Single Cell RNA Sequencing Reveals Gene Expression Continuums Along the Spatial Hierarchy of the Pulmonary Circulation

Stefanie Sveiven[1]*, Carsten Knutsen[1]*, Fabio Zanini[2,3,4], David Cornfield[5,6], Cristina Alvira[1]**

1. Division of Critical Care Medicine, Department of Pediatrics, University of California San Francisco, San Francisco, CA 94158, USA

2. School of Clinical Medicine, UNSW Sydney, 2052, NSW, Australia

3. UNSW Cellular Genomics Futures Institute, 2052, NSW, Australia

4. Evolution & Ecology Research Centre, 2052, NSW, Australia

5. Center for Excellence in Pulmonary Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

6. Division of Pulmonary, Asthma and Sleep Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA 94305, USA

*** These authors contributed equally.**

**** Corresponding author: cristina.alvira@ucsf.edu**

Abstract

Objective: To establish a novel strategy that leverages scRNAseq data to map individual pulmonary endothelial cells to vessels of a specific size and location, thereby providing functional and spatial context to transcriptomic identity.
Methods and Results: We applied a method of endothelial enrichment and deep single cell RNA sequencing to create a high resolution, transcriptomic dataset from the developing mouse lung. We developed an analytical framework to assign vessel-size scores and categorize individual endothelial cells (EC) and mural cells along a continuum of vessel sizes. We delineated a continuum of proximal arterial through distal venous EC states by uncovering transcriptional signatures associated with vessel size, spanning micro- to macrovascular zones. Our data recapitulated previously established zonally defined signaling axes, including *Cxcl12* and *Cxcr4* in arterioles, and identified localization of disease relevant markers such as *Esr2*. This vessel-size informed framework was robust across species and revealed how spatial EC heterogeneity underlies key processes in lung development and injury.

Conclusions: These findings provide a comprehensive transcriptional map of pulmonary endothelial cells across the vascular continuum, offering valuable insights into spatial inferences and mechanistic insights within single cell RNA sequencing data sets that may help pave the way for targeted therapeutic strategies to treat pulmonary vascular diseases and expansion to tissues outside the lung.

Introduction

Recent advances in single cell RNA sequencing (scRNAseq) have transformed our understanding of cellular biology, revealing a previously underappreciated degree of tissue-specific diversity across cellular subtypes, including insights into the diversity of pulmonary vascular cells during development and response to injury[1–5]. Although dimensionality reduction and clustering algorithms are powerful tools for interpreting large scale scRNAseq data, they prioritize global variance and shared transcriptional programs, often fragmenting continuous cell states into discrete clusters[6]. This is a significant limitation

because, despite shared origins or functions, lineage subtypes contribute uniquely to tissue physiology and pathology[7,8]. In the pulmonary circulation, EC exist along a vascular continuum, experiencing differential exposure to key microenvironmental signals such as blood flow, pressure, pulsatility, and oxygenation depending on their precise location along the arterial-capillary-venous axis[9]. Simplified categorization of EC broadly as artery, vein, or capillary likely obscures the transcriptomic and phenotypic variation of cells across these physiological axes.

The entire cardiac output passes through the lungs with each heartbeat, allowing the totality of deoxygenated blood to pass through the alveolar capillary network to ensure distribution of oxygen rich blood to the remaining organs. Lining the pulmonary circulation, pulmonary arterial (PAEC), capillary and venous (PVEC) EC play distinct roles, serving to modulate vascular tone to optimize ventilation and perfusion matching, performing gas exchange while maintaining a tight barrier to prevent fluid extravasation into the alveolus, and regulation of leukocyte trafficking and immune responses to pathogens entering through the alveolar space. In addition to their primary role in blood transport, EC also inform local cellular niches through angiocrine signaling– paracrine communication via cell surface and secreted factors that influence the behavior of neighboring cells[9]. Although single cell transcriptomics has identified distinct gene expression signatures exhibited by broad pulmonary EC subtypes, even within a larger subpopulation, PAEC and PVEC located at proximal versus distal positions of the circulation encounter differences in physiologic cues and microenvironments. As many pulmonary vascular diseases selectively affect specific locations within the circulation, developing computation methodologies to identify changes in pulmonary EC at specific locations will greatly enhance our ability to contextualize alterations during lung development and disease, such as muscularization in resistance arterioles in pulmonary arterial hypertension, or remodeling of venules in pulmonary veno-occlusive disease[10,11]. Moreover, computational methods to infer cell-cell communication have greatly advanced our understanding of how endothelial phenotype influences the behavior of surrounding cells yet are currently limited by lack clarity regarding the true proximity of specific cells that will dictate communication capability.

Using a strategy of endothelial enrichment and deep single cell RNA sequencing to create a high resolution, transcriptomic dataset from the developing mouse lung, we developed an analytical framework to assign vessel-size scores and categorize individual endothelial cells (EC) and mural cells along a continuum of vessel size. We delineated a continuum of arterial to venous, and macro- to microvascular zones using these sized-based transcriptional signatures. Our strategy allowed identification of signaling axes previously established with spatial methodologies purely with transcriptomic data, and disease relevant markers such as *Esr2*. This vessel-size informed framework is robust across development and species and reveals how spatial EC heterogeneity underlies key processes in lung development and injury.

Materials and Methods

Sample preparation- At birth, neonatal C57BL-6 (Charles River) mice were housed in room air (normoxia; N) or 80% O_2 (hyperoxia; H) for 72hr before euthanasia (N=6 per exposure; 3M 3F). Lungs were perfused with HBSS, excised, finely minced, pooled, and digested in Liberase TM (0.2mg/mL) plus DNaseI (0.01mg/mL) for 10min at 37.5 in a bacteriological shaker. The suspension was triturated 15 times. Liberase was inactivated using 2mM EDTA and cold FBS prior to pelleting the cell suspension. Cells were treated with 1X RBC Lysis for 5min then passed over a 30µm cell strainer before microbead enrichment (SmartStrainer, Miltenyi). In a parallel set of experiments, lungs were pressure fixed at 20cm H_2O and paraffin embedded (FFPE) for *in situ* validation experiments.

Enrichment by magnetic associated cell sorting (MACS)- Immune cells were depleted using anti-CD45 coated Dynabeads (Invitrogen). The supernatant was pelleted, counted, and resuspended with blocking buffer (anti-mouse CD16/CD32 and anti-rat IgG; 1:100) for 10min. Mouse anti-BST1–APC (PVEC

marker, Biolegend) and rat anti-CD31–biotin (pan-EC marker, BD Pharminogen) antibodies were added to the cells (2µg/10M cells) and incubated on a HulaMixer for 20 minutes. To ensure sufficient PVEC for our analysis, following the manufacturer's protocol, cells were first enriched for BST1 using the MACS anti-APC microbeads and then the flow through was enriched for CD31 using anti-biotin microbeads (Miltenyi MACS).

Chromium GEMX 3' v4 gene expression- Following the manufacturers protocol, enriched populations were then adjusted to 1500 cell/µL to obtain the 20,000 cell target input for gel bead emulsion formation (10X Genomics). Samples that met QC standards (Bioanalyzer, Stanford Protein and Nucleic Acid core) were used for barcoding, library construction and sequenced by NovoSeq X at a depth of 1B paired reads per sample (NovoGene Corporation Inc.).

scRNASeq Analysis: Sequencing reads were aligned to the Grcm39 mouse genome using Cellranger (v9.0.0). SoupX was used to remove ambient RNA signal[12]. Gene expression count tables were turned into anndata objects and processed using scanpy when not otherwise specified[13]. Cells that had more than five median absolute deviations of either unique genes or UMIs detected were removed. Cells that had more than three median absolute deviations of percent mitochondrial or ribosomal UMIs were removed[6]. Scrublet was used to automatically detect and remove doublets. Counts were normalized and log transformed[14]. PCA was run prior to embedding and clustering. The Leiden algorithm was used for clustering, and Uniform Manifold Approximation and Projection for embedding. Lineage identity was assigned using canonical markers: *Cdh5* (endothelial); *Epcam* (epithelial); *Ptprc* (immune); and *Col1a1* (mesenchymal). Cell typing was performed after each lineage was re-clustered and embedded as detailed above. Cell types were assigned a second time after regressing out cell cycle genes[6]. Pseudotime analysis and branching was performed using Palanitr[15], and RNA velocity determined using velocyto to align reads, and scVelo to calculate trajectories[16,17]. Vessel size scoring was performed by finding shared genes within the top 50 genes positively or negatively correlating with pseudotime in PAEC or PVEC, using Cap1 EC as the root. Scores were generated for positively and negatively correlated genes separately, and the negative correlation score subtracted from the positive score, then scaled to be between 0-1. Vessel size categories were assigned from quartiles in the vessel size score. For more detail see https://github.com/CarstenKnutsen/Vessel_size_manuscript

In situ Validation by RNA fluorescence in situ hybridization and immunofluorescence: RNAScope MultiPlex v2 Assay (Advanced Cellular Diagnostics) and immunofluorescence was performed on 5µm FFPE lung sections from P3 mice. RNAScope probes were purchased from ACD and fluorophores from Akoya: Mm-*Fbln2*, Mm-*Tmem100,* Mm-*Pecam1*. Fluorescence images were captured using a Zeiss Axio Observer 7 equipped with Apotome for optical sectioning using a 20X objective.

Results

Pulmonary endothelial cells exhibit a transcriptomic continuum from which size and arterial-venous zonation can be inferred.

We performed single cell RNA sequencing from pulmonary cells obtained from mice on postnatal day 3 (P3), a period of rapid growth of the pulmonary vasculature. We enriched for PVEC (BST1+) or total EC (CD31+) resulting in 11,724 of the total 32,946 cells being of endothelial identity, with a median of 4,586 genes detected per cell **(Fig. 1A, B)**. We included mice exposed to room air or acute hyperoxia (80% O_2), a model that impairs both pulmonary angiogenesis and alveolarization.

Deoxygenated blood flows through the pulmonary artery and its branching network of progressively smaller arteries, ultimately reaching capillaries at the alveoli where gas exchange occurs. Oxygenated blood then flows through a mirrored network of small to large veins back through the pulmonary vein to the left atrium. We included capillary type 1 cells (Cap1) alongside PAEC and PVEC due to their proximity to macrovessels at the pre- and post-capillary transition zones[1,18]. To achieve higher resolution

of these macro-microvascular transitions, we re-embedded these three populations resulting in a central node of Cap1 with PAEC and PVEC clusters extending from either side **(Fig 1C)**. To investigate cellular trajectories within these populations we performed RNA velocity (**Fig. 1D**). RNA velocity vectors indicate that Cap1 cells near the cell type transitions progress towards either PAEC or PVEC. Further, diffusion-based pseudotime and branching analyses conferred this relationship, revealing two mirrored trajectories from Cap1 toward PAEC and PVEC with pseudotime progressively increasing away from the Cap1 node **(Fig. 1E)**. Using these psuedotime branches, we inferred arterio-venous zonation, with the Cap1 cluster exhibiting features of an arterial-venous continuum; one half expressed high levels of aterial markers (*Hey1*, *Pcsk5*), while the other half expressed venous markers (*Nr2f2*, *Ackr3*) (**Fig. 1F, G**).

We hypothesized that this transcriptional continuum paralleled the structural branching of the lung vasculature, with arteries and veins progressively decreasing in diameter approaching the distal capillary network, and that vessel size therefore be inferred from transcriptomic data. To assess whether specific genes may indicate vessel size, we identified the top 50 genes that correlated positively and negatively with pseudotime along the Cap1-arterial and Cap1-venous trajectories. Roughly one-third of the top positively and negatively correlated genes from each comparison were shared among the macrovessel EC pseudotime correlations (**Fig 1H**). Some of these genes were known macrovessel markers such as elastin, *Eln,* and von Willebrand Factor, *Vwf*, and microvascular markers such as proto-oncogene, *Kit*, and transmembrane protein 100, *Tmem100*[18,19]. Using the expression levels of these shared pseudotime-correlated genes, we calculated vessel size scores, and set four binned-categories by quartile, representing 'large', 'medium', 'small', and 'capillary' based on score ranges. We projected these vessel-size scores and categories onto the UMAP and found that the terminal ends of the macrovascular clusters exhibited the highest vessel-size scores, revealing a progressive distribution of vessel sizes(**Fig 1I, J)**.

We then identified signals localized along the arterio-venous continuum using these sizes. Consistent with our previous report, proximal large arteries showed highest expression of the Wnt-inhibitor *Dkk2* and transcription factor *Sox6*[3]. Small arteries were enriched for the HIF-responsive remodeling gene *Depp1* and the autocrine angiogenesis regulator *Stc1.* The glucagon-like peptide-1 receptor *Glp1r*, and *Kit*–a known Cap1 marker– were most highly expressed in small arteries, veins, and capillaries[1,3,18]. Among PVEC with small size scores, muscarinic cholinergic receptor 3 *Chrm3*, a regulator of vasodilation and blood pressure is enriched[20]. *Fads2b*, a fatty acid desaturase, was also enriched in PVEC– particularly those with medium-size scores– and is implicated in fatty acid metabolism, a pathway linked to acute lung injury pathogenesis[21]. PVEC ascribed to large veins were enriched for the tumor suppressor *Moxd1*, and *Ptger3,* prostaglandin E receptor 3, which regulates lung edema by modulating vascular permeability[22]. These data highlight the spatial heterogeneity of gene expression along the vascular continuum (**Fig 1K)**.

Given this clear transcriptional pattering identifying size within the developing pulmonary endothelium, we sought to determine if this paradigm extended to the mural populations. Vascular smooth muscle cells (VSMC) surround large arteries and veins and pericytes stabilize capillaries, thus, we performed pseudotime analyses using pericytes as the root cell[23–25]. Like macrovascular EC, VSMC exhibit advanced pseudotime when compared to capillary-associated pericytes (**Fig. 1L**). Strikingly, genes correlated with advancing pseudotime in VSMC versus pericytes included genes shared by the macrovascular-capillary comparisons (**Fig. 1M**). For example, *Mgp* (matrix Gla protein), a gene with highest expression in the largest arteries and veins, was also most highly expressed in VSMC with the greatest pseudotime values (**Fig. 1N**).

We validated this transcriptomic scoring of vessel size, and performed RNA fluorescence *in situ* hybridization targeting *Pecam1* to label endothelial cells alongside either small vessel marker, *Tmem100*, and large vessel marker, *Fbln2* (**Fig. 1O**). *Tmem100* had limited expression in medium-large vessels while diffusely expressed in the smaller vessels. *Fbln2* had enriched expression in medium-large vessels, and diminished expression in the smaller vessels.

Vessel size gradients are developmentally and disease relevant and are conserved across species

We next aimed to contextualize the significance of our vessel-size analysis by validating existing size-related findings in development and disease. We first examined the CXCL12-CXCR4-ACKR3 autocrine signaling axis, which is critical for arteriolar patterning during lung development (**Fig. 2A**)[26,27]. Consistent with published data, we found the highest *Cxcl12* and its' cognate receptor *Cxcr4* expression in PAEC. Notably, the expression of *Ackr3*, a decoy receptor that clears excess CXCL12, is highest among large veins, underscoring the spatially essential role of PVEC in restricting this signaling to the pulmonary (not systemic) circulation (**Fig. 2B**). This axis is also implicated in the pathogenesis of pulmonary hypertension (PH), where elevated CXCL12 activates CXCR4 to remodel of small arteries, increasing pulmonary resistance. We examined the size-specific expression of *Esr2* and found highest expression in EC located along the Cap1-Artery arm corresponding to small and medium arteries (**Fig 2C**). Analysis of *Esr2* across the arterial pseudotime branch reveals peak expression among cells with intermediate pseudotime values, corroborating that expression of *Esr2* is highest among small-medium arteries (**Fig 2D**)[28]. All together, these data validate and amplify the importance of understanding size specific vascular expression in the lung development and disease.

We next evaluated whether this analytical framework could give novel insight into size-related differences in a distinct vascular response to injury. We previously showed that PVEC exhibit a unique proliferative response to hyperoxia[3]. We compared the proliferation scores using roughly 100 proliferation-associated genes among PVEC within each of the three size scores in normoxia and hyperoxia. We found that PVEC from small and medium veins exhibited the greatest increase in proliferation in hyperoxia (**Fig 2E).** We specifically compared the differential expression of 5 genes that promote proliferation in each size of PVEC, and found the greatest increase in hyperoxia mediated expression was found in PVEC from small-medium vessels (**Fig 2F**). These data suggest that although pro-proliferative transcription is enhanced in EC from veins of all sizes, the most pronounced increase in proliferation is likely occurring in small and medium veins.

Finally, we evaluated the applicability of our approach to human lung data using three different datasets that including samples from neonatal, pediatric, and adult lungs[29–31]. We re-analyzed and embedded PVEC, Cap1, and PAEC in each dataset, which formed similar UMAP patterning to our mouse data (**Fig 2 G,K,O**). We then applied the strategy to approximate vessel size and found that the terminal ends of the venous and arterial clusters exhibited the highest size scores, which gradually declined towards the Cap1 cluster (**Fig. 2H-I,L-M,P-O**). Of note, we found that many of the size-related genes were consistent between mouse and human, including *ELN,* which was highest in human lung macrovessels and *COL4A1,* highest in human lung Cap1(**Fig 2J,N,R**). Together, these data validate the ability to apply our framework for attributing vessel-size scores to individual arterial, venous, and Cap1 EC across heterogenous scRNAseq datasets, and that specific genes expressed along the vascular hierarchy are conserved across development and species (**Fig. 2G-R**).

Discussion

Endothelial cell subtype identity is likely shaped by positional cues, mechanical forces, and other physiologic signals that vary by vessel size, providing a physiologically relevant axis that we hypothesized was evident within robust scRNAseq data. By integrating vessel-size-informed analyses, gradients of gene expression linked to structure can be revealed, a feature central to both normal vascular development and disease pathogenesis. Our model accurately inferred vessel size gradients in EC and EC-adjacent mural populations across multiple datasets that were previously unresolved. To optimally apply single cell transcriptomes within the broader spatial and physiologic context of the lung, the pulmonary vasculature must be viewed as a continuum of cells across multiple spectra of features (size,

oxygenation, flow, pressure. Prior single cell transcriptomic simplify large datasets through segregation of cell types based on predefined marker genes, an approach likely to oversimplify vascular identities which exist along a continuum of lineage-specific, yet spatially and transcriptionally graded-cell states[5]. By assuming that distinct transcriptomic signatures exist among EC along this vascular hierarchy, our approach enabled efficient inference of spatially relevant vascular features. Other approaches, such as spatial transcriptomics or multiplexed immunohistochemistry, offer limited resolution and are restricted to a single tissue section or a handful of pre-selected proteins, whereas transcriptome-wide high-throughput scRNAseq enable unbiased discovery and broader functional inference. Furthermore, the conservation of vessel gradients of expression across both mouse and human datasets supports the robustness and translational relevance of this approach.

Our findings align with and provide resolution on prior reports of size-specific differential gene across development and disease. For example, Chandrasekaran et al., clearly demonstrate the critical role of CXCL12 (stromal cell derived factor 1) and its receptor CXCR4 (CXC chemokine receptor 4) in arterial patterning, particularly in the development and organization of distal arteries[15]. Their work highlights the importance of CXCL12-CXCR4 signaling axis in guiding EC positioning and vessel branching during lung vascular development. Their scRNAseq data demonstrates the expression of *Cxcl12* among distal, small arteries which we validated using our methodology and dataset, revealing that small arteries express both ligand, *Cxcl12* and receptor, *Cxcr4*. We also highlight *Ackr3* expression among larger veins, clearing excess CXCL12 prior to systemic dissemination of the blood and restricting this signaling access to the lung. Estrogen acts upstream of this axis by activating estrogen receptor α (ERα; *Esr1*/*Esr2)* and increasing CXCL12 production, as the promoter region of the *Cxcl12* gene contains estrogen response elements (ERE) [32]. Notably, estrogen plays context-specific, contrasting roles in PH pathogenesis. Estrogen can both confer susceptibility to PH development, and attenuate PH by promoting endothelial repair[33–36]. In mouse models, knock-out of Esr2 ameliorates muscularization of small arteries[37]. Thus, CXCL12-CXCR4 contribution to maladaptive angiogenesis might be mediated by estrogen receptor β (ESR2)[35,37]. Chen et al. demonstrate that deletion of *Esr2* in mice ameliorates the development of PAH in a genetic model, specifically the muscularization of small vessels. In the analyses of our data, *Esr2* expression is enriched in PAEC with pseudotime scores associated with small-medium arteries.

Zonally-defined transcriptomic patterns of EC proliferation enhance our understanding of heterogenous EC function in lung development and injury. In early vascular development, the pulmonary vasculature branches in a proximal-to-distal pattern. Subsequently, EC proliferation predominately in the microvasculature supports distal vascular and epithelial expansion while proximal vessels undergo stabilization and maturation. Recent studies, including our own, highlight the unique ability of PVEC to exhibit heightened proliferation in response to injuries, such as hyperoxia or infection[3,38]. In our prior work, hyperoxia uniquely increased proliferation in PVEC; however, the resolution was insufficient to infer the size of the proliferating veins[26]. Our current analytical approach refined the location of proliferative activity, revealing enrichment in small veins, where they may contribute to re-establishing lost capillary populations following hyperoxic injury– a function recently attributed to lung PVEC by Wong et al[27].

There are some notable limitations to our proposed methods. First, achieving this level of resolution required endothelial-enriched libraries and increased sequencing depth enabled by newer technologies with higher target cell capture efficiency. Even with enrichment and deep sequencing, the diversity of EC from different vessel sizes remains constrained by sampling. In larger tissues like human lung, sampling from proximal versus distal regions impact the range of vessel sizes captured, thereby limiting resolution of the full vascular hierarchies. Although vessel-size scoring offers valuable insights, it is based on transcriptomic proxies that may not fully reflect anatomical scale or hemodynamic properties. Moreover, EC transcriptional profiles are also shaped by factors such as flow dynamics and oxygenation–variables not addressed in our current model. However, because these variables are also spatially zoned, we anticipate they could serve as complementary axes for future analyses.

Signaling gradients are fundamental to organ development– providing spatial cues that guide cell fate decisions and tissue organization. Thus, it is unsurprising that vascular EC exhibit a transcriptomic continuum that reflects their position along physiologic axes. Importantly, cellular microenvironments composed of diverse cell types also exist along these axes. Defining the hierarchy of cell-cell interactions along vascular zonation provides a more complete understanding of the complex, spatially organized cellular crosstalk that drives both development and disease. Incorporating multimodal datasets such as single cell assay for transposase-accessible chromatin combined with scRNAseq may illuminate the transcriptional regulation underlying these gradients, offering deeper mechanistic insights into cell-state specification along physiologic axes of interest. Further, given that the vascular supply to all organs is comprised of an arterial to capillary to venous structure, similar approaches to robust single cell datasets from other organs and tissues will likely allow discrimination of similar hierarchical system that may be resolved using our analytical approach.

Significance

There are broad, translational implications for the framework that we have demonstrated in this report. The conservation of pulmonary vessel-size-associated gene expression and vascular hierarchy between mouse and human supports relevance for human vascular disease modeling. In addition to validating disease models, this analytical approach enables researchers to glean deeper insights from the limited scRNAseq datasets available from rare human samples. By resolving vascular features at the level of vessel size, this method facilitates the identification of transcriptional programs specific to small, medium, and large vessels– thereby uncovering potential therapeutic targets for diseases that preferentially affect vessels of specific sizes.

Data Availability

Data generated for this manuscript is shared at https://figshare.com/account/articles/29079218?file=54578297. Human neonatal data was sourced from https://cellxgene.cziscience.com/collections/28e9d721-6816-48a2-8d0b-43bf0b0c0ebc. Human 3-year-old data was sourced from https://explore.data.humancellatlas.org/projects/01aacb68-4076-4fd9-9eb9-aba0f48c1b5a. Adult human data was sourced from https://figshare.com/articles/dataset/Tabula_Sapiens_v2/27921984.

References

1. Gillich A, Zhang F, Farmer CG, Travaglini KJ, Tan SY, Gu M, Zhou B, Feinstein JA, Krasnow MA, Metzger RJ. Capillary cell-type specialization in the alveolus. *Nature*. 2020;586(7831):785-789. doi:10.1038/s41586-020-2822-7

2. Schupp JC, Adams TS, Cosme C, Raredon MSB, Yuan Y, Omote N, Poli S, Chioccioli M, Rose KA, Manning EP, Sauler M, DeIuliis G, Ahangari F, Neumark N, Habermann AC, Gutierrez AJ, Bui LT, Lafyatis R, Pierce RW, Meyer KB, Nawijn MC, Teichmann SA, Banovich NE, Kropski JA, Niklason LE, Pe'er D, Yan X, Homer RJ, Rosas IO, Kaminski N. Integrated Single-Cell Atlas of Endothelial Cells of the Human Lung. *Circulation*. 2021;144(4):286-302. doi:10.1161/CIRCULATIONAHA.120.052318

3. Zanini F, Che X, Knutsen C, Liu M, Suresh NE, Domingo-Gonzalez R, Dou SH, Zhang D, Pryhuber GS, Jones RC, Quake SR, Cornfield DN, Alvira CM. Developmental diversity and unique sensitivity to injury of lung endothelial subtypes during postnatal growth. *iScience*. 2023;26(3):106097. doi:10.1016/j.isci.2023.106097

4. Zepp JA, Morley MP, Loebel C, Kremp MM, Chaudhry FN, Basil MC, Leach JP, Liberti DC, Niethamer TK, Ying Y, Jayachandran S, Babu A, Zhou S, Frank DB, Burdick JA, Morrisey EE. The

genomic, epigenomic, and biophysical cues controlling the emergence of the lung alveolus. *Science*. 2021;371(6534):eabc3172. doi:10.1126/science.abc3172

5. Zepp JA, Morrisey EE. Cellular crosstalk in the development and regeneration of the respiratory system. *Nat Rev Mol Cell Biol*. 2019;20(9):551-566. doi:10.1038/s41580-019-0141-3

6. Heumos L, Schaar AC, Lance C, Litinetskaya A, Drost F, Zappia L, Lücken MD, Strobl DC, Henao J, Curion F, Schiller HB, Theis FJ. Best practices for single-cell analysis across modalities. *Nat Rev Genet*. 2023;24(8):550-572. doi:10.1038/s41576-023-00586-w

7. Ardini-Poleske ME, Clark RF, Ansong C, Carson JP, Corley RA, Deutsch GH, Hagood JS, Kaminski N, Mariani TJ, Potter SS, Pryhuber GS, Warburton D, Whitsett JA, Palmer SM, Ambalavanan N, LungMAP Consortium. LungMAP: The Molecular Atlas of Lung Development Program. *Am J Physiol Lung Cell Mol Physiol*. 2017;313(5):L733-L740. doi:10.1152/ajplung.00139.2017

8. Kalucka J, de Rooij LPMH, Goveia J, Rohlenova K, Dumas SJ, Meta E, Conchinha NV, Taverna F, Teuwen LA, Veys K, García-Caballero M, Khan S, Geldhof V, Sokol L, Chen R, Treps L, Borri M, de Zeeuw P, Dubois C, Karakach TK, Falkenberg KD, Parys M, Yin X, Vinckier S, Du Y, Fenton RA, Schoonjans L, Dewerchin M, Eelen G, Thienpont B, Lin L, Bolund L, Li X, Luo Y, Carmeliet P. Single-Cell Transcriptome Atlas of Murine Endothelial Cells. *Cell*. 2020;180(4):764-779.e20. doi:10.1016/j.cell.2020.01.015

9. Sharma A, Niethamer TK. Specialized Pulmonary Vascular Cells in Development and Disease. *Annual Review of Physiology*. 2025;87(Volume 87, 2025):229-255. doi:10.1146/annurev-physiol-022724-105226

10. Hansmann G, Sallmon H, Roehr CC, Kourembanas S, Austin ED, Koestenberger M. Pulmonary hypertension in bronchopulmonary dysplasia. *Pediatr Res*. 2021;89(3):446-455. doi:10.1038/s41390-020-0993-4

11. Montani D, Lau EM, Dorfmüller P, Girerd B, Jaïs X, Savale L, Perros F, Nossent E, Garcia G, Parent F, Fadel E, Soubrier F, Sitbon O, Simonneau G, Humbert M. Pulmonary veno-occlusive disease. *European Respiratory Journal*. 2016;47(5):1518-1534. doi:10.1183/13993003.00026-2016

12. Young MD, Behjati S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *GigaScience*. 2020;9(12):giaa151. doi:10.1093/gigascience/giaa151

13. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology*. 2018;19(1):15. doi:10.1186/s13059-017-1382-0

14. Wolock SL, Lopez R, Klein AM. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst*. 2019;8(4):281-291.e9. doi:10.1016/j.cels.2018.11.005

15. Characterization of cell fate probabilities in single-cell data with Palantir | Nature Biotechnology. Accessed May 9, 2025. https://www.nature.com/articles/s41587-019-0068-4

16. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lönnerberg P, Furlan A, Fan J, Borm LE, Liu Z, van Bruggen D, Guo J, He X, Barker R, Sundström E, Castelo-Branco G, Cramer P, Adameyko I, Linnarsson S, Kharchenko PV. RNA velocity of single cells. *Nature*. 2018;560(7719):494-498. doi:10.1038/s41586-018-0414-6

17. Bergen V, Soldatov RA, Kharchenko PV, Theis FJ. RNA velocity—current challenges and future perspectives. *Molecular Systems Biology*. 2021;17(8):e10282. doi:10.15252/msb.202110282

18. Vila Ellis L, Cain MP, Hutchison V, Flodby P, Crandall ED, Borok Z, Zhou B, Ostrin EJ, Wythe JD, Chen J. Epithelial Vegfa Specifies a Distinct Endothelial Population in the Mouse Lung. *Dev Cell*. 2020;52(5):617-630.e6. doi:10.1016/j.devcel.2020.01.009

19. Somekawa S, Imagawa K, Hayashi H, Sakabe M, Ioka T, Sato GE, Inada K, Iwamoto T, Mori T, Uemura S, Nakagawa O, Saito Y. Tmem100, an ALK1 receptor signaling-dependent gene essential for arterial endothelium differentiation and vascular morphogenesis. *Proceedings of the National Academy of Sciences*. 2012;109(30):12064-12069. doi:10.1073/pnas.1207210109

20. Saternos HC, Forero KV, Meqdad MA, Buqaileh R, Sunderman CL, Gallagher G, Messer WS, Mohieldin AM, Mucci CA, Kumariya S, Osman IA, Burkett JP, AbouAlaiwi WA. Muscarinic acetylcholine receptor 3 localized to primary endothelial cilia regulates blood pressure and cognition. *Sci Rep*. 2025;15(1):3745. doi:10.1038/s41598-025-87212-3

21. Lu X, Li G, Liu Y, Luo G, Ding S, Zhang T, Li N, Geng Q. The role of fatty acid metabolism in acute lung injury: a special focus on immunometabolism. *Cell Mol Life Sci*. 2024;81(1):120. doi:10.1007/s00018-024-05131-4

22. Jiang T, Samapati R, Klassen S, Lei D, Erfinanda L, Jankowski V, Simmons S, Yin J, Arenz C, Dietrich A, Gudermann T, Adam D, Schaefer M, Jankowski J, Flockerzi V, Nüsing R, Uhlig S, Kuebler WM. Stimulation of the EP3 receptor causes lung oedema by activation of TRPC6 in pulmonary endothelial cells. *European Respiratory Journal*. 2022;60(4). doi:10.1183/13993003.02635-2021

23. Payne LB, Hoque M, Houk C, Darden J, Chappell JC. Pericytes in Vascular Development. *Curr Tissue Microenviron Rep*. 2020;1(3):143-154. doi:10.1007/s43152-020-00014-9

24. Yuan K, Agarwal S, Chakraborty A, Condon DF, Patel H, Zhang S, Huang F, Mello SA, Kirk OI, Vasquez R, de Jesus Perez VA. Lung Pericytes in Pulmonary Vascular Physiology and Pathophysiology. *Compr Physiol*. 2021;11(3):2227-2247. doi:10.1002/cphy.c200027

25. Zanini F, Che X, Suresh NE, Knutsen C, Klavina P, Xie Y, Domingo-Gonzalez R, Liu M, Kum A, Jones RC, Quake SR, Alvira CM, Cornfield DN. Hyperoxia prevents the dynamic neonatal increases in lung mesenchymal cell diversity. *Sci Rep*. 2024;14(1):2033. doi:10.1038/s41598-023-50717-w

26. Chandrasekaran P, Negretti NM, Sivakumar A, Liberti DC, Wen H, Peers de Nieuwburgh M, Wang JY, Michki NS, Chaudhry FN, Kaur S, Lu M, Jin A, Zepp JA, Young LR, Sucre JMS, Frank DB. CXCL12 defines lung endothelial heterogeneity and promotes distal vascular growth. *Development*. 2022;149(21):dev200909. doi:10.1242/dev.200909

27. Yi D, Liu B, Wang T, Liao Q, Zhu MM, Zhao YY, Dai Z. Endothelial Autocrine Signaling through CXCL12/CXCR4/FoxM1 Axis Contributes to Severe Pulmonary Arterial Hypertension. *Int J Mol Sci*. 2021;22(6):3182. doi:10.3390/ijms22063182

28. Favre J, Vessieres E, Guihot AL, Proux C, Grimaud L, Rivron J, Garcia MC, Réthoré L, Zahreddine R, Davezac M, Fébrissy C, Adlanmerini M, Loufrani L, Procaccio V, Foidart JM, Flouriot G, Lenfant F, Fontaine C, Arnal JF, Henrion D. Membrane estrogen receptor alpha (ERα) participates

in flow-mediated dilation in a ligand-independent manner. *Elife*. 2021;10:e68695. doi:10.7554/eLife.68695

29. Bhattacharya S, Myers JA, Baker C, Guo M, Danopoulos S, Myers JR, Bandyopadhyay G, Romas ST, Huyck HL, Misra RS, Dutra J, Holden-Wiltse J, McDavid AN, Ashton JM, Al Alam D, Potter SS, Whitsett JA, Xu Y, Pryhuber GS, Mariani TJ. Single-Cell Transcriptomic Profiling Identifies Molecular Phenotypes of Newborn Human Lung Cells. *Genes (Basel)*. 2024;15(3):298. doi:10.3390/genes15030298

30. Wang A, Chiou J, Poirion OB, Buchanan J, Valdez MJ, Verheyden JM, Hou X, Kudtarkar P, Narendra S, Newsome JM, Guo M, Faddah DA, Zhang K, Young RE, Barr J, Sajti E, Misra R, Huyck H, Rogers L, Poole C, Whitsett JA, Pryhuber G, Xu Y, Gaulton KJ, Preissl S, Sun X, NHLBI LungMap Consortium. Single-cell multiomic profiling of human lungs reveals cell-type-specific and age-dynamic control of SARS-CoV2 host genes. Morrisey EE, De Langhe S, eds. *eLife*. 2020;9:e62522. doi:10.7554/eLife.62522

31. Quake SR, Consortium TTS. Tabula Sapiens reveals transcription factor expression, senescence effects, and sex-specific features in cell types from 28 human organs and tissues. Published online December 4, 2024:2024.12.03.626516. doi:10.1101/2024.12.03.626516

32. Zhu Y, Sullivan LL, Nair SS, Williams CC, Pandey AK, Marrero L, Vadlamudi RK, Jones FE. Coregulation of estrogen receptor by ERBB4/HER4 establishes a growth-promoting autocrine signal in breast tumor cells. *Cancer Res*. 2006;66(16):7991-7998. doi:10.1158/0008-5472.CAN-05-4397

33. Filipe C, Lam Shang Leen L, Brouchet L, Billon A, Benouaich V, Fontaine V, Gourdy P, Lenfant F, Arnal JF, Gadeau AP, Laurell H. Estradiol accelerates endothelial healing through the retrograde commitment of uninjured endothelium. *Am J Physiol Heart Circ Physiol*. 2008;294(6):H2822-2830. doi:10.1152/ajpheart.00129.2008

34. Toutain CE, Filipe C, Billon A, Fontaine C, Brouchet L, Guéry JC, Gourdy P, Arnal JF, Lenfant F. Estrogen receptor alpha expression in both endothelium and hematopoietic cells is required for the accelerative effect of estradiol on reendothelialization. *Arterioscler Thromb Vasc Biol*. 2009;29(10):1543-1550. doi:10.1161/ATVBAHA.109.192849

35. Umar S, Iorga A, Matori H, Nadadur RD, Li J, Maltese F, van der Laarse A, Eghbali M. Estrogen Rescues Preexisting Severe Pulmonary Hypertension in Rats. *Am J Respir Crit Care Med*. 2011;184(6):715-723. doi:10.1164/rccm.201101-0078OC

36. Wang L, Moonen JR, Cao A, Isobe S, Li CG, Tojais NF, Taylor S, Marciano DP, Chen PI, Gu M, Li D, Harper RL, El-Bizri N, Kim YM, Stankunas K, Rabinovitch M. Dysregulated Smooth Muscle Cell BMPR2-ARRB2 Axis Causes Pulmonary Hypertension. *Circ Res*. 2023;132(5):545-564. doi:10.1161/CIRCRESAHA.121.320541

37. Chen X, Austin ED, Talati M, Fessel JP, Farber-Eger EH, Brittain EL, Hemnes AR, Loyd JE, West J. Estrogen Inhibition Reverses Pulmonary Arterial Hypertension and Associated Metabolic Defects. *Eur Respir J*. 2017;50(2):1602337. doi:10.1183/13993003.02337-2016

38. Wong J, Zhao G, Adams-Tzivelekidis S, Wen H, Chandrasekaran P, Michki SN, Gentile ME, Singh M, Kass-Gergi S, Mendoza M, Holcomb NP, Li X, Tang AT, Negretti NM, Sucre JMS, Frank DB,

Vaughan AE. Dynamic behavior and lineage plasticity of the pulmonary venous endothelium. *Nat Cardiovasc Res*. 2024;3(12):1584-1600. doi:10.1038/s44161-024-00573-2

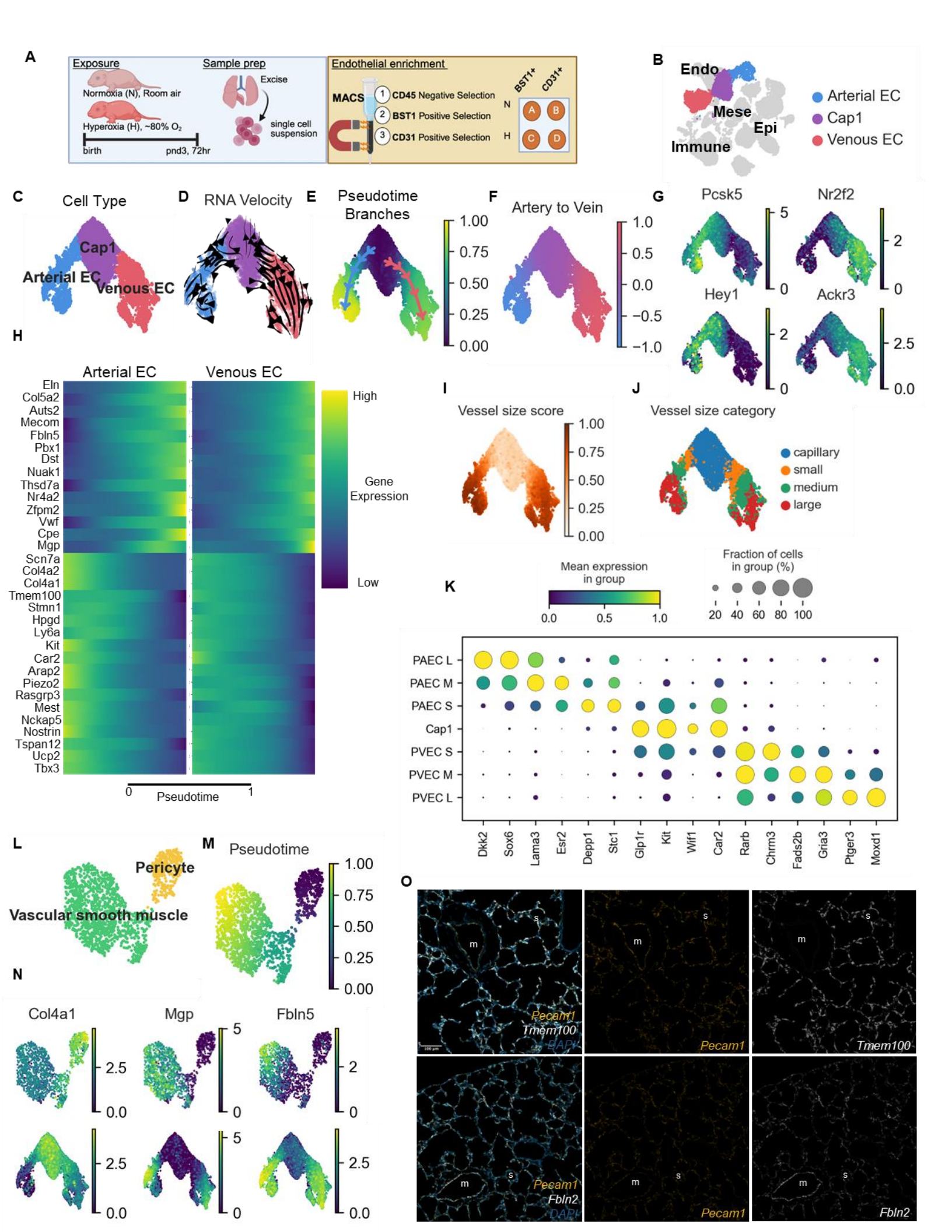
A
Exposure
Normoxia (N), Room air
Hyperoxia (H), ~80% O2
birth
pnd3, 72hr
Sample prep
Excise
single cell suspension
Endothelial enrichment
MACS
1 CD45 Negative Selection
2 BST1 Positive Selection
3 CD31 Positive Selection
BST1+
CD31+
N
H
A
B
C
D
B
Endo
Mese
Epi
Immune
Arterial EC
Cap1
Venous EC
C
Cell Type
Cap1
Arterial EC
Venous EC
D
RNA Velocity
E
Pseudotime Branches
1.00
0.75
0.50
0.25
0.00
F
Artery to Vein
1.0
0.5
0.0
−0.5
−1.0
G
Pcsk5
Nr2f2
Hey1
Ackr3
5
0
2
0
2
0
2.5
0.0
H
Arterial EC
Venous EC
Eln
Col5a2
Auts2
Mecom
Fbln5
Pbx1
Dst
Nuak1
Thsd7a
Nr4a2
Zfpm2
Vwf
Cpe
Mgp
Scn7a
Col4a2
Col4a1
Tmem100
Stmn1
Hpgd
Ly6a
Kit
Car2
Arap2
Piezo2
Rasgrp3
Mest
Nckap5
Nostrin
Tspan12
Ucp2
Tbx3
High
Gene Expression
Low
0
Pseudotime
1
I
Vessel size score
1.00
0.75
0.50
0.25
0.00
J
Vessel size category
capillary
small
medium
large
K
Mean expression in group
0.0
0.5
1.0
Fraction of cells in group (%)
20
40
60
80
100
PAEC L
PAEC M
PAEC S
Cap1
PVEC S
PVEC M
PVEC L
Dkk2
Sox6
Lama3
Esr2
Depp1
Stc1
Glp1r
Kit
Wif1
Car2
Rarb
Chrm3
Fads2b
Gria3
Ptger3
Moxd1
L
Pericyte
Vascular smooth muscle
M
Pseudotime
1.00
0.75
0.50
0.25
0.00
N
Col4a1
Mgp
Fbln5
2.5
0.0
5
0
2
0
2.5
0.0
5
0
2.5
0.0
O
s
m
Pecam1
Tmem100
DAPI
Pecam1
Tmem100
Pecam1
Fbln2
DAPI
Pecam1
Fbln2

Figure 1 scRNAseq data of pulmonary vascular endothelial cells exist on spectrums of size and localization

- **A.** Mice were exposed from birth to room air (21% O2) or hyperoxia (80%). Lung homogenate was enriched for either PVEC (BST1+) or all EC (CD31+).
- **B.** UMAP of all cells showing lineage, black text overlaying location of lineage: Endo:Endothelial, Epi:Epithelial, Imm:Immune, Mese:Mesenchymal. Vascular cell types are shown in color, Arterial (blue), Cap1 (Purple), or Venous (Red) EC
- **C.** UMAP of vascular EC cell types alone
- **D.** UMAP colored by vascualr EC cell type with RNA velocity vectors overlayed. Arrow of vector points in the direction of cellular trajectory
- **E.** UMAP colored by pseudotime (low to high, purple to yellow) with predicted branches overlayed, Arterial (blue), Venous (red).
- **F.** UMAP of arterio-venous zonation based off of pseudotime branches (blue to red, artery to vein)
- **G.** UMAP of arterial and venous zonation markers colored by gene expression (low to high, purple to yellow)
- **H.** Heatmaps of top shared genes correlated with pseudotime in both arterial and venous branches. Psuedotime increases from left to right in each heatmap. Gene expression is displayed by color expression (low to high, purple to yellow)
- **I.** UMAP of cell type scoring based off expression of cells in panel H(low to high, white to orange).
- **J.** UMAP colored by vessel size category, based off of quartiles from panel I
- **K.** Dotplot showing localized gene expression along the vascular hierarchy. Average gene expression is indicated by color (low to high, purple to yellow). Proportion of population expressing the gene is indicated by the size of the circle.
- **L.** UMAP of mural populations colored by celltype
- **M.** UMAP of pseudotime across mural populations (low to high, purple to yellow)
- **N.** UMAPs of select shared gene expression gradients (low to high, purple to yellow) across pseudotime in mural (top row) and vascular endothelial (bottom row) celltypes.
- **O.** Multiplexed in situ hybridization to detect expression of the endothelial marker gene *Pecam1* and either small vessel gene *Tmem100* (white) or *Fbln2* (orange) in lung tissue from mice at P3. Calibration bar, 100 µm

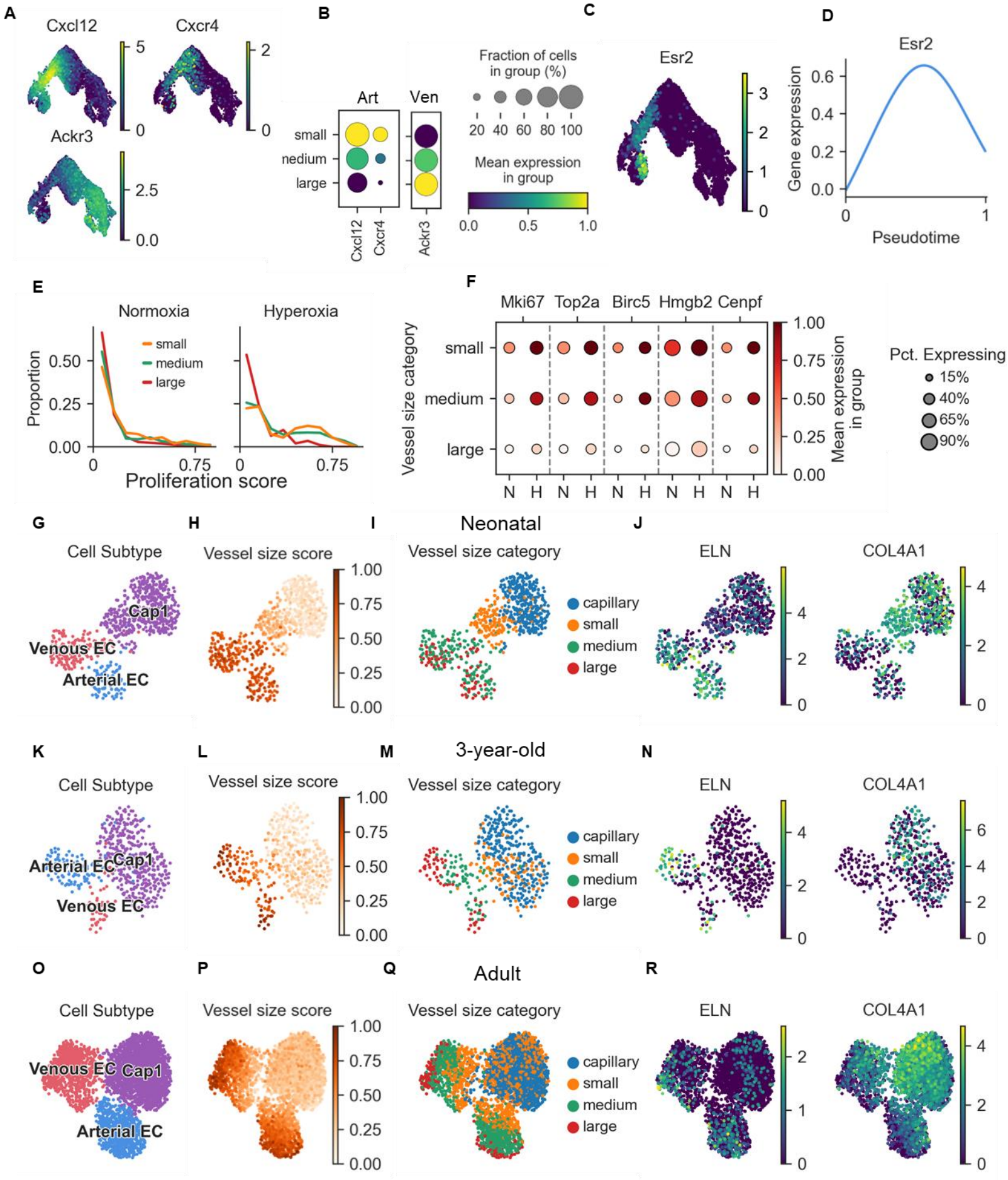
A
Cxcl12
Cxcr4
Ackr3
B
Art
Ven
small
nedium
large
Cxcl12
Cxcr4
Ackr3
Fraction of cells in group (%)
20 40 60 80 100
Mean expression in group
0.0 0.5 1.0
C
Esr2
D
Esr2
Gene expression
Pseudotime
E
Normoxia
Hyperoxia
small
medium
large
Proportion
Proliferation score
F
Mki67 Top2a Birc5 Hmgb2 Cenpf
Vessel size category
small
medium
large
N H N H N H N H N H
Mean expression in group
Pct. Expressing
15%
40%
65%
90%
G
Cell Subtype
Cap1
Venous EC
Arterial EC
H
Vessel size score
I
Neonatal
Vessel size category
capillary
small
medium
large
J
ELN
COL4A1
K
Cell Subtype
Arterial EC
Cap1
Venous EC
L
Vessel size score
M
3-year-old
Vessel size category
capillary
small
medium
large
N
ELN
COL4A1
O
Cell Subtype
Venous EC
Cap1
Arterial EC
P
Vessel size score
Q
Adult
Vessel size category
capillary
small
medium
large
R
ELN
COL4A1

Figure 2 Vascular continuums show reveal patterns across development, disease, and species

- **A.** UMAPs colored by gene expression (low to high, purple to yellow) of the CXCL12-CXCR4-ACKR3 signaling axis
- **B.** Dotplots Cxcl12-Cxcr 4 in artery by vessel size, and Ackr3 in vein by vessel size. Average gene expression is indicated by color (low to high, purple to yellow). Proportion of population expressing the gene is indicated by the size of the circle.
- **C.** UMAPs colored by gene expression of *Esr2* (low to high, purple to yellow)
- **D.** Histogram of *Esr2* gene expression across the arterial pseudotime branch
- **E.** Histogram showing distribution of proliferation score of PVEC by size in normoxia and hyperoxia. small:Orange, medium:Green, large:Red
- **F.** Dotplot of select proliferation marker genes in PVEC by vessel size in normoxia and hyperoxia. Average gene expression is indicated by color (low to high, white to red). Proportion of population expressing the gene is indicated by the size of the circle.
- **G.** UMAP of neonatal human vascular endothelial cells colored by cell type
- **H.** UMAP of neonatal human vascular endothelial cells vessel size scoring (low to high, white to orange).
- **I.** UMAP of neonatal human vascular endothelial cells colored by vessel size category
- **J.** UMAP of neonatal human vascular endothelial cells gene expression (low to high, purple to yellow) of *ELN*, and *COL4A1*
- **K.** UMAP of 3-year-old human vascular endothelial cells colored by cell type
- **L.** UMAP of 3-year-old human vascular endothelial cells vessel size scoring (low to high, white to orange).
- **M.** UMAP of 3-year-old human vascular endothelial cells colored by vessel size category
- **N.** UMAP of 3-year-old human vascular endothelial cells gene expression (low to high, purple to yellow) of *ELN*, and *COL4A1*
- **O.** UMAP of adult human vascular endothelial cells colored by cell type
- **P.** UMAP of adult human vascular endothelial cells vessel size scoring (low to high, white to orange).
- **Q.** UMAP of adult human vascular endothelial cells colored by vessel size category
- **R.** UMAP of adult human vascular endothelial cells gene expression (low to high, purple to yellow) of *ELN*, and *COL4A1*