within the bona fide tissue space. This is particularly valuable in organs under development, as it enables a thorough

understanding of the spatiotemporal dynamics governing organogenesis. Here, we observed that SOX2^{high} proximal epithelial cells, SOX9^{high} distal epithelial cells, and airway smooth muscle cells exhibited the most exclusive physical interactions among themselves compared to other cell types. This robust interaction pattern suggests that these cells have the most densely interconnected compartment compared to other cell types present within the mesenchymal tissue space, emphasizing the presence of distinct anatomical structures, and the presence of homotypic interactions in these distinct structures. Overall, mesenchyme and endothelium domains were spatially the most congregated cellular domains, whereas immune, proliferating mesenchyme and lymphatic endothelium domains were spatially the most dispersed ones, interacting more with cells of other cell types. Our analyses revealed a dramatic change in the composition and complexity of the spatial network of cells covering the tissue from week 6 to week 8.5, and showed also how this complex architecture is preserved from week 8.5 onwards.

The developing lung undergoes dramatic tissue growth and remodeling, during which proliferation and apoptosis occurs[28]. Investigating the degree and change of proliferation patterns across different cell types revealed a gradual decrease in proliferation. SOX2^{high} and SOX9^{high} epithelial cells were the cell types with the highest fraction of proliferating cells in week 6 and their degree of proliferation decreased over time indicating an extensive branching in the airways in earlier stages of development, confirming a similar pattern in proliferation reported for slightly later stages of lung development[10]. In animal models, hyperproliferation of epithelial and mesenchymal cells during fetal development is reported to disrupt distal airway formation and mesenchymal and vascular remodeling[29]. Compromised proliferation is also indicated in the context of rare congenital anomalies such as pulmonary hypoplasia[30]. Thus, our protein level data reporting the degree of proliferation provides a valuable resource to benchmark proliferation patterns in the developing human lung.

Our data revealed that SOX9^{high} distal tips contain a higher fraction of proliferating cells as compared to SOX2^{high} proximal regions, confirming that the lung epithelial network expands mainly through the outgrowth of the distal SOX9^{high} epithelial cells. Our data also showed that the degree of proliferation in smaller SOX2^{high} proximal airways is higher than in larger proximal airways, suggesting that proliferation at these stalk regions locally contributes to the branch elongation. Collectively, epithelial proliferation followed a graded pattern along the distal to

proximal axis of the organ. Following epithelial cells, endothelial cells were the second most proliferative cell type among others. This finding suggests a synchronized progression of airway branching with vascularization, supporting previous research[31].

Immune cells do not only play an important role in lung homeostasis[32] but they also contribute to organ development. Here we show how immune cells get diversified over a span of only seven weeks and how they gradually build a spatial network to cover the entire tissue. We identified macrophages as the most abundant immune cell type in early developmental human lung. Interestingly, our study pinpointed, to our knowledge for the first time, that human lung macrophages become MRC1^{+} (CD206^{+}) between post conceptional week 6 and 8.5 and complements previous studies reporting that human neonates are born with MRC1^{+} lung macrophages[33,34]. Compared to macrophages and ILC & T cells, NK and B cells were less abundant during the developmental period of 6 to 13 weeks. Previous transcriptomics-based studies reported that these cells indeed become prominent first after week 15[6]. Our data also revealed that while the fraction of proliferating macrophages decreased during the investigated period, the fraction of proliferating B cells increased and B cells contained the highest fraction of proliferating cells. This aligns with previous reports stating that although B cells are rare, they express proliferation markers[11,35].

Our study also revealed the presence of immune cells, predominantly macrophages, in a ring-shaped arrangement encircling the developing vessels. This was previously described in prenatal murine lung and was observed to disappear right after birth[35]. This transient accumulation of lung macrophages around the developing vessels, temporally restricted only to prenatal development, suggests their function in modulating vascular growth and remodeling. Our data imply for the first time for humans the same possible function for macrophages in the developing human lung. In addition to macrophages, we also identified the presence of CD3 expressing cells in these regions, indicating a possibly broader spectrum of immune cell types than previously recognized. Previous studies, not on prenatal but adult murine lungs[36], reported the presence of group 2 innate lymphoid cells (ILC2s) encircling arteries. While our panel did not allow us to distinguish ILC2 cells from other ILC and T cells, our observations align with the literature[35,36], suggesting a potential involvement of also these cells in modulating vascular growth and remodeling in developing human lung.

In summary, our work provides a protein-level characterization of the spatial diversity of the developing human lung and serves as a novel hypothesis-generating resource that the pulmonary and developmental biology community can explore to further understand the human lung formation and function. This dataset, now openly accessible for in-depth exploration in an interactive portal (https://hdca-sweden.scilifelab.se/tissues-overview/lung/), provides a new avenue to unravel the developmental origins of health and disease in the human lung.

ACKNOWLEDGEMENTS

We acknowledge the service of the Karolinska Institutet Developmental Tissue Bank for providing the prenatal human tissues, and the Spatial Proteomics Unit at SciLifeLab Stockholm for technical support and providing access to instruments. This work was supported mainly by grants to EL from the Knut and Alice Wallenberg Foundation (KAW 2018.0172) and the Erling Persson Foundation for the Human Developmental Cell Atlas (HDCA) project. BA acknowledges Starting Grant from the Swedish Research Council-Vetenskapsrådet (2022-04732). JNH acknowledges a post-doctoral fellowship from the Wenner-Gren Foundations and an EMBO Postdoctoral Fellowship (ALTF 556-2022). Part of the computations and data storage was enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX partially funded by the Swedish Research Council (2018-05973). This publication is part of the Human Cell Atlas – www.humancellatlas.org/publications/.

COMPETING INTERESTS

JL is a scientific consultant for 10X Genomics. SMS is a cofounder of Spatialist AB.

DATA AVAILABILITY

Data is made openly accessible at HDCA Sweden portal:
https://hdca-sweden.scilifelab.se/tissues-overview/lung/

AUTHOR CONTRIBUTIONS

SS, BA, AMC, EL developed the methodology for the study; ES, XL, ZA coordinated and/or performed the isolation and processing of human embryonic specimen; SS carried out the immunofluorescent experimental work; FB, MM, JNH developed tools for analysis of imaging data; SS, BA analyzed the data; SMS provided support for data analysis; SS, BA, AS, CS, EL interpreted the data; SS, BA created the figures; SS, BA wrote the original manuscript; AS, JNH, EL revised the manuscript; BA, EL supervised the project; EL, ES, JL, MN, CS, SL acquired funding. All authors read and commented on the manuscript.

TABLE AND FIGURES

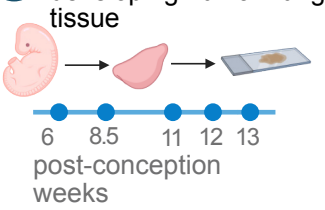
Table 1. Content of the 30-plex antibody panel used for protein profiling in developing human lung tissue. The information source for PDPN is data from Nikolić *et al.*, 2017[23]; for CD123 Jambusaria *et al.*, 2020[37]; for SOX9 Ma *et al.*, 2018[38], and for TTF1 Yatabe *et al.*, 2002[39]. The source for all other targets is Schupp *et al.*, 2020[22].

	Antibody-Barcode	Corresponding gene(s)	In-house conjugation	Targeted main cell type/state
1	CD3-BX015	CD3D, CD3E, CD3G	No	Immune (T-cell, Innate lymphoid cell)
2	CD4-BX021	CD4	No	Immune (T-cell)
3	CD19-BX003	CD19	No	Immune (B-cell)
4	CD31-BX032	PECAM1	No	Endothelial
5	CD34-BX035	CD34	No	Endothelial
6	CD45-BX001	PTPRC	No	Immune
7	CD90-BX022	THY1	No	Mesenchymal (Fibroblast)
8	Podoplanin-BX023	PDPN	No	Lymphatic endothelial & Epithelial
9	HLA-DR-BX026	HLA-DRA, HLA-DRB1, ...	No	Immune (Antigen presenting cell)
10	Ki67-BX047	MKI67	No	Proliferation
11	Pan-Cytokeratin-BX019	KRT8, KRT18, ...	No	Epithelial
12	EPCAM-BX042	EPCAM	Yes	Epithelial
13	CD123-BX054	IL3RA	Yes	Immune & Endothelial
14	SOX2-BX024	SOX2-OT	Yes	Epithelial
15	CD144-BX016	CDH5	Yes	Endothelial
16	ACTA2-BX028	ACTA2	Yes	Mesenchymal (Smooth muscle)
17	CD68-BX010	CD68	Yes	Immune (Macrophage)
18	CD44-BX020	CD44	Yes	Immune
19	DCN-BX013	DCN	Yes	Mesenchymal
20	E-Cadherin-BX014	CDH1	Yes	Epithelial
21	WT1-BX006	WT1	Yes	Mesenchymal
22	CD163- BX005	CD163	Yes	Immune (Macrophage)
23	COL1A1-BX002	COL1A1	Yes	Mesenchymal
24	CD56-RX029	NCAM1	Yes	Immune (NK cell)
25	MRC1-BX030	MRC1	Yes	Immune (Macrophage)
26	Vimentin-BX045	VIM	Yes	Mesenchymal
27	PRX-BX052	PRX	Yes	Endothelial
28	CLDN5-BX041	CLDN5	Yes	Endothelial
29	SOX9-BX033	SOX9	Yes	Epithelial
30	TTF1-BX007	NKX2.1	Yes	Epithelial

Figure 1

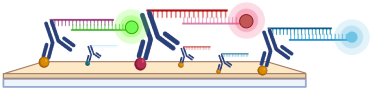
A

1. Collection & sectioning of developing human lung tissue



2. Building & validation of antibody panel

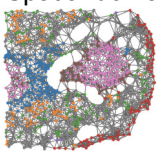
3. Multiplexed IF with antibody panel



4. Segmentation of imaging data

Sample age (weeks)	Segmented nr. of cells
6	33,277
8.5	49,396
11	867,796
12	397,191
13	994,603

5. Cell type composition
Cellular adjacency patterns
Spatial domains



6. Accessible interactive browser

B

Literature

In-house scRNA-seq data

Selection of 72 antibodies

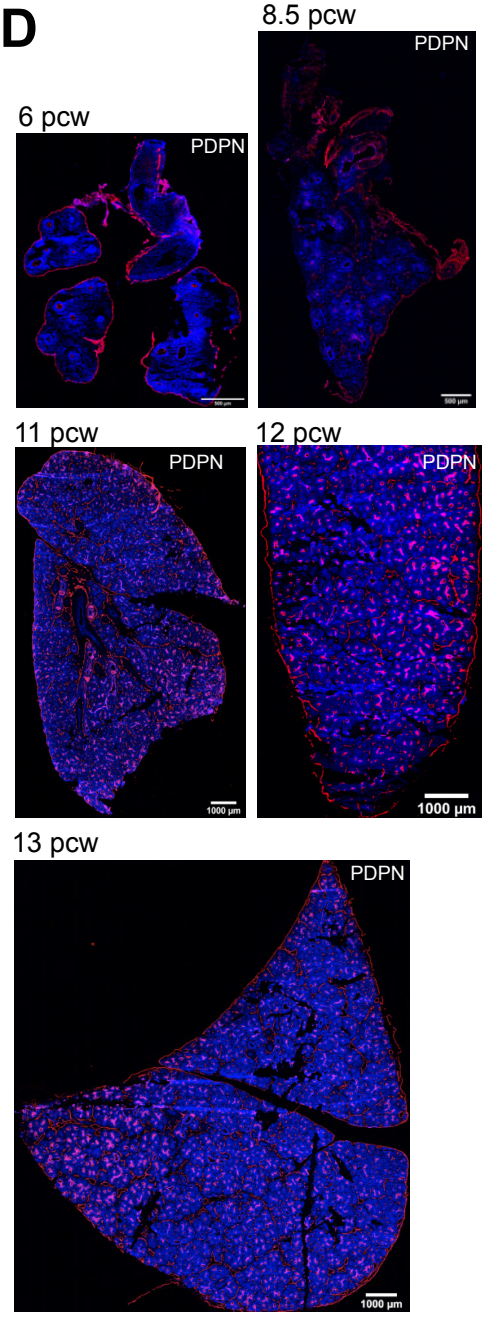
1) Regular IF
2) Conjugation validation
3) Multiplexed IF

Final panel of 30 antibodies

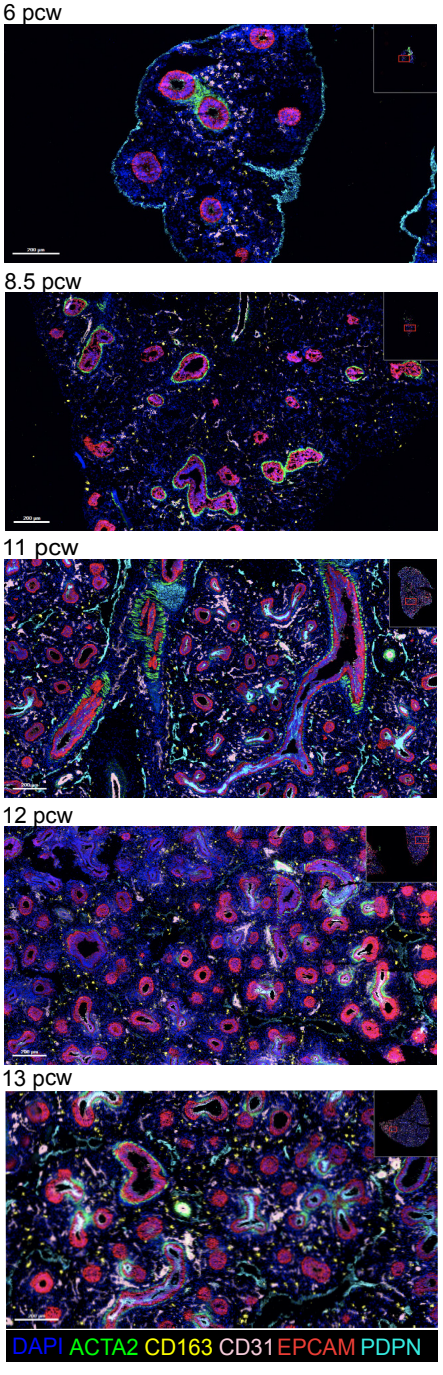
C

Epithelial	Endothelial	Mesenchymal	Immune		Proliferation
EPCAM	CD31	CD90	CD3	CD56	Ki67
Pan-CK	CD34	ACTA2	CD4	CD44	
SOX2	PDPN	DCN	CD19	CD68	
SOX9	CD144	WT1	CD45	CD163	
E-Cadherin	PRX	COL1A1	HLA-DR	MRC1	
TTF1	CLDN5	VIM	CD123		

D



E



F

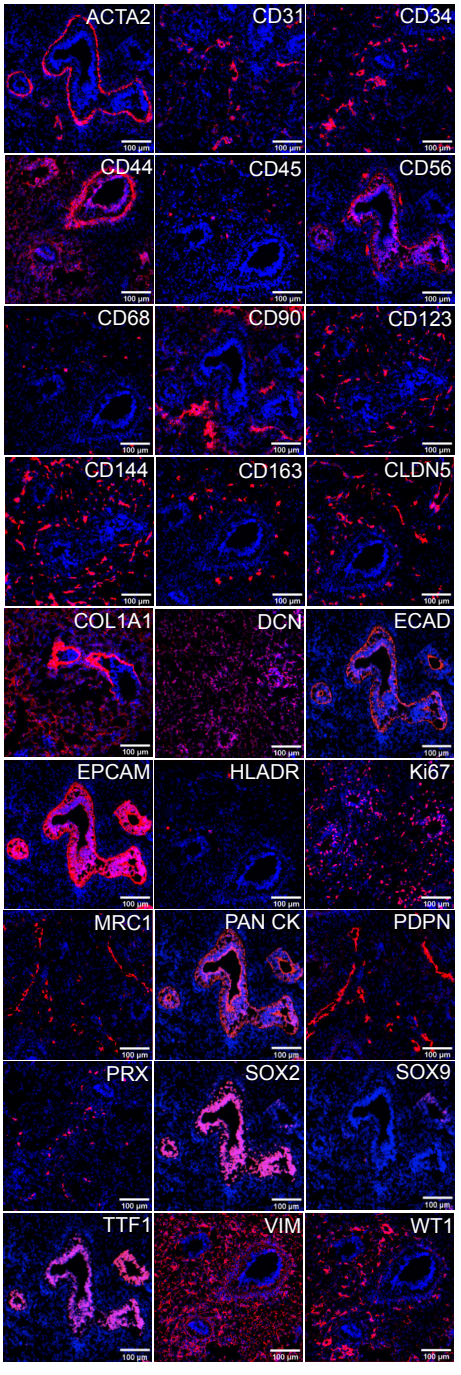


Figure 1. Overview of experimental approach for spatial proteomic analysis of the developing human lung.

A) Samples from 6- to 13-week-old human lungs were sectioned. Each selected antibody was tested for immunofluorescence (IF) staining, and then conjugated with DNA barcodes and validated. High-parametric imaging was performed, and imaging data was segmented for downstream analysis to characterize the cell type composition, cell type adjacency patterns, spatial domains and their temporal changes in the developing human lung. (Figure was created with BioRender.com.)

B) Initially, a total of 73 antibodies were selected based on literature and in-house single-cell transcriptomics data[7]. After various validation steps, 30 of these antibodies were included in the final panel used for multiplexed imaging.

C) The content of the 30-plex antibody panel used for multiplexed imaging.

D) Whole tissue images of lung samples at week 6 to 13 showing PDPN in red, DAPI in blue. From week 8.5 onwards, PDPN expression is observed in lymphatic endothelium, as well as in epithelium. Scale bars for weeks 6 and 8.5 correspond to 500 μm, and for weeks 11 to 13 to 1000 μm.

E) Images for five antibodies overlaid to show main structures and cell types: ACTA (smooth muscle marker) in green; CD163 (macrophage marker) in yellow; CD31 (endothelial marker) in pink; EPCAM (epithelial marker) in red; PDPN (lymphatic endothelial marker) in cyan, and DAPI (blue). Scale bars correspond to 200 μm.

F) Staining patterns of all panel markers in selected regions of the 8.5-week-old lung (except CD3, CD4 and CD19). DAPI in blue. Scale bars correspond to 100 μm.

Figure 2

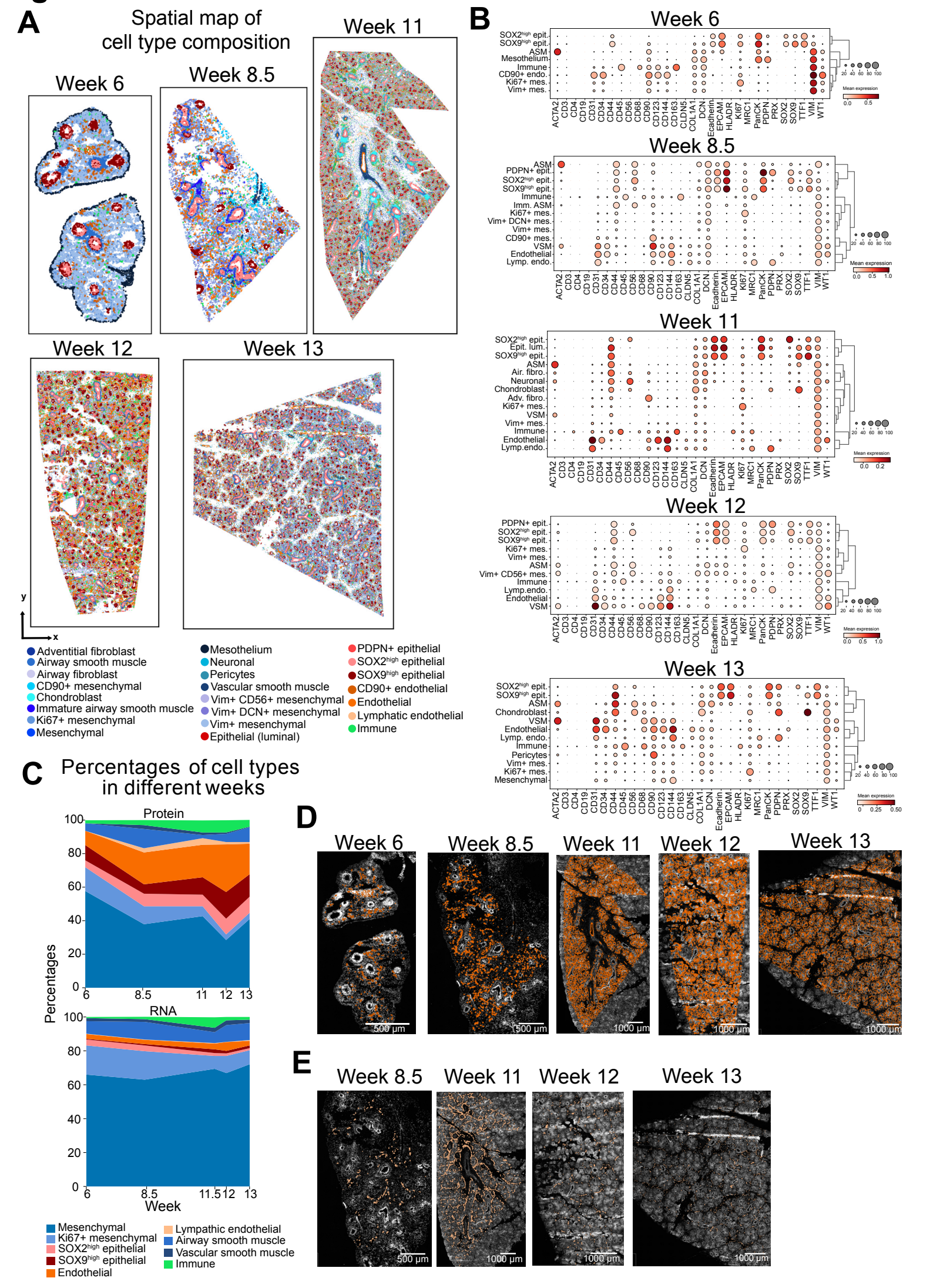

A
Spatial map of cell type composition
Week 6
Week 8.5
Week 11
Week 12
Week 13
y
x
Adventitial fibroblast
Airway smooth muscle
Airway fibroblast
CD90+ mesenchymal
Chondroblast
Immature airway smooth muscle
Ki67+ mesenchymal
Mesenchymal
Mesothelium
Neuronal
Pericytes
Vascular smooth muscle
Vim+ CD56+ mesenchymal
Vim+ DCN+ mesenchymal
Vim+ mesenchymal
Epithelial (luminal)
PDPN+ epithelial
SOX2high epithelial
SOX9high epithelial
CD90+ endothelial
Endothelial
Lymphatic endothelial
Immune
B
Week 6
Week 8.5
Week 11
Week 12
Week 13
Mean expression
C
Percentages of cell types in different weeks
Protein
RNA
Percentages
Week
Mesenchymal
Ki67+ mesenchymal
SOX2high epithelial
SOX9high epithelial
Endothelial
Lympathic endothelial
Airway smooth muscle
Vascular smooth muscle
Immune
D
Week 6
Week 8.5
Week 11
Week 12
Week 13
500 µm
1000 µm
E
Week 8.5
Week 11
Week 12
Week 13
500 µm
1000 µm

Figure 2. Overview of the identified main cell types in developing human lung.

A) Single cells colored by cell type identified by clustering are visualized in xy-scatterplots

B) Dot plot shows the average marker expression in each cell type cluster. The size of the circles indicates the percentage of cells expressing the marker in each cluster. Dark red color indicates high and white color indicates low expression.

C) Percentage of main cell types identified in the present multiplexed protein imaging dataset and single-cell transcriptomics dataset[7].

D) Endothelial cluster in each week plotted on DAPI images. Scale bars corresponds to 500 μm for week 6 and 8.5, and 1000 μm for week 11,12 and 13.

E) Lymphatic endothelial cluster in week 8.5, 11, 12 and 13 plotted on DAPI images. Scale bars corresponds to 500 μm for week 6 and 8.5, and 1000 μm for week 11,12 and 13.

Figure 3

A

Cell type adjacency patterns

Week 6

Week 8.5

Week 11

Week 12

Week 13

Adventitial fibroblast
Airway smooth muscle
Airway fibroblast
CD90+ mesenchymal
Chondroblast
Immature airway smooth muscle
Ki67+ mesenchymal
Mesenchymal
Mesothelium
Neuronal
Pericytes
Vascular smooth muscle
Vim+ CD56+ mesenchymal
Vim+ DCN+ mesenchymal
Vim+ mesenchymal
Epithelial (luminal)
PDPN+ epithelial
SOX2^{high} epithelial
SOX9^{high} epithelial
CD90+ endothelial
Endothelial
Lympathic endothelial
Immune

B

Week 6

Z-scored enrichment: 100, 0, -100

SOX2^{high} epit.
SOX9^{high} epit.
ASM
Mesothelium
Vim+ mes.
CD90+ endo.
Immune
Ki67+ mes.

Week 8.5

Z-scored enrichment: 200, 0, -200

SOX9^{high} epit.
CD90+ mes.
VSM
Lym. endo.
Endothelial
Immune
Vim+ mes.
Ki67+ mes.
Vim+ DCN+ mes.
ASM
Imm. ASM
PDPN+ epit.
SOX2^{high} epit.

Week 11

Z-scored enrichment: 1000, 500, 0, -500, -1000

SOX2^{high} epit.
Chondroblast
Epit. lum.
SOX9^{high} epit.
VSM
ASM
Air. fibro.
Neuronal
Lymp. endo.
Adv. fibro.
Immune
Endothelial
Ki67+ mes.
Vim+ mes.

Week 12

Z-scored enrichment: 500, 0, -500

SOX9^{high} epit.
ASM
SOX2^{high} epit.
PDPN+ epit.
Endothelial
VSM
Vim+ CD56+ mes.
Lym. endothelial
Immune
Ki67+ mes.
Vim+ mes.

Week 13

Z-scored enrichment: 1000, 500, 0, -500, -1000

Chondroblast
ASM
VSM
Lymp. endo.
Pericytes
Endothelial
Vim+ mes.
Immune
Ki67+ mes.
Mesenchymal
SOX2^{high} epit.
SOX9^{high} epit.

Figure 3. Cell type adjacency patterns in the developing human lung.

A) Cell-cell adjacency graphs representing the entire tissue and a zoomed-in region, where each cell is represented by a spot in the color of its annotated cell type and physically neighboring cells are connected via edges.

B) Heatmaps with dendrograms summarizing enrichment of cellular adjacency based on the Z-scores. Dark red color indicates cellular adjacency patterns with high enrichment Z-score and dark blue color indicates cellular adjacency patterns with low enrichment Z-score.

Figure 4

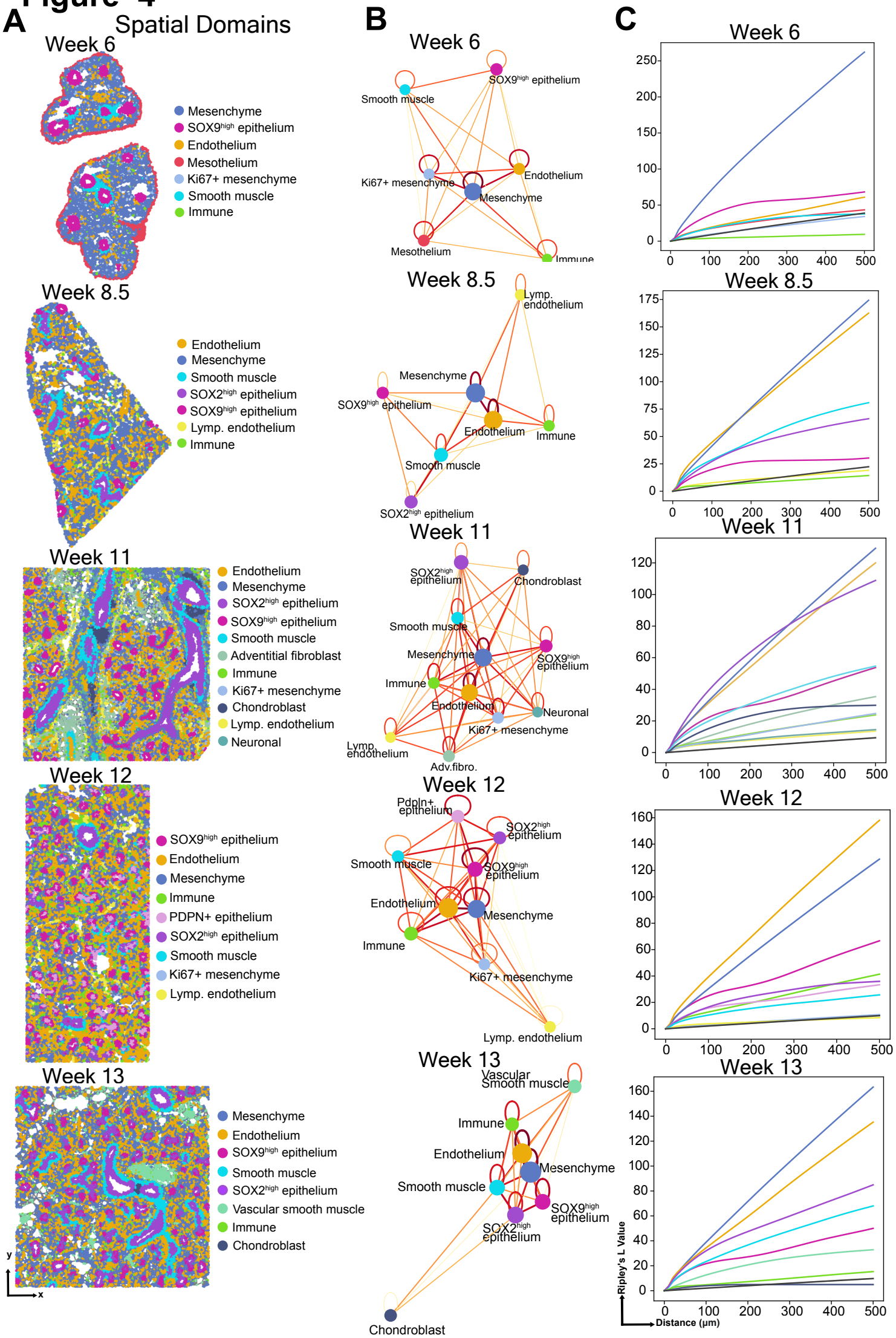

A
Spatial Domains
Week 6
Mesenchyme
SOX9high epithelium
Endothelium
Mesothelium
Ki67+ mesenchyme
Smooth muscle
Immune
Week 8.5
Endothelium
Mesenchyme
Smooth muscle
SOX2high epithelium
SOX9high epithelium
Lymp. endothelium
Immune
Week 11
Endothelium
Mesenchyme
SOX2high epithelium
SOX9high epithelium
Smooth muscle
Adventitial fibroblast
Immune
Ki67+ mesenchyme
Chondroblast
Lymp. endothelium
Neuronal
Week 12
SOX9high epithelium
Endothelium
Mesenchyme
Immune
PDPN+ epithelium
SOX2high epithelium
Smooth muscle
Ki67+ mesenchyme
Lymp. endothelium
Week 13
Mesenchyme
Endothelium
SOX9high epithelium
Smooth muscle
SOX2high epithelium
Vascular smooth muscle
Immune
Chondroblast
y
x
B
Week 6
SOX9high epithelium
Smooth muscle
Ki67+ mesenchyme
Endothelium
Mesenchyme
Mesothelium
Immune
Week 8.5
Lymp. endothelium
Mesenchyme
SOX9high epithelium
Endothelium
Immune
Smooth muscle
SOX2high epithelium
Week 11
SOX2high epithelium
Chondroblast
Smooth muscle
Mesenchyme
SOX9high epithelium
Immune
Endothelium
Neuronal
Ki67+ mesenchyme
Lymp. endothelium
Adv.fibro.
Week 12
Pdpln+ epithelium
SOX2high epithelium
Smooth muscle
SOX9high epithelium
Endothelium
Mesenchyme
Immune
Ki67+ mesenchyme
Lymp. endothelium
Week 13
Vascular Smooth muscle
Immune
Endothelium
Mesenchyme
Smooth muscle
SOX9high epithelium
SOX2high epithelium
Chondroblast
C
Week 6
Week 8.5
Week 11
Week 12
Week 13
Ripley's L Value
Distance (μm)

Figure 4. Unsupervised identification of topographical structures in human developing lung and network diagrams summarizing the proximity between the identified spatial domains.

A) In each week, seven to eleven different spatial domains were identified, annotated and plotted as xy-scatterplots for each week.

B) Network model summarizing the degree of adjacency between the different spatial domains identified at each week. The line color indicates the frequency of adjacency patterns between the domains, where more frequent adjacency between two different domains is shown in red and less frequent adjacency is shown in yellow.

C) Summary of Ripley's L statistics describing the spatial domain characteristics in each week, where a higher Ripley's L value indicates a more congregated and a lower value indicates a more dispersed spatial distribution of a domain across the tissue. In each plot, dark gray line represents the expected value corresponding to complete spatial randomness.

Figure 5

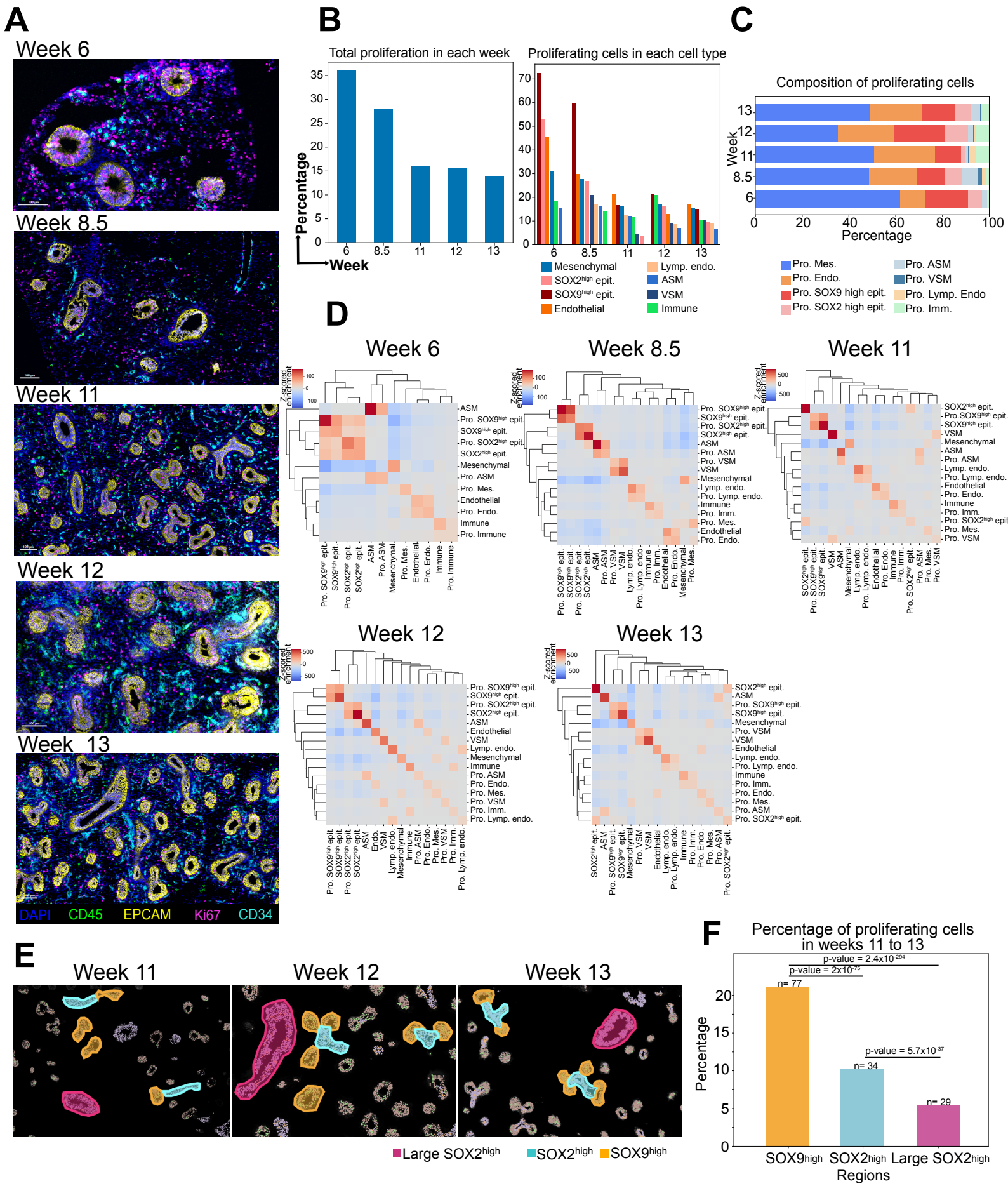

A
Week 6
Week 8.5
Week 11
Week 12
Week 13
DAPI CD45 EPCAM Ki67 CD34
B
Total proliferation in each week
Percentage
Week
Proliferating cells in each cell type
Mesenchymal
SOX2high epit.
SOX9high epit.
Endothelial
Lymp. endo.
ASM
VSM
Immune
C
Composition of proliferating cells
Week
Percentage
Pro. Mes.
Pro. Endo.
Pro. SOX9 high epit.
Pro. SOX2 high epit.
Pro. ASM
Pro. VSM
Pro. Lymp. Endo
Pro. Imm.
D
Week 6
Week 8.5
Week 11
Week 12
Week 13
Z-scored enrichment
E
Week 11
Week 12
Week 13
Large SOX2high
SOX2high
SOX9high
F
Percentage of proliferating cells in weeks 11 to 13
p-value = 2.4x10-294
p-value = 2x10-75
p-value = 5.7x10-37
n= 77
n= 34
n= 29
Percentage
SOX9high
SOX2high
Large SOX2high
Regions

Figure 5. Cell type specific proliferation patterns in the developing human lung.

A) Snapshots of Ki67+ proliferating cells in human lungs at different developmental weeks. DAPI in blue, CD34 (endothelial marker) in cyan, CD45 (immune marker) in green, EPCAM in yellow, and Ki67 (proliferation marker) in red. Scale bar corresponds to 100 μm.

B) The barplot on the left shows the fraction of proliferating cells among all cells analyzed in each week. The barplot on the right shows the fraction of proliferating cells, separately, for each annotated cell type and week.

C) The barplot shows the cell type composition of proliferating cells in each week.

D) Heatmaps with dendrograms summarizing enrichment of cellular adjacency patterns between and among proliferating and non-proliferating cell types based on the Z-scores. Dark red color indicates cellular adjacency patterns with high enrichment Z-score and dark blue color indicates cellular adjacency patterns with low enrichment Z-score.

E) Examples of $SOX9^{high}$ (orange), relatively small $SOX2^{high}$ (cyan) and relatively large $SOX2^{high}$ (magenta) epithelial airway structures in 11- to 13-week-old lungs, selected for analysis of proliferation patterns.

F) Bar plot shows the fraction of proliferating cells in these three tissue structures at weeks 11 to 13. The x-axis represents the three different structures and the y-axis represents the percentage of proliferating cells. The statistical significance of the differences in proliferation degree are summarized with Fisher's exact test p-values.

Figure 6

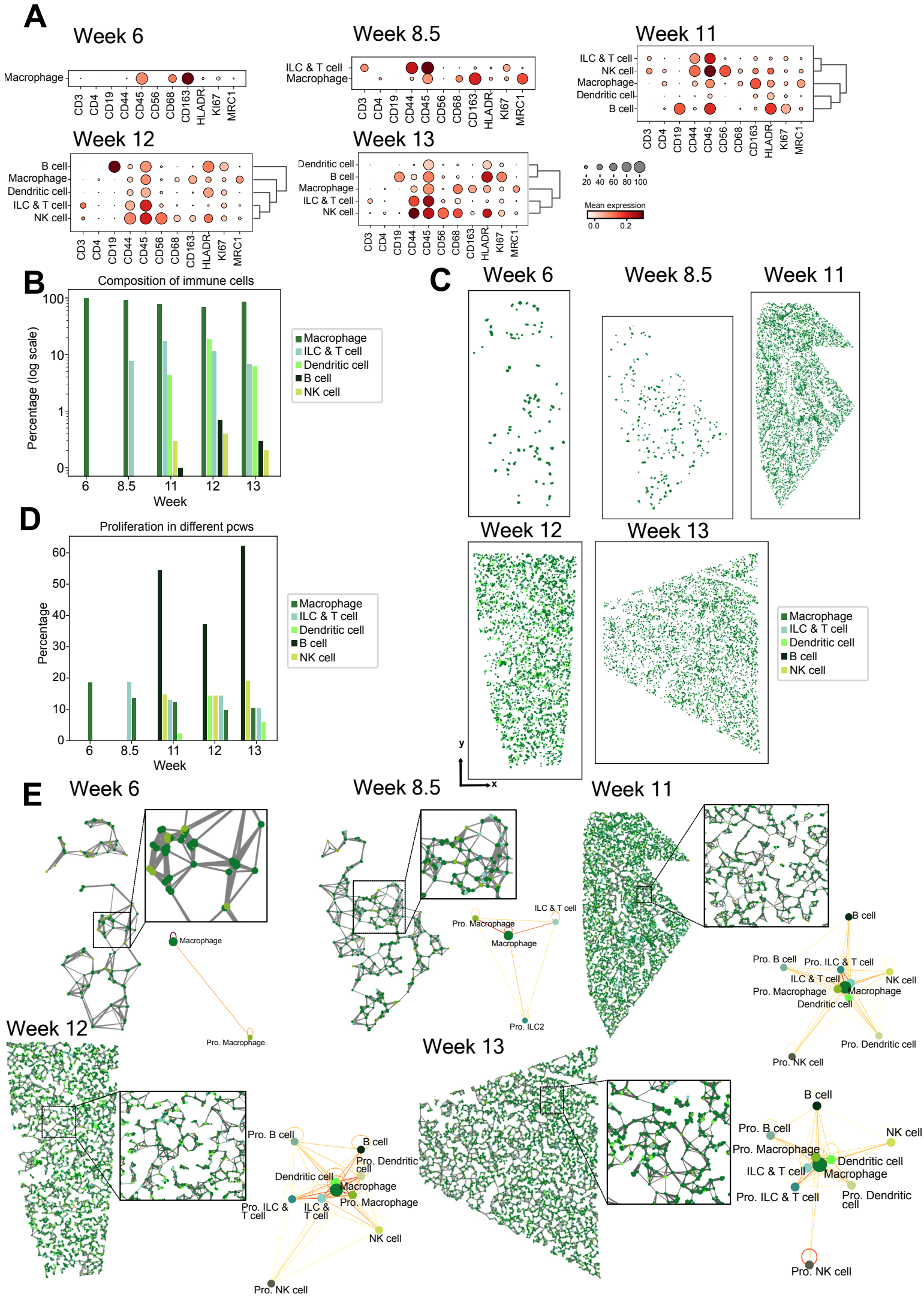

A
Week 6
Week 8.5
Week 11
Week 12
Week 13
Macrophage
ILC & T cell
Dendritic cell
B cell
NK cell
CD3 CD4 CD19 CD44 CD45 CD56 CD68 CD163 HLADR KI67 MRC1
20 40 60 80 100
Mean expression
0.0 0.2
B
Composition of immune cells
Percentage (log scale)
100
10
1
0
6 8.5 11 12 13
Week
Macrophage
ILC & T cell
Dendritic cell
B cell
NK cell
C
Week 6
Week 8.5
Week 11
Week 12
Week 13
y
x
D
Proliferation in different pcws
Percentage
60
50
40
30
20
10
0
6 8.5 11 12 13
Week
E
Week 6
Week 8.5
Week 11
Week 12
Week 13
Macrophage
Pro. Macrophage
ILC & T cell
Pro. ILC2
B cell
Pro. B cell
Pro. ILC & T cell
NK cell
Dendritic cell
Pro. Dendritic cell
Pro. NK cell
Pro. ILC & T cell
ILC & T cell

Figure 6. Immune cells in developing human lung.

A) Dot plots show marker expression in annotated immune cell sub-clusters at each week. The size of the circles indicates the percentage of cells expressing the marker in each cluster. Dark red color indicates high and white color indicates low expression.

B) Bar plot shows the composition of immune cell subtypes. The y-axis (in log scale) represents the percentage of each immune cell subtype and the x-axis represents the week.

C) Scatterplots summarize the distribution of each immune cell subtype based on cell coordinates.

D) Bar plot shows the percentage of proliferating cells in each immune cell subtype at each week.

E) For each week, the left panel shows cellular adjacency among immune cell subtypes on the tissue and the right panel shows the network model summarizing the degree of adjacency between the immune cell subtypes. The line color indicates the frequency of adjacency patterns between the domains, where more frequent adjacency between two different domains is shown in red and less frequent adjacency is shown in yellow.

Figure 7

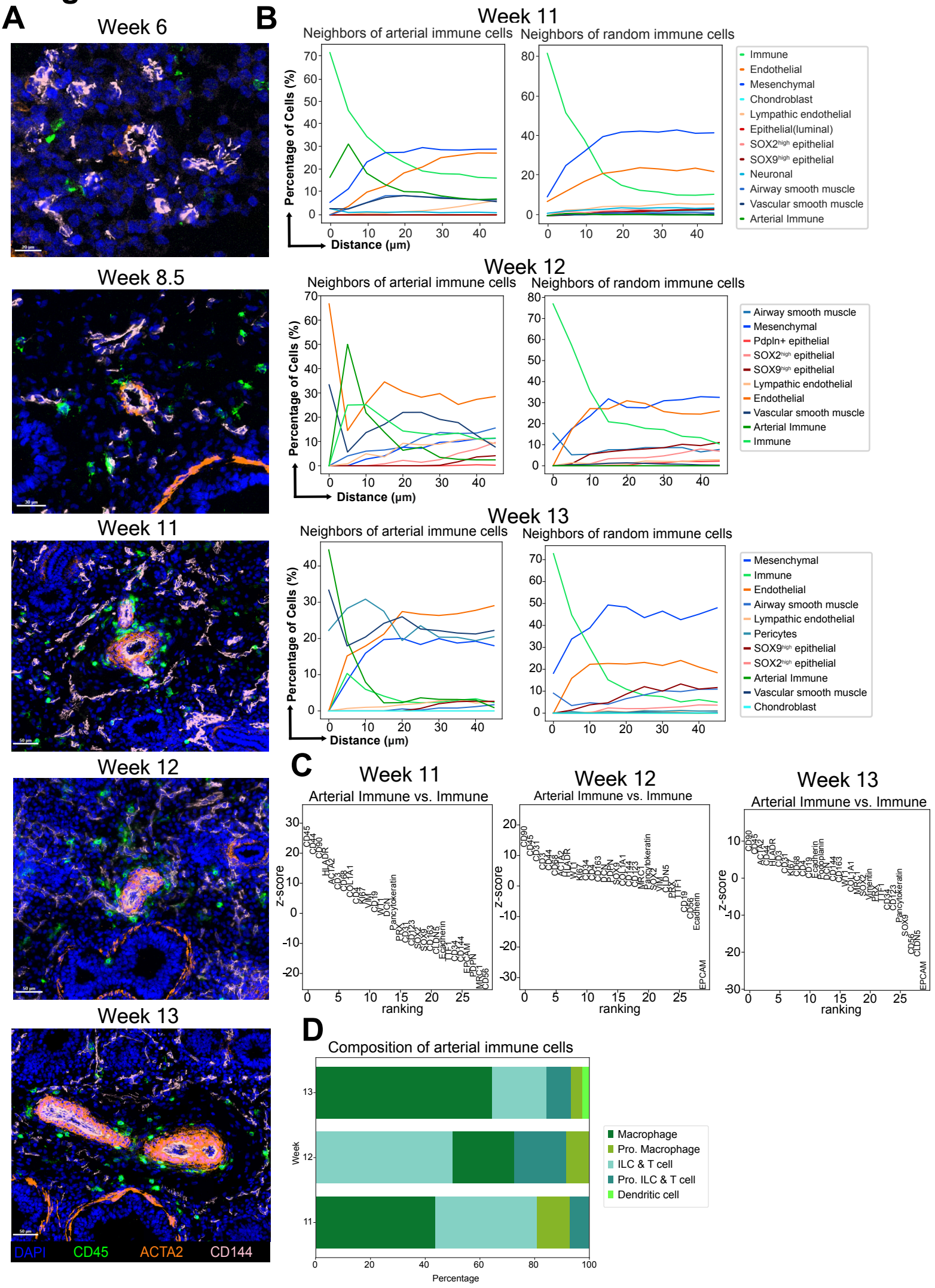

A
Week 6
Week 8.5
Week 11
Week 12
Week 13
DAPI
CD45
ACTA2
CD144
B
Week 11
Neighbors of arterial immune cells
Neighbors of random immune cells
Percentage of Cells (%)
Distance (μm)
Immune
Endothelial
Mesenchymal
Chondroblast
Lympathic endothelial
Epithelial(luminal)
SOX2high epithelial
SOX9high epithelial
Neuronal
Airway smooth muscle
Vascular smooth muscle
Arterial Immune
Week 12
Neighbors of arterial immune cells
Neighbors of random immune cells
Percentage of Cells (%)
Distance (μm)
Airway smooth muscle
Mesenchymal
Pdpln+ epithelial
SOX2high epithelial
SOX9high epithelial
Lympathic endothelial
Endothelial
Vascular smooth muscle
Arterial Immune
Immune
Week 13
Neighbors of arterial immune cells
Neighbors of random immune cells
Percentage of Cells (%)
Distance (μm)
Mesenchymal
Immune
Endothelial
Airway smooth muscle
Lympathic endothelial
Pericytes
SOX9high epithelial
SOX2high epithelial
Arterial Immune
Vascular smooth muscle
Chondroblast
C
Week 11
Arterial Immune vs. Immune
Week 12
Arterial Immune vs. Immune
Week 13
Arterial Immune vs. Immune
z-score
ranking
D
Composition of arterial immune cells
Week
Percentage
Macrophage
Pro. Macrophage
ILC & T cell
Pro. ILC & T cell
Dendritic cell

Figure 7. Immune cells encircling vasculature observed in later developmental weeks. **A)** Snapshots of immune cells and their spatial organization around arteries at different weeks. DAPI in blue, CD45 (immune marker) in green, ACTA2 (vascular smooth muscle marker) in orange and CD144 (endothelial marker) in pink. Scale bar corresponds to 20 μm for week 6, 30 μm for week 8.5 and 50 μm for weeks 11 to 13.

B) Line plots summarizing the cell type and count of neighbors of immune cells detected around arteries (left) and of neighbors of randomly selected other immune cells (right) within a radius of 50 μm for weeks 11 to 13. The x-axis represents the distance in μm and the y-axis represents the percentage of neighboring cells.

C) Differential marker expression analysis between immune cells detected around arteries and randomly selected immune cells for weeks 11 to 13. The x-axis shows the ranking of markers and the y-axis shows the Z-score for differential marker expression analysis.

D) Composition of the subtype of immune cells encircling arteries.

REFERENCES

1 Nikolić, M. Z., Sun, D. & Rawlins, E. L. Human lung development: recent progress and new challenges. *Development* **145**, doi:10.1242/dev.163485 (2018).

2 Niki, D. J. U. *et al.* Early origins of lung disease: towards an interdisciplinary approach. *European Respiratory Review* **29**, 200191, doi:10.1183/16000617.0191-2020 (2020).

3 Travaglini, K. J. *et al.* A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619-625, doi:10.1038/s41586-020-2922-4 (2020).

4 Sikkema, L. *et al.* An integrated cell atlas of the lung in health and disease. *Nature Medicine* **29**, 1563-1577, doi:10.1038/s41591-023-02327-2 (2023).

5 Haniffa, M. *et al.* A roadmap for the Human Developmental Cell Atlas. *Nature* **597**, 196-205, doi:10.1038/s41586-021-03620-1 (2021).

6 He, P. *et al.* A human fetal lung cell atlas uncovers proximal-distal gradients of differentiation and key regulators of epithelial fates. *Cell* **185**, 4841-4860.e4825, doi:10.1016/j.cell.2022.11.005 (2022).

7 Sountoulidis, A. *et al.* A topographic atlas defines developmental origins of cell heterogeneity in the human embryonic lung. *Nature Cell Biology* **25**, 351-365, doi:10.1038/s41556-022-01064-x (2023).

8 L, R. *et al.* A single-cell atlas of human fetal lung development between 14 and 19 weeks of gestation. *European Respiratory Journal* **60**, 145, doi:10.1183/13993003.congress-2022.145 (2022).

9 Henry, Q. *et al.* Early human fetal lung atlas reveals the temporal dynamics of epithelial cell plasticity. *bioRxiv*, 2023.2010.2027.564403, doi:10.1101/2023.10.27.564403 (2023).

10 Danopoulos, S. *et al.* Human lung branching morphogenesis is orchestrated by the spatiotemporal distribution of ACTA2, SOX2, and SOX9. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **314**, L144-L149, doi:10.1152/ajplung.00379.2017 (2017).

11 Barnes, J. L. *et al.* Early human lung immune cell development and its role in epithelial cell fate. *Science Immunology* **8**, eadf9988, doi:10.1126/sciimmunol.adf9988 (2023).

12 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).

13 Weigert, M., Schmidt, U., Haase, R., Sugawara, K. & Myers, G. in *2020 IEEE Winter Conference on Applications of Computer Vision (WACV).* 3655-3662.

14 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **13**, 731-740, doi:10.1038/nmeth.3901 (2016).

15 Isaac, V., Sergei, R., Fabian, J. T., Philipp, A. & Wolf, F. A. anndata: Annotated data. *bioRxiv*, 2021.2012.2016.473007, doi:10.1101/2021.12.16.473007 (2021).

16 Hickey, J. W., Tan, Y., Nolan, G. P. & Goltsev, Y. Strategies for Accurate Cell Type Identification in CODEX Multiplexed Imaging Data. *Frontiers in Immunology* **12** (2021).

17 Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology* **19**, 15, doi:10.1186/s13059-017-1382-0 (2018).

18 Zhang, Y., Parmigiani, G. & Johnson, W. E. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics and Bioinformatics* **2**, lqaa078, doi:10.1093/nargab/lqaa078 (2020).

19 Palla, G. *et al.* Squidpy: a scalable framework for spatial omics analysis. *Nature Methods* **19**, 171-178, doi:10.1038/s41592-021-01358-2 (2022).

20 Kim, J. *et al.* Unsupervised discovery of tissue architecture in multiplexed imaging. *Nature Methods* **19**, 1653-1661, doi:10.1038/s41592-022-01657-2 (2022).

21 Pielawski, N. *et al.* TissUUmaps 3: Improvements in interactive visualization, exploration, and quality assessment of large-scale spatial omics data. *Heliyon* **9**, e15306, doi:https://doi.org/10.1016/j.heliyon.2023.e15306 (2023).
22 Schupp, J. C. *et al.* Integrated Single-Cell Atlas of Endothelial Cells of the Human Lung. *Circulation* **144**, 286-302, doi:10.1161/CIRCULATIONAHA.120.052318 (2021).
23 Nikolić, M. Z. *et al.* Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids. *eLife* **6**, e26575, doi:10.7554/eLife.26575 (2017).
24 Warburton, D. *et al.* Lung organogenesis. *Curr Top Dev Biol* **90**, 73-158, doi:10.1016/s0070-2153(10)90003-3 (2010).
25 Bullwinkel, J. *et al.* Ki-67 protein is associated with ribosomal RNA transcription in quiescent and proliferating cells. *J Cell Physiol* **206**, 624-635, doi:10.1002/jcp.20494 (2006).
26 Yin, G., Zhao, C. & Pei, W. Crosstalk between macrophages and innate lymphoid cells (ILCs) in diseases. *International Immunopharmacology* **110**, 108937, doi:https://doi.org/10.1016/j.intimp.2022.108937 (2022).
27 Denisenko, E. *et al.* Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biology* **21**, 130, doi:10.1186/s13059-020-02048-6 (2020).
28 Riccio, V. D., Van Tuyl, M. & Post, M. Apoptosis in Lung Development and Neonatal Lung Injury. *Pediatric Research* **55**, 183-189, doi:10.1203/01.PDR.0000103930.93849.B2 (2004).
29 Yu, S. *et al.* Fetal and postnatal lung defects reveal a novel and required role for Fgf8 in lung development. *Developmental Biology* **347**, 92-108, doi:https://doi.org/10.1016/j.ydbio.2010.08.013 (2010).
30 Sanford, E. L. *et al.* MiR-449a Affects Epithelial Proliferation during the Pseudoglandular and Canalicular Phases of Avian and Mammal Lung Development. *PLoS One* **11**, e0149425, doi:10.1371/journal.pone.0149425 (2016).
31 Hislop, A. A. Airway and blood vessel interaction during lung development. *J Anat* **201**, 325-334, doi:10.1046/j.1469-7580.2002.00097.x (2002).
32 Lloyd, C. M. & Marsland, B. J. Lung Homeostasis: Influence of Age, Microbes, and the Immune System. *Immunity* **46**, 549-561, doi:https://doi.org/10.1016/j.immuni.2017.04.005 (2017).
33 Evren, E. *et al.* Distinct developmental pathways from blood monocytes generate human lung macrophage diversity. *Immunity* **54**, 259-275.e257, doi:10.1016/j.immuni.2020.12.003 (2021).
34 Bharat, A. *et al.* Flow Cytometry Reveals Similarities Between Lung Macrophages in Humans and Mice. *Am J Respir Cell Mol Biol* **54**, 147-149, doi:10.1165/rcmb.2015-0147LE (2016).
35 Domingo-Gonzalez, R. *et al.* Diverse homeostatic and immunomodulatory roles of immune cells in the developing mouse lung at single cell resolution. *Elife* **9**, doi:10.7554/eLife.56890 (2020).
36 Dahlgren, M. W. *et al.* Adventitial Stromal Cells Define Group 2 Innate Lymphoid Cell Tissue Niches. *Immunity* **50**, 707-722.e706, doi:10.1016/j.immuni.2019.02.002 (2019).
37 Jambusaria, A. *et al.* Endothelial heterogeneity across distinct vascular beds during homeostasis and inflammation. *eLife* **9**, e51413, doi:10.7554/eLife.51413 (2020).
38 Ma, Q. *et al.* Regeneration of functional alveoli by adult human SOX9+ airway basal cell transplantation. *Protein & Cell* **9**, 267-282, doi:10.1007/s13238-018-0506-y (2018).
39 Yatabe, Y., Mitsudomi, T. & Takahashi, T. TTF-1 Expression in Pulmonary Adenocarcinomas. *The American Journal of Surgical Pathology* **26** (2002).