

# A molecular atlas of cell types and zonation in the brain vasculature

Michael Vanlandewijck<sup>1,3\*</sup>, Liquan He<sup>2\*</sup>, Maarja Andaloussi Mäe<sup>3</sup>, Johanna Andrae<sup>3</sup>, Koji Ando<sup>3</sup>, Francesca Del Gaudio<sup>4</sup>, Khayrun Nahar<sup>3</sup>, Thibaud Lebouvier<sup>3,5</sup>, Bárbara Laviña<sup>3</sup>, Leonor Gouveia<sup>3</sup>, Ying Sun<sup>6</sup>, Elisabeth Raschperger<sup>1</sup>, Markus Räsänen<sup>7</sup>, Yvette Zarb<sup>8</sup>, Naoki Mochizuki<sup>9,10</sup>, Annika Keller<sup>8</sup>, Urban Lendahl<sup>4</sup> & Christer Betsholtz<sup>1,3</sup>

**Cerebrovascular disease is the third most common cause of death in developed countries, but our understanding of the cells that compose the cerebral vasculature is limited. Here, using vascular single-cell transcriptomics, we provide molecular definitions for the principal types of blood vascular and vessel-associated cells in the adult mouse brain. We uncover the transcriptional basis of the gradual phenotypic change (zonation) along the arteriovenous axis and reveal unexpected cell type differences: a seamless continuum for endothelial cells versus a punctuated continuum for mural cells. We also provide insight into pericyte organotypicity and define a population of perivascular fibroblast-like cells that are present on all vessel types except capillaries. Our work illustrates the power of single-cell transcriptomics to decode the higher organizational principles of a tissue and may provide the initial chapter in a molecular encyclopaedia of the mammalian vasculature.**

The profound role of vascular dysfunction in ischaemic stroke and the vascular-specific expression of genes that are mutated in individuals with neurological diseases illustrate the critical importance of vascular health for brain function<sup>1,2</sup>. The brain vasculature harbours a unique specialization—the blood-brain barrier (BBB)—that is necessary for neuronal function<sup>3,4</sup> but represents a hurdle for the pharmacological treatment of brain diseases<sup>5,6</sup>. While specialization of the vascular endothelium is known to be a key feature of the BBB, the contribution of vascular mural cells (pericytes and smooth muscle cells (SMCs)) and vessel-associated cell types such as astrocytes is less clear.

The brain vasculature displays an arteriovenous hierarchy similar to those of other vascular beds. Differences in blood flow, pressure and chemical composition are paralleled along the arteriovenous axis by morphological and functional differences, such as junction structure and permeability, that have been elucidated using electron microscopy<sup>7,8</sup>. Systematic efforts to unravel the molecular basis of vascular arteriovenous specialization have, to our knowledge, not been undertaken, and only a handful of endothelial arteriovenous markers have been identified in various species, organs and developmental stages<sup>9–13</sup>. Mural arteriovenous differentiation involves distinct cellular morphologies ranging from the multiple-layered SMC coats of large arteries to the singular, longitudinally oriented or stellate pericytes of capillaries and venules. However, as for endothelial cells, the molecular underpinnings of these differences are largely unknown. Whereas the contractility of vascular SMCs is crucial for the regulation of blood pressure and flow, the physiological role of pericytes is unclear<sup>14</sup>. Pericyte identification is problematic owing to a lack of specific markers<sup>14</sup>, and criteria for pericyte definition remain debated<sup>15</sup>.

Considering the vast medical and pharmacological importance of the brain vasculature and the lack of a molecular understanding of its

constituent cell types, we set out to transcriptionally profile the principal cell types of the brain vasculature using single-cell RNA sequencing (scRNA-seq).

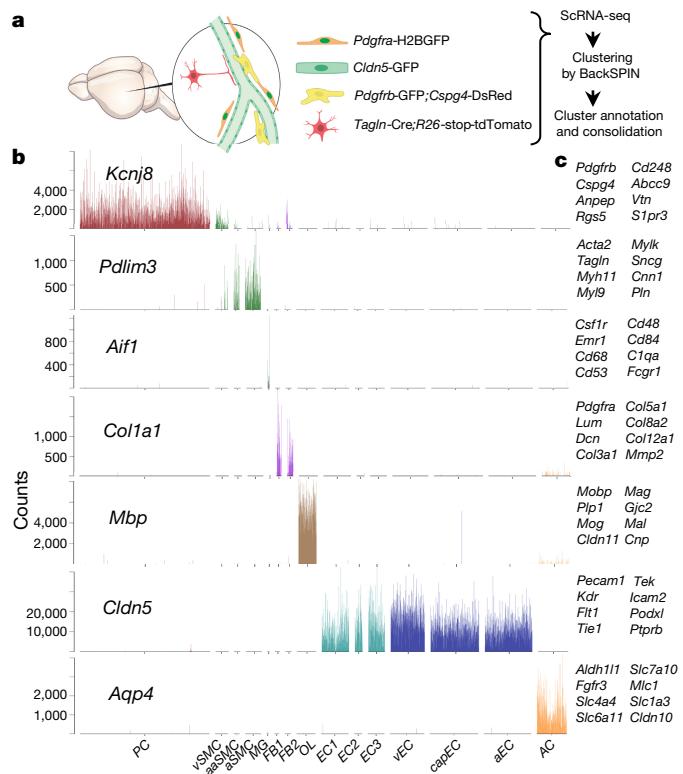
## Cell classes of the brain vasculature

We used a set of transgenic reporter mice to capture all major vascular and vessel-associated cell types from the adult brain (Fig. 1a). Single cells were sorted into 384-well plates (Extended Data Fig. 1a) and profiled using SmartSeq2<sup>16</sup>. About 3,500 single cell transcriptomes were clustered using the BackSPIN clustering algorithm<sup>17</sup> and annotated (Extended Data Fig. 1b–m), leading to the cluster map shown in Fig. 1b. Annotations were guided by the expression of canonical cell class markers, as exemplified in Fig. 1c. Primary bar-plot data and average expression scores for the cell types (Extended Data Fig. 2a) are available gene-by-gene at <http://betsholtzlab.org/VascularSingleCells/database.html>. This dataset provides molecular definitions for the major vascular and perivascular cell classes of the adult mouse brain. In addition to endothelial and mural cells, we obtained clusters corresponding to astrocytes, oligodendrocyte precursors (*Pdgfra*<sup>+</sup>), intermediate stage oligodendrocytes (*Opalin*<sup>+</sup>)<sup>18</sup>, perivascular (*Mrc1*<sup>+</sup>) and interstitial (*Cx3cr1*<sup>+</sup>) microglia<sup>17,19</sup> and two types of fibroblast-like cells (Extended Data Fig. 2b–d).

## Endothelial arteriovenous zonation

Gradual cellular phenotypic changes along an anatomical axis are referred to as cellular ‘zonation’<sup>20</sup>. Three of the endothelial clusters (venous (vEC), capillary (capEC) and arterial (aEC)), containing a total of 1,100 endothelial cells, showed a biased distribution of known arteriovenous markers and were re-ordered by SPIN (sorting points into neighbourhoods) into a single one-dimensional range. We found

<sup>1</sup>Karolinska Institutet/AstraZeneca Integrated Cardio Metabolic Centre (KI/AZ ICMC), Blickagången 6, SE-141 57 Huddinge, Sweden. <sup>2</sup>Department of Neurosurgery, Tianjin Medical University General Hospital, Tianjin Neurological Institute, Key Laboratory of Post-Neuroinjury Neuro-Repair and Regeneration Central Nervous System, Ministry of Education and Tianjin City, Tianjin 300052, China. <sup>3</sup>Department of Immunology, Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Dag Hammarskjölds väg 20, SE-751 85 Uppsala, Sweden. <sup>4</sup>Department of Cell and Molecular Biology, Karolinska Institutet, Von Eulers väg 3, SE-171 77 Stockholm, Sweden. <sup>5</sup>Inserm U1171, University of Lille, CHU, Memory Center, Distalz, F-59000 Lille, France. <sup>6</sup>Department of Bioinformatics, Zhongyuan Union Genetic Technology Co., Ltd., No.45, the 9th East Road, Tianjin Airport Economic Area, Tianjin 300304, China. <sup>7</sup>Wihuri Research Institute and Translational Cancer Biology Program, Biomedicum Helsinki, University of Helsinki, Haartmaninkatu 8, P.O. Box 63, FI-00014 Helsinki, Finland. <sup>8</sup>Division of Neurosurgery, Zürich University Hospital, Zürich University, Zürich, CH-8091, Switzerland. <sup>9</sup>Department of Cell Biology, National Cerebral and Cardiovascular Center Research Institute, Suita, Osaka, Japan. <sup>10</sup>AMED-CREST, National Cerebral and Cardiovascular Center, Suita, Osaka, Japan.  
\*These authors contributed equally to this work.



