Article

Capillary cell-type specialization in the alveolus

Astrid Gillich[1,2,3], Fan Zhang[3], Colleen G. Farmer[4], Kyle J. Travaglini[1,2,3], Serena Y. Tan[5], Mingxia Gu[2,3,6], Bin Zhou[7], Jeffrey A. Feinstein[2,3,6], Mark A. Krasnow[1,2,3]✉ & Ross J. Metzger[2,3,6]✉

In the mammalian lung, an apparently homogenous mesh of capillary vessels surrounds each alveolus, forming the vast respiratory surface across which oxygen transfers to the blood[1]. Here we use single-cell analysis to elucidate the cell types, development, renewal and evolution of the alveolar capillary endothelium. We show that alveolar capillaries are mosaics; similar to the epithelium that lines the alveolus, the alveolar endothelium is made up of two intermingled cell types, with complex 'Swiss-cheese'-like morphologies and distinct functions. The first cell type, which we term the 'aerocyte', is specialized for gas exchange and the trafficking of leukocytes, and is unique to the lung. The other cell type, termed gCap ('general' capillary), is specialized to regulate vasomotor tone, and functions as a stem/progenitor cell in capillary homeostasis and repair. The two cell types develop from bipotent progenitors, mature gradually and are affected differently in disease and during ageing. This cell-type specialization is conserved between mouse and human lungs but is not found in alligator or turtle lungs, suggesting it arose during the evolution of the mammalian lung. The discovery of cell type specialization in alveolar capillaries transforms our understanding of the structure, function, regulation and maintenance of the air–blood barrier and gas exchange in health, disease and evolution.

Lungs have evolved complex and diverse architectures that combine a large surface with an exquisitely thin barrier for efficient exchange of oxygen and carbon dioxide between air and blood. In mammalian lungs, gas exchange occurs in tightly packed alveoli, the terminal airspaces of the bronchial tree, which are surrounded by walls that contain a dense network of capillaries (Fig. 1a). The discovery of alveoli and their associated capillaries by Malpighi in the seventeenth century inaugurated what became the study of the structural basis of gas exchange, providing the foundation for modern respiratory physiology and pulmonary medicine[2,3]. Efforts to understand the cellular structure of the barrier have focused mainly on epithelial cells, beginning with the recognition that alveoli are lined by a continuous epithelium composed of intermixed alveolar type 1 (AT1) and AT2 cell types[4]. AT1 cells are large, thin and highly extended cells that comprise 95% of the respiratory surface across which diffusion occurs, whereas cuboidal AT2 cells secrete surfactant that prevents alveolar collapse[1]. Although much progress has been made in understanding the development, maintenance and repair of the alveolar epithelium[5–7], the cells of the alveolar endothelium—the other side of the air–blood barrier—have received less attention.
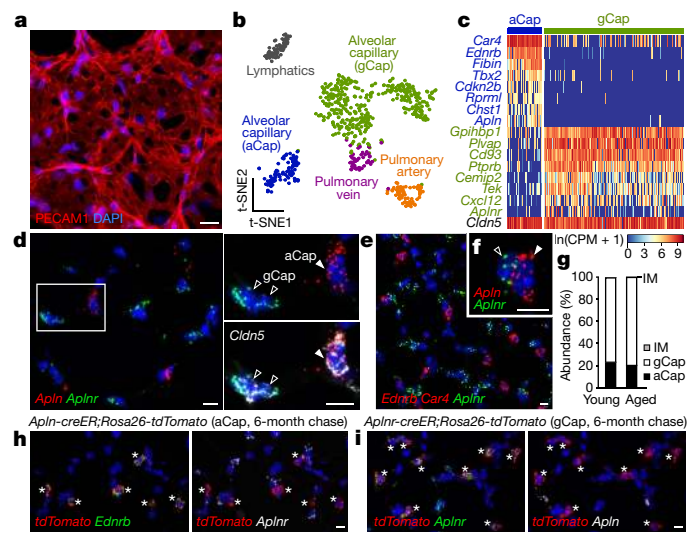
Intermingled alveolar capillary cell types

We systematically defined the cellular diversity of the pulmonary endothelium in the adult mouse lung by single-cell RNA sequencing (scRNA-seq) and mapping, and identified two molecularly distinct populations of capillary cells in the alveolus (Fig. 1b–d, Extended Data Fig. 1a–g, Supplementary Tables 1, 2). The two subsets are interspersed throughout the gas-exchange region, creating apparently random diversity within the alveolar capillary network, and their relative abundance changes little with age (Fig. 1e–g, Extended Data Fig. 1h–j). To test whether the capillary populations are stable cell types or interconverting cell states, we used complementary genetic strategies to permanently label each subset (using either *Apln-creER* or *Aplnr-creER*) and analysed the expression of subset markers in labelled cells after 48 hours, 1 month or 6 months to determine whether the labelled population continues to express the markers, or whether cells turn them off and start to express markers of the other population. We found minimal (0.2% at 6 months) interconversion (Fig. 1h, i, Extended Data Fig. 1k–n), indicating that the populations are not transient cell states. We conclude that the alveolar capillary network is composed of two intermingled, stable cell types, which we call gCap (general capillary cells) and aCap (aerocytes; see below).

Aerocytes are specialized for gas exchange

We used sparse cell labelling and deep imaging to visualize individual capillary cells in three dimensions. aCap cells are complex, large cells (spanning more than 100 μm; 21 × 10³ μm³ mean volume) with ramified

Fig. 1 | Two stable, intermingled alveolar capillary cell types. a, Alveolar capillaries in adult mouse lung immunostained for PECAM1. **b**, t-distributed stochastic neighbour embedding (t-SNE) plot of endothelial cell populations annotated in scRNA-seq data for adult mouse lung[13]. **c**, Heat map of expression of capillary subset markers and the general endothelial marker *Cldn5* in individual capillary cells. CPM, counts per million. **d–f**, Single-molecule fluorescent in situ hybridization (smFISH) for the capillary subset markers *Apln* (**d, f**) or *Ednrb* and *Car4* (aCap) (**e**), and *Aplnr* (gCap) (**d–f**), in adult mouse lung. Images in **d** (right) and **f** show individual aCap and gCap cells. **g**, Relative abundance of aCap cells, gCap cells and cells that co-express aCap and gCap markers (intermediate (IM) cells) in lungs from 3-month-old (young) and 24-month-old (aged) mice (data shown as mean; *n* = 500 cells scored per mouse; 2 mice per group). **h, i**, Co-expression of *tdTomato* lineage label (asterisks) and aCap marker *Ednrb* but not gCap marker *Aplnr* (**h**), or gCap marker *Aplnr* but not aCap marker *Apln* (**i**), in lungs collected six months after mature aCap (**h**) or gCap (**i**) cells were lineage-labelled. Blue, DAPI. Scale bars, 10 μm.

extensions that surround pores (mean of 6 pores per cell, range 2–9), giving cells the appearance of Swiss cheese (Fig. 2a, c, Extended Data Fig. 2f–h, Supplementary Video 1). The cells have a variety of sizes and shapes, and a single cell frequently spans multiple alveoli. Morphological complexity of this kind has also been described for AT1 cells[1].

gCap cells have a related but less extreme morphology. They are smaller (spanning less than 40 μm; 4×10^3 μm^3 mean volume), have fewer pores (mean of 3 pores per cell, range 1–6) and are less extensively branched, rarely spanning multiple alveoli (Fig. 2b, c, Extended Data Fig. 2e, g, h, Supplementary Video 2). The two cell types fit together to form multicellular tubes (Fig. 2d, e, Supplementary Video 3). The mean surface area of aCap cells is four to five times greater than that of gCap cells (Extended Data Fig. 2g), but they are fourfold less abundant (Fig. 1g), hence each contributes about half of the total capillary surface area. The morphologies of both types—especially aCap cells—are distinct from the morphologies of capillary cells elsewhere in the lung, within the bronchial circulation, and in other organs (Extended Data Fig. 3a–g), reflecting the unique architecture and function of the pulmonary circulation. Comparison of their molecular diversity (Extended Data Fig. 3h–l) suggests that capillary cells in other organs are more similar to gCap cells (supporting the name 'general' capillary), whereas aCap cells are unique to the lung.

Capillaries are asymmetrically positioned within alveolar walls such that only some of the endothelium is tightly apposed to squamous AT1 cells to form thin regions of the gas-exchange surface in which the barrier to diffusion is minimized, whereas other ('thick') regions are separated from the epithelium by stromal cells and connective tissue[3] (Fig. 2f). To look for differences in the localization of the cell types in these structurally distinct regions, we performed immuno-electron
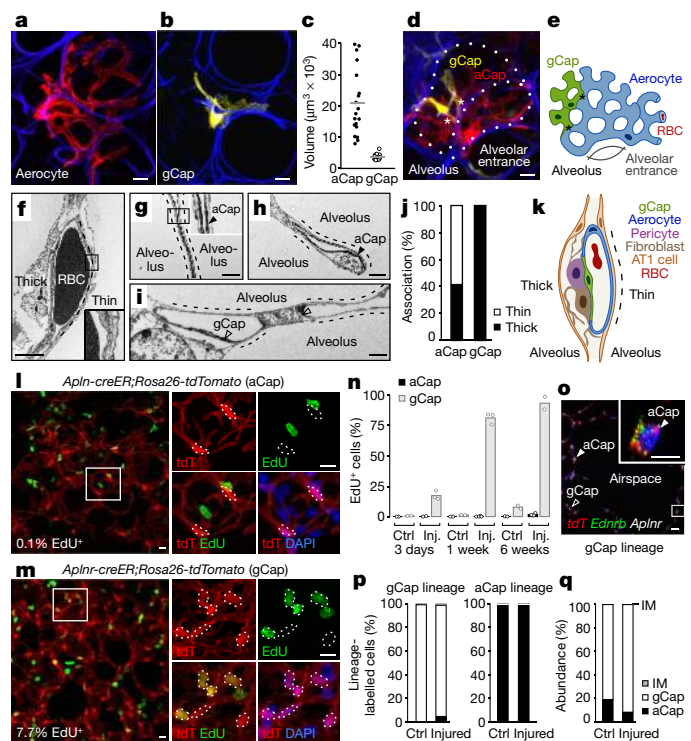
Fig. 2 | Specialized alveolar capillary cell types in gas exchange and capillary renewal. a, b, Single aCap (aerocyte) (**a**) or gCap (**b**) cells in adult *Apln-creER; Rosa26-Confetti* (**a**) or *Aplnr-creER; Rosa26-Confetti* (**b**) lungs. **c**, Quantification of individual cell volumes. Bar indicates mean (19 aCap and 17 gCap cells scored from *n* = 2 mice). **d**, aCap and gCap cells in adult *Cdh5-creER; Rosa26-Confetti* lung form multicellular tubes (asterisks) within capillaries surrounding a single alveolus (dotted outline). Blue, elastin fibres (**a, b, d**). **e**, Schematic of alveolar capillary network. Asterisks, multicellular tubes. RBC, red blood cell. **f–i**, Transmission electron micrographs of adult mouse alveolar walls. **f**, Thick and thin regions of the air–blood barrier. **g–i**, *Apln-creER; Rosa26-tdTomato* (**g, h**) or *Aplnr-creER; Rosa26-tdTomato* (**i**) lungs immunostained for tdTomato (heavy black stain). Labelled aerocytes (**g, h**) but not gCap cells (**i**) are associated with thin regions (dashed lines). **j**, Quantification of the percentage of each labelled cell type associated with thick or thin regions (*n* = 2 mice of each genotype; 21 labelled aCap cells and 24 labelled gCap cells scored). **k**, Schematic representation of the air–blood barrier. **l, m**, Analysis of the proliferation of lineage-labelled aCap (**l**) or gCap (**m**) cells during adult homeostasis, detected by cumulative EdU incorporation for six weeks. The mean percentage of EdU$^+$ aCap cells (**l**) or gCap cells (**m**) is shown at the bottom left (*n* = 400–4,000 cells scored per lung in *n* = 2 mice of each genotype). tdT, tdTomato. **n**, Quantification of the fraction of EdU$^+$ lineage-labelled aCap or gCap cells during the indicated intervals after elastase-induced injury (Inj.) or mock injury with saline as control (Ctrl) (data shown as mean; *n* = 200–1,600 cells scored per lung in 2–4 mice of each genotype per time point and treatment group; see Methods for exact sample sizes). **o**, smFISH for lineage label (*tdTomato*), aCap (*Ednrb*) and gCap (*Aplnr*) markers in gCap-lineage-labelled lung six weeks after elastase injury. Blue, DAPI (**l, m, o**). **p**, Quantification of the fraction of lineage-labelled gCap (left) or aCap (right) cells expressing aCap (*Ednrb*), gCap (*Aplnr* or *Ptprb*) or aCap and gCap markers (IM) six weeks after elastase injury in injured and uninjured (Ctrl) regions (*n* = 500–1,000 cells scored per region; 3 injured and uninjured regions scored in *n* = 2 mice of each genotype). **q**, Relative abundance of capillary cell types in injured and uninjured (Ctrl) regions, six weeks after elastase administration (*n* = 800–5,600 cells scored per region; 3 injured and uninjured regions; *n* = 3 mice). Scale bars, 10 μm (**a, b, d, l, m, o**); 2 μm (**f–i**).

microscopy on lungs in which aCap and gCap cells were separately labelled. We found that thin regions are composed entirely of aCap cells, whereas gCap cells are positioned in contact with stromal cells in thick regions (Fig. 2g–j, Extended Data Fig. 4a–c). Because of their

close association with AT1 cells within thin regions of the respiratory surface (Fig. 2k) and their expansive morphology, which reflects a specialized role in gas exchange analogous to AT1 cells, we term aCap cells 'aerocytes'.

gCap cells are capillary stem cells

Little is known about how alveolar capillaries are maintained throughout life and repaired after alveolar damage[7]. To examine the behaviours of the capillary cell types in alveolar homeostasis, we first analysed proliferation by cumulative labelling with 5-ethynyl-2′-deoxyuridine (EdU) for six weeks in mice in which either gCap cells or aerocytes were genetically labelled. Capillary cell turnover was slow[8] but, notably, proliferation was almost entirely restricted to gCap cells (7.7% EdU[+] gCap cells; Fig. 2l, m). We detected extremely rare, solitary EdU[+] aerocytes (2 of 4,401 cells), whereas EdU[+] gCap cells were present as clusters of up to 10 cells, indicative of focal proliferation.

Acute lung injury can induce the proliferation of alveolar capillary cells[8,9]. To investigate the role of the cell types in capillary repair, we used a mouse model of emphysema[10] (Extended Data Fig. 5a, b) in which elastase-induced alveolar damage is accompanied, we found, by robust capillary cell proliferation. Lineage-labelled gCap cells proliferated as early as day 3 after elastase instillation (17% EdU[+]), with almost all (93%) of the gCap cells in injured regions being EdU[+] at 6 weeks (Fig. 2n, Extended Data Fig. 5c–f). Aerocytes rarely proliferated even after injury (0.2% EdU[+] at 3 days, 1.7% at 6 weeks).
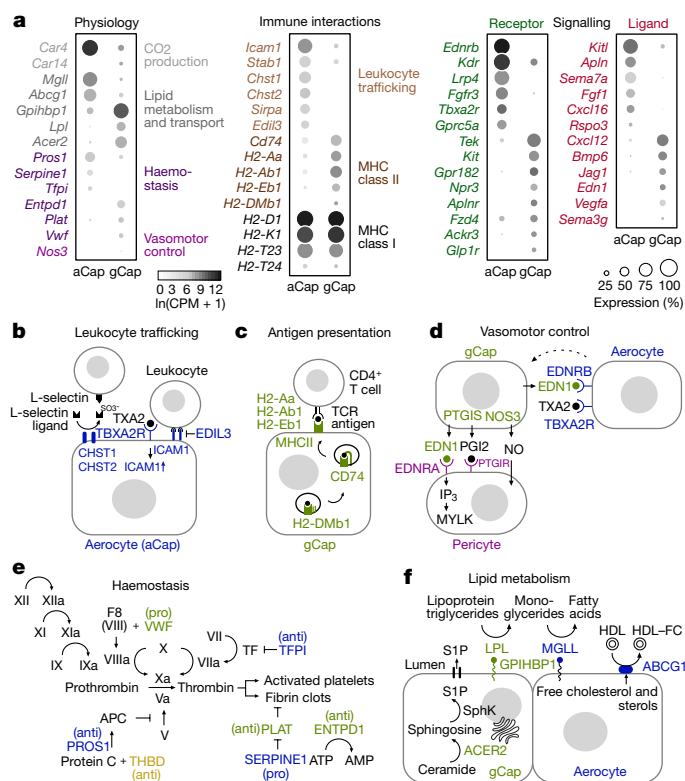
We examined the fate of lineage-labelled gCap cells after injury and found labelled aerocytes as well as gCap cells, demonstrating that aerocytes are generated from gCap cells during repair (Fig. 2o, p). We also detected rare, lineage-labelled aerocytes in the absence of injury, after extended chases (3.4% lineage-labelled aCap cells at 14 months; Extended Data Fig. 1n)—indicating that aerocytes are generated intermittently from gCap cells during homeostasis. We conclude that gCap cells function as specialized stem/progenitor cells that replenish the alveolar capillary endothelium during maintenance and repair.

Even after six weeks of recovery from injury, the cellular composition of the alveolar capillary network is altered (Fig. 2q), suggesting that repair is abnormal or incomplete at this stage. Aberrant or insufficient repair may underlie vascular changes in lung diseases such as emphysema and interstitial lung disease, as well as respiratory distress syndromes that accompany severe injury or virus-induced alveolar damage—as in coronavirus disease 2019 (COVID-19)[11,12]. Understanding the behaviour of gCap cells, and the signals that activate their proliferation and reprogramming, may offer a strategy to restore the normal pattern.

Molecular functions of capillary cell types

We used scRNA-seq profiles[13] to discover common and additional type-specific functions of capillary cells. We found only a small number of genes (*Scn7a*, *Mapt*) that were expressed by all (or most), and only, alveolar capillary cells, suggesting that few if any molecular functions are carried out by both capillary cell types but not by other lung endothelial cells (Extended Data Fig. 1a, b). By contrast, we identified many genes with roles in physiology, immune interactions and signalling, the expression of which differed between the cell types, revealing further specialization (Fig. 3a, Extended Data Fig. 4d, Supplementary Table 2).

Some functions appear to be unique to one cell type. Aerocytes are the likely site of leukocyte trafficking—which is primarily a capillary function in the lung[14]—as they specifically express adhesion and leukocyte-sequestration genes (Fig. 3a, b). gCap cells, in contrast, express genes that encode MHC class II components, suggesting that they present antigens (Fig. 3a, c). gCap cells may also have a specialized role in vasomotor control (see below; Fig. 3a, d). Other functions appear to be distributed across both cell types. The two cell types produce
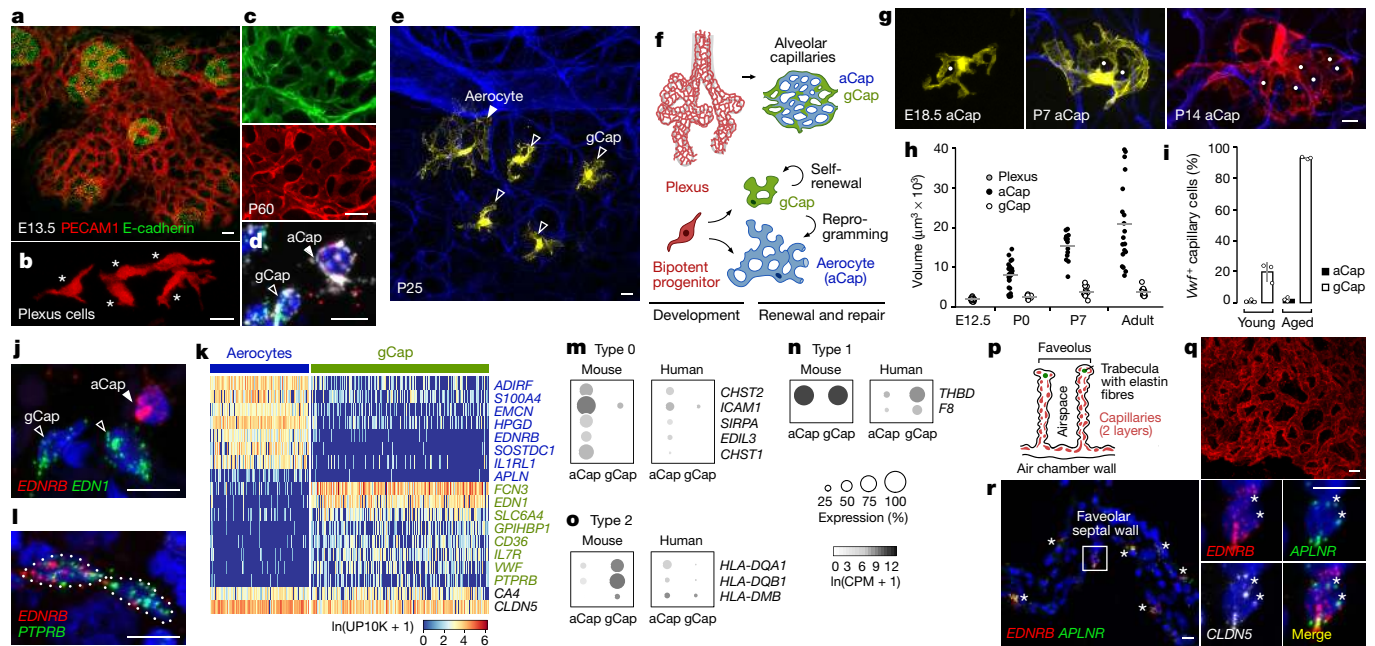
distinct pro- and anticoagulants, suggesting that they have different roles in haemostasis (Fig. 3a, e), and they may cooperate in lipid metabolism, forming an 'assembly line' that produces fatty acids (Fig. 3a, f).

Our analysis also revealed that aerocytes and gCap cells can signal to one another. Aerocytes are a source of ligands (for example, apelin (encoded by *Apln*), kit ligand (*Kitl*)) that signal through cognate receptors (*Aplnr*, *Kit*) that are displayed by gCap cells; conversely, gCap cells produce ligands (for example, endothelin 1 (*Edn1*), vascular endothelial growth factor A (*Vegfa*)) with cognate receptors (*Ednrb*, *Kdr*) on aerocytes (Fig. 3a, Extended Data Fig. 4d, e). Such bidirectional signalling indicates that the two cell types can regulate each other.

We also identified specialized signalling interactions with other alveolar cell types. Aerocytes express *Ednrb* and *Kdr*, suggesting that they interact with AT1 cells (Fig. 3a, Extended Data Fig. 4d, e). gCap cells express a vasoconstrictor (*Edn1*) that can signal to endothelin receptor type A (*Ednra*) on pericytes; they also express endothelial nitric oxide synthase (*Nos3*) and prostaglandin I2 synthase (*Ptgis*), making them a source of vasodilators (Fig. 3a, d, Extended Data Fig. 4d, e).

Fig. 3 | Molecular functions of alveolar capillary cell types. a, Dot plots showing the expression of selected differentially expressed genes in mouse aerocytes and gCap cells. **b**–**f**, Diagrams of proposed specialized functions of alveolar capillary cell types. Aerocyte genes, blue; gCap genes, green; aerocyte and gCap genes, yellow; pericyte genes[16], purple. **b**, Leukocyte trafficking. **c**, Antigen presentation. TCR, T cell receptor. **d**, Vasomotor control. gCap cells express EDN1, eNOS (encoded by *Nos3*) and PTGIS, making them a unique source of vasomodulators. EDN1 can signal to EDNRA (expressed on pericytes) or to EDNRB (expressed on aerocytes), which may feed back to gCap cells (dashed arrow) to regulate vasodilator production. IP$_3$, inositol trisphosphate; NO, nitric oxide; PGI2, prostaglandin I2. **e**, Haemostasis. Roman numerals indicate coagulation factors. TF, tissue factor. APC, activated protein C. Pro, procoagulants; anti, anticoagulants. **f**, Lipid metabolism. Lipoprotein lipase (LPL), anchored to the lumen of gCap cells by GPIHBP1, converts circulating lipoprotein triglycerides to monoglycerides, which are broken down to free fatty acids by monoacylglycerol lipase (MGLL) expressed by aerocytes. HDL, high-density lipoprotein; HDL–FC, HDL bound to free cholesterol; S1P, sphingosine 1-phosphate; SphK, sphingosine kinase.

Fig. 4 | Development and evolution of specialized alveolar capillary cell types. a, Plexus surrounding airways in immunostained E13.5 mouse lung. **b**, Sparsely labelled plexus cells in E12.5 *Aplnr-creER;Rosa26-Confetti* lung immunostained for red fluorescent protein (RFP). **c**, Near-complete labelling of alveolar capillaries in postnatal day (P)60 *Aplnr-creER;Rosa26-tdTomato* lungs from mice that received tamoxifen at E12.5. Endomucin, green; tdTomato, red. **d**, smFISH to detect *tdTomato* (white) in aCap (*Apln*, red) and gCap (*Aplnr*, green) cells in P60 *Aplnr-creER;Rosa26-tdTomato* lung that was lineage-labelled at E12.5. **e**, Clone in P25 *Aplnr-creER;Rosa26-Confetti* lung composed of aCap and gCap cells derived from a single yellow fluorescent protein (YFP)-expressing plexus cell that was labelled at E14.5. **f**, Schematics depicting the origin of both alveolar capillary cell types from single bipotent cells in the embryonic plexus. **g**, Individual aerocytes in *Apln-creER;Rosa26-Confetti* lungs at the indicated stages. Dots, pores. **h**, Quantification of individual cell volumes at the indicated stages (*n* > 10 cells scored for each cell type at each time point from *n* = 2 mice; see Methods for exact cell number; adult aCap and gCap cells from Fig. 2c). Bar, mean value. **i**, Quantification of *Vwf*-expressing capillary cells in lungs from 3-month-old (young) and 24-month old (aged) mice (mean ± s.d.; *n* > 500 cells scored per mouse; 3 mice per age group). **j**, smFISH for capillary cell-type markers in alveolar capillary cells in lung tissue from a 75-year-old man. **k**, Heat map of expression of cell-type markers in individual human capillary cells[16]. UP10K, unique molecular identifiers per ten thousand. **l**, Human adenocarcinoma vessel, containing cells that co-express *EDNRB* (aCap marker) and *PTPRB* (gCap marker). See also Extended Data Fig. 9a. **m–o**, Dot plots showing expression in aCap and gCap cells for selected conserved genes (type 0, leukocyte trafficking; **m**); genes with species-specific specialized expression (type 1, haemostasis; **n**), or genes that switch cell type between species (mouse[13], human[16]) (type 2, antigen presentation; **o**). See Extended Data Fig. 10c–e. **p**, Schematic depicting alligator lung faveolus. **q**, Faveolar capillary network in alligator lung immunostained for claudin 5 (CLDN5). **r**, Co-expression (asterisks) of the mammalian alveolar capillary cell-type markers *EDNRB* (aCap) and *APLNR* (gCap) in faveolar capillary cells (*CLDN5*, white) in alligator lung. Blue, DAPI (**d**, **j**, **l**, **r**) or elastin fibres (**e**, **g**). Scale bars, 10 μm.

This indicates that gCap cells regulate vasomotor tone through interactions with pericytes. These specialized signalling relationships reflect distinct associations of the capillary cell types with surrounding cells, revealing functional compartmentalization within the alveolus (Extended Data Fig. 4f).

Development and ageing of alveolar capillaries

To determine when the capillary cell types emerge during development, we investigated their origin in the embryonic lung, in which a dense vascular plexus surrounds branching airways (Fig. 4a). Using lineage tracing, we found that this plexus—which is composed of small (2×10^3 μm³ mean volume), simple, proliferating endothelial cells—gives rise to both subsets of alveolar capillary cells (Fig. 4b–d, Extended Data Fig. 6a–c, Supplementary Data 1, 2). The near-complete labelling of the capillary network suggests that the plexus is the major or sole source of aerocytes and gCap cells. To determine whether individual plexus cells can give rise to both capillary cell types, we performed a clonal analysis. Clones contained both aerocytes and gCap cells, demonstrating that plexus cells are bipotent (Fig. 4e, f, Extended Data Fig. 6d–f, Supplementary Video 4).

Aerocytes first emerge at embryonic day (E)17.5 and begin to acquire their Swiss-cheese-like morphology during embryonic development (Fig. 4g, h, Extended Data Figs. 2, 6g, h). This is consistent with a recent study that identified emerging aerocytes and characterized their gene expression and morphology[15]. The emergence of capillary cell types is gradual and asynchronous, and both cell types continue to mature molecularly and morphologically after birth (Fig. 4g, h, Extended Data Figs. 2, 6h–j, 7, Supplementary Data 3, 4). These results reveal a remarkable transformation of the plexus into the alveolar capillary network, beginning in the embryonic lung.

Endothelial cell phenotypes change with age, and may contribute to age-related disease. We detected a widespread induction of von Willebrand factor (encoded by *Vwf*), which is considered a marker of endothelial dysfunction, in the lungs of aged mice. *Vwf* is induced specifically in gCap cells, but not in aerocytes (Fig. 4i, Extended Data Fig. 6k–m), indicating that the cell types are differentially regulated during ageing.

Human alveolar capillary cell types

We identified both cell types intermingled within human alveolar capillary networks (Fig. 4j, k, Extended Data Fig. 8a–i, Supplementary Video 5), indicating these cell types have been conserved in mammalian evolution[16]. Similarly to mice, aerocytes emerge during embryonic development in humans (Extended Data Fig. 8j).

The mosaic pattern of capillary cells, however, is lost or altered in lung tumours. In human adenocarcinoma vessels, we observed abundant intermediate cells that co-express markers of both cell types (Fig. 4l, Extended Data Fig. 9a). Cell composition is also altered in mouse adenomas; tumour vessels are composed of gCap cells and intermediate cells, with few or no aerocytes (Extended Data Fig. 9b–f).

As in mice, we identified genes with key roles in physiology, immune interactions and signalling that exhibited specialized expression in the human cell types (Extended Data Fig. 10a, b, Supplementary Table 3). Many genes with cell-type specificity in mice are also expressed by the corresponding human cell type[16] (Fig. 4m, Extended Data Fig. 10c). However, we also identified mouse–human differences. Some genes show specialized expression in only one species, including genes involved in functions that are distributed between the cell types (Fig. 4n, Extended Data Fig. 10d). For other genes, the cell type with specialized expression switches between mouse and human, presumably reflecting species-specific functional differences (Fig. 4o, Extended Data Fig. 10e).

Our analysis suggests that some cell-type-specific functions are conserved between mice and humans. For example, specialized leukocyte-trafficking genes are restricted to aerocytes, and gCap cells may regulate vasomotor tone in both species (Fig. 4m, Extended Data Fig. 10a–c). But capillary cell types can also gain (or lose) functions: in mice, gCap cells preferentially express MHC class II genes, whereas in humans, these genes are expressed by aerocytes (Fig. 4o, Extended Data Fig. 10a, b, e).

Evolution of capillary cell specialization

To investigate the evolutionary origins of the cell types, we examined capillary cell diversity in lungs from the American alligator (*Alligator mississippiensis*) and the western painted turtle (*Chrysemys picta bellii*)—reptiles from distinct phylogenetic groups (Extended Data Fig. 11a). Alligator and turtle respiratory systems are, in many ways, representative of the ancestral amniote condition[17], and gas exchange occurs across a thick air–blood barrier in faveoli, which are surrounded by capillary nets that resemble those of mouse and human alveoli (Fig. 4p, q, Extended Data Fig. 11b–d, g–k, m, Supplementary Video 6). We found that in each species, lung capillary cells express markers of both mammalian cell types (Fig. 4r, Extended Data Fig. 11e, f, l). However, in contrast to alveolar capillary cells, alligator and turtle faveolar capillary cells co-express these cell-type markers, suggesting that they may lack the cell specialization of mammalian lungs.

Discussion

Here we show that the alveolar capillary endothelium, like the alveolar epithelium, is composed of two intermingled cell types. Such cell-type specialization may have evolved to optimize gas exchange within the complex environment of the alveolus, balancing and integrating structure and function. Aerocytes and AT1 cells are both large, complex cells that are tightly apposed in the thinnest regions of the gas-exchange surface—separated only by a shared, compositionally unique basement membrane[18]—which facilitates diffusion. This specialized interface may be particularly important in lung injury. In pulmonary oedema, seen early in acute respiratory distress syndrome (ARDS), fluid accumulates in thick regions, initially protecting thin regions and preserving gas exchange[19,20]. The alveolar capillary cell types arise—like the epithelium—from bipotent progenitors, through distinct maturation programs[6]. Aerocytes first emerge as AT1 differentiation begins, highlighting coordination between the two cell types critical for gas exchange[15]. During adult life, the alveolar endothelium is maintained and repaired by gCap cells, which—like AT2 cells—are 'bifunctional' stem/progenitor cells[5,6] that also serve physiological functions. Separating the progenitor function from aerocytes and AT1 cells may be another mechanism for preserving the gas-exchange surface. Capillary changes could underlie pathologies in the lung (Fig. 4i, l, Extended Data Fig. 9) and other organs, making it essential to now explore and map the full heterogeneity of capillary cell types, states and specializations in health, ageing and disease; to identify changes in capillary composition, which cell types change and how they change; and to investigate the consequences of capillary changes for organ function.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2822-7.

1. Hsia, C. C., Hyde, D. M. & Weibel, E. R. Lung structure and the intrinsic challenges of gas exchange. *Compr. Physiol.* **6**, 827–895 (2016).
2. Malpighi, M. Dissertationes Epistolicæ de Pulmonibus. In *Opera Omnia* 320–332 (Pieter van der Aa, 1687) Available at https://www.biodiversitylibrary.org/bibliography/566#.
3. Weibel, E. R. Morphological basis of alveolar-capillary gas exchange. *Physiol. Rev.* **53**, 419–495 (1973).
4. Bertalanffy, F. D. & Leblond, C. P. Structure of respiratory tissue. *Lancet* **266**, 1365–1368 (1955).
5. Barkauskas, C. E. et al. Type 2 alveolar cells are stem cells in adult lung. *J. Clin. Invest.* **123**, 3025–3036 (2013).
6. Desai, T. J., Brownfield, D. G. & Krasnow, M. A. Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* **507**, 190–194 (2014).
7. Basil, M. C. et al. The cellular and physiological basis for lung repair and regeneration: past, present, and future. *Cell Stem Cell* **26**, 482–502 (2020).
8. Liu, Q. et al. c-kit+ cells adopt vascular endothelial but not epithelial cell fates during lung maintenance and repair. *Nat. Med.* **21**, 866–868 (2015).
9. Niethamer, T. K. et al. Defining the role of pulmonary endothelial cell heterogeneity in the response to acute lung injury. *eLife* **9**, e53072 (2020).
10. Hamakawa, H. et al. Structure–function relations in an elastase-induced mouse model of emphysema. *Am. J. Respir. Cell Mol. Biol.* **45**, 517–524 (2011).
11. Golden, A. & Bronk, T. T. Diffuse interstitial fibrosis of lungs; a form of diffuse interstitial angiosis and reticulosis of the lungs. *AMA Arch. Intern. Med.* **92**, 606–614 (1953).
12. Ackermann, M. et al. Pulmonary vascular endothelialitis, thrombosis, and angiogenesis in Covid-19. *N. Engl. J. Med.* **383**, 120–128 (2020).
13. The Tabula Muris Consortium. Single-cell transcriptomics of 20 mouse organs creates a *Tabula Muris*. *Nature* **562**, 367–372 (2018).
14. Lien, D. C. et al. Physiological neutrophil sequestration in the lung: visual evidence for localization in capillaries. *J. Appl. Physiol.* **62**, 1236–1243 (1987).
15. Vila Ellis, L. et al. Epithelial Vegfa specifies a distinct endothelial population in the mouse lung. *Dev. Cell* **52**, 617–630 (2020).
16. Travaglini, K. J. et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* (in press).
17. Lambertz, M., Grommes, K., Kohlsdorf, T. & Perry, S. F. Lungs of the first amniotes: why simple if they can be complex? *Biol. Lett.* **11**, 20140848 (2015).
18. Vaccaro, C. A. & Brody, J. S. Structural features of alveolar wall basement membrane in the adult rat lung. *J. Cell Biol.* **91**, 427–437 (1981).
19. Bachofen, M. & Weibel, E. R. Structural alterations of lung parenchyma in the adult respiratory distress syndrome. *Clin. Chest Med.* **3**, 35–56 (1982).
20. Szidon, J. P., Pietra, G. G. & Fishman, A. P. The alveolar-capillary membrane and pulmonary edema. *N. Engl. J. Med.* **286**, 1200–1204 (1972).

Article

Methods

Mice

The following mouse strains were used: C57BL/6 (C57BL/6NCrl, Charles River Laboratories, strain 027) was the wild-type strain. *Apln-creER* (*Apln^tm1.1(cre/ERT2)Bzsh*)[21] (provided by B. Zhou), *Aplnr-creER (Tg(Aplnr-cre/ERT2)#Krh*)[22]) (provided by K. Red-Horse), *Cdh5-creER* (*Tg(Cdh5-cre/ERT2)1Rha*)[23] (provided by R. Adams) and *Sftpc-creER* (*Sftpc^tm1(cre/ERT2,rtTA)Hap*)[24] (provided by H. Chapman) were used for conditional expression of Cre recombinase. *Rosa26-tdTomato* (*Gt(ROSA)26Sor^tm14(CAG-tdTomato)Hze*)[25] (The Jackson Laboratory, strain 007914), which expresses cytoplasmic tdTomato after recombination, and *Rosa26-Confetti (Gt(ROSA)26Sor^tm1(CAG-Brainbow2.1)Cle*)[26] (The Jackson Laboratory, strain 017492), which expresses membrane targeted Cerulean CFP, nuclear GFP, cytoplasmic EYFP or cytoplasmic RFP after recombination, were used as Cre reporters. *Kras^LSL-G12D* (*Kras^tm4Tyj/J*)[27] (The Jackson Laboratory, strain 008179) was used to express a constitutively active form of KRAS from the endogenous locus after Cre-mediated recombination. All experimental mice and embryos were heterozygous (or hemizygous) for indicated alleles. Only female mice and embryos were used for experiments with *Apelin-creER*, as *Apelin* is X-linked[21]. Noon of the day a vaginal plug was detected was considered as E0.5. The day a litter was born was considered as P0. For induction of Cre recombinase activity, tamoxifen (Sigma, T5648) was dissolved in corn oil and administered by intraperitoneal (i.p.) injection unless otherwise noted. Adult lungs were perfused, inflated with 2% low melting point agarose (Invitrogen), and collected as previously described[6]. Heart, brain, small intestine, thyroid and kidney were collected after perfusion of the left heart with Ca^{2+}- and Mg^{2+}-free phosphate-buffered saline, pH 7.4 (PBS; Gibco). Postnatal (P7) retinas[28] and embryonic lungs[29] were collected and prepared as previously described.

Tamoxifen dose, administration and tissue collection schedules for individual experiments (unless noted elsewhere) were as follows:

For cell-type stability pulse-chase experiments (presented in Fig. 1h, i, Extended Data Fig. 1k–n): 4 mg tamoxifen to adults, lungs collected after 48 h, 1, 6 or 14 months.

For cell morphology (sparse labelling) experiments: 2 mg (Fig. 2a, b, Extended Data Fig. 2e, f), 0.5 mg (Extended Data Fig. 3a) or 1 mg (Fig. 2d, Extended Data Fig. 3b–f) tamoxifen to adults, collected after 5–7 days; 2 mg (Fig. 4b) or 0.5 mg (Extended Data Fig. 2a) tamoxifen to pregnant dams at E11.5, collected at E12.5, 2 mg tamoxifen at E17.5, collected at E18.5 (Fig. 4g, Extended Data Fig. 2b), 0.5 mg tamoxifen at E18.5, collected at P0 (Extended Data Fig. 2c) or 3 mg tamoxifen at E18.5, collected at P0 (Extended Data Fig. 2d); 0.2 mg tamoxifen at P5 by intragastric injection, collected at P7 (Extended Data Fig. 3g) or 0.05 mg tamoxifen by intragastric injection at P5, collected at P7 or P14 (Fig. 4g).

For lineage-tracing experiments: 4 mg tamoxifen administered to pregnant dams; lungs collected from progeny at P21 or P60 (Fig. 4c, d, Extended Data Fig. 6c, Supplementary Data 1); 0.5 mg tamoxifen administered at P7, collected at P21 (Extended Data Fig. 6f).

For maximal labelling experiments: two or three 4 mg tamoxifen doses (administered 48 h apart) to *Apln-creER; Rosa26-tdTomato* (Fig. 2g, h, Extended Data Fig. 4a, b) or *Aplnr-creER; Rosa26-tdTomato* (Fig. 2i, Extended Data Fig. 4a, c) adult mice, collected 5–14 days after the first dose.

To induce adenoma formation: 4 mg tamoxifen to *Sftpc-creER; Kras^LSL-G12D/+* adult mice (Extended Data Fig. 9b, c, f), collected three weeks later.

Mice were housed and bred in the animal facility at Stanford University in accordance with Institutional Animal Care and Use Committee guidance, and were maintained on a 12-h light–dark cycle with food and water provided ad libitum. Adult mice were 2–6 months old, unless otherwise noted. All mouse experiments were approved by the Stanford University Institutional Animal Care and Use Committee.

Human tissue

De-identified healthy human adult lung tissue from 69- and 75-year-old men and a 66-year-old woman was obtained from the Stanford Tissue Bank. De-identified aborted human fetal lung tissue (17 and 23 weeks) was obtained in collaboration with the Stanford Family Planning Research Team, Department of Obstetrics and Gynecology, Division of Family Planning Services and Research, Stanford University School of Medicine. De-identified human tissue representing well-differentiated invasive lung adenocarcinoma (confirmed by pathological evaluation by S.Y.T.) from a 41-year-old woman was obtained from archival diagnostic material in collaboration with the Stanford Department of Pathology, Stanford University School of Medicine. Tissue collection and use in research were approved by the Stanford Institutional Review Board.

Alligators

Lungs were collected from juvenile (body mass, 1.3 kg) and adult (body mass, 14 kg) American alligators (*A. mississippiensis*, Daudin; male), acquired from the Rockefeller Wildlife Refuge. Lungs were inflated with sterile PBS or 10% formalin for smFISH, or 4% paraformaldehyde (PFA; Electron Microscopy Sciences (EMS)) in PBS for immunostaining. Experiments were approved by the University of Utah Animal Care and Use Committee.

Turtles

Lungs were collected from two adult (body mass, 262 g and 281 g) western painted turtles (*C. p. bellii*; male; The Turtle Source). Lungs were inflated with 10% formalin for smFISH, or 4% PFA in PBS for immunostaining. Experiments were approved by the University of Utah Institutional Animal Care and Use Committee.

Immunostaining

Immunostaining was performed using previously published protocols[6,29] with modifications for adult mouse, human, alligator and turtle tissues as described below. Adult mouse, alligator and turtle lungs and human lung tissue pieces were fixed in 4% PFA in PBS at 4 °C for 2–3 h, then dehydrated through a PBS and methanol series into 100% methanol and stored at −20 °C. Immediately before sectioning, tissue was rehydrated through a methanol and PBT (PBT: 0.1% Tween-20 in PBS) series into PBT. Sections (350 μm) were cut from adult mouse lung lobes on a vibratome (Leica Biosystems). Alligator and turtle lungs, and human lung pieces, were manually cut with a platinum coated double-edge razor blade (EMS) into rough sections 0.5–3 mm thick. Sections were incubated with primary antibody for three nights and secondary antibody for two nights. Sections stained using peroxidase-conjugated secondary antibodies were incubated in tyramide reagents (Perkin Elmer; 1:100) for 45 min. Stained sections were post-fixed in 4% PFA in PBS at 4 °C for 1 h, dehydrated into methanol and cleared in benzyl alcohol:benzyl benzoate (1:2; BABB), or cleared in Vectashield (Vector Laboratories) for confocal imaging.

Immunostaining of the human adenocarcinoma sample was performed on a formalin-fixed paraffin-embedded tumour section using the BOND automated staining system with ER2 epitope retrieval solution and the BOND Polymer Refine Detection system (Leica Biosystems), which includes a haematoxylin counterstain. Adjacent sections were used for immunostaining and smFISH.

Primary antibodies used, at indicated concentrations, were: CD34 (BD Biosciences, 347660; 1:160); claudin 5 (Abcam, ab53765; 1:300); E-cadherin (BD Biosciences, 610181; 1:100); endomucin (Invitrogen, eBioV.7C7, 14-5851-82; 1:300); integrin α8 (R&D, AF4076; reconstituted to 1 mg/ml, used at 1:500); PECAM1 (rat anti-mouse; BD Biosciences, 553370; 1:5,000 for staining embryonic lung, 1:500 for staining adult lung); PECAM1 (mouse anti-human; R&D, BBA7; reconstituted to 0.5 mg/ml in PBS, used at 1:200); tdTomato (Rockland, 600-401-379; 1:300); and VE-cadherin (R&D, AF938; reconstituted to 0.5 mg/ml in PBS, used at 1:300).

Secondary antibodies used, at indicated concentrations, were: donkey anti-goat IgG, Alexa Fluor 568 conjugated (Invitrogen, A11057; 1:250); horse anti-mouse IgG, peroxidase conjugated (Vector Laboratories, PI-2000; 1:150); goat anti-rabbit IgG, peroxidase conjugated (Vector Laboratories, PI-1000; 1:125-1:250); goat anti-rabbit IgG, Alexa 568 conjugated (Invitrogen, A11036; 1:250); goat anti-rat IgG, Alexa 488 conjugated (Invitrogen, A11006; 1:250), for embryonic lung; donkey, anti-rat IgG, Alexa 647 conjugated (Jackson Immunoresearch, 712-605-153; 1:250); goat anti-rat IgG, biotin conjugated (Vector Laboratories, BA-9401; 1:250), for embryonic lung; goat anti-rat IgG, peroxidase conjugated (Vector Laboratories, PI-9401; 1:250), for adult lung. DAPI (4′,6-diamidino-2-phenylindole, dihydrochloride, Invitrogen, D1306; reconstituted in PBS, used at 2 µg/ml), to stain nuclei, and/or Alexa Fluor 350 hydrazide (Invitrogen, A10439; reconstituted to 0.5 mg/ml in PBS, used at 1:100) or Alexa Fluor 633 hydrazide (Invitrogen, A30634; reconstituted to 0.5 mg/ml in PBS, used at 1:500–1:1,000), to visualize elastin fibres, were added along with secondary antibody.

smFISH

Mouse, alligator and turtle lungs, inflated as described above, human lung or adult mouse kidney tissue were fixed in 10% neutral buffered formalin (Fisher Scientific) for 24 h at room temperature and transferred to 70% ethanol (made up in PBS) following 3 brief washes in PBS for embedding in paraffin. Sections were cut at 6 µm. smFISH was performed using a proprietary high-sensitivity RNA amplification and detection technology (RNAscope, Advanced Cell Diagnostics), according to the manufacturer's instructions using the indicated proprietary probes, the RNAscope Multiplex Fluorescent Reagent Kit (v.2) and TSA Plus reagents (Perkin Elmer; 1:500 dilution for Cy3 and Cy5 dyes, 1:250 dilution for FITC) or Opal dyes (Akoya Biosciences, 1:500 dilution for Opal 570 and 620 dyes, 1:250 dilution for Opal 520 and 690 dyes). After smFISH, sections were incubated in DAPI (used at 2 µg/ml in PBS) for 5 min to counterstain nuclei and mounted in Prolong Gold antifade reagent (Invitrogen). Proprietary (Advanced Cell Diagnostics) probes used were: mouse, Mm-Aplnr (436171), Mm-Vwf (499111), Mm-Ednrb (473801, 473801-C2, 473801-C3), Mm-Apln (415371-C2), Mm-Ptprb (481391-C2), Mm-H2-Ab1 (414731-C2), Mm-Car4 (468421-C3), Mm-Gpihbp1 (540631-C3), Mm-Cldn5 (491611-C3), Mm-Pecam1 (316721-C3), tdTomato (317041-C3); human, Hs-EDN1 (459381), Hs-PTPRB (588141), Hs-CA4 (438561), Hs-EDNRB (528301, 528301-C2), Hs-CLDN5 (517141-C2, 517141-C3), Hs-VWF (560461-C3), Hs-APLN (449971-C3); alligator, Ami-APLNR (576071), Ami-PTPRB (828711), Ami-EDNRB (576081-C2), Ami-CA4 (828621-C2), Ami-CLDN5 (576091-C3); western painted turtle, Cpi-APLNR (828481), Cpi-EDNRB (828471-C2), Cpi-CLDN5 (828461-C3).

For quantification of capillary cell-type abundance, aCap and gCap cells were detected using probes for *Ednrb* or *Ptprb*, respectively, and the pan-endothelial probe *Cldn5*. *Cldn5*-expressing alveolar cells with 2 or more *Ednrb* puncta and 0–1 *Ptprb* puncta were classified as aCap; cells with 2 or more *Ptprb* puncta and 0–1 *Ednrb* puncta were classified as gCap; and cells with 2 or more *Ednrb* and 2 or more *Ptprb* puncta were classified as capillary intermediate (IM) cells. A total of 500 *Cldn5*-expressing alveolar cells were scored per lung in 5–10 random fields of view taken with a Plan-Apochromat 25× objective (Carl Zeiss Microscopy), using Volocity software (Quorum Technologies). For quantification of capillary cell-type distribution, capillary cells were scored in the last generation of alveoli immediately adjacent to the pleura and in intra-acinar regions of left and right cranial lobes from 3-month-old mice. For quantification of capillary cell-type abundance in mouse adenomas, capillary cells were scored in sections from tumours with intra-acinar (rather than pleural) location and round, compact morphologies with clear boundaries between tumour and surrounding alveolar tissue. For quantification of *Vwf* induction with age, *Ptprb*+*Cldn5*+ alveolar cells were classified as gCap and *Ptprb*−*Cldn5*+ alveolar cells were classified as aCap. Cells with three or more *Vwf*

puncta were scored as positive. For quantification of aerocyte emergence in the fetal human lung, *CA4*+*APLN*+*EDNRB*+ triple-positive cells with five or more puncta for each transcript were classified as emerging aerocytes. For quantification of capillary fate conversion upon elastase injury, aCap and gCap cells were detected using probes for *Ednrb* (aCap) and *Ptprb* or *Aplnr* (gCap) in injured areas, identified as regions with enlarged airspaces and remodelled elastin fibres, in 3–4-month-old *Apln-creER;Rosa26-tdTomato* and *Aplnr-creER;Rosa26-tdTomato* lungs collected six weeks after elastase instillation. For quantification of capillary cell-type abundance in the human lung, *CLDN5*+ cells with 2 or more *EDNRB* puncta and 0–1 *PTPRB* (or *EDN1*) puncta were classified as aCap; *CLDN5*+ cells with 2 or more *PTPRB* (or *EDN1*) puncta and 0–1 *EDNRB* puncta were classified as gCap; and *CLDN5*+ cells with 2 or more *EDNRB* and 2 or more *PTPRB* (or *EDN1*) puncta were classified as 'IM'.

Histology

Haematoxylin and eosin (H&E) staining was performed using standard protocols on formalin-fixed, paraffin-embedded alligator, turtle and fetal human lung tissue, processed as described above for smFISH. Adjacent sections were used for H&E staining and smFISH.

Electron microscopy and ultrastructural analysis

To visualize capillaries within alveolar walls, perfused and inflated adult mouse lungs were fixed in 2% glutaraldehyde in PBS for 1 h at room temperature. Tissue was manually cut with a platinum-coated double-edge razor blade (EMS) into rough sections. Samples were post-fixed in Karnovsky's fixative (2% glutaraldehyde (EMS) and 4% PFA (EMS) in 0.1M sodium cacodylate (EMS) pH 7.4) for 1 h, and incubated in cold aqueous 1% osmium tetroxide (EMS), washed, stained in 1% uranyl acetate for 2 h, dehydrated into 100% ethanol, infiltrated with Embed 812 resin (EMS) and cured at 65 °C overnight. Sections (75–90 nm) were collected on formvar/carbon-coated slot copper grids, observed in the JEM-1400 transmission electron microscope (JEOL) with a 120-kV beam and imaged with an Orius SC1000 (Gatan) digital camera.

For immuno-electron microscopy, perfused and inflated *Apln-creER; Rosa26-tdTomato* and *Aplnr-creER;Rosa26-tdTomato* adult lungs, were fixed in 4% PFA and 0.1% glutaraldehyde (EMS) in PBS for 3 h at 4 °C. Sections (200 µm) were cut on a vibratome and then immunostained for tdTomato as described above. DAB (3,3′-diaminobenzidine)–nickel (Vector Laboratories, SK-4100) was used as a substrate for the peroxidase conjugated to the secondary antibody. Sections were incubated in DAB–nickel working solution (prepared following the manufacturer's instructions) for 6–15 min at room temperature. After washing, samples were processed for electron microscopy as described, omitting uranyl acetate staining. To determine the percentage of aerocytes or gCap cells associated with thin or thick regions of the air–blood barrier, capillaries with complete lumens and containing immunolabelled endothelial cells were identified on sections viewed by electron microscopy. Labelled endothelial cells (*n* = 2 mice of each genotype, 21 labelled aCap cells and 24 labelled gCap cells) were scored as being associated either with the thin region (defined as regions in which the endothelial cell is tightly apposed to the epithelium) or the thick region (defined as regions in which the endothelial cell is clearly separated from the epithelium by stromal cells or connective tissue fibres). Samples were observed by electron microscopy at multiple magnifications to confirm endothelial cell labelling, AT1 cell identity and separation between endothelium and epithelium. As a control, we also scored the association of unlabelled capillary cells. Notably, we found that in capillaries containing labelled gCap cells, all endothelial cells associated with thin regions (*n* = 12 scored) were unlabelled, consistent with the conclusion that only aCap cells are associated with thin regions. Some sections were scored by an investigator blinded to the genotype of the sample, and similar results were obtained. Representative electron micrographs (Fig. 2g–i) were pseudocoloured in Adobe Illustrator (Extended Data Fig. 4a).

Article

Pulse-chase labelling experiments

To determine the stability of the two alveolar capillary cell populations in the adult mouse lung, aCap cells were labelled using the *Apln-creER* knock-in allele combined with the *Rosa26-tdTomato* Cre reporter and gCap cells were labelled with the *Aplnr-creER* bacterial artificial chromosome (BAC) transgenic allele combined with the *Rosa26-tdTomato* Cre reporter. Co-expression of *tdTomato* and a marker for the respective capillary populations (*Apln* or *Ednrb* for aCap, *Aplnr* or *Ptprb* for gCap) was detected after a 1-, 6- or 14-month chase by smFISH. Around 500–2,000 lineage-labelled cells were scored per lung in at least 3 random fields of view taken with a Plan-Apochromat 25× oil objective (Carl Zeiss Microscopy), using Volocity software (Quorum Technologies).

The fidelity of *Apln-creER* was confirmed by dosing an *Apln-creER; Rosa26-tdTomato* mouse with 4 mg tamoxifen followed by a 48-h chase. Tamoxifen-dependent Cre recombination was observed only in *Ednrb*-expressing cells ($n = 604$ scored cells; 598 cells co-expressed *tdTomato* and *Ednrb*, but not *Aplnr*; 6 cells co-expressed *tdTomato*, *Ednrb* and *Aplnr*. Because the *Apln-creER* knock-in allele is a loss of function allele, we used *Ednrb* rather than *Apln* as an aCap marker).

The fidelity of *Aplnr-creER* was established by dosing an *Aplnr-creER; Rosa26-tdTomato* mouse with 4 mg tamoxifen followed by a 48-h chase. Tamoxifen-dependent Cre recombination was observed only in *Aplnr*-expressing cells ($n = 1,879$ scored cells; 1,860 cells co-expressed *tdTomato* and *Aplnr*, but not *Apln*; 19 cells co-expressed *tdTomato*, *Aplnr* and *Apln*).

Sparse labelling of endothelial cells and analysis of cell morphology

To visualize single endothelial cells in the lung and other organs, mice carrying inducible creER alleles were administered limiting doses of tamoxifen (see 'Mice' for details). Organs were collected as described, fixed in 2% PFA in PBS for 5 h at 4 °C and cut into 200–250 μm sections on a vibratome. To preserve endogenous fluorescence for imaging, tissue was not dehydrated into methanol. For some experiments, lung sections were stained with Alexa Fluor 350 hydrazide (Invitrogen, A10439; 1:100) or Alexa Fluor 633 hydrazide (Invitrogen, A30634; 1:1,000–1:5,000) to visualize elastin fibres[30]. To visualize the vasculature, tamoxifen-dosed *Cdh5-creER; Rosa26-Confetti* mice were injected with 0.2 ml DyLight 649-labelled *Lycopersicon esculentum* (Tomato) lectin (Vector Laboratories, DL-1178; 1 mg/ml) and humanely euthanized after 5 min. Sections were cleared and mounted in CUBIC1[31] for confocal imaging. Sparse labelling, with individual fluorescent cells well separated from other cells, was verified. Volume and surface area of individual plexus cells ($n = 18$ cells at E12.5 from $n = 2$ mice), aCap cells ($n = 23$ cells at P0; $n = 14$ (volume and surface area); $n = 16$ (pores) at P7; $n = 19$ cells at 4 months from $n = 2$ mice) and gCap cells ($n = 12$ cells at P0; $n = 11$ cells at P7; $n = 17$ cells at 4 months from $n = 2$ mice) were automatically calculated from computed three-dimensional (3D) surfaces using Imaris software (Bitplane). Pore number was determined by counting using the original confocal stack viewed in 3D in Imaris.

Clonal analysis

Lungs from P25 *Aplnr-creER; Rosa26-Confetti* mice ($n = 2$), the dams of which were administered limiting doses of tamoxifen (0.06 mg administered by i.p. injection) at E14.5, were collected as described in 'Mice', fixed in 4% PFA in PBS for 2 h at 4 °C and cut into 300-μm sections on a vibratome. Sections were stained with Alexa Fluor 633 hydrazide to visualize elastin fibres as above. To preserve endogenous fluorescence for imaging, sections were cleared and mounted in CUBIC1[31]. The tamoxifen dose was chosen to yield small numbers of well-separated clones for each fluorescent reporter, such that some sections did not contain any clones. Only YFP- and RFP-expressing clones were analysed, because cell number and type could not be identified in nuclear GFP- and membrane targeted Cerulean CFP-expressing clones.

After confocal imaging, clones were visualized in 3D in Imaris. Cell-type composition was scored on the basis of cell morphology. For some clones, cell-type identity could not be definitively assigned to all cells in the clone, especially if cell boundaries could not be determined. In these cases, the number of cells of a given type that could be identified is given as a lower limit in Extended Data Fig. 6e. (For example, in clone 28, which contains 40 cells, there are at least 4 aCap and 30 gCap cells.)

Detection of capillary cell proliferation by EdU labelling

The synthetic deoxyribonucleoside analogue EdU (Carbosynth, NE08701) was administered to adult *Apln-creER; Rosa26-tdTomato* and *Aplnr-creER; Rosa26-tdTomato* mice in drinking water at 0.2 mg/ml 3 weeks after tamoxifen injection (two doses of 4 mg administered by i.p. injection 48 h apart) for 6 weeks. Lungs were collected as described, fixed in 4% PFA in PBS for 5 h at 4 °C and cut into 300-μm sections on a vibratome. EdU was detected using click chemistry to covalently attach Alexa Fluor 647 azide to EdU alkyne incorporated into DNA during the S phase of the cell cycle (Invitrogen, C10340; Click-iT EdU Alexa Fluor 647 Imaging Kit), by incubating vibratome sections in Click-iT reaction cocktail for 4 h at room temperature.

Alveolar injury with elastase

Elastase solution (0.8 μg/μl) was prepared by dissolving elastase (Worthington, LS002292) in PBS and stored at −20 °C. A single dose of elastase (40 μg) or PBS was delivered by intratracheal instillation into the lungs of avertin (1.2% in PBS)-anaesthetized adult *Apln-creER; Rosa26-tdTomato* and *Aplnr-creER; Rosa26-tdTomato* mice 3 weeks after tamoxifen injection (two doses of 4 mg administered by i.p. injection 48 h apart) to allow tamoxifen clearance. EdU was administered as described above starting the night before elastase injury for three days, one week or six weeks. Lungs were collected as described, fixed in 4% PFA in PBS for 5 h at 4 °C and cut into 300-μm sections on a vibratome. EdU was detected using click chemistry as described above. Sections were incubated in Alexa Fluor 488 hydrazide (Invitrogen, A10436; reconstituted to 0.5 mg/ml in PBS, used at 1:100) and DAPI (2 μg/ml) in 0.1% Triton-X-100 in PBS overnight at 4 °C. After washing, sections were cleared and imaged in CUBIC1[31]. EdU incorporation was analysed in lineage-labelled cells in *Apln-creER; Rosa26-tdTomato* lungs (3 days, control: $n = 2$ mice, 296 and 297 cells counted; 1 week, control: $n = 2$ mice, 404 and 295 cells counted; 6 weeks, control: $n = 2$ mice, 424 and 3,977 cells counted; 3 days, elastase: $n = 2$ mice, 767 and 1,018 cells counted; 1 week, elastase: $n = 4$ mice, 431, 877, 253 and 223 cells counted; 6 weeks, elastase: $n = 3$ mice, 1252, 420 and 1,054 cells counted) and in *Aplnr-creER; Rosa26-tdTomato* lungs (3 days, control: $n = 2$ mice, 1,416 and 1,569 cells counted; 1 week, control: $n = 2$ mice, 1,286 and 1,129 cells counted; 6 weeks, control: $n = 2$ mice, 1,172 and 3,390 cells counted; 3 days, elastase: $n = 3$ mice, 1,286, 810 and 1,267 cells counted; 1 week, elastase: $n = 3$ mice, 844, 623 and 707 cells counted; 6 weeks, elastase: $n = 2$ mice, 967 and 1,246 cells counted). In elastase-treated lungs, only cells in injured regions were scored.

Acquisition and processing of images

The image of the immunostained embryonic lung in Fig. 4a was captured using a Zeiss Axioskop microscope and the MRC-1000 Laser Scanning Confocal Imaging System (Bio-Rad) and processed in ImageJ (NIH). All other fluorescent samples were imaged on an LSM 780 (Carl Zeiss Microscopy) or LSM 880 equipped with Airyscan (Carl Zeiss Microscopy) confocal microscope, and images were processed in Zen (Carl Zeiss Microscopy) or Imaris (Bitplane) software. Two-dimensional (2D) confocal images presented are maximum intensity projections of z-stacks. H&E stains were imaged on a slide scanning system (Philips), and images were processed using QuPath software[32].

Analysis of scRNA-seq data

Processed scRNA-seq Smart-Seq2 data for adult (3-month-old) mouse lung were obtained from the Tabula Muris resource[13]

(https://tabula-muris.ds.czbiohub.org) as gene count tables with de-multiplexed and aligned reads. Cells with fewer than 500 detected genes or 50,000 reads were excluded. Data were log-transformed: ln(CPM + 1). Expression profiles of cells were clustered using the R software package Seurat[33] (v.2.3). Highly variable genes were selected using the 'FindVariableGenes' function (dispersion (mean/variance) z-score > 0.5) for linear dimensionality reduction using principal component analysis. The number of principal components was selected by inspection of the plot of variance explained. Cells were clustered by constructing a shared nearest neighbour graph and clusters were visualized by t-SNE. Lung endothelial cells (n = 693 cells, Fig. 1b, c, Fig. 3, Extended Data Fig. 1a, b, Extended Data Fig. 3h, Extended Data Fig. 4d, e, Extended Data Fig. 10a) were identified by *Cldn5*, *Pecam1* and *Cdh5* expression and subclustered. Artery (n = 76 cells), vein (n = 54 cells) and lymphatics (n = 68 cells) clusters were annotated using canonical markers (*Gja5* and *Bmx* for arteries; *Nr2f2* for veins; *Vwf* for arteries and veins; *Pdpn* and *Prox1* for lymphatics). For annotation of the remaining clusters (aCap, n = 101 cells; gCap, n = 394 cells), cluster markers were identified by differential expression analysis using the Wilcoxon rank sum test with Bonferroni correction (P < 0.01) as implemented in the 'FindMarkers' function in Seurat, and specific cluster markers were selected for localization of the cells by smFISH (*Apln* or *Ednrb* for aCap, *Aplnr* or *Ptprb* for gCap). The apparent separation of the cluster we annotated as gCap into two populations (see t-SNE plot in Fig. 1b) is probably due to batch effects, on the basis of the correlation of gene expression differences with technical differences, and was not analysed further.

Processed scRNA-seq MARS-Seq data for embryonic (E12.5–E19.5) and postnatal (P0, P1, P7 and 2-month-old) mouse lung from a previous study[34] were obtained from the Gene Expression Omnibus (GEO) (accession number GSE119228). Cells with fewer than 500 unique molecular identifiers were discarded. Cells from adult lung processed differently than tissue from embryonic and postnatal time points were also eliminated. Data were log-transformed: ln(UP10K + 1) and analysed with Seurat as described above. Cells were initially clustered separately at each developmental stage and endothelial cells were identified by *Cldn5*, *Pecam1* and *Cdh5* expression. Annotated endothelial cells were combined from all time points (n = 4,378 cells) and clustered again. Contaminant haematopoietic cells were eliminated by filtering out *Ptprc*-expressing cells (log-transformed expression levels > 0.5). Artery, vein and lymphatics clusters were identified using canonical markers (*Gja5* and *Bmx* for arteries; *Nr2f2* for veins; *Vwf* for arteries and veins; *Pdpn* and *Prox1* for lymphatics) following correction for cell cycle as described previously[35] and eliminated from the analysis, resulting in a dataset of 3,094 plexus and capillary cells (64 cells at E12.5, 404 cells at E16.5, 296 cells at E18.5, 117 cells at E19.5, 1,016 cells at P0, 426 cells at P1, 438 cells at P7 and 333 cells at 2 months (adult); see Extended Data Figs. 6h–j, 7, Supplementary Data 2–4).

Single-cell trajectories were constructed for plexus and capillary cells (n = 3,094) using Monocle2[36]. Mature aCap and gCap markers, identified by differential gene expression analysis in Seurat, were used for ordering of the cells, which produced a tree-shaped branched developmental trajectory (Extended Data Fig. 6h), with plexus cells located along the stem before the branchpoint, mature aCap cells at the tip of one branch and mature gCap cells at the tip of the second branch. Genes with branch-dependent expression were identified by branched expression analysis modelling (BEAM) (n = 1,119 genes; q-value < 0.05), as implemented in the 'BEAM' function in Monocle2[36]. Genes that vary as a function of pseudotime were identified by differential expression analysis (n = 3,734 genes; q-value < 0.05), as implemented in the 'differentialGeneTest' function in Monocle2. The genes identified by the two approaches (n = 4,129) were clustered hierarchically and plotted as a branched heat map (Extended Data Fig. 7b, Supplementary Data 4) to visualize groups of genes that co-vary across pseudotime.

Processed scRNA-seq Smart-Seq2 data for adult (3-month-old) mouse heart and brain were obtained from the Tabula Muris resource[13] as

Seurat objects with annotated clusters. Brain and heart endothelial cells were identified by *Cldn5*, *Pecam1*, *Esam* and *Cdh5* expression and subclustered. Artery, vein and lymphatics clusters were identified using canonical markers (*Gja5*, *Bmx* and *Vegfc* for arteries; *Vwf* for arteries and veins; *Il1r1* for veins; *Pdpn* and *Prox1* for lymphatics)[37] and excluded from the analysis. The remaining cells were annotated as capillaries based on expression of *Car4*[38,39].

Processed scRNA-seq droplet (10X) data for adult (1-, 3-, 18-, 21- and 30-month-old) mouse lung, kidney and mammary gland were obtained from the Tabula Muris Senis resource[40] (https://tabula-muris-senis. ds.czbiohub.org). Scanpy single-cell objects were imported in R and analysed with Seurat as described above. Cells with fewer than 500 detected genes or 1,000 unique molecular identifiers were discarded. Kidney, mammary gland and lung endothelial cells were identified by *Pecam1*, *Cdh5* and *Esam* (for kidney and mammary gland) or *Cldn5* (for lung) expression and subclustered. Contaminant haematopoietic cells were eliminated by filtering out *Ptprc*-expressing cells (log-transformed expression levels > 0.5). Contaminant stromal and epithelial cells were eliminated in the mammary gland dataset by filtering out cells that express *Col1a1*, *Pdgfrb* or *Epcam*. Artery, vein and lymphatics clusters were identified using canonical markers as above and excluded. Lung capillary cells were annotated using aerocyte (*Apln*, *Car4*, *Ednrb*, *Tbx2*) and gCap markers (*Aplnr*, *Gpihbp1*, *Lpl*). Clusters with *Ehd3*, *Sost* and *Meg3* expression were annotated as glomerular capillaries in the kidney dataset. Clusters with expression of *Gpihbp1*, *Rbp7* and *Car4*, and little or no expression of *Vwf*, were annotated as capillaries in the mammary gland dataset.

Alveolar aCap and gCap signature scores (defined as the sum of normalized and scaled gene expression values for the significant alveolar aCap or gCap marker genes: Bonferroni corrected P value < 0.01; average normalized expression fold change > 1; % expression > 40) were calculated for annotated lung (n = 495), heart (n = 753) and brain (n = 441) capillary endothelial cells in the Tabula Muris Smart-Seq2 data[13] (Extended Data Fig. 3h), and for annotated lung (n = 2,050), glomerular (n = 126) and mammary gland (n = 173) capillary endothelial cells in the Tabula Muris Senis droplet data[40] (Extended Data Fig. 3i).

Smart-Seq2 and droplet (10X) data for adult human lung (patient 1, 75-year-old man)[16] were analysed as described above. Endothelial cells were identified by *CLDN5* expression, subclustered and annotated using markers (*GJA5* for arteries, *ACKR1* for vein, *PDPN* for lymphatics, *EDNRB* for aerocytes, *EDN1* for gCap cells, *COL15A1* for bronchial vessels; droplet: n = 1,497 cells; 211 artery cells, 154 vein cells, 33 lymphatic cells, 230 bronchial endothelial cells, 315 aerocytes, 554 gCap cells; Fig. 4k, m–o, Extended Data Figs. 8b, c, 10; Smart-Seq2, n = 599 cells, Extended Data Fig. 10).

To identify differences in cell-type-specific expression between mouse and human, we first identified genes that are differentially expressed (P < 0.01, Wilcoxon rank sum test with Bonferroni correction) between the alveolar capillary cell types in each species using scRNA-seq Smart-Seq2 data for adult mouse lung endothelial cells obtained from the Tabula Muris resource[13] and Smart-Seq2 data for adult human lung endothelial cells (patient 1, 75-year-old man)[16]. Lists of differentially expressed genes (see Supplementary Tables 2, 3) were then compared to identify genes specifically expressed in the same cell type in the two species (type 0), genes specifically expressed in one cell type in one species but not the other (type 1), or genes specifically expressed in different cell types in the two species (type 2). Selected genes of each type are shown in Fig. 4m–o, Extended Data Fig. 10c–e. A complete analysis of the evolutionary changes between alveolar capillary cell types in mouse versus human lungs is presented in supplementary table 7 in ref.[16].

Statistics and reproducibility
Data analysis and statistical tests were performed using R software (v.3.5.1). Data are represented as mean ± standard deviation (s.d.) for

Article

sample sizes larger than two. For comparison of two groups, a two-sided Wilcoxon rank sum test was conducted at 5% significance level. For Fig. 1a, the image is representative of $n = 10$ mice; for Fig. 1d, smFISH was performed on samples from $n = 5$ mice; for Fig. 1e, smFISH was performed on samples from $n = 2$ mice; for Fig. 1f, smFISH was performed on samples from $n = 5$ mice; for Fig. 1h, smFISH was performed on samples from two lobes of $n = 1$ mouse; for Fig. 1i, smFISH was performed on samples from two lobes of $n = 1$ mouse; for Fig. 2a, the image is representative of $n = 5$ mice; for Fig. 2b, the image is representative of $n = 5$ mice; for Fig. 2d, the image is representative of $n = 2$ mice; for Fig. 2f–i, each image is representative of $n = 2$ mice; for Fig. 2l, images are representative of $n = 3$ mice; for Fig. 2m, images are representative of $n = 2$ mice; for Fig. 2o, smFISH was performed on samples from $n = 2$ mice; for Fig. 4a, the image is representative of $n = 10$ embryos; for Fig. 4b, the image is representative of $n = 2$ embryos; for Fig. 4c, images are representative of $n = 2$ mice; for Fig. 4d, smFISH was performed on samples from $n = 2$ mice; for Fig. 4e, clones were examined in $n = 2$ mice, see clone table in Extended Data Fig. 6e; for Fig. 4g, images are representative of $n = 2$ mice at each time point; for Fig. 4j, smFISH was performed on samples from $n = 3$ humans; for Fig. 4l, smFISH was performed on samples from a single adenocarcinoma; for Fig. 4q, the image is representative of multiple lung regions from a single alligator; for Fig. 4r, smFISH was performed on samples from $n = 2$ alligators (1 juvenile; 1 adult); for Extended Data Fig. 1c, smFISH was performed on samples from $n = 5$ mice; for Extended Data Fig. 1d, smFISH was performed on samples from $n = 2$ mice; for Extended Data Fig. 1e, e', smFISH was performed on samples from $n = 2$ mice; for Extended Data Fig. 1f, g, smFISH was performed on samples from $n = 1$ mouse; for Extended Data Fig. 1j, smFISH was performed on samples from $n = 2$ mice at each age; for Extended Data Fig. 1k, l, smFISH was performed on samples from two lobes of $n = 1$ mouse of each genotype; for Extended Data Fig. 2a–d; images are representative of $n = 2$ embryos or mouse pups at each time point; for Extended Data Fig. 2e, f, images are representative of $n = 5$ mice; for Extended Data Fig. 3a, the image is representative of $n = 2$ mice; for Extended Data Fig. 3b–f, images are representative of $n = 4$ mice; for Extended Data Fig. 3g, the image is representative of $n = 2$ mice; for Extended Data Fig. 3j–l, smFISH was performed on samples from $n = 2$ mice; for Extended Data Fig. 4a–c, images are representative of $n = 2$ mice of each genotype; for Extended Data Fig. 5a, the images are representative of $n = 2$ mice at each time point; for Extended Data Fig. 5c, the images are representative of $n = 2$ mice of each genotype; for Extended Data Fig. 5d, the images are representative of $n = 2$ *Apln-creER; Rosa26-tdTomato* mice and $n = 3$ *Aplnr-creER; Rosa26-tdTomato* mice; for Extended Data Fig. 5e, the images are representative of $n = 4$ *Apln-creER; Rosa26-tdTomato* mice and $n = 3$ *Aplnr-creER; Rosa26-tdTomato* mice; for Extended Data Fig. 5f, the images are representative of $n = 3$ *Apln-creER; Rosa26-tdTomato* mice and $n = 2$ *Aplnr-creER; Rosa26-tdTomato* mice; for Extended Data Fig. 6a, b, smFISH was performed on samples from $n = 2$ embryos; for Extended Data Fig. 6c, smFISH was performed on samples from two lobes of $n = 1$ mouse; for Extended Data Fig. 6d, e, clones were examined in $n = 2$ mice; for Extended Data Fig. 6f, f', f'', smFISH was performed on samples from two lobes of $n = 1$ mouse; for Extended Data Fig. 6g, smFISH was performed on samples from $n = 3$ embryos; for Extended Data Fig. 6k, l, smFISH was performed on samples from $n = 3$ mice at each age; for Extended Data Fig. 8a, the images are representative of $n = 5$ human lungs; for Extended Data Fig. 8d–h, smFISH was performed on samples from $n = 3$ humans; for Extended Data Fig. 8j, j', histology and smFISH were performed on samples from a single human fetal lung from each time point (17 weeks and 23 weeks; see figure legend); for Extended Data Fig. 9a, a', a'', immunostaining and smFISH were performed on samples from a single adenocarcinoma; for Extended Data Fig. 9b, b', c, c', f, smFISH was performed on samples from $n = 2$ mice; for Extended Data Fig. 11c, d, images are representative of multiple regions from a single alligator lung; for Extended Data Fig. 11e, f, smFISH was performed on samples

from $n = 2$ alligators (1 juvenile, 1 adult); for Extended Data Fig. 11g, the image is representative of multiple regions from a single alligator lung; for Extended Data Fig. 11i, the image is representative of multiple regions from a single turtle lung; for Extended Data Fig. 11j, k, images are representative of $n = 2$ turtles; for Extended Data Fig. 11l, l', smFISH was performed on samples from $n = 2$ turtles; for Extended Data Fig. 11m, the image is representative of $n = 2$ turtles; for Supplementary Data 1, images are representative of $n = 2$ mice. For all graphs, the number of biologically independent samples is reported in the legend, or in the 'Alveolar injury with elastase' (for Fig. 2n) or 'Sparse labelling of endothelial cells and analysis of cell morphology' (for Fig. 2c, 4h, Extended Data Fig. 2g, h) sections of the Methods. Sample size calculations were not performed. Mice of the appropriate genotype or age were allocated into experimental groups (control versus elastase injury) at random. The investigators were not blinded to sample allocation.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The scRNA-seq datasets analysed are available at the GEO under the accession numbers GSE109774 (Tabula Muris[13], https://tabula-muris.ds.czbiohub.org), GSE132042 (Tabula Muris Senis[40], https://tabula-muris-senis.ds.czbiohub.org) and GSE119228 (a previous study[34]) or on Synapse under the accession number syn21041850 (Human Lung Cell Atlas[16], https://hlca.ds.czbiohub.org). Source data are provided with this paper.

21. Tian, X. et al. Subepicardial endothelial cells invade the embryonic ventricle wall to form coronary arteries. *Cell Res.* **23**, 1075–1090 (2013).
22. Chen, H. I. et al. The sinus venosus contributes to coronary vasculature through VEGFC-stimulated angiogenesis. *Development* **141**, 4500–4512 (2014).
23. Wang, Y. et al. Ephrin-B2 controls VEGF-induced angiogenesis and lymphangiogenesis. *Nature* **465**, 483–486 (2010).
24. Chapman, H. A. et al. Integrin α6β4 identifies an adult distal lung epithelial population with regenerative potential in mice. *J. Clin. Invest.* **121**, 2855–2862 (2011).
25. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).
26. Snippert, H. J. et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134–144 (2010).
27. Jackson, E. L. et al. Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev.* **15**, 3243–3248 (2001).
28. Pitulescu, M. E., Schmidt, I., Benedito, R. & Adams, R. H. Inducible gene targeting in the neonatal vasculature and analysis of retinal angiogenesis in mice. *Nat. Protocols* **5**, 1518–1534 (2010).
29. Metzger, R. J., Klein, O. D., Martin, G. R. & Krasnow, M. A. The branching programme of mouse lung development. *Nature* **453**, 745–750 (2008).
30. Shen, Z., Lu, Z., Chhatbar, P. Y., O'Herron, P. & Kara, P. An artery-specific fluorescent dye for studying neurovascular coupling. *Nat. Methods* **9**, 273–276 (2012).
31. Susaki, E. A. et al. Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. *Nat. Protocols* **10**, 1709–1727 (2015).
32. Bankhead, P. et al. QuPath: open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).
33. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
34. Cohen, M. et al. Lung single-cell signaling interaction map reveals basophil role in macrophage imprinting. *Cell* **175**, 1031–1044 (2018).
35. Su, T. et al. Single-cell analysis of early progenitor cells that build coronary arteries. *Nature* **559**, 356–362 (2018).
36. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
37. Chen, M. B et al. Brain endothelial cells are exquisite sensors of age-related circulatory cues. *Cell Rep.* **30**, 4418–4432 (2020).
38. Ghandour, M. S., Langley, O. K., Zhu, X. L., Waheed, A. & Sly, W. S. Carbonic anhydrase IV on brain capillary endothelial cells: a marker associated with the blood–brain barrier. *Proc. Natl Acad. Sci. USA* **89**, 6823–6827 (1992).
39. Sender, S. et al. Localization of carbonic anhydrase IV in rat and human heart muscle. *J. Histochem. Cytochem.* **46**, 855–861 (1998).
40. The Tabula Muris Consortium. A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. *Nature* **583**, 590–595 (2020).
41. Fleming, R. E., Crouch, E. C., Ruzicka, C. A. & Sly, W. S. Pulmonary carbonic anhydrase IV: developmental regulation and cell-specific expression in the capillary endothelium. *Am. J. Physiol.* **265**, L627–L635 (1993).

42. Beigneux, A. P. et al. Glycosylphosphatidylinositol-anchored high-density lipoprotein-binding protein 1 plays a critical role in the lipolytic processing of chylomicrons. *Cell Metab*. **5**, 279–291 (2007).

43. Davies, B. S. et al. GPIHBP1 is responsible for the entry of lipoprotein lipase into capillaries. *Cell Metab*. **12**, 42–52 (2010).

44. Esnouf, M. P. Biochemistry of blood coagulation. *Br. Med. Bull*. **33**, 213–218 (1977).

45. Davie, E. W., Fujikawa, K. & Kisiel, W. The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry* **30**, 10363–10370 (1991).

46. Nikolić, M. Z., Sun, D. & Rawlins, E. L. Human lung development: recent progress and new challenges. *Development* **145**, dev163485 (2018).

47. Goldenberg, N. M. & Kuebler, W. M. Endothelial cell regulation of pulmonary vascular tone, inflammation, and coagulation. *Compr. Physiol*. **5**, 531–559 (2015).

48. Kreisel, D. et al. Cutting edge: MHC class II expression by pulmonary nonhematopoietic cells plays a critical role in controlling local inflammatory responses. *J. Immunol*. **185**, 3809–3813 (2010).

49. Milani, A. Beiträge zur Kenntnis der Reptilienlunge. II. *Zool. Jahrb. Abt. Anat. Ontogenie Tiere* **10**, 93–156 https://www.biodiversitylibrary.org/item/41831#page/103/mode/1up (1897).

50. Sanders, R. K. & Farmer, C. G. The pulmonary anatomy of *Alligator mississippiensis* and its similarity to the avian respiratory system. *Anat. Rec*. **295**, 699–714 (2012).

51. Maina, J. N. & West, J. B. Thin and strong! The bioengineering dilemma in the structural and functional design of the blood-gas barrier. *Physiol. Rev*. **85**, 811–844 (2005).

Author contributions A.G., M.A.K. and R.J.M. conceived and designed the project, analysed the data and wrote the manuscript. A.G. and R.J.M. performed experiments, assisted by F.Z. A.G. and R.J.M. performed the computational analyses of scRNA-seq data. C.G.F. contributed to the design and interpretation of the evolutionary analysis, collected alligator and turtle lung tissue and contributed to the manuscript. K.J.T. made initial observations about endothelial proliferation and elastin fibre remodelling after elastase injury and contributed to the design and execution of the elastase injury experiments. S.Y.T. provided tissue from a patient with adenocarcinoma and pathological analysis. M.G. provided fetal human lung tissue. B.Z. provided tissue and a mouse line. J.A.F. provided expertise and infrastructure, and supported the work. All authors reviewed the manuscript.

Article
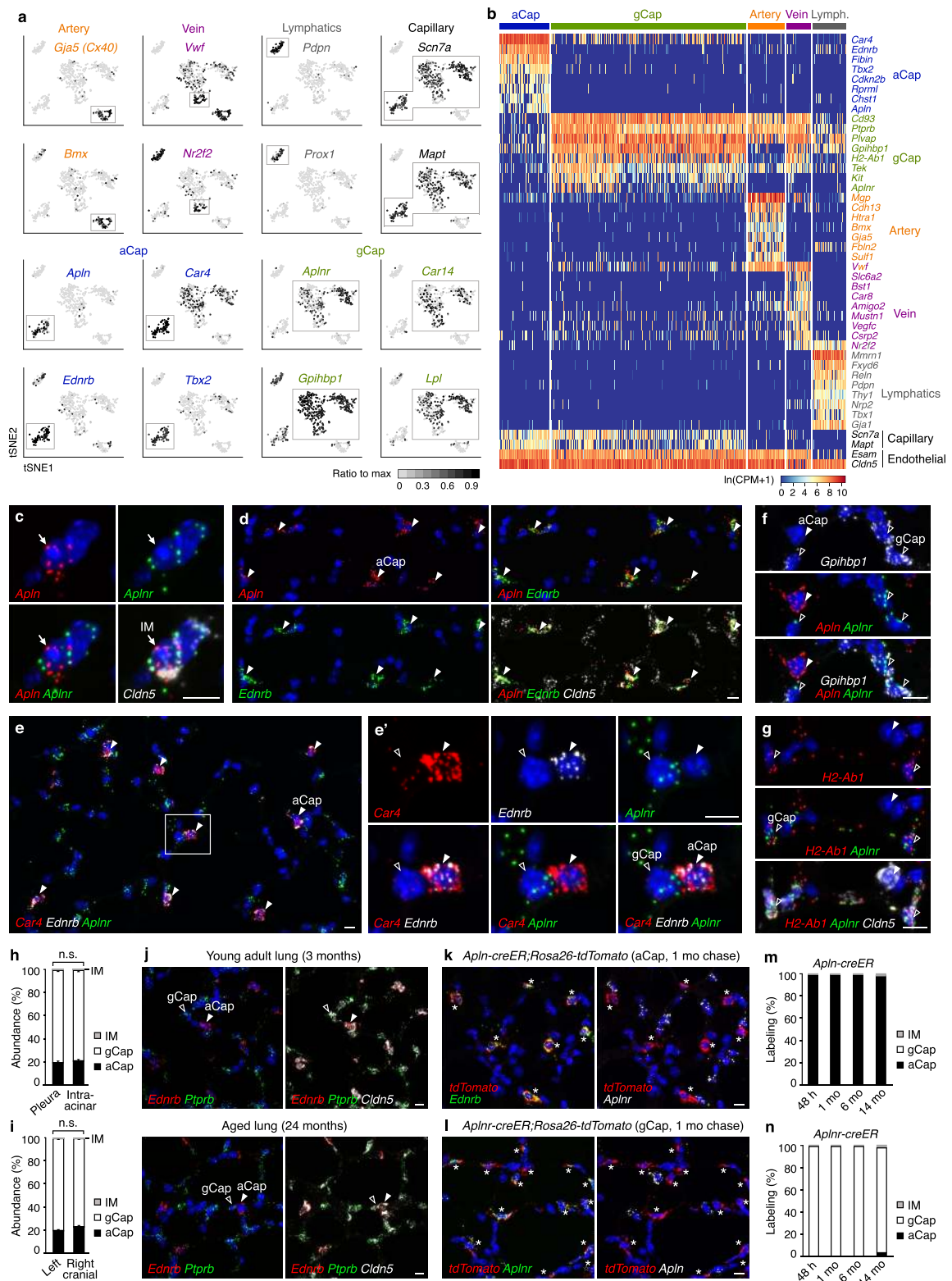
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Molecular characterization and mapping of the two alveolar capillary cell types. a, Feature plots showing log-transformed expression of marker genes used to annotate five molecularly distinct clusters (see t-SNE plot in Fig. 1b) derived by unsupervised clustering of lung endothelial cells in the Tabula Muris scRNA-seq Smart-seq2 data[13] for adult mouse lung. We identified two endothelial cell clusters as capillaries based on expression of carbonic anhydrases (*Car4*, *Car14*) that catalyse the conversion of bicarbonate to carbon dioxide[41], *Gpihbp1*, a lipoprotein-binding protein that localizes to the luminal membrane of alveolar capillaries[42], and lipoprotein lipase (*Lpl*), which is transported to the capillary lumen by Gpihbp1[43]. **b**, Heat map showing log-transformed expression of selected cluster markers in individual pulmonary artery, vein, lymphatics and alveolar capillary (aCap, gCap) endothelial cells identified in the Tabula Muris Smart-Seq2 data[13]. **c**, Co-expression of alveolar capillary cell type markers *Apln* (aCap) and *Aplnr* (gCap) in a single alveolar capillary intermediate (IM) endothelial cell (arrow; marked by *Cldn5*) in adult mouse lung, as detected by smFISH. **d**, Co-expression of aCap markers (*Apln* and *Ednrb*) in a subset (aCap cells) of alveolar endothelial cells (marked by *Cldn5*) in adult mouse lung. **e**, Co-expression of aCap markers *Ednrb* and *Car4* but not *Aplnr*, a gCap marker, in a subset (aCap cells) of adult mouse lung alveolar endothelial cells. **e′**, Boxed region from **e** at higher magnification. Filled arrowhead indicates aCap cell expressing high levels of *Car4*. Open arrowhead points to gCap cell expressing low levels of *Car4*. **f**, **g**, Co-expression of gCap markers *Gpihbp1* (**f**) or *H2-Ab1* (**g**), a major histocompatibility complex (MHC) class II gene, and *Aplnr* in a subset (gCap cells) of adult mouse lung alveolar endothelial cells. **h**, **i**, Quantification of the relative abundance (in % of capillary endothelial cells) of the two alveolar capillary populations (aCap, gCap) and rare cells (IM) that co-express aCap and gCap markers at the pleura and in intra-acinar regions (**h**), or in different lobes (**i**, left versus right cranial) in lungs from 3-month old mice. Data shown as mean ± s.d.; *n* = 500 cells scored per mouse; *n* = 3 mice; *P* values comparing cell type abundance by two-sided Wilcoxon rank sum test: aCap (**h**, 0.2; **i**, 0.2), gCap (**h**, 0.4; **i**, 0.2), IM (**h**, 0.7; **i**, 0.8). n.s., not significant. **j**, aCap (marked by *Ednrb*) and gCap (marked by *Ptprb*) cells in lungs from 3- and 24-month old mice. Endothelial cells are marked by *Cldn5*. **k**, **l**, Co-expression of *tdTomato* lineage label (asterisks) and either aCap marker *Ednrb* but not gCap marker *Aplnr* in an *Apln-creER; Rosa26-tdTomato* lung (**k**) or gCap marker *Aplnr* but not aCap marker *Apln* in an *Aplnr-creER; Rosa26-tdTomato* lung (**l**) harvested one month after mature aCap (**k**) or gCap (**l**) cells were lineage-labelled. Lineage-labelled cells (asterisks) continue to express the marker (*Ednrb*, aCap; *Aplnr*, gCap) of the labelled population. **m**, **n**, Quantification of percent of lineage-labelled cells in *Apln-creER; Rosa26-tdTomato* (**m**) or *Aplnr-creER; Rosa26-tdTomato* (**n**) lungs that continue to express the aCap marker, gCap marker, or markers of both cell types (IM) after 48 h, 1, 6, or 14 months (500–2,000 cells scored at each time point from multiple regions in each of two separate lobes from one lung). Blue (**c**–**g**, **j**–**l**), DAPI. Scale bars, 10 μm.
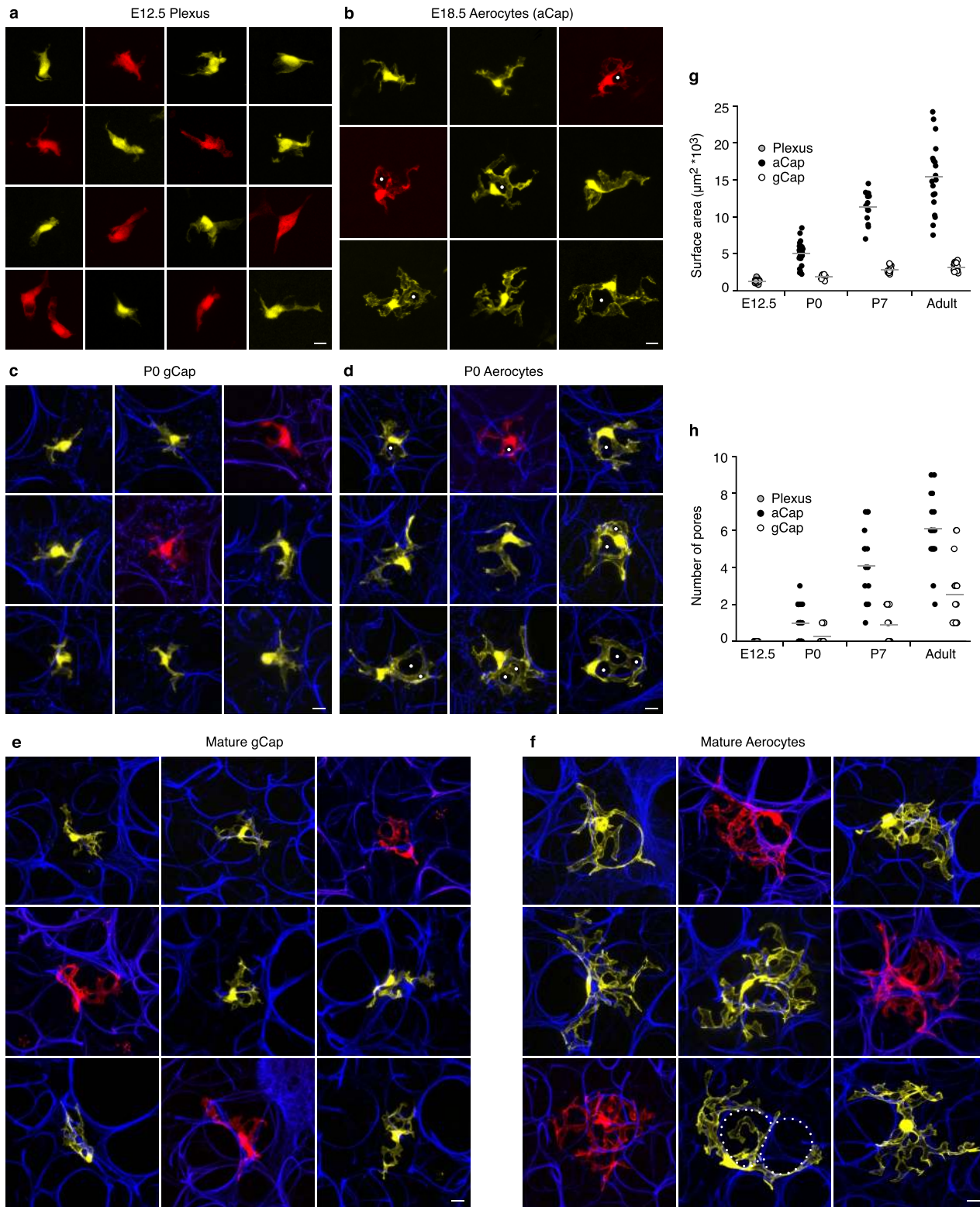
Extended Data Fig. 2 | Alveolar capillary cell morphologies. a–d, Additional examples of single cytoplasmic YFP- or RFP-expressing plexus cells at E12.5 (**a**), aerocytes at E18.5 (**b**), gCap cells at P0 (**c**) and aerocytes at P0 (**d**). Plexus and gCap cells were labelled in *Aplnr-creER; Rosa26-Confetti* lungs; aerocytes were labelled in *Apln-creER; Rosa26-Confetti* lungs. Pores, present in some aerocytes by E18.5, but not plexus or P0 gCap cells, are marked by dots. **e**, **f**, Additional examples of single gCap cells (**e**) or aerocytes (**f**) expressing cytoplasmic RFP or YFP in adult *Aplnr-creER; Rosa26-Confetti* (**e**) or *Apln-creER; Rosa26-Confetti* (**f**)

lungs. Dotted circles outline two alveoli spanned by a single aerocyte (**f**). Note that gCap cells are less morphologically diverse than aerocytes. **g**, **h**, Quantification of individual plexus, aerocyte (aCap), or gCap cell surface area (**g**) and pore number (**h**) at indicated times. *n* > 10 cells scored for each cell type at each time point from *n* = 2 lungs; see Methods for exact cell number. Blue, elastin fibres highlighting the alveolar entrance. Grey bar, mean value. Scale bars, 10 µm.

Capillary cells in different organs

a, Bronchial capillary cell
b, Brain capillary cell
c, Heart capillary cell
d, Intestinal capillary cell
e, Glomerular capillaries
e'
f, Thyroid capillary cells
g, P7 retinal tip cell

h, Lung / Heart / Brain / Lung, heart, brain — gCap signature score vs aCap signature score

i, Lung / Kidney / Mammary gland / Lung, kidney, mammary gland — gCap signature score vs aCap signature score

j, Glomerular capillaries — *Ednrb* (aCap) *Ptprb* (gCap)
j', *Ednrb* (aCap) / *Ptprb* (gCap) / Merge / *Pecam1*
k, Glomerular capillaries — *Ednrb* (aCap) *Gpihbp1* (gCap)
k', *Ednrb* (aCap) / *Gpihbp1* (gCap) / Merge
l, Glomerular capillaries — *Apln* (aCap) *Aplnr* (gCap)
l'

Extended Data Fig. 3 | Capillary cells in the lung and other organs.
a, Bronchial (systemic) capillary cell immunostained for tdTomato (red) in adult *Aplnr-creER;Rosa26-tdTomato* lung. Airway epithelium is immunostained for E-cadherin (blue). **b–f**, Single capillary cells labelled with membrane targeted CFP (**b**) or cytoplasmic YFP (**c–f**) in brain (**b**), heart (**c**), small intestine (**d**), kidney glomerulus (**e**) or thyroid (**f**), in which capillaries are arranged in baskets similar to pulmonary alveoli, from an adult *Cdh5-creER;Rosa26-Confetti* mouse. Capillaries labelled with tomato lectin (blue, **b–f**). **e'**, Cell in boxed region in **e** at higher magnification, along with other examples of labelled glomerular endothelial cells. **g**, Single endothelial tip cell expressing cytoplasmic YFP in developing (P7) *Apln-creER;Rosa26-Confetti* mouse retina. Filled arrowheads, filopodia. **h**, **i**, Scatter plots of alveolar aCap and gCap signature scores assigned to annotated lung, heart and brain capillary endothelial cells identified in the Tabula Muris scRNA-seq Smart-seq2 data[13] (**h**) or annotated lung, kidney glomerular and mammary gland capillary

endothelial cells identified in the Tabula Muris Senis scRNA-seq droplet data for 1-, 3-, 18-, 21- and 30-month-old mice[40] (**i**). Lung capillary cells segregate into two clusters, with annotated gCap cells having high gCap and low aCap signature scores, and annotated aCap cells having low gCap and high aCap signature scores. Heart (**h**), brain (**h**), glomerular (**i**) and mammary gland (**i**) capillary cells have high gCap and low aCap signature scores and each form a single cluster near gCap but not aCap cells, suggesting that they are more similar to gCap cells and any heterogeneity within these other capillary cell populations is different from that in the lung. **j–l**, Some aCap and gCap markers (aCap, *Ednrb* (**j**, **k**); gCap, *Ptprb* (**j**) or *Gpihbp1* (**k**)) are broadly co-expressed in glomerular endothelial cells (*Pecam1*) in adult mouse kidney whereas expression of others (aCap, *Apln*, or gCap, *Aplnr*; **l**) is detected only in small numbers of cells or not at all. Dotted lines outline glomeruli. **j'**, **k'**, **l'**, Boxed regions from **j–l** at higher magnification. Arrows (**j'**, **k'**) point to cells co-expressing the markers. Blue (**j–l**), DAPI. Scale bars, 10 μm.

Article

a, Aerocyte (thin) / Aerocyte (thick) / gCap (thick) / Aerocyte (thin)

b, Aerocytes / Alveolar fibroblasts / Aerocytes + fibroblasts — tdTomato / Integrin α8 / tdTomato Integrin α8

c, gCap / Alveolar fibroblasts / gCap + fibroblasts — tdTomato / Integrin α8 / tdTomato Integrin α8

d

	Aerocyte (aCap)	gCap
Respiration	Car4	Car14
Lipid metab., transport	Mgll, Abcg1	Gpihbp1, Lpl, Acer2
Hemostasis, platelet interactions	Serpine1 (pro), Pros1 (anti), Tfpi (anti)	Vwf (pro), Entpd1 (anti), Plat (anti)
Vasomotor control	Ednrb, Tbxa2r	Edn1 (vasoconstrictor), Nos3 (vasodilator)
Leukocyte trafficking	Icam1, Chst1/2, Sirpa, Edil3, Stab1, Tbxa2r	
Antigen presentation		Cd74, H2-Aa, H2-Ab1, H2-Eb1, H2-DMb1
Signaling	Apln, Kitl, Ednrb, Kdr	Aplnr, Kit, Edn1, Vegfa
Transcription factors	Tbx2, Tbx3, Hopx, Meox1, Prdm1	Sox17, Hey1, Gata3

e

AT1 → aCap ⇌ gCap ⇌ Fibroblast (F) or pericyte (P)

| Ligand | Receptor |

f

Extended Data Fig. 4 | Functional compartmentalization of the alveolus.
a, Transmission electron micrographs of alveolar walls from adult mouse lungs, pseudocoloured to highlight cells and connective tissue (extracellular matrix (ECM) or fibres). *Apln-creER;Rosa26-tdTomato* (to label aerocytes) or *Aplnr-creER;Rosa26-tdTomato* (to label gCap cells) lungs were immunostained for tdTomato (detected with DAB and NiCl$_2$, heavy black stain). Labelled aerocytes but not gCap cells are associated with thin regions of the air–blood barrier (delimited by dashed lines), where the epithelium is tightly apposed to the endothelium. Some thin regions do not contain labelled aerocytes as aerocyte labelling using *Apelin-creER* is inefficient. Thick regions, where the epithelium is separated from the endothelium by fibres or other cells, can be associated with either aerocytes or gCap cells. AT1 cell, alveolar type 1 epithelial cell; AT2 cell, alveolar type 2 epithelial cell. Micrographs in top left, middle and bottom panels are shown without pseudocolouring in Fig. 2g–i. **b**, **c**, gCap cells associate with fibroblasts. Alveoli with labelled aerocytes

(aCap) in *Apln-creER;Rosa26-tdTomato* (**b**) or gCap cells in *Aplnr-creER; Rosa26-tdTomato* (**c**) lungs immunostained for integrin α8 to show alveolar fibroblasts[16]. Asterisks mark aerocytes that do not associate with fibroblasts (in **b**) and gCap cells that do (in **c**). Example regions of alveoli not covered by fibroblasts are dotted. Note that the dotted region in **b** is occupied by an aerocyte not overlaid by fibroblasts, whereas fibroblasts overlay gCap cells in **c**. **d**, Selected genes with known functions, which are differentially expressed between the two alveolar capillary cell types in adult mouse lung[13]. Pro, procoagulants; anti, anticoagulants[44,45]. **e**, Summary of signalling interactions between capillary cell types and surrounding cells in the mouse alveolus. Arrows indicate direction of signalling. **f**, Schematic representation of an alveolus highlighting proposed specialized functions and signalling interactions of alveolar capillary cell types. AT1, alveolar type 1 epithelial cell; AT2, alveolar type 2 epithelial cell. Scale bars, 2 μm (**a**); 10 μm (**b**, **c**).
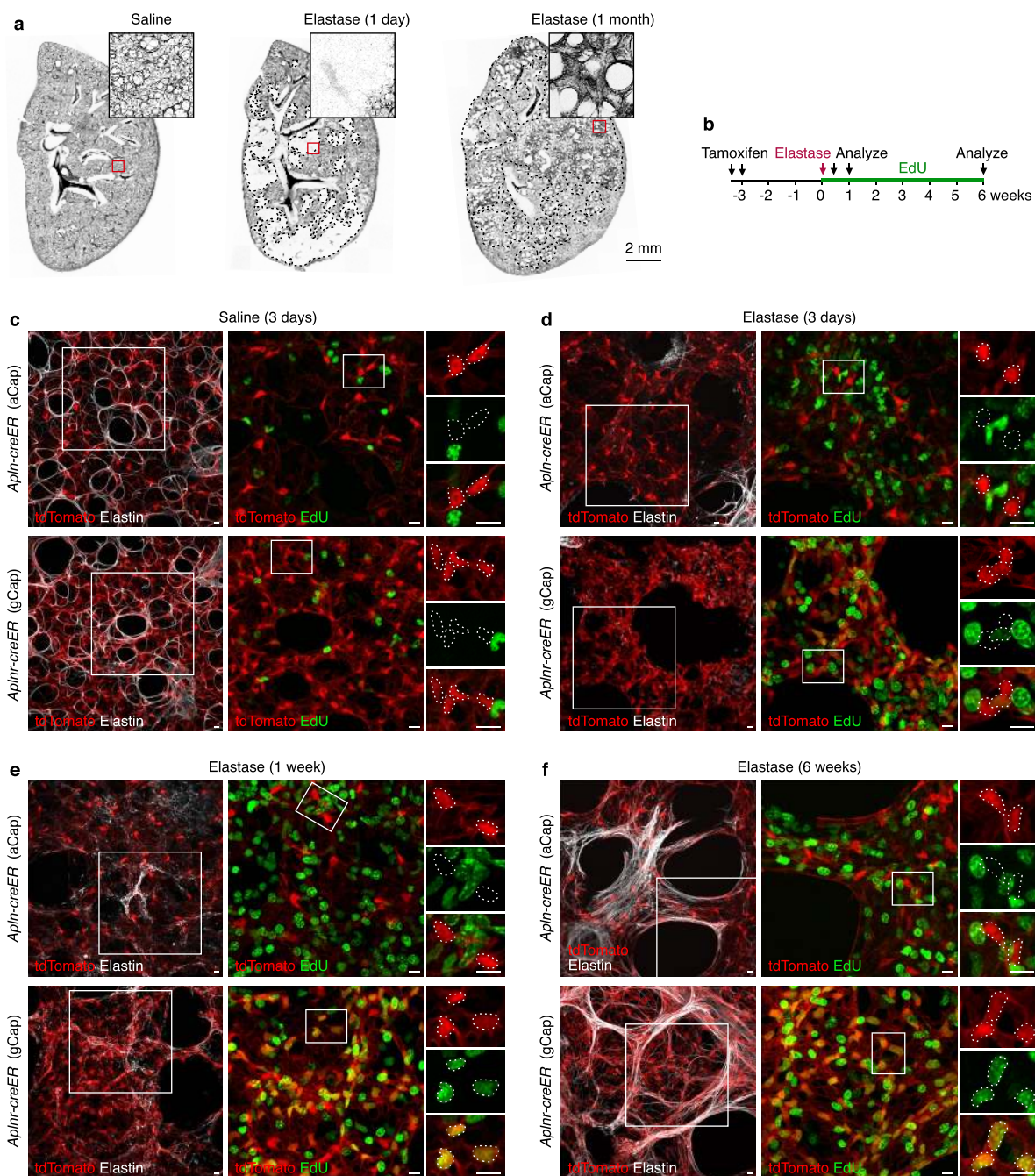
Extended Data Fig. 5 | Proliferation of gCap cells after elastase-induced injury. a, Elastin fibres (black), labelled with fluorescent hydrazide, in lungs of wild-type mice treated with elastase or mock-treated with saline as control. Loss of elastin fibres in injured areas (dashed outlines) is apparent one day after intratracheal instillation of elastase. At one month after injury, regions containing abnormally thickened elastin fibres and enlarged airspaces are evident. Boxed areas are shown at higher magnification. **b**, Scheme for detecting proliferation of alveolar capillary cells after elastase injury. aCap and gCap cells were lineage-labelled in adult *Apln-creER; Rosa26-tdTomato* (for aCap cells) or *Aplnr-creER; Rosa26-tdTomato* (for gCap cells) mice three weeks before elastase administration, then EdU was administered in drinking water and capillary cells were analysed at the indicated times after injury. **c–f**, Proliferation analysed by cumulative EdU incorporation in lineage-labelled (red) aCap cells (top panels in **c–f**) and gCap cells (bottom panels in **c–f**) at 3 days (**c**, **d**), 1 week (**e**) or 6 weeks (**f**) after instillation of saline (as control; **c**) or elastase (**d–f**). Boxed areas are shown at higher magnification, with cell bodies of individual capillary cells outlined. After treatment with elastase, gCap cells proliferated in injured areas. White, elastin fibres. Scale bars (**c–f**), 10 μm.

a Embryonic plexus — E12.5 — Airway lumen — *Aplnr* — Airway lumen — *Aplnr Pecam1*

b Embryonic plexus — E12.5 — Airway lumen — *Apln* — Airway lumen — *Pecam1 Ednrb Apln*

c *Apln-creER;Rosa26-tdTomato* — P21 — aCap — gCap — *Ednrb Aplnr tdTomato* — **c'**

d Aerocyte — gCap — P25 *Aplnr-creER;Rosa26-Confetti* — Aerocyte (aCap) — gCap

Aplnr-creER;Rosa26-tdTomato — **f'** gCap — **f''** aCap — *Apln Aplnr* — *tdTomato* — *Merge*

f P21 — *Apln Aplnr tdTomato* — f'' — f'

e

Clone No.	No. cells	Lobe	Color	Span (µm)	Arrangement	Structure	aCap	gCap	LV
1	4	RAc	RFP	100	Coherent	gCap	0	4	0
2	4	L	YFP	100	Coherent	Mixed	2	2	0
3	4	L	YFP	100	Dispersed	Mixed	1	3	0
4	5	RCr	YFP	150	Dispersed	Mixed	1	4	0
5	6	L	YFP	150	Dispersed	Mixed	3	3	0
6	7	L	RFP	150	Dispersed	Mixed	2	5	0
7	7	RCr	YFP	150	Dispersed	Mixed	2	5	0
8	7	RCd	YFP	150	Dispersed	Mixed	2	5	0
9	7	RCr	YFP	200	Dispersed	Mixed	2	5	0
10	7	RAc	RFP	250	Dispersed	Mixed	1	6	0
11	8	L	RFP	150	Dispersed	Mixed	≥1	≥5	0
12	9	L	YFP	200	Dispersed	Mixed	2	7	0
13	10	L	RFP	150	Coherent	Mixed	≥2	≥7	0
14	11	RCr	RFP	100	Coherent	Mixed	3	8	0
15	11	L	YFP	250	Dispersed	Mixed	4	7	0
16	12	RCd	RFP	150	Coherent	Mixed	≥4	≥4	0
17	13	L	RFP	200	Coherent	Mixed	≥4	≥7	0
18	14	RCr	RFP	200	Dispersed	Mixed	2	12	0
19	15	RCr	RFP	250	Dispersed	Mixed	3	12	0
20	17	RCd	RFP	250	Dispersed	Mixed	4	9	4
21	18	RCr	RFP	300	Dispersed	Mixed	2	16	0
22	19	L	RFP	400	Dispersed	Mixed	≥1	≥5	9
23	23	L	YFP	200	Dispersed	Mixed	≥4	≥8	0
24	28	RCr	YFP	300	Dispersed	Mixed	≥2	≥10	0
25	29	RCd	RFP	350	Dispersed	Mixed	≥4	≥10	0
26	30	RCr	RFP	350	Dispersed	Mixed	≥3	≥20	0
27	37	L	RFP	250	Coherent	Mixed	≥8	≥20	0
28	40	L	YFP	350	Dispersed	Mixed	≥4	≥30	0
29	72	RCr	RFP	350	Dispersed	Mixed	≥1	≥20	5

g Emergence of alveolar capillary cell types — E16.5 — *Car4* — *Ednrb* — Air-space — *Apln* — *Merge* — E17.5 — *Car4* — *Ednrb* — Air-space — *Apln* — *Merge*

h Emergence and maturation — E12.5 — E18.5 — P0 — Adult — Pseudotime → — Plexus — Emerging/maturing aCap — Emerging/maturing gCap — Mature aCap — Mature gCap

i Molecular maturation — Plexus — Emerging/maturing aCap — Mature aCap — *Mki67, Top2a, Aplnr, Car4, Ednrb, Tbx3, Tbx2, Apln, Prdm1, Sirpa, Chst1, Chst2, Fibin, Serpine1, AW112010, Adamtsl5, Rprml, Scn7a, Cldn5* — Dev. time → — ln(UP10K+1) — 0 2 4 6

j Maturing gCap (P0-P7) — Mature gCap (adult) — *H2-Ab1* — *Cd74* — *H2-DMb1* — Pseudo-time — Ratio to max — 0 0.5 1

k Young adult lung — gCap — Vessel — Alveoli — *Vwf* — *Vwf* — *Vwf-Ptprb* — *Vwf Ptprb Cldn5* — **k'**

l Aged lung — Vessel — Alveoli — *Vwf* — gCap — aCap — *Vwf* — *Vwf Ptprb* — *Vwf Ptprb Cldn5* — **l'**

m Young capillaries — Aged capillaries — *Vwf* — *Vwf* — *Vwf Ptprb* — *Vwf Ptprb*

Extended Data Fig. 6 | See next page for caption.

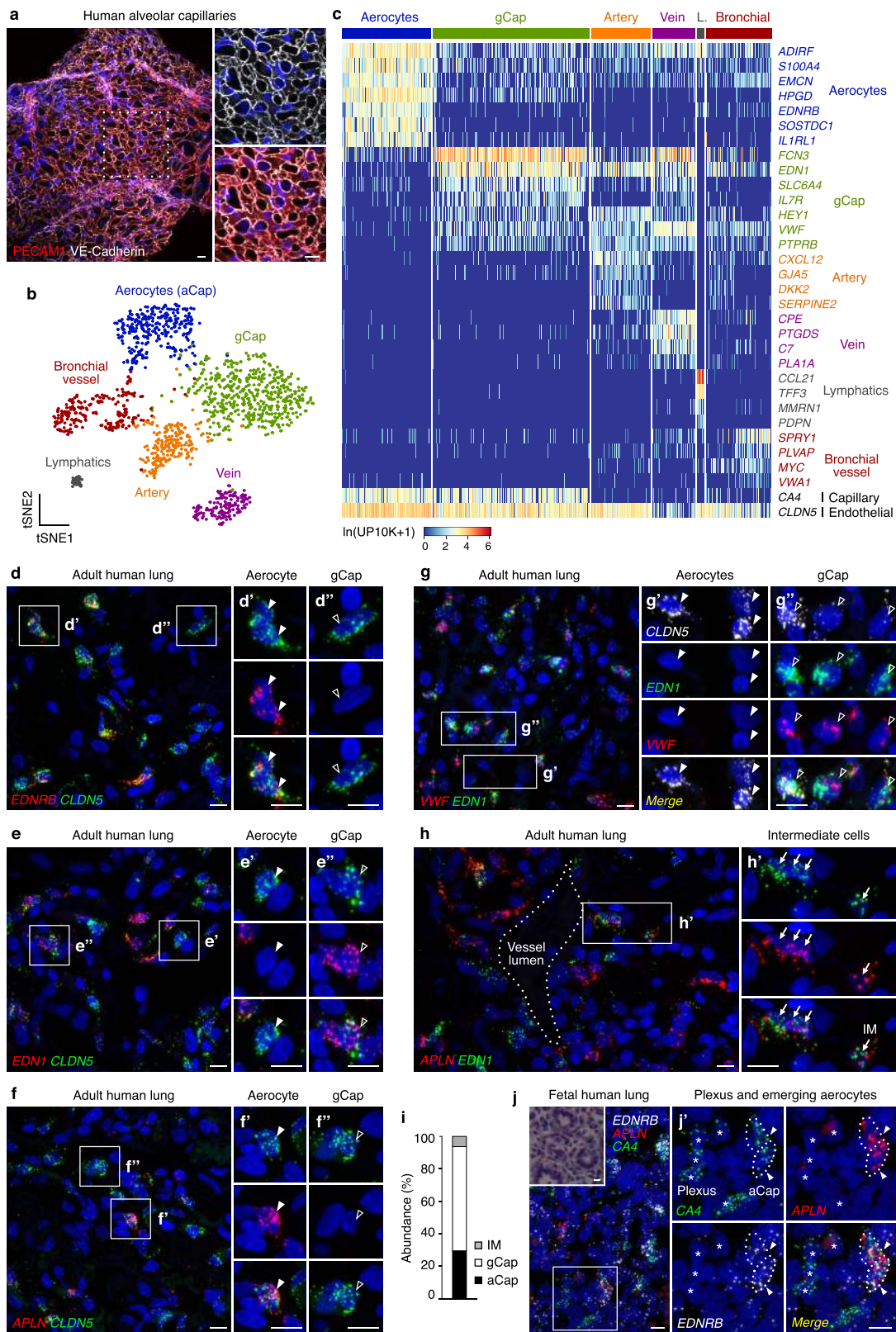**Extended Data Fig. 6 | Development, maturation and ageing of specialized alveolar capillary cell types. a**, Expression of *Aplnr* throughout plexus (labelled by *Pecam1*) surrounding a developing airway at E12.5, detected by smFISH. **b**, Expression of mature aerocyte markers *Apln* and *Ednrb* in plexus (labelled by *Pecam1*) surrounding a developing airway at E12.5. Plexus cells expressing *Apln* but not *Ednrb* are marked by asterisks. *Ednrb* is expressed at high levels in *Pecam1*-negative (stromal) cells, but rarely and only at low levels in *Pecam1*-positive endothelial cells (arrow) at this stage. **c**, The subpopulation of *Apln*-expressing cells within the *Aplnr*+ plexus gives rise to both capillary cell types, as shown by expression of *tdTomato* transcripts in both aerocyte (filled arrowheads, marked by *Ednrb*) and gCap (open arrowheads, marked by *Aplnr*) cells of P21 *Apln-creER;Rosa26-tdTomato* lung lineage-labelled at E12.5, demonstrating that even plexus cells that express an aerocyte marker are uncommitted. **c′**, Lineage-labelled aerocyte (filled arrowhead) and gCap (open arrowheads) cells shown at higher magnification. **d**, Alveolar capillary clone in a P25 *Aplnr-creER;Rosa26-Confetti* lung composed of both aCap and gCap cells (higher magnification at right) derived from a single YFP-expressing plexus cell labelled at E14.5. **e**, Composition of *Aplnr-creER;Rosa26-Confetti* clones induced at E14.5 and analysed at P25, ordered by clone size. All analysed clones, with the exception of clone number 1, contained both aCap and gCap cells. Some clones also contained cells located in larger vessels (LV). Both coherent clones (in which all cells are touching) and dispersed clones were observed, suggesting that there can be cell movement during capillary development. The number of cells of each type was scored from 3D renderings of confocal z-stacks, as described in the Methods. Clone 4 is shown in Fig. 4e; clone 28 is shown in Extended Data Fig. 6d and Supplementary Video 4. RCr, right cranial lobe; RAc, right accessory lobe; RCd, right caudal lobe; L, left lobe. **f**, The *Aplnr*+ population remains uncommitted even after birth. The population labelled at P7 in an *Aplnr-creER;Rosa26-tdTomato* lung gives rise to both aCap and gCap cells, as shown by co-expression of *tdTomato* lineage label and either aCap marker *Apln* or gCap marker *Aplnr* at P21, detected by smFISH. **f′**, **f″**, Lineage-labelled aerocyte and gCap cells shown at higher magnification. **g**, Expression of aerocyte markers *Car4*, *Ednrb* and *Apln* in developing mouse lung at indicated stages. Aerocytes (asterisks), identified by marker co-expression, begin to emerge at E17.5. Insets, cells in boxed regions at higher magnification. **h**, Plexus and capillary cells identified in scRNA-seq data for developing mouse lung[34] arranged as tree-shaped branched developmental trajectory inferred using Monocle2[36]. Cells collected at the indicated stages are coloured to show plexus located along the stem before the branchpoint, maturing and mature aerocytes along the bottom branch, and maturing and mature gCap cells along the top branch. Note that cells from a single late embryonic or early postnatal time point are found at multiple positions along the trajectory. This analysis is consistent with that presented in a previous study[15]. Each cell type undergoes distinct and asynchronous molecular and morphological maturation (see Extended Data Fig. 2), suggesting that the two capillary cell specialization programs are under separate genetic control. **i**, Heat map showing log-transformed expression levels in individual plexus and capillary cells for selected genes differentially expressed during aerocyte emergence and maturation. Many aerocyte markers are expressed in emerging aerocytes; others are only expressed in mature aerocytes. **j**, Developmental trajectory plot (top) and feature plots showing log-transformed expression for selected gCap markers expressed only in mature gCap cells (upper branch tip). **k**, **l**, Expression of *Vwf* in the gas-exchange region at 3 (**k**) and 24 months (**l**), as detected by smFISH. **k′**, **l′**, Boxed regions from **k** and **l** shown at higher magnification. Endothelial cells are marked by *Cldn5* expression. At 24 months, *Vwf* is induced in gCap cells (marked by *Ptprb*, open arrowheads), but not aerocytes (filled arrowhead). **m**, Schematic representation of *Vwf* induction in gCap cells with ageing. At 24 months, *Vwf* is more broadly expressed in alveolar capillaries than at 3 months but only in gCap cells (marked by *Ptprb*). Blue, DAPI (**a**–**c**, **f**, **g**, **k′**, **l′**). Scale bars, 10 μm.
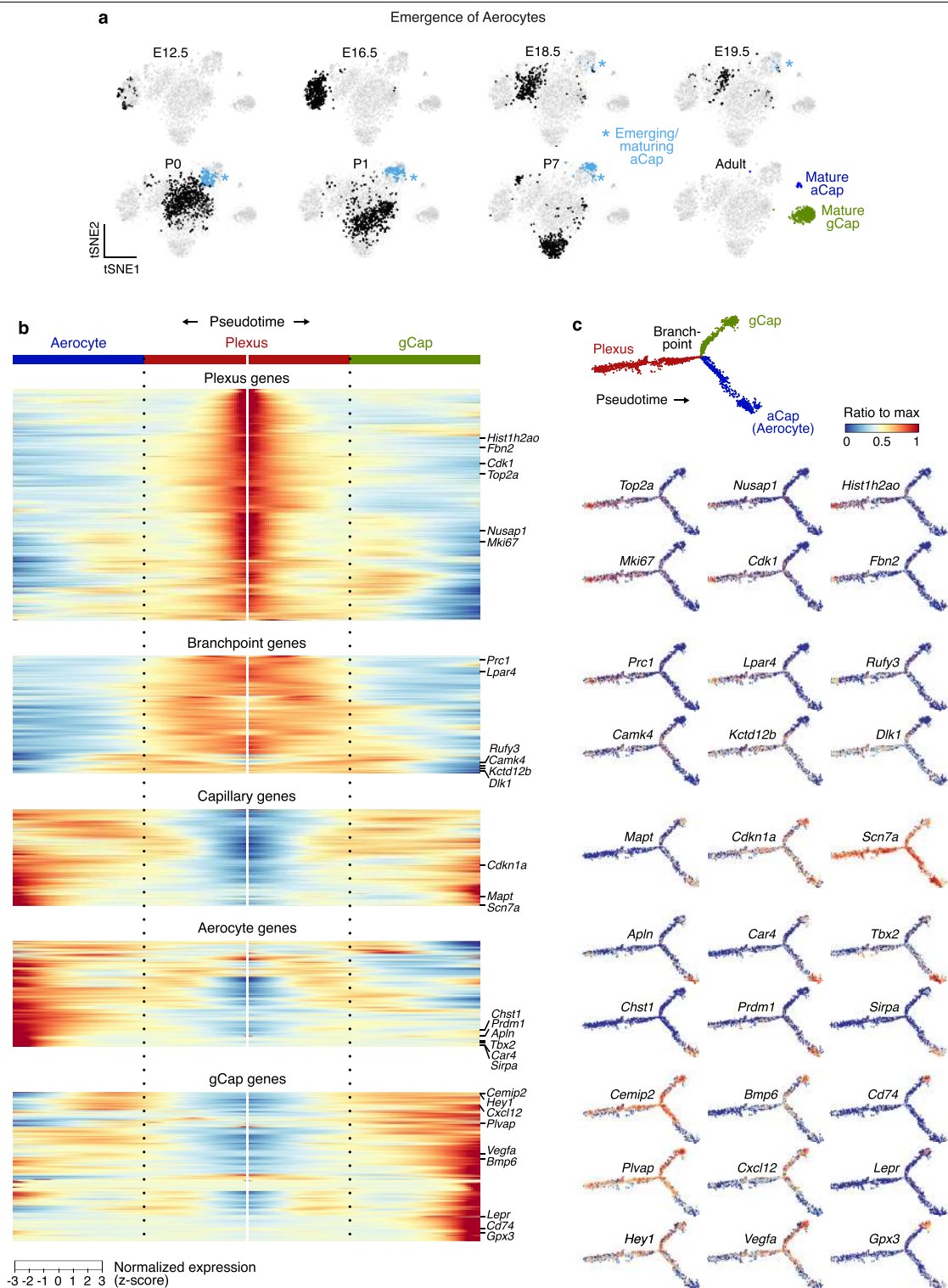
Extended Data Fig. 7 | Emergence and maturation of the alveolar capillary cell types in developing mouse lung. a, t-SNE plots with plexus and capillary endothelial cells from embryonic, postnatal and adult stages ($n = 3,094$ cells) identified in scRNA-seq MARS-Seq data for developing mouse lung[34]. Cells from each indicated stage are coloured black, except emerging/maturing aerocytes (aCap), first evident at E18.5 in this dataset, light blue; mature aerocytes, dark blue; mature gCap cells, dark green. See also feature plots in Supplementary Data 3. **b**, Branched heat map showing gene expression changes during the transformation of plexus into mature alveolar capillary cell types. Normalized expression values (z-scores) are plotted for genes with branch-dependent expression, identified by branched expression analysis modelling (BEAM) ($n = 1119$ genes at q-value < 0.05), and genes that vary as a function of pseudotime, identified by differential expression analysis as implemented in Monocle2 ($n = 3,734$ genes at q-value < 0.05), in plexus and capillary cells. Cells are ordered by ascending pseudotime values with plexus in the middle, mature aerocytes (aCap) on the left and mature gCap cells on the right. Genes ($n = 4,129$) are grouped by expression pattern and selected genes are indicated on the heat map. Branchpoint genes change their expression where the stem splits into aCap and gCap branches. Capillary genes are expressed in both aCap and gCap cells, but not plexus. An uncropped version of the heat map is shown in Supplementary Data 4. **c**, Developmental trajectory plot (top) and feature plots (below) showing log-transformed expression for selected genes in plexus (cells located along the stem before the branchpoint), aerocytes (bottom branch) and gCap cells (top branch).

a Human alveolar capillaries

PECAM1 VE-Cadherin

b

Aerocytes (aCap)

gCap

Bronchial vessel

Lymphatics

Artery

Vein

tSNE2

tSNE1

c

Aerocytes | gCap | Artery | Vein | L. | Bronchial

ADIRF
S100A4
EMCN
HPGD   Aerocytes
EDNRB
SOSTDC1
IL1RL1
FCN3
EDN1
SLC6A4
IL7R
HEY1   gCap
VWF
PTPRB
CXCL12
GJA5
DKK2   Artery
SERPINE2
CPE
PTGDS
C7   Vein
PLA1A
CCL21
TFF3   Lymphatics
MMRN1
PDPN
SPRY1
PLAVP
MYC   Bronchial vessel
VWA1
CA4 | Capillary
CLDN5 | Endothelial

ln(UP10K+1)
0 2 4 6

d Adult human lung | Aerocyte | gCap

EDNRB CLDN5

e Adult human lung | Aerocyte | gCap

EDN1 CLDN5

f Adult human lung | Aerocyte | gCap

APLN CLDN5

g Adult human lung | Aerocytes | gCap

CLDN5
EDN1
VWF
Merge

VWF EDN1

h Adult human lung | Intermediate cells

Vessel lumen

IM

APLN EDN1

i

Abundance (%)

100
80
60
40
20
0

▨ IM
☐ gCap
■ aCap

j Fetal human lung | Plexus and emerging aerocytes

EDNRB
APLN
CA4

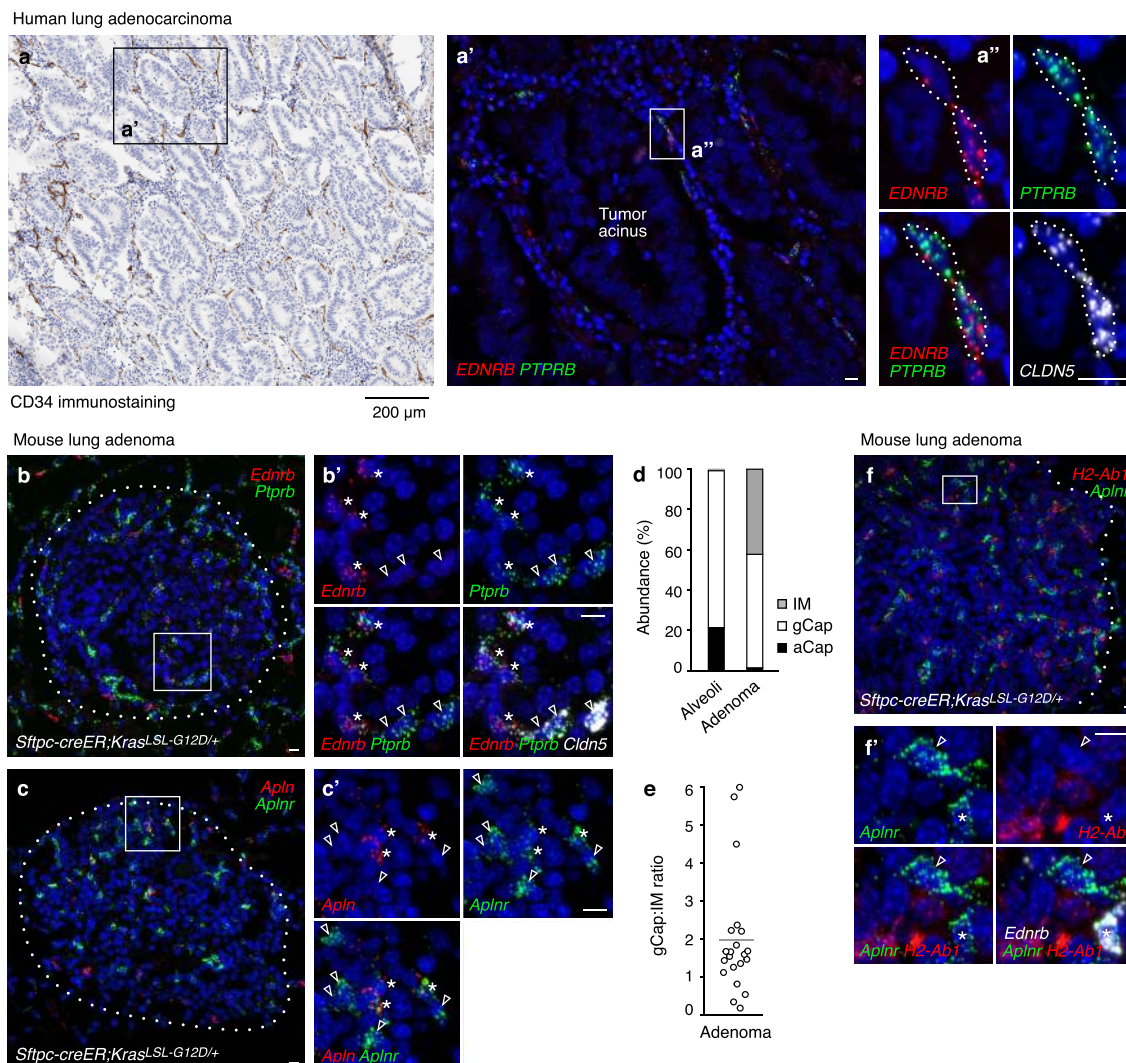Plexus   aCap

CA4   *APLN*

EDNRB   *Merge*

Extended Data Fig. 8 | See next page for caption.

Article

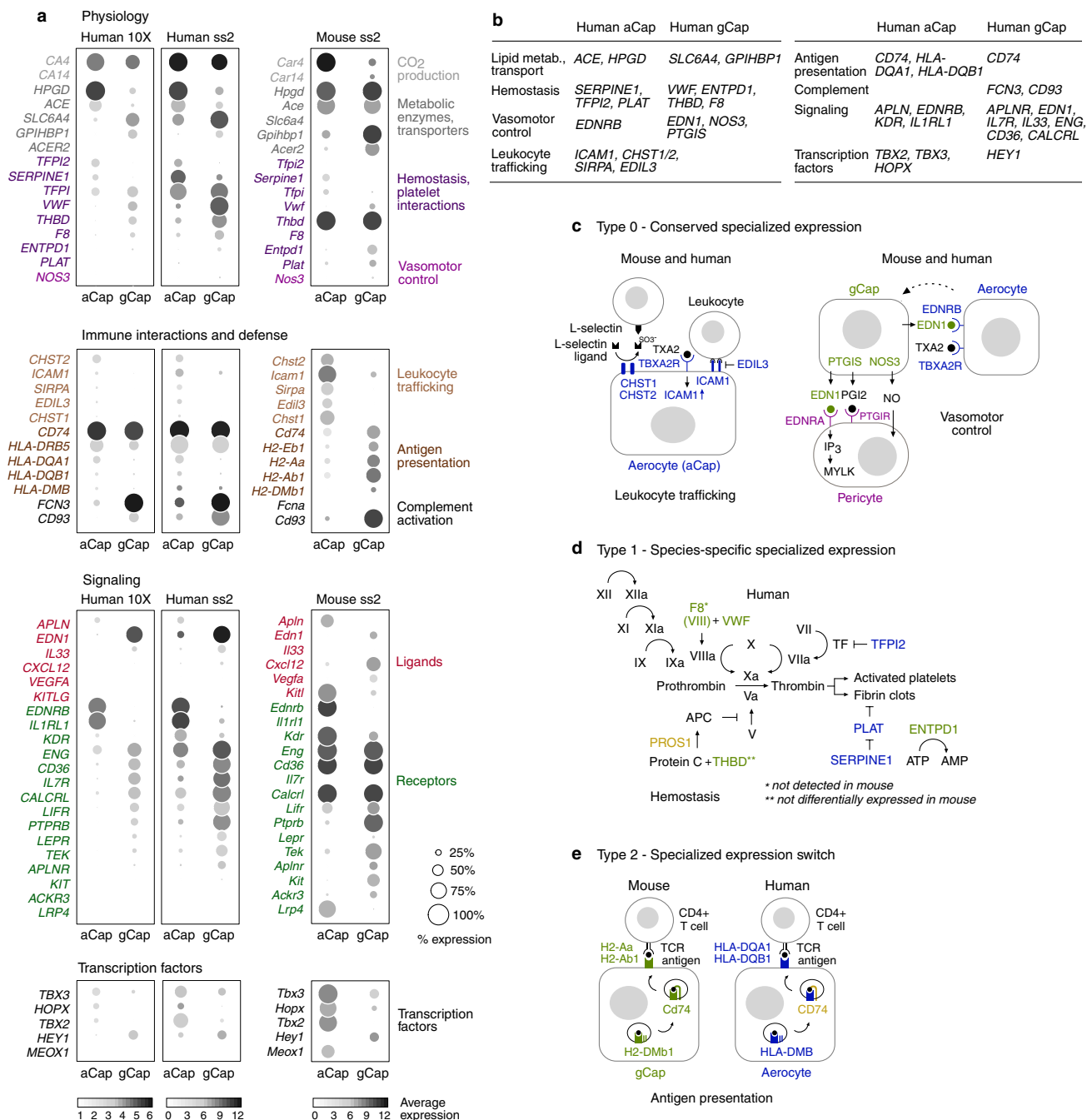**Extended Data Fig. 8 | Specialized alveolar capillary cell types in the human lung. a**, Alveolar capillary network in adult human lung immunostained for PECAM1 and VE-cadherin. Boxed region shown at higher magnification. **b**, t-SNE plot of annotated artery, vein, lymphatics, bronchial endothelial, aerocyte (aCap) and gCap cell clusters identified in scRNA-seq droplet data for adult human lung[16] (75-year-old man). **c**, Heat map showing log-transformed expression of selected cluster markers in individual endothelial cells. L., lymphatics. **d**–**g**, Expression of aerocyte (*EDNRB*, **d**, or *APLN*, **f**) and gCap (*EDN1*, **e**, **g** or *VWF*, **g**) markers in alveolar endothelial cells (marked by *CLDN5*, **d**–**f**), detected by smFISH in adult human lung. **d'**–**g'**, **d''**-**g''**, Aerocytes and gCap cells shown at higher magnification. **h**, As in the mouse lung (see Extended Data Fig. 1c), occasional human alveolar capillary cells co-express aerocyte (*APLN*) and gCap (*EDN1*) markers. **h'**, Boxed region shown at higher magnification with clustered capillary intermediate cells (arrows) near a large vessel (dotted line in **h**) and a nearby single cell. **i**, Quantification of the relative abundance (in % of capillary endothelial cells) of the two alveolar capillary populations (aCap, gCap) and intermediate (IM) cells that co-express aCap and gCap markers in adult human lung ($n = 2$ individuals; 69- and 75-year-old men; 500–600 capillary cells scored per lung; data as mean). **j**, **j'** Co-expression of aerocyte markers *EDNRB*, *APLN* and *CA4* in emerging aerocytes (dotted outlines) but not plexus cells (asterisks) in fetal human lung (23 weeks gestational age, corresponding to E16.5–17.5 in mouse[46]). At 23 weeks, 6% of *CA4*-expressing cells also express *APLN* and *EDNRB* at high levels (emerging aerocytes; 5 or more puncta per cell), compared to 0% at 17 weeks (500 *CA4*+ cells scored at each time point from one lung). Inset, adjacent haematoxylin and eosin (H&E)-stained section. **j'**, Boxed region from **j** at higher magnification. Blue (**a**, **d**–**h**, **j**), DAPI. Scale bars, 10 μm.

Human lung adenocarcinoma

CD34 immunostaining

200 µm

Tumor acinus

EDNRB *PTPRB*

EDNRB *PTPRB* *CLDN5*

Mouse lung adenoma

Ednrb *Ptprb*

Sftpc-creER;Kras^LSL-G12D/+

Ednrb *Ptprb* *Ednrb* *Ptprb* *Cldn5*

Apln *Aplnr*

Sftpc-creER;Kras^LSL-G12D/+

Apln *Aplnr*

Mouse lung adenoma

H2-Ab1 *Aplnr*

Sftpc-creER;Kras^LSL-G12D/+

Aplnr *H2-Ab1*

Aplnr *H2-Ab1* *Ednrb* *Aplnr* *H2-Ab1*

Extended Data Fig. 9 | Altered capillary cell patterns in mouse and human lung tumours. a, Human lung adenocarcinoma with tumour vessels immunostained for CD34 (brown), an endothelial marker that is expressed by both aCap and gCap cells[16]. Blue, haematoxylin counterstain. **a'**, Co-expression of alveolar capillary cell type markers *EDNRB* (aCap) and *PTPRB* (gCap) in vessels surrounding tumour acinus (boxed in **a**), as detected in adjacent section by smFISH. **a''**, Boxed region from **a'** showing two endothelial cells (marked by *CLDN5*) co-expressing aerocyte and gCap markers (dotted outlines) at higher magnification. Merged image also shown in Fig. 4l. **b**–**c'**, Co-expression (asterisks) of alveolar capillary cell type markers (*Ednrb*, **b**, **b'** or *Apln*, **c**, **c'** (aCap) and *Ptprb*, **b**, **b'** or *Aplnr*, **c**, **c'** (gCap)) in a subset of endothelial cells (marked by *Cldn5*, **b'**) of mouse lung adenomas induced by conditional expression of an activating *Kras* mutation in AT2 cells[6]. Open arrowheads point to gCap cells. **b'**, **c'**, Boxed regions in **b** and **c** shown at higher magnification.
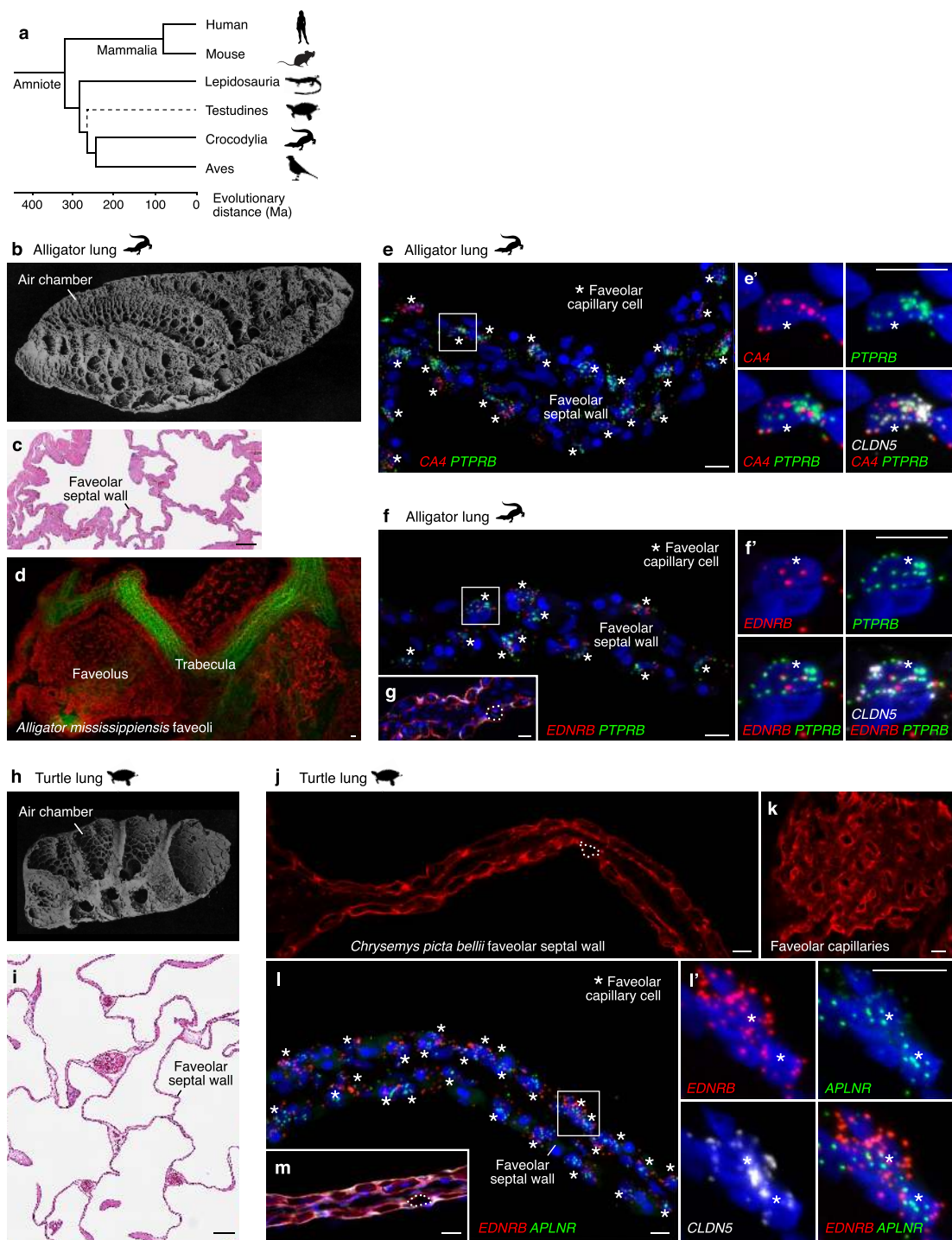
d, Quantification of the relative abundance (in % of capillary endothelial cells) of the two alveolar capillary populations (aCap, gCap) and intermediate cells (IM) that co-express aCap and gCap markers in adenomas (n = 1,332 cells scored in 21 tumour sections from n = 2 mice) and alveoli (n = 11,219 cells scored in n = 5 mice at 3 months). **e**, Scatter plot of gCap:IM ratios in individual adenomas (n = 21). Grey bar, mean value. The fraction of intermediate cells varies from tumour to tumour even within the same lung, perhaps reflecting different stages of tumour development. **f**, **f'**, Expression of MHC class II gene *H2-Ab1* (normally expressed in gCap cells; see Extended Data Fig. 1g) is lost in gCap cells (open arrowhead) in mouse adenomas, suggesting that they may lose their antigen presentation function. *H2-Ab1* is also not expressed by the abundant tumour capillary cells that co-express *Aplnr* and *Ednrb* (asterisk). **f'**, Boxed region in **f** shown at higher magnification. Scale bars, 200 µm (**a**); 10 µm (**a'**–**c'**, **f**, **f'**).

Extended Data Fig. 10 | Conserved and species-specific specialization of alveolar capillary cells. a, Dot plots showing log-transformed average expression levels and percent expression in aerocytes (aCap) and gCap cells for selected genes with known physiological and immune functions, ligands and receptors or co-receptors and transcription factors, which are differentially expressed by the two capillary cell types in adult mouse lung (Tabula Muris scRNA-seq Smart-Seq2 data[13]) or human lung (Human Lung Cell Atlas[16]; droplet (10X) or Smart-Seq2 data; 75-year-old man). **b**, Selected differentially expressed genes in the human lung with known functions. **c**–**e**, Diagrams of selected

proposed specialized alveolar capillary functions. Selected genes involved in leukocyte trafficking and vasomotor control[47] (see legend to Fig. 3d) show the same specialized expression pattern in mouse and human (**c**; type 0), whereas some genes involved in haemostasis show specialized expression only in the human lung (**d**; type 1), and some antigen-presenting genes[48] show specialized expression that switches cell type between mouse and human (**e**; type 2). Blue, genes expressed in aerocytes; green, genes expressed in gCap cells; yellow, genes expressed in both aerocytes and gCap cells; purple, genes expressed in pericytes[16].

a

Amniote

Mammalia
— Human
— Mouse

— Lepidosauria
--- Testudines
— Crocodylia
— Aves

Evolutionary distance (Ma)
400 300 200 100 0

b Alligator lung

Air chamber

c

Faveolar septal wall

d

Faveolus   Trabecula

Alligator mississippiensis faveoli

e Alligator lung

★ Faveolar capillary cell

Faveolar septal wall

CA4 PTPRB

e'

CA4   *PTPRB*

CA4 PTPRB   *CLDN5 CA4 PTPRB*

f Alligator lung

★ Faveolar capillary cell

Faveolar septal wall

g

EDNRB PTPRB

f'

EDNRB   *PTPRB*

EDNRB PTPRB   *CLDN5 EDNRB PTPRB*

h Turtle lung

Air chamber

i

Faveolar septal wall

j Turtle lung

Chrysemys picta bellii faveolar septal wall

k

Faveolar capillaries

l

★ Faveolar capillary cell

Faveolar septal wall

m

EDNRB APLNR

l'

EDNRB   *APLNR*

CLDN5   *EDNRB APLNR*

Extended Data Fig. 11 | See next page for caption.

Article

Extended Data Fig. 11 | Faveolar capillary cell types in alligator and turtle.
a, Cladogram illustrating the phylogenetic relationships between amniote taxa with different lung structures. The relationship between testudines and other diapsids is unresolved (dotted line). Ma, million years. **b**, **h**, Internal anatomy of alligator (*A. mississippiensis*; **b**) and turtle (*Emys orbicularis*; **h**) lungs, reproduced from a classical study[49]. Alligator and turtle lungs have multiple separate airspaces ('chambers'), smooth muscle, a branched vasculature and terminal airspaces known as faveoli, in which gas exchange occurs across a thick air–blood barrier[50,51]. **c**, **i**, Gas-exchange regions of American alligator (*A. mississippiensis*; **c**) and western painted turtle (*C. p. bellii*, **i**) lungs, stained with haematoxylin and eosin. **d**, Faveolar capillary network in American alligator lung immunostained for CLDN5 (red; endothelium). Green, elastin fibres. **e**, Co-expression (asterisks) of alveolar capillary cell type markers *CA4* (aCap) and *PTPRB* (gCap) in faveolar endothelial cells (marked by *CLDN5*) in juvenile American alligator lung as detected by smFISH. **e′**, Faveolar capillary cell in boxed region from **e** shown at higher magnification. Note that faveolar capillary cells—reminiscent of the intermediate cells found in adult mouse and human lungs—appear to differ in gene expression from the developmental precursors identified in the mouse vascular plexus, in which aCap and gCap markers are only rarely co-expressed (Supplementary Data 2). **f**, Co-expression (asterisks) of alveolar capillary cell-type markers *EDNRB* (aCap) and *PTPRB* (gCap) in faveolar endothelial cells (marked by *CLDN5*) in adult American alligator lung. **f′**, Faveolar capillary cell in boxed region from **f** shown at higher magnification. **g**, Alligator faveolar septum immunostained for CLDN5 (red; endothelium) and E-cadherin (white; epithelium). Dotted line, capillary lumen. **j**, **k**, Faveolar capillary network in western painted turtle lung immunostained for CLDN5 (red) to label endothelial cells. Neighbouring faveoli are separated by septa with a capillary layer on each side of the septal wall (**j**). A capillary lumen is outlined (**j**, dotted line). **l**, Co-expression (asterisks) of alveolar capillary cell type markers (*EDNRB*, aCap and *APLNR*, gCap) in faveolar endothelial cells (marked by *CLDN5*) in adult turtle lung. **l′**, Two faveolar capillary cells in boxed region from **l** shown at higher magnification. **m**, Faveolar septum in turtle lung immunostained for CLDN5 (red) to label endothelial cells in the capillary layer on each side of the septal wall and E-cadherin (white) to label respiratory epithelium. A capillary lumen is outlined (dotted line). Blue, DAPI. Scale bars, 200 μm (**c**, **i**), 10 μm (**d–g**, **j–m**).

nature research

Corresponding author(s):	Mark Krasnow, Ross Metzger
Last updated by author(s):	07/10/2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Zen 2.3 SP1 (Carl Zeiss) and MRC-1000 Laser Scanning Confocal Imaging System software (LaserSharp OS/2 version; Bio-Rad) were used to collect confocal images. Electron micrographs were collected using DigitalMicrograph version 2.30.542.0 (Gatan).
Data analysis	Confocal image processing and analysis was performed using freely (ImageJ 1.440 [NIH]) or commercially available software (Zen 2.3 SP1 [Carl Zeiss], Imaris 9.3.0 [Bitplane], Volocity [Quorum Technologies]). QuPath v0.2.0 was used to process images of scanned slides. Filtering, clustering, and annotation of cells from published scRNAseq datasets was performed using R software package Seurat (version 2.3). Single cell trajectories were constructed using Monocle2. Electron micrographs were pseudocolored using Illustrator CC (Adobe, version 22.1).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The scRNAseq datasets analyzed are available in the GEO repository under the accession numbers GSE109774 (Tabula Muris, https://tabula-muris.ds.czbiohub.org), GSE132042 (Tabula Muris Senis, https://tabula-muris-senis.ds.czbiohub.org), and GSE119228 (Cohen et al.), or on Synapse under the accession number syn21041850 (Human Lung Cell Atlas, https://hlca.ds.czbiohub.org). Source data used in all graphs are provided in the Source Data file.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculations were performed. Sample size was determined on the basis of previous experience and the nature of the experiment. Where possible, multiple lobes and regions within them were analyzed, and large numbers of cells scored to detect potentially rare events such as cell proliferation.
Data exclusions	No data were excluded.
Replication	For mouse experiments, biological replicates using different animals were performed, except for capillary stability/commitment experiments where only a single animal was available for each time point and multiple lung regions were examined. Single cell data for human lung from three individuals were used and validated by in situ hybridization in tissue from three other individuals. In situ hybridization experiments were performed on tissue from two alligators (one juvenile, one adult) and two adult turtles. Due to limited tissue availability, in situs were performed on tissue from two fetal human lungs (one at each of two different developmental stages); cells in multiple regions were scored. Consistent results were obtained betwen all replicates.
Randomization	Mice of the appropriate genotype or age were allocated into experimental groups (control vs elastase injury) at random. Only female animals or embryos were used for experiments with the Apelin-CreER knock-in line, since Apelin is X-linked. For other experiments, randomization was not relevant as the experiments did not involve allocation of samples into control vs treatment groups.
Blinding	Due to the nature of sample collection, and the sometimes obvious differences between samples, blinding was not used, except in rare cases noted in the Methods in which samples were scored blind.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Primary antibodies used: CD34 (BD Biosciences, 347660); Claudin5 (Abcam, ab53765); E-Cadherin (BD Biosciences, 610181); Endomucin (Invitrogen, eBioV.7C7, 14-5851-82); Integrin alpha8 (R&D, AF4076); Pecam1 (rat anti-mouse; BD Biosciences, 553370); Pecam1 (mouse anti-human; R&D, BBA7); tdTomato (Rockland, 600-401-379); VE-Cadherin (R&D, AF938). Secondary antibodies used: Donkey anti-goat IgG, Alexa Fluor 568 conjugated (Invitrogen, A11057); Horse anti-mouse IgG, peroxidase conjugated (Vector Laboratories, PI-2000); Goat anti-rabbit IgG, peroxidase conjugated (Vector Laboratories, PI-1000); Goat anti-rabbit IgG, Alexa 568 conjugated (Invitrogen, A11036); Goat anti-rat IgG, Alexa 488 conjugated (Invitrogen, A11006), for embryonic lung; Donkey, anti-rat IgG, Alexa 647 conjugated (Jackson Immunoresearch, 712-605-153); Goat anti-rat IgG, biotin conjugated (Vector Laboratories, BA-9401; 1:250), for embryonic lung; Goat anti-rat IgG, peroxidase conjugated (Vector Laboratories, PI-9401), for adult lung.
Validation	The following primary antibodies have been validated for use in mouse or human, as appropriate, by the manufacturers. See the manufacturers' websites (listed below) for details. In our experiments, stainings with these antibodies were consistent with the expected patterns. CD34 (BD Biosciences, 347660):https://www.bdbiosciences.com/eu/applications/research/clinical-research/oncology-research/

blood-cell-disorders/surface-markers/human/purified-mouse-anti-human-cd34-my10/p/347660
Endomucin (Invitrogen, eBioV.7C7, 14-5851-82): https://www.thermofisher.com/antibody/product/Endomucin-Antibody-clone-eBioV-7C7-V-7C7-Monoclonal/14-5851-82
Integrin alpha8 (R&D, AF4076): https://www.rndsystems.com/products/mouse-rat-integrin-alpha8-antibody_af4076
Pecam1 (rat anti-mouse; BD Biosciences, 553370): https://www.bdbiosciences.com/ds/pm/tds/553370.pdf
Pecam1 (mouse anti-human; R&D, BBA7): https://www.rndsystems.com/products/human-cd31-pecam-1-antibody-9g11_bba7
tdTomato (Rockland, 600-401-379): https://rockland-inc.com/store/Antibodies-to-GFP-and-Antibodies-to-RFP-600-401-379-O4L_24299.aspx
VE-Cadherin (R&D, AF938): https://www.rndsystems.com/products/human-ve-cadherin-antibody_af938

The following primary antibodies were used to stain turtle and/or alligator tissue. Although they have not been validated specifically for use in these species, they have been used extensively in other species, and in our experiments, stainings with these antibodies were consistent with the expected patterns.
Claudin5 (Abcam, ab53765): https://www.citeab.com/antibodies/723501-ab53765-anti-claudin-5-antibody
E-Cadherin (BD Biosciences, 610181): https://www.bdbiosciences.com/us/applications/research/stem-cell-research/cancer-research/human/purified-mouse-anti-e-cadherin-36e-cadherin/p/610181

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Mice were housed and bred in the animal facility at Stanford University in accordance with Institutional Animal Care and Use Committee (IACUC) guidance, and were maintained on a 12hr light/dark cycle with food and water provided ad libitum. The following mouse strains were used: C57BL/6 (C57BL/6NCrl, Charles River Laboratories, strain code: 027) was the wild-type strain. Apln-CreER (Aplntm1.1(cre/ERT2)Bzsh) (provided by Bin Zhou), Aplnr-CreER (Tg(Aplnr-cre/ERT2)#Krh) (provided by Kristy Red-Horse), Cdh5-CreER (Tg(Cdh5-cre/ERT2)1Rha) (provided by Ralf Adams), and Sftpc-CreER (Sftpctm1(cre/ERT2,rtTA)Hap) (provided by Harold Chapman) were used for conditional expression of Cre recombinase. Rosa26-tdTomato (Gt(ROSA)26Sortm14(CAG-tdTomato)Hze) (The Jackson Laboratory, strain #007914) and Rosa26-Confetti (Gt(ROSA)26Sortm1(CAG-Brainbow2.1)Cle) (The Jackson Laboratory, strain #017492) were used as Cre reporters. KrasLSL-G12D (Krastm4Tyj/J) (The Jackson Laboratory, strain # 008179) was used to express a constitutively active form of KRAS from the endogenous locus following Cre-mediated recombination. Both male and female mice were used, except only female animals/embryos were used for experiments with Apelin-CreER, since Apelin is X-linked. Early postnatal (P0-P7) and adult (2-6 months, 24 months) mice were used.
Wild animals	Alligators (Alligator mississippiensis; male; exact age unknown) were captured by hand by the staff of the Louisiana Department of Wildlife and Fisheries at the Rockefeller Wildlife Refuge (Grand Chenier, Louisiana), and transported to the University of Utah in an air conditioned vehicle. Turtles (Chrysemys picta bellii; male; exact age unknown) were captured by hand and received from a commercial supplier (The Turtle Source, Florida; http://www.theturtlesource.com) and transported to the University of Utah by a combination of air and land travel through the United Parcel Service. Animals were housed at the University of Utah and euthanized using methods of euthanasia conforming to the recommendations published by the AVMA Panel on Euthanasia and determined in consultation with a clinical veterinarian at the University of Utah.
Field-collected samples	No field-collected samples were used.
Ethics oversight	Mouse experiments were approved by the Stanford University Institutional Animal Care and Use Commitee; experiments using alligators and turtles were approved by the University of Utah Institutional Animal Care and Use Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Normal human adult lung tissue from 69 and 75 year old males, 66 year old female; aborted human fetal lung tissue (17 and 23 weeks); well-differentiated invasive lung adenocarcinoma from a 41 year old female.
Recruitment	De-identified normal human adult lung tissue was obtained from the Stanford Tissue Bank. De-identified aborted human fetal lung tissue was obtained through a collaboration with the Stanford Family Planning Research Team, Department of Obstetrics & Gynecology, Division of Family Planning Services and Research, Stanford University School of Medicine. De-identified human tissue representing well-differentiated invasive lung adenocarcinoma was obtained from archival diagnostic material in collaboration with the Stanford Department of Pathology, Stanford University School of Medicine.
Ethics oversight	Tissue collection and use in research were approved by the Stanford Institutional Review Board.

Note that full information on the approval of the study protocol must also be provided in the manuscript.