**Figure 1 | Procedures and cell clusters.** **a**, Method flowchart. **b**, Bar plots of representative cell type-specific markers. See Extended Data Fig. 1 for cell type abbreviations. **c**, More cell type-specific markers used for cluster annotation. See <http://betsholtzlab.org/VascularSingleCells/database.html> for high-resolution images and statistics.

that known arterial<sup>9–12</sup> (*Bmx*, *Efnb2*, *Vegf* and *Sema3g*) and venous<sup>13</sup> (*Nr2f2*) markers peaked at opposite ends of this range (Fig. 2a), with gradual changes in expression that suggested zonation. A total of 1,798 transcripts were significantly differentially expressed across the range (Supplementary Table 1), with the 500 most significant ones (Supplementary Table 2) distributed into six patterns with putative arterial, venous or capillary predominance, or combinations thereof (Fig. 2b). Again, the nested patterns of gradually changing gene expression in the cells across the range were indicative of cellular zonation.

To confirm that the SPIN range represented a cell order matching the arteriovenous axis, we analysed the tissue expression of eight selected markers. *Bmx* was confirmed as a brain arterial marker using *Bmx-lacZ* mice (Extended Data Fig. 3a). *Mfsd2a*, encoding a BBB-specific lipid transporter<sup>21</sup>, peaked at the middle of the SPIN range (Fig. 2a), matching the concentration of MFSD2A protein in capillary endothelial cells (Fig. 2c, d and Extended Data Fig. 3b). *Tfrc* (encoding the transferrin receptor (TFRC)) peaked in the middle-left part of the SPIN range (Fig. 2a), matching the TFRC staining of capillaries and veins, but not arteries (Fig. 2c, d and Extended Data Fig. 3c). Similar correlations were found between the expression of *Vwf*, *Vcam1* (artery and vein, but not capillary) and *Slc16a1* (capillary–venous) transcripts and the localization of the corresponding proteins von Willebrand factor (VWF), vascular cell adhesion molecule 1 (VCAM1) and solute carrier 16a1 (SLC16A1) to different vessel types (Fig. 2a, c, d and Extended Data Figs 3d, 4a, b). In addition, RNA *in situ* hybridization (ISH) was used to localize *Gkn3* and *Slc38a5* expression to arteries and veins, respectively, consistent with the SPIN distribution (Fig. 2a and Extended Data Fig. 4c–e). The arteriovenous zonation of these transcripts was also consistent with ISH data deposited in the Allen Brain Atlas<sup>22</sup>, which also validated the zonal expression of *Slc6a6* and *Slc7a5* (Extended Data Fig. 5). Collectively, the gradual changes in expression and nested expression patterns of groups of genes across the SPIN range

suggest that endothelial arteriovenous zonation represents a seamless continuum of transcriptional states (schematically illustrated in Extended Data Fig. 6a).

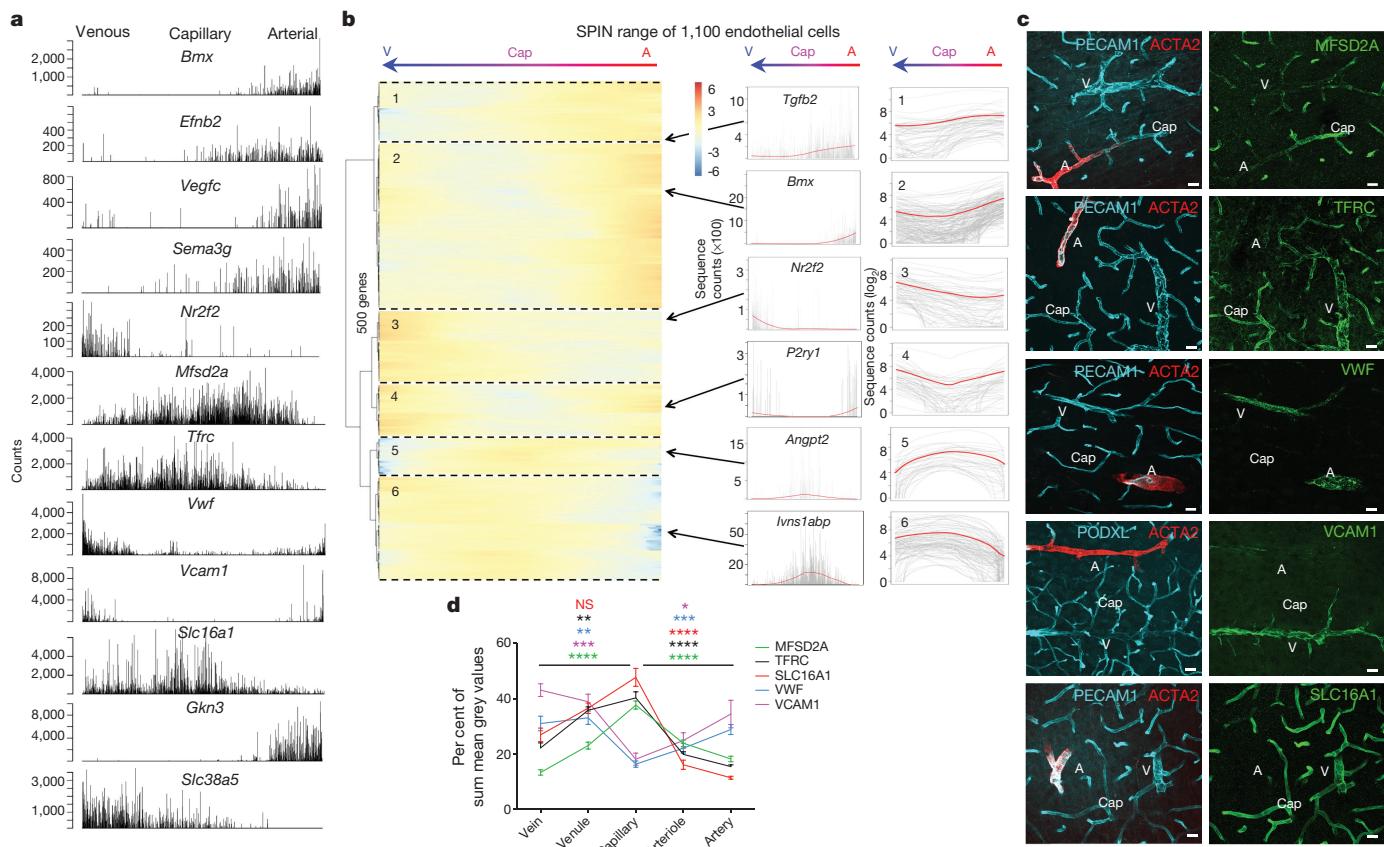
As examples of how specific gene or protein classes distribute along the arteriovenous axis, we found zonal and nested distribution of 120 transcription factor transcripts and 145 transmembrane transporter transcripts (Extended Data Fig. 6b, c), providing insight into transcriptional regulation of arteriovenous zonation and how this results in endothelial specialization. We noted a particular difference in distribution: whereas transcription factors were overrepresented at arterial locations, transporters dominated in capillaries and veins, indicating a concentration of BBB-associated trans-endothelial molecular transport functions at the latter locations.

BackSPIN also assigned endothelial cell clusters that fell outside the arteriovenous zone. These clusters (EC1–EC3) (Fig. 1b) showed high expression of ribosomal protein transcripts of the Rpl and Rps classes, suggesting that these cells are (more) active in protein synthesis. Whether EC1–EC3 cells are located in a specific brain region or distributed evenly across brain regions is currently unclear. The presence of arteriovenous markers in all EC1–EC3 clusters suggests that these cells are distributed across all vessel types.

### Mural arteriovenous zonation

To study mural cell zonation, we first compared the distribution of known mural cell transcripts across the BackSPIN clusters with the corresponding protein expression *in situ*. *Cnn1*, which was specific to the arterial SMC (aSMC) cluster (Fig. 3a), matched the localization of CNN1 (calponin 1) protein to arteries with diameters larger than 13 µm (Fig. 3b and Extended Data Fig. 6d). *Acta2* and *Tagln*, which were highly expressed in the aSMC and arteriole (aa)SMC clusters (Fig. 3a), matched the expression of α-smooth muscle actin (ACTA2) and smooth muscle protein 22-α (TAGLN) in arteries and arterioles with diameters larger than 8 µm (Extended Data Figs 6d, 7a). Expression of *Acta2* and *Tagln* was weak in the venous (v)SMC cluster and almost absent from the PC (pericyte) cluster, consistent with weak staining for ACTA2 and TAGLN in large veins and undetectable levels in capillaries and venules (Extended Data Fig. 7a). The broad distribution of *Cspg4* and *Pdgfrb* across all mural clusters (Fig. 3a) correlated with the broad expression of DsRed driven by the *Cspg4* promoter (*Cspg4-DsRed*) and GFP driven by the *Pdgfrb* promoter (*Pdgfrb-GFP*) across all vessel types (Extended Data Fig. 7b). Consistent with the mouse data, we found zebrafish *tagln:EGFP* predominantly in brain arteries and arterioles, whereas *pdgfrb*-driven GFP was detected in all vessels (Extended Data Fig. 8a), suggesting that mural cell zonation has been evolutionarily conserved. Consistent with the specific expression of *Abcc9* in the mouse PC and vSMC clusters, the newly established zebrafish line *TgBAC(abcc9:Gal4FF)* showed specific labelling of brain mural cells in capillaries and veins (Fig. 3b and Extended Data Fig. 8b).

Figure 3c illustrates the distribution of the 1,708 most variable mural transcripts (Supplementary Table 3) across 1,385 mural cells arranged by SPIN. Notably, the order of mural cell types according to transcriptional relatedness (PC–vSMC–aaSMC–aSMC) did not match the anatomical organization along the arteriovenous axis. Moreover, hierarchical clustering indicated two predominant gene expression patterns: one cluster of transcripts that was expressed highly in the PC and vSMC clusters but at low levels in the aaSMC and aSMC clusters, and another cluster of transcripts with the opposite pattern. This suggests that there are two distinct subclasses of mural cells (Fig. 3c and Extended Data Fig. 6e): one in which pericytes occur in a continuum with vSMC through gradual loss of PC markers and acquisition of (low levels of) SMC markers, and the other in which aaSMCs are in continuum with aSMCs through the progressive acquisition of aSMC markers. The abrupt transition from aaSMCs to pericytes at the arteriole–capillary boundary is illustrated when the expression of the 1,708 genes are plotted in the correct arteriovenous order (Fig. 3d). High-resolution imaging of ACTA2 and PDGFRB expression in brain tissue shows that



**Figure 2 | Endothelial zonation.** **a**, Zonal expression of transcripts across endothelial cells sorted by SPIN. **b**, Left, heat map of relative gene expression (red, high; blue, low) reveals six major patterns: artery (A; 1 and 2), vein (V; 3), artery and vein (A+V; 4), capillary (Cap; 5) and vein and capillary (V+Cap; 6). Middle, examples of individual genes. Arrows mark position in heat map. Grey bars show individual cells and red curves average expression. Right, curve plots for each group, average in red. **c**, Immunofluorescent staining of cerebral cortex for indicated markers.

**d**, Quantification of immunofluorescent results. For each protein the number of mice and of image fields was MFSD2A 4, 20; TFRC 5, 36; SLC16A1 4, 21; VWF 4, 20; VCAM1 3, 20, respectively. All proteins show statistically significant zonation (Kruskal–Wallis,  $P < 0.0001$ ). Dunn's multiple comparison test demonstrates statistically significant differences between vein and capillary and between artery and capillary ( $****P < 0.0001$ ,  $***P < 0.001$ ,  $**P < 0.01$ ,  $*P < 0.05$ , NS, not significant). Scale bars, 20  $\mu\text{m}$ .

the transition from aaSMC to pericyte takes place between one cell and the next at this location (Extended Data Fig. 6f). These data indicate that mural cell arteriovenous zonation is punctuated into two separated phenotypic continua, in contrast to the single arteriovenous continuum of endothelial cells.

The zonal distribution of transcription factors and transporters in mural cells (Extended Data Fig. 7c, d) implies that there are gene regulatory and physiological differences between mural cell categories, and the numerous pericyte-enriched novel transcription factors will be interesting targets for future analysis. Notably, aSMCs were enriched for expression of immediate early genes (IEGs), including *Fos*, *Fosb*, *Jun*, *Junb* and *Egr1* (Extended Data Fig. 7e), which may reflect a rapid (within minutes) response of aSMCs to removal of natural physical conditions such as high shear and pressure. It was recently reported that the dissociation of muscle stem cells induces an IEG response<sup>23</sup>, demonstrating that this may be a common confounder in single cell analysis.

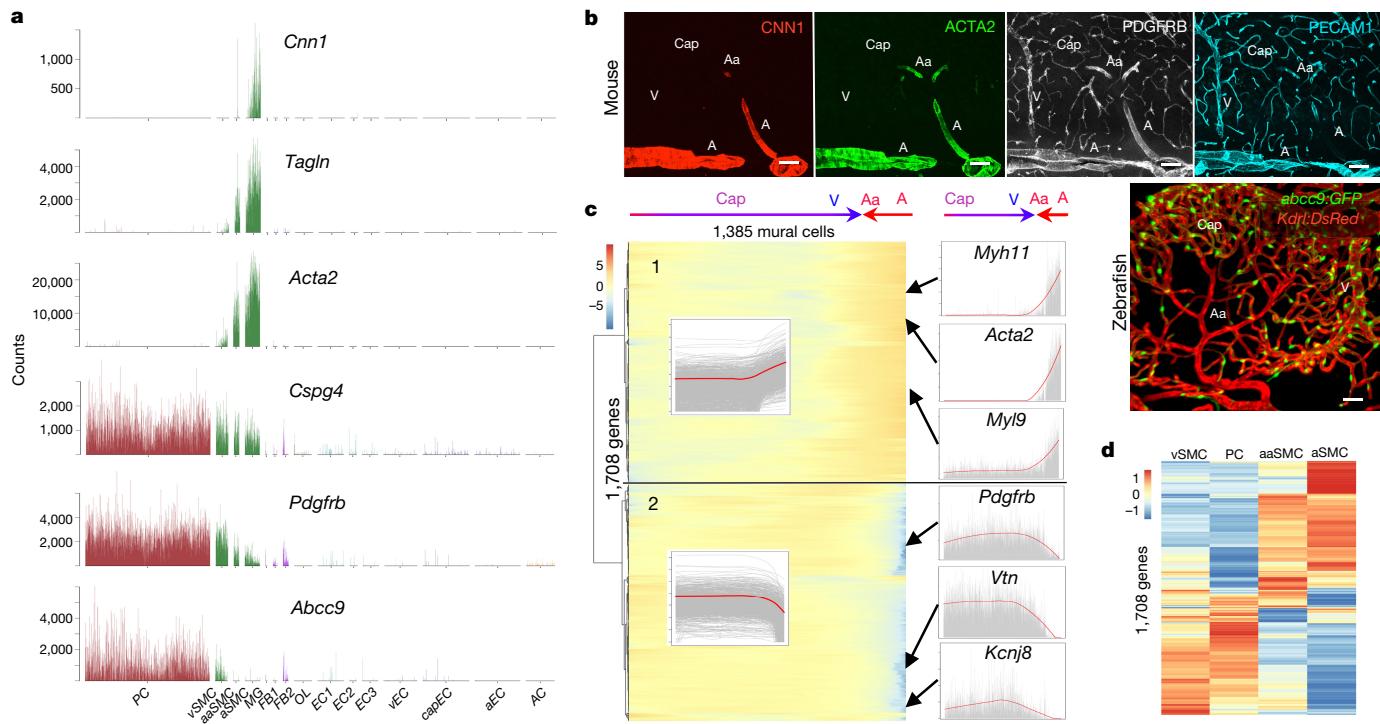
### Brain pericytes molecularly defined

The 1,088 transcriptomes of the PC cluster provide an opportunity to molecularly define the brain pericyte. Closer analysis revealed that the vast majority (about 99%) of pericytes were free from contamination by any of the other analysed cell classes. The remaining roughly 1% of the pericytes contained canonical endothelial transcripts at 5–10% of their average levels in endothelial cells (Extended Data Fig. 9a). We assume that these pericytes were contaminated by endothelial cell fragments

remaining attached to the pericytes during isolation. Using protocols in which cell dissociation conditions were deliberately relaxed, we found abundant endothelial–pericyte transcript contamination (data not shown). This type of contamination is also noticeable in previously published data<sup>17,24</sup>.

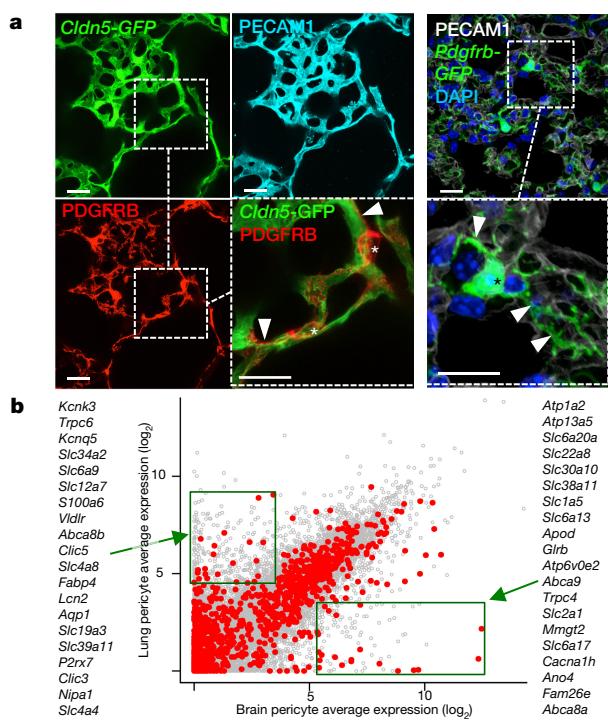
The non-endothelial-contaminated population of pericytes did not form obvious subclusters, and a manual gene search of the BackSPIN data did not reveal the existence of brain pericyte subtypes. A distinguishing criterion for the PC cluster was absent (or extremely low) *Acta2* expression. The lack of ACTA2 signal in venules suggests that capillaries and venules harbour transcriptionally indistinguishable pericytes. The distribution of *abcc9* in zebrafish (Fig. 3b and Extended Data Fig. 8b) further supports this notion.

The lack of heterogeneity among pericytes in the brain prompted us to investigate possible heterogeneity between organs. For a first insight into pericyte organotropism, we identified pericytes in the adult lung of *Pdgfrb*-GFP mice as strongly GFP-positive cells located on alveolar capillaries (Fig. 4a) and isolated these cells for scRNA-seq. Altogether, about 1,500 lung cells were sequenced and clustered by BackSPIN. In addition to pericytes, we isolated and annotated endothelial cells, SMCs and fibroblast- or cartilage-like cell types (which are not further discussed here) following the same strategy as for the brain cells. We identified a distinct cluster of lung pericytes based on the presence of the canonical pericyte markers *Pdgfrb*, *Cspg4*, and *Des* and the simultaneous absence of SMC markers (*Acta2* and *Tagln*) and fibroblast markers (*Pdgfra*, *Lum* and *Dcn*) (Extended Data Fig. 9b). In addition,



**Figure 3 | Mural cell zonation.** **a**, Bar plots of selected markers (see Extended Data Fig. 1 for cell type abbreviations). **b**, Immunofluorescence (mice,  $n=3$ ) and reporter gene (zebrafish) expression of selected genes and proteins from **a** (see also Extended Data Fig. 7a, b). **c**, Cells from the PC, vSMC, aaSMC and aSMC clusters re-sorted into a SPIN range

(*x*-axis) and hierarchical clustering (*y*-axis) shows two major patterns with preferences for aaSMC and aSMC (1) and PC and vSMC (2). Right: three examples from each group. Aa, arterioles. **d**, Expression heat map (red, high; blue, low) with cell clusters in anatomical order. All images from cerebral cortex. Scale bars, 50  $\mu$ m.

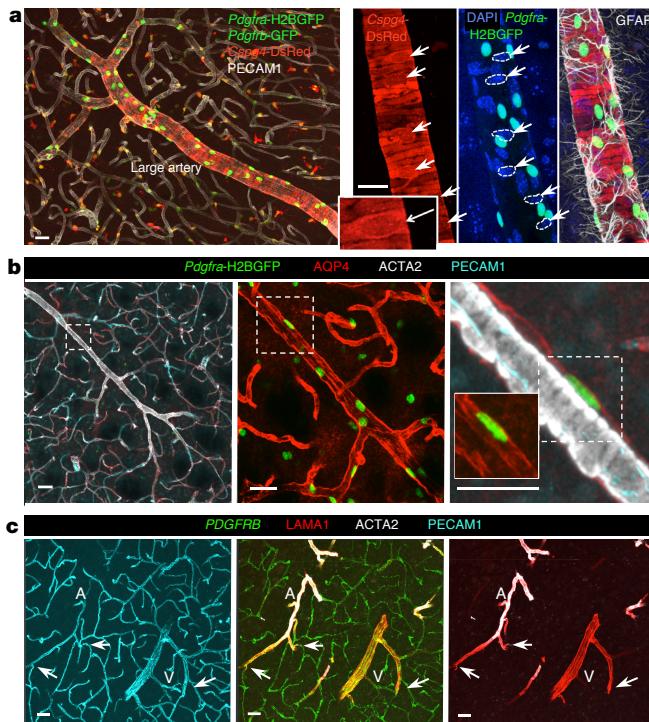


**Figure 4 | Pericyte organotypicity.** **a**, *Cldn5*-GFP and *PECAM1* immunofluorescence mark endothelial cells, and *Pdgfrb*-GFP and PDGFRB immunofluorescence mark pericytes in mouse lung. Insets show pericytes at high magnification. Asterisks, cell bodies; arrowheads, pericyte processes. Scale bars, 20  $\mu$ m. **b**, Transporters (red) in brain and lung pericytes. The 20 most differentially expressed transporters are indicated for each cell type.

lung pericytes shared the following markers with brain pericytes: *Vtn*, *Higd1b*, *S1pr3*, *Mcam*, *Ifitm1*, *Baiap3* and *Ehd3* (Extended Data Fig. 9b and <http://betsholtzlab.org/VascularSingleCells/database.html>). By extending the comparison to all genes, we identified substantial organotypic differences between brain and lung pericytes. For example, the commonly used marker for brain pericytes CD13 (encoded by *Anpep*) was not expressed by lung pericytes (Extended Data Fig. 9b). Differences were also apparent through *t*-distributed stochastic neighbour embedding (*t*-SNE) analysis, which clearly separated brain and lung pericytes (Extended Data Fig. 9c). A noticeable difference between brain and lung pericytes involved the SLC, ABC and ATP transporters, many of which (for example, *Atp13a5*) were abundant in brain pericytes (Extended Data Fig. 9d, e) but low or absent in lung pericytes (Fig. 4b). The gene ontology (GO) term ‘transmembrane transporter activity’ was associated with genes that were overexpressed in brain pericytes when compared with lung pericytes (data not shown). Beyond providing evidence for organotypic specialization of pericytes, this observation suggests that brain pericytes are directly involved in molecular transport at the BBB.

### Perivascular fibroblast-like cells

We found a population of *Pdgfra*-driven histone 2B-fused GFP (H2BGFP)-positive cells (hereinafter referred to as brain fibroblast-like cells) in close association with arteries, arterioles, veins and venules. These cells were non-overlapping with endothelial and mural cells and were located between the vessel wall and the AQP4- and GFAP-positive astrocyte end-feet (Fig. 5a, b and Extended Data Fig. 10a, b). By contrast, *Pdgfra*-H2BGFP-positive oligodendrocyte progenitors, which are abundant in the brain parenchyma and occasionally apposed to blood vessels, were always located outside the AQP4-positive astrocyte end-feet (Fig. 5b, Extended Data Fig. 10b and Supplementary Video 1). Thus, the fibroblast-like cells reside within the perivascular Virchow–Robin space, which has been implicated in fluid exchange between the



**Figure 5 | Brain fibroblast-like cells.** **a**, *Pdgfra*-H2BGFP-positive cells along arteries in cerebral cortex. Arrows indicate SMC nuclei. **b**, *Pdgfra*-H2BGFP-positive cells in arterioles are located between SMCs and astrocyte end-feet. **c**, LAMA1-positive ECM sleeve surrounds arteries, arterioles, veins and venules, but not capillaries. White arrows indicate transition from LAMA1-positive to LAMA1-negative vessel. Observations from 5 (a, c) and 2 (b) mice. Scale bars, 20 μm.

brain and the cerebrospinal fluid<sup>25</sup>. In contrast to mural cells, which are firmly embedded in the basement membrane, the fibroblast-like cells were only loosely adhered to the vessels and required relaxed digestion conditions to remain attached. Both BackSPIN and t-SNE analysis provided distinct clustering of brain fibroblast-like cells from all other vascular cell types (Fig. 1b and Extended Data Figs 1j, k, 9c). Subtypes of these cells could be distinguished (FB1 and FB2), but are not discussed further here.

We assigned the 50 most specific markers for the FB, PC, a/aaSMC and EC clusters by correlation with *Pdgfra*, *Abcc9*, *Acta2* and *Cldn5*, respectively (Extended Data Fig. 11a). Many of the FB-specific transcripts encode structural components, modifiers or receptors for the extracellular matrix (ECM) (Extended Data Fig. 11b), including fibrillar (for example, *Col1a1*, *Col1a2*, *Col3a1*, *Col5a1*, *Col5a2* and *Col5a3*) and non-fibrillar collagens (for example, *Col6a1*, *Col6a2*, *Col6a3*, *Col8a1*, *Col8a2*, *Col11a1*, *Col12a1*, *Col13a1*, *Col15a1*, *Col16a1*, *Col23a1* and *Col26a1*) (Extended Data Fig. 11c, d), collagen-modifying enzymes (for example, lysyl oxidases) and proteins involved in collagen fibril spacing, such as the small leucin-rich proteoglycans lumican (*Lum*) and decorin (*Dcn*) (Extended Data Fig. 11e, f). GO terms associated with ECM, collagen, cell adhesion and migration were significantly ( $P < 1 \times 10^{-10}$ ) associated with the fibroblast-like cells; the other vascular cell types (PC, a/aaSMC and EC) instead showed enrichment for terms associated with cation transporter, muscle contraction and cell junction, respectively (Extended Data Fig. 11a). Of the 50 brain FB-specific transcripts, 45 were also expressed by lung fibroblasts (Extended Data Fig. 11b) pointing to a strong resemblance between these cells. However, the brain FB cells also expressed the endothelial marker *Cdh5*, the epithelial marker *Lama1* and the mesothelial marker *Efemp1*, suggesting that they may have mixed mesenchymal and epithelial characteristics (Extended Data Fig. 11b and <http://betsholtzlab.org/VascularSingleCells/database>).

The specific expression of *Lama1* allowed us to further study the anatomical distribution of brain fibroblasts. We found that LAMA1 antibodies decorate a continuous ECM sleeve located around arteries, arterioles, venules and veins, but not around capillaries, matching the perivascular distribution of *Pdgfra*-H2BGFP-positive cells (Fig. 5c and Extended Data Fig. 10c).

## Discussion

Here, we have provided molecular definitions of the principal classes of brain vascular cells through their genome-wide quantitative transcriptomes. We have also unravelled a molecular blueprint for zonation along the arteriovenous axis, and illustrated how endothelial and mural cells differ in this regard. Single-cell analysis was crucial for this information, which would have been lost in the average profiles provided through bulk RNA sequencing. The examples of transcription factors and transporters illustrate that the zonal distribution of any class of proteins can now be assessed at <http://betsholtzlab.org/VascularSingleCells/database.html>. Our data also expand by around 100-fold the number of brain arteriovenous markers and pinpoint several arteriovenous landmarks that can be applied for vessel type identification in the brain. The cellular correlates of these landmarks are schematically summarized in Extended Data Fig. 6a.

Despite their functional importance and implication in a number of processes, pericytes remain poorly defined. Their proposed roles in vessel stabilization<sup>26</sup>, BBB formation<sup>27,28</sup>, immune cell guidance<sup>29</sup>, blood flow regulation<sup>30</sup>, tissue repair<sup>31</sup> and pathological scarring<sup>32</sup> are all clouded by ambiguities regarding cell type identification<sup>33–35</sup>. The molecular definition of brain and lung pericytes reveals several new markers common for these two types of pericyte, but also an extensive organotypicity<sup>36</sup>. It should be remembered that the cell type definitions presented herein are applicable only to the normal adult mouse brain and lung, and that heterogeneity may be expected in development and disease, topics that can now be systematically addressed.

The molecular identification of fibroblast-like cells and elucidation of their anatomical location provide insights into the cellular anatomy of the brain vasculature. The expression of markers on these cells matches that of scar-forming cells in spinal cord injury<sup>32,33</sup> and recently described perivascular cells with proposed pericyte-like characteristics<sup>18</sup>. Our analysis instead reveals a resemblance to lung fibroblasts combined with the expression of certain epithelial, endothelial and mesothelial markers. The location of these cells within a perivascular space that has been implicated in the glymphatic system<sup>37</sup> warrants investigations into their involvement in brain fluid homeostasis.

While it will now be important to map the transcriptional changes that occur in vascular cells during disease, the data available herein and at <http://betsholtzlab.org/VascularSingleCells/database.html> may already provide clues to cell types involved in brain diseases that feature neurovascular dysfunction, such as Alzheimer's disease, certain monogenic neurological diseases and brain tumours. For these conditions, our data pinpoint pericytes and fibroblast-like cells as common cellular culprits besides endothelial cells (Extended Data Fig. 12), suggesting that cellular involvement in neurovascular pathologies is more complex than anticipated.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 21 June 2017; accepted 10 January 2018.

Published online 14 February 2018.

- Iadecola, C. The pathobiology of vascular dementia. *Neuron* **80**, 844–866 (2013).
- Zhao, Z., Nelson, A. R., Betsholtz, C. & Zlokovic, B. V. Establishment and dysfunction of the blood-brain barrier. *Cell* **163**, 1064–1078 (2015).
- Abbott, N. J., Rönnbäck, L. & Hansson, E. Astrocyte–endothelial interactions at the blood-brain barrier. *Nat. Rev. Neurosci.* **7**, 41–53 (2006).

4. Zlokovic, B. V. The blood-brain barrier in health and chronic neurodegenerative disorders. *Neuron* **57**, 178–201 (2008).
5. Phoenix, T. N. et al. Medulloblastoma genotype dictates blood brain barrier phenotype. *Cancer Cell* **29**, 508–522 (2016).
6. de Lange, E. C. & Hammarlund-Udenaes, M. Translational aspects of blood-brain barrier transport and central nervous system effects of drugs: from discovery to patients. *Clin. Pharmacol. Ther.* **97**, 380–394 (2015).
7. Majno, G. & Palade, G. E. Studies on inflammation. 1. The effect of histamine and serotonin on vascular permeability: an electron microscopic study. *J. Biophys. Biochem. Cytol.* **11**, 571–605 (1961).
8. Simionescu, M., Simionescu, N. & Palade, G. E. Segmental differentiations of cell junctions in the vascular endothelium. The microvasculature. *J. Cell Biol.* **67**, 863–885 (1975).
9. Ekman, N. et al. *Bmx* tyrosine kinase is specifically expressed in the endocardium and the endothelium of large arteries. *Circulation* **96**, 1729–1732 (1997).
10. Wang, H. U., Chen, Z. F. & Anderson, D. J. Molecular distinction and angiogenic interaction between embryonic arteries and veins revealed by ephrin-B2 and its receptor Eph-B4. *Cell* **93**, 741–753 (1998).
11. Hogan, B. M. et al. Vegfc/Flt4 signalling is suppressed by Dll4 in developing zebrafish intersegmental arteries. *Development* **136**, 4001–4009 (2009).
12. Kutschera, S. et al. Differential endothelial transcriptomics identifies semaphorin 3G as a vascular class 3 semaphorin. *Arterioscler. Thromb. Vasc. Biol.* **31**, 151–159 (2011).
13. Hirashima, M. & Suda, T. Differentiation of arterial and venous endothelial cells and vascular morphogenesis. *Endothelium* **13**, 137–145 (2006).
14. Armulik, A., Genové, G. & Betsholtz, C. Pericytes: developmental, physiological, and pathological perspectives, problems, and promises. *Dev. Cell* **21**, 193–215 (2011).
15. Attwell, D., Mishra, A., Hall, C. N., O'Farrell, F. M. & Dalkara, T. What is a pericyte? *J. Cereb. Blood Flow Metab.* **36**, 451–455 (2016).
16. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
17. Zeisel, A. et al. Brain structure: Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
18. Marques, S. et al. Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. *Science* **352**, 1326–1329 (2016).
19. Goldmann, T. et al. Origin, fate and dynamics of macrophages at central nervous system interfaces. *Nat. Immunol.* **17**, 797–805 (2016).
20. Halpern, K. B. et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352–356 (2017).
21. Nguyen, L. N. et al. Mfsd2a is a transporter for the essential omega-3 fatty acid docosahexaenoic acid. *Nature* **509**, 503–506 (2014).
22. Lein, E. S. et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
23. van den Brink, S. C. et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14**, 935–936 (2017).
24. Zhang, Y. et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* **34**, 11929–11947 (2014).
25. Iliff, J. J. et al. A paravascular pathway facilitates CSF flow through the brain parenchyma and the clearance of interstitial solutes, including amyloid  $\beta$ . *Sci. Transl. Med.* **4**, 147ra111 (2012).
26. Lindahl, P., Johansson, B. R., Levéen, P. & Betsholtz, C. Pericyte loss and microaneurysm formation in PDGF-B-deficient mice. *Science* **277**, 242–245 (1997).
27. Armulik, A. et al. Pericytes regulate the blood-brain barrier. *Nature* **468**, 557–561 (2010).
28. Daneman, R., Zhou, L., Kebede, A. A. & Barres, B. A. Pericytes are required for blood-brain barrier integrity during embryogenesis. *Nature* **468**, 562–566 (2010).
29. Probstl, D. et al. Pericytes support neutrophil subendothelial cell crawling and breaching of venular walls *in vivo*. *J. Exp. Med.* **209**, 1219–1234 (2012).
30. Hall, C. N. et al. Capillary pericytes regulate cerebral blood flow in health and disease. *Nature* **508**, 55–60 (2014).
31. Crisan, M. et al. A perivascular origin for mesenchymal stem cells in multiple human organs. *Cell Stem Cell* **3**, 301–313 (2008).
32. Göritz, C. et al. A pericyte origin of spinal cord scar tissue. *Science* **333**, 238–242 (2011).
33. Soderblom, C. et al. Perivascular fibroblasts form the fibrotic scar after contusive spinal cord injury. *J. Neurosci.* **33**, 13882–13887 (2013).
34. Hill, R. A. et al. Regional blood flow in the normal and ischemic brain is controlled by arteriolar smooth muscle cell contractility and not by capillary pericytes. *Neuron* **87**, 95–110 (2015).
35. Guimaraes-Camboa, N. et al. Pericytes of multiple organs do not behave as mesenchymal stem cells *in vivo*. *Cell Stem Cell* **20**, 345–359 (2017).
36. Nolan, D. J. et al. Molecular signatures of tissue-specific microvascular endothelial cell heterogeneity in organ maintenance and regeneration. *Dev. Cell* **26**, 204–219 (2013).
37. Xie, L. et al. Sleep drives metabolite clearance from the adult brain. *Science* **342**, 373–377 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This study was supported by AstraZeneca AB (C.B., U.L.), the Swedish Research Council (C.B.: 2015-00550; U.L.: K2014-64X-20097-09-5), the European Research Council (C.B.: AdG294556), the Leducq Foundation (C.B., A.K.: 14CVDO2), Swedish Cancer Society (C.B.: 150735; U.L.: CAN 2016/271), Knut and Alice Wallenberg Foundation (C.B.: 2015.0030), Hjärnfonden (U.L.), Swiss National Science Foundation (A.K.: 31003A\_159514/1) and the Synapsis Foundation (A.K.). We thank C. Olsson, H. Leksell, P. Peterson, J. Chmielnikova, K. Gaengel, BioVis (Uppsala), Center for Microscopy and Image Analysis (Zurich) and Eukaryotic Single Cell Genomics facility (Science for Life Laboratory) for technical help, and K. Alitalo, P. Soriano, D. Silver and L. Sorokin for reagents.

**Author Contributions** M.V., L.H., N.M., U.L. and C.B. conceived and designed the project; M.V., M.A.M., J.A., K.A., F.D.G., K.N., T.L., B.L., E.R., L.G., Y.Z., M.R., A.K. and C.B. performed experiments; L.H. performed bioinformatic analysis; L.H. and Y.S. constructed the online database; C.B., L.H. and M.V. analysed the bioinformatic data; C.B. and U.L. wrote the manuscript with substantial input from M.V., L.H., M.A.M., J.A. and K.A. All authors reviewed the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to C.B. ([christer.betsholtz@ki.se](mailto:christer.betsholtz@ki.se)).

**Reviewer Information** *Nature* thanks D. McDonald and the other anonymous reviewer(s) for their contribution to the peer review of this work.

## METHODS

**Mice.** The following mouse strains were used: *C57Bl6* (The Jackson Laboratory, *C57BL6/J*), *Bmx<sup>LacZ</sup>* (*Bmx<sup>tm1Ali</sup>*)<sup>38</sup>, *Cspg4-DsRed* (The Jackson Laboratory, *Tg(Cspg4-DsRed.T1)1Alik/J*), *Pdgfrb-GFP* (Gensat.org, *Tg(Pdgfrb-eGFP)JN169Gsat/Mmucd*)<sup>39</sup>, *Pdgfra-H2BGFP* (*Pdgfra<sup>tm11(EGFP)Sor</sup>*, gift from P. Soriano), *Cldn5-GFP* (*Tg(Cldn5-GFP)Cbet/U*), *SM22-Cre* (*Tg(Tagln-cre)1Her/J* and *R26-stop-tdT*Tomato (*B6;129S6-G(ROSA)26Sor<sup>tm1(CAG-tdT)Hze</sup>*)). All mice were back-crossed on a *C57BL6/J* genetic background. Adult mice of either sex aged 10–19 weeks were used for all experiments. To generate the brain single-cell dataset we used the following mice: two mice of *Pdgfrb-GFP/Cspg4-DsRed* genotype, two mice of *Cldn5-GFP/Cspg4-DsRed* genotype, three mice of *Pdgfra-H2BGFP* genotype and three mice of *SM22-Cre;R26-stop-tdT*Tomato genotype. The lung single-cell dataset was generated from two *Pdgfrb-GFP/Cspg4-DsRed* mice, one *Pdgfrb-GFP* mouse and two *Cldn5-GFP/Cspg4-DsRed* mice. No statistical method was used to predetermine sample size. Animal experiment protocols were approved by the Uppsala Ethical Committee on Animal Research (permit numbers C224/12 and C115/15) and the Stockholm North Animal Ethics committee (Stockholms Norra Djurförsöksetiska Nämnd), permit N150/14. All animal experiments were carried out in accordance with their guidelines.

**Single cell isolation.** The mice were killed by cervical dislocation, after which the brain was surgically removed and placed in ice-cold Dulbecco's modified Eagle's medium (DMEM, ThermoFisher Scientific) supplemented with penicillin/streptomycin (ThermoFisher Scientific cat# 15140122). For all brain single-cell isolations, the entire brain, without olfactory bulbs, was used.

**Brain cell isolation.** Brain mural cells (*Pdgfrb-GFP/Cspg4-DsRed* positive) and brain endothelial cells (*Cldn5-GFP* positive) were isolated with a modified version of the Neural Tissue Dissociation kit (P) (Mylenyi Biote, Cat# 130-092-628). The brain tissue was mechanically and enzymatically dissociated before removal of myelin, which is detrimental for efficient sorting. Myelin was removed using magnetic bead separation (Myelin Removal Beads II, Mylenyi Biote, cat# 130-096-733) according to the manufacturer's protocol. The final cell pellet was resuspended in 500–1,000 µl FACS buffer (DMEM without phenol red (ThermoFisher Scientific), supplemented with 2% fetal bovine serum (FBS, ThermoFisher Scientific). For a detailed protocol, see Protocol Exchange<sup>40</sup>.

**Perivascular single-cell isolation.** In order to enrich for perivascular *Pdgfra-H2BGFP*-positive cells from the brain before sorting, the brain vasculature was isolated with Dynabeads (Dynabeads Sheep Anti-Rat IgG, ThermoFisher Scientific, cat# 11035) coupled to rat anti-mouse CD31 (BD Pharmingen, cat# 553370) before the tissue was digested into single cells. A detailed description of the protocol can be found in Protocol Exchange<sup>41</sup>. After microvascular isolation (and before single-cell dissociation), the microvascular fragments were microscopically inspected to verify the presence of microvascular fragments with *Pdgfra-H2BGFP*-positive cells still attached (Leica Inverted microscope, Leica Microsystems). The final pellet was resuspended in 500 µl FACS buffer.

**Lung single-cell isolation.** To isolate single lung cells, mice were killed as described above and the lungs were removed. Single cells were isolated using a combination of mechanical and enzymatic dissociation. The primary enzymes used were collagenase type 2 (ThermoFisher Scientific, cat# C6885) and enzymes 3 and 4 from the Neural Tissue Dissociation kit (P), described above. For detailed instructions, see Protocol Exchange<sup>42</sup>. The obtained cell pellet was resuspended in 500 µl FACS buffer.

**Fluorescence-activated cell sorting (FACS).** Prior to FACS, all single-cell suspensions were strained over a 35-µm mesh (Cell-Strainer capped tubes, Corning, cat# 352235). Note that FACS was used merely to enrich and capture the cells of interest, and not for identification. For this reason, all FACS settings were set very broadly to avoid FACS-induced bias. All cell suspensions were sorted on a BD FACSAria III (BD Biosciences) at the BioVis core facility of the Department of Immunology, Genetics and Pathology (IGP), Uppsala University. A 100-µm nozzle and 20-psi PBS sheet fluid pressure were used. First, cells were selected using forward scatter area/side scatter area (FSC-A/SSC-A, linear scale) with a very wide gate setting, to accommodate all possible cell sizes and morphologies with as little bias as possible. For the same reason, no doublet discrimination was implemented, since this might favour the enrichment of cells with a uniform, round morphology. Next, fluorescent events based on the parent FSC-A/SSC-A gate were selected: GFP was excited with a 488-nm laser, and emission was detected through a 530/30 filter, while DsRed or TdTomato was exited with a 561-nm laser and emission detected through a 582/15 filter. As controls, wild-type *C57Bl6/J* mice lacking fluorescent reporters were used. Single cells, positive for GFP alone (*Cldn5-GFP;Cspg4-DsRed* and *Pdgfra-H2BGFP* mice), positive for both GFP and DsRed (*Pdgfrb-GFP;Cspg4-DsRed* mice) or positive for TdTomato (*SM22-Cre;R26-stop-tdT*Tomato) were sorted straight into 2.3 µl lysis buffer pre-dispensed into a 384-well plate. For all sorts, the single cell mask was used. Prior to cell collection, special care was taken to ensure that cells were sorted directly into the lysis buffer without touching the well wall:

1) The waste collection bin was moved to the right as much as possible (while still catching the waste stream) so that the sort angle could be reduced to 8–10 units. 2) Sort accuracy was optimized by sorting 100 BD FACS Accudrop beads (BD Biosciences) per well on the lid of an empty 384-well plate, followed by visual inspection of sort drop location. If needed, the plate position was recalibrated to ensure the best sort accuracy. All FACS data were analysed using FlowJo v10.1 (FlowJo, LLC) and FACSDiva v8.0.2 (BD Biosciences).

**Smart-seq2 library prep and sequencing.** Single-cell libraries were prepared as described previously<sup>43</sup>, with the following specifications: 0.0025 µl of a 1:40,000 diluted ERCC spike-in concentration stock was used, and all cDNA was amplified with 22 PCR cycles before QC control with a Bioanalyzer (Agilent Biosystems). The libraries were sequenced on a HiSeq2500 at the National Genomics Infrastructure (NGI), Science for Life Laboratory, Sweden, with single 50-bp reads (dual indexing reads). All single-cell transcriptome data were generated at the Eukaryotic Single-cell Genomics facility at Science for Life Laboratory in Stockholm, Sweden.

**Antibodies and immunofluorescence staining.** Mice under full anaesthesia were euthanized by transcardial perfusion with Hanks balanced salt solution (HBSS, GIBCO, cat# 14025092) followed by 4% buffered formaldehyde (Histolab, cat# 02178). Brains were removed and postfixed in 4% buffered formaldehyde for 4 h (or overnight) at 4 °C. Sagittal and coronal vibratome sections (50–150 µm) were incubated in blocking/permeabilization solution (1% bovine serum albumin, 2.5% donkey serum, 0.5% Triton X-100 in PBS) overnight at 4 °C, followed by incubation in primary antibody solution for two nights at 4 °C, and subsequently in secondary antibody solution, overnight at 4 °C. Sections were mounted in ProLong Gold Antifade mountant (Life Technologies, cat# P36930). Images were taken with a Leica TCS SP8 confocal microscope (Leica Microsystems). An overview of the antibodies used is presented in Supplementary Information Table 4. All images were acquired with a Leica TCS SP8 confocal microscope (Leica Microsystems) and Leica Application Suite 3.2.1.9702. Images were adjusted for brightness and contrast using Leica Application Suite 3.2.1.9702, Fiji v1.51 s and Adobe PhotoShop CC 2015. All confocal images are represented as maximum intensity projections unless stated otherwise (Leica Application Suite 3.2.1.9702 and Fiji v1.51 s).

**X-gal staining.** *Bmx<sup>LacZ</sup>* knock-in reporter mice<sup>38</sup> were used to visualize *Bmx* expression in mouse brains with X-gal staining. First, the mice were perfused with PBS, followed by 4% PFA. Dissected brains were postfixed in 4% PFA for 60 min at 4 °C, washed and cut with a vibratome into 50-µm-thick coronal sections. The brain sections were washed 3 times for 20 min with 0.1 M phosphate buffer containing 0.02% NP-40, 0.01% deoxycholate and 2 mM MgCl<sub>2</sub>. Next, the sections were incubated in staining solution containing 1 mg/ml X-gal, 5 mM K<sub>3</sub>Fe(CN)<sub>6</sub>, 5 mM K<sub>4</sub>Fe(CN)<sub>6</sub>, 2 mM MgCl<sub>2</sub>, 0.02% NP-40, 0.01% deoxycholate and 0.1 M phosphate buffer for 21 h at 37 °C. The brain slices were washed extensively with PBS, blocked for 7 h at 4 °C in 1% BSA, 0.5% Triton X-100 in PBS, before adding primary antibody solution. The rest of the protocol was continued as described in the previous paragraph.

**RNA *in situ* hybridization.** RNA *in situ* hybridization was performed using the RNAscope technology (Advanced Cell Diagnostics) following the manufacturer's protocol with minor modifications. In brief, fresh-frozen brains were cut into 20-µm sagittal sections and mounted on SuperFrost Plus glass slides. After dehydration, slides were subjected to RNAscope Multiplex Fluorescent Assay. First, slides were incubated in Pretreat 3 for 30 min at room temperature. After that, RNAscope probes, *Gkn3-C1*, *Slc38a5-C1*, *Acta2-C2* and *Pecam1-C3*, were hybridized for 2 h at 40 °C and the remainder of the assay protocol was implemented. The fluorescent signal emanating from RNA probes was visualized and captured using a Leica TCS SP8 confocal microscope (Leica Microsystems). All *in situ* hybridization images presented are 2D maximum intensity projections of ~4-µm z-stacks obtained from cerebral cortex.

**Quantifications.** **Vessel diameter.** Vessel diameter and length were measured in 75–150-µm vibratome brain sections obtained from *C57Bl6/J* and *Pdgfra-H2BGFP* mice. Four to seven areas from cerebral cortex in three animals were analysed. Blood vessels were visualized by CD31 and perivascular cells by TAGLN, ACTA2, CNN1 or LAMA1 immunostaining. Z-stacks were captured using a Leica TCS SP8 confocal microscope (Leica Microsystems). Vessel diameter and length were measured manually using Leica Application Suite 3.2.1.9702 software.

**Endothelial zonation.** Quantifications were performed on 8-bit confocal images obtained from samples visualized by immunohistochemistry or RNA *in situ* hybridization (as described above). For quantifying immunohistochemical images, the ROI manager of the Fiji software was used to mark and measure the mean grey values of the vascular segments of interest (vein, venule, capillary, arteriole, and artery). The sum of all mean grey values per field was set as 100% and the percentage of mean grey values per vascular segment was calculated for each image. A Kruskal–Wallis test with Dunn's multiple comparison test was performed to demonstrate statistically significant differences in protein expression along the arteriovenous axis. The data were presented using GraphPad Prism 6 software. For

the quantification of RNA *in situ* hybridization images, the ROI manager of the Fiji software was used to mark and measure the mean grey values of the vascular segments of interest (vein, artery) and presented using GraphPad Prism 6 software. An unpaired *t*-test with Welsh correction was used to demonstrate statistical significance in gene expression between arteries and veins.

**Zebrafish. BAC recombination.** The pRedET plasmid (GeneBridge) was introduced into *Escherichia coli* containing a CH211-58C15 BAC clone encoding the *abcc9* gene (BacPAC Resources) by electroporation (1,800 V, 25 µF, 200 Ω). Next, two *Tol2* long terminal repeats in opposing directions flanking an ampicillin resistance cassette were amplified by PCR using *Tol2\_amp* as a template and were inserted into the BAC vector backbone. Then, cDNA encoding IRES-Gal4FF together with a kanamycin resistance cassette (IRES\_Gal4FF\_KanR) were amplified by PCR using pCS2\_mCherry\_IRES\_Gal4FF\_KanR plasmid as a template and inserted at the C terminus of the *abcc9* gene (ENSDART00000079987.4). Primers to amplify the IRES\_Gal4FF\_KanR PCR product were as follows:

5'-atggaggcaggaggacggctgttgcattttgtcaagccgacatgtAGATGAGTAAC TTGTACAAAG-3' and 5'-aaaatggctttatgtatctgttaaggccaaaatggtaaaatgggg TCAGAAGAACCTCGTCAAGAAGGCG-3' (lowercase, homology arm to BAC vector; uppercase, primer binding site to the template plasmid).

**Plasmids.** The *Tol2\_amp* and pCS2\_Gal4FF\_KanR vector for BAC recombination were kindly provided by S. Schulte-Merker (University of Münster, Germany). To construct the plasmid encoding pCS2-mCherry\_IRES\_Gal4FF\_KanR, the pCS2\_Gal4FF\_KanR vector had the following cDNA inserted: a PCR-amplified IRES using p3E-IRES-nlsEGFP in *Tol2*Kit<sup>44</sup> as a template and pmCherry-C1 (Clontech, Takara Bio Inc.).

**Transgenic and mutant zebrafish lines.** The *Tol2* transposase mRNA was transcribed *in vitro* with SP6 RNA polymerase from a NotI-linearized pCS-TP vector using the mMESSAGE mMACHINE kit (Ambion). To generate the *TgBAC(abcc9:Gal4FF)<sup>ncv34</sup>* zebrafish line, the corresponding BAC DNA (described above) was co-injected with *Tol2* transposase mRNA into one-cell stage embryos of AB wild type. *Tg(UAS:GFP)* and *Tg(UAS:RFP)* fish lines<sup>45</sup> were kindly provided by K. Kawakami (National Institute of Genetics, Japan). *Tg(fli1a:Myr-mCherry)<sup>ncv1</sup>*, *TgBAC(pdgfrb:Gal4FF)<sup>ncv</sup>* and *TgBAC(tagln:EGFP)<sup>ncv25</sup>* fish lines were previously described<sup>44–47</sup>. *Tg(kdr:DsRed2)<sup>pD27</sup>* was obtained from zebrafish international resource centre<sup>48</sup>. Throughout the text, all transgenic lines used in this study are simply described without their line numbers. For example, *TgBAC(abcc9:Gal4FF)<sup>ncv34</sup>* is abbreviated to *TgBAC(abcc9:Gal4FF)*.

**Raw sequencing data processing.** The samples were analysed by first demultiplexing the fastq files with deindexer (<https://github.com/ws6/deindexer>) using the Nextera index adapters and the 384-well layout. Individual fastq files were then mapped to the mouse reference genome (mm10) using the STAR aligner version 2.4.2a<sup>49</sup>. Two-pass alignment was chosen to have improved performance of *de novo* splice junction reads, filtered for only uniquely mapping reads. The expression values were computed per gene as described<sup>50</sup>, using uniquely aligned reads and correcting for the uniquely alignable positions using MULTo<sup>51</sup>. The count matrix shows the individual counts aligning to each gene per cell. Cells with total aligned reads less than 100,000 were removed, and also outlier cells with maximum Spearman's correlation with other cells less than 0.3 were filtered out. After quality filtering, 3,186 brain cells and 1,504 lung cells were selected for further analysis. In order to compare the gene expression counts between different cells, the total gene counts for each cell were normalized to 500,000. Genes that were identified in less than three cells in the brain or lung were filtered out in the respective tissue analysis.

**Cell type classification.** In order to classify the cells into different cell types, the BackSPIN algorithm was used<sup>17</sup>. BackSPIN was run with the following parameters: -d 6 -g 3 -c 5, splitting the cells into six levels. The 3,186 brain cells were divided into 38 groups. BackSPIN iteratively splits the cells and the genes (assigning a proportion of the genes to each new branch), thus clusters that receive many genes expressed at background/noise levels continue to split at subsequent levels based on noise. Other clusters may instead stop splitting prematurely, because genes that were differentially expressed within the cluster had been assigned to another branch. The pericyte/vSMC cluster stopped splitting on the second level, because the SMC genes distinguishing pericytes from vSMC in this cluster had been assigned to another node which contained the aa/aSMCs and therefore expressed higher levels of SMC genes. For this reason, the pericyte cluster was re-run using BackSPIN including all genes, which led to clear separation of the pericytes and the vSMC. For other clusters, no other differences were found above noise by manual reading of all barplots, or, when such differences were found (such as in the oligodendrocyte and microglia clusters), they contained too few cells and were left without further splitting, or merged into an ancestral node. In the end, the 3,186 brain cells were classified into 15 clusters. Among them, we identified a small astrocyte cluster of six cells. To achieve a better representation of brain vessel-associated astrocyte gene expression profiles for comparison with the other vessel-associated cell types, we added a set of 250 pure brain astrocyte cells that

were obtained in a parallel study in which cells were sorted from SM22-Cre;R26-stop-*tdTomato* mice. Although this project was focused on smooth muscle cells, it fortuitously also generated a significant number of *Aqp4*-positive astrocytes, apparently caused by ectopic reporter activation in this cell type. The full dataset, which will be published separately, was obtained and analysed using the same methodologies as described herein, including mouse genetic background. The 1,504 lung cells were divided into 46 groups by BackSPIN. After manual inspection, they were classified into 17 clusters. In order to better visualize expression differences between clusters, we created bar plots using average read count from all cells (even those with no expression).

**Differential gene expression analysis.** After cell type classification, the edgeR method was applied to compare the gene expression differences between different clusters<sup>52</sup>. In the pairwise comparisons of the three brain endothelial cell clusters (vEC, capEC and aEC), 1,798 genes were identified as being significantly differentially expressed (selection criteria: >twofold difference, *P* < 0.05 and represented in ≥10 cells in these three clusters). One-way ANOVA analysis was also applied to the 1,798 genes to compare their differential expression among the three brain endothelial cell clusters. Multiple test correction was implemented using the false discovery rate (FDR) method<sup>53</sup> in R. The top 500 genes with the smallest ANOVA *P* values were selected for further analysis in Fig. 2b. In the pairwise comparisons of the four brain mural cell clusters (aSMC, aaSMC, vSMC and PC), 1,708 genes were identified (selection criteria: >twofold difference, *P* < 0.05 and represented in >20% cells in the corresponding clusters).

**t-SNE and heat map visualization.** The 2D t-SNE visualization was performed using the *t*-distributed stochastic neighbour embedding method for R software with default parameters (R tsne package version 0.1-3). The normalized gene count matrix for all genes was used as input. To visualize the gene expression level in cells on t-SNE map, the plotexptsne function in RaceID package (version 1.0) was adapted<sup>54</sup>. The heat map visualization was performed using the pheatmap packages (version 1.0.8) in R. The genes were clustered using Pearson correlation distance and the ward.D2 cluster method was used to build the cluster dendrogram.

**Zonation analysis.** The cells included in the zonation analysis were first sorted using the SPIN method, and then a smooth curve was fitted to the gene expression counts for each gene by loess function using the default parameters in R software.

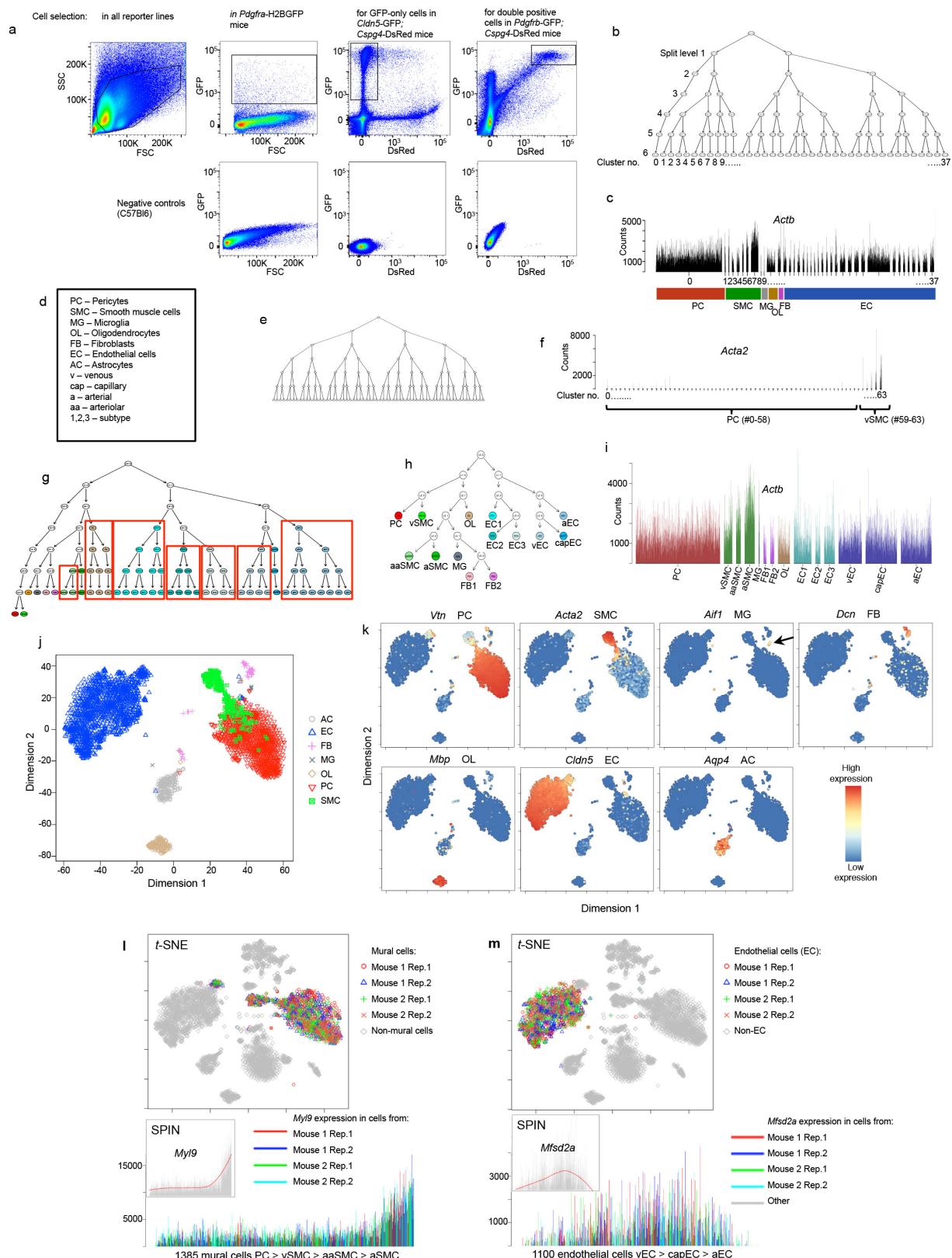
**Transcription factor and transporter analysis.** To identify the transcription factors in the mural cells and endothelial cells, we made use of a previously reported list of 1,507 transcription factors on the mouse genome<sup>55</sup>. To identify transporter related genes, we made use of gene ontology annotation ([www.geneontology.org](http://www.geneontology.org)) and all genes annotated as transporter activity (GO:0005215) or its child terms were extracted.

**Gene ontology (GO) analysis.** The GO analysis was performed using GOSTat packages (version 2.40.0) in R software, with the associated annotation packages GO.db (version 3.4.0) and org.Mm.eg.db (version 3.4.0).

**Data availability.** All sequence data used in this study have been deposited in the NCBI Gene Expression Omnibus database and are accessible through accession numbers GSE98816, GSE99058 and GSE99235. The searchable database is freely available at <http://betsholtzlab.org/VascularSingleCells/database.html>.

38. Rajantie, I. et al. Bmx tyrosine kinase has a redundant function downstream of angiopoietin and vascular endothelial growth factor receptors in arterial endothelium. *Mol. Cell. Biol.* **21**, 4647–4655 (2001).
39. He, L. et al. Analysis of the brain mural cell transcriptome. *Sci. Rep.* **6**, 35108 (2016).
40. Vanlandewijck, M. et al. Primary isolation of vascular cells from murine brain for single cell sequencing. *Protoc. Exch.* <https://doi.org/10.1038/protex.2017.159> (2018).
41. Vanlandewijck, M., Andrae, J., Gouveia, L. & Betsholtz, C. Isolation of vessel-associated pdgf-H2B GFP positive cells from murine brain. *Protoc. Exch.* <https://doi.org/10.1038/protex.2018.005> (2018).
42. Vanlandewijck, M., Andrae, J., Gouveia, L. & Betsholtz, C. Preparation of single cell suspensions from the adult mouse lung. *Protoc. Exch.* <https://doi.org/10.1038/protex.2018.006> (2018).
43. Picelli, S. et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013).
44. Kwan, K. M. et al. The Tol2kit: a multisite gateway-based construction kit for *Tol2* transposon transgenesis constructs. *Dev. Dyn.* **236**, 3088–3099 (2007).
45. Asakawa, K. et al. Genetic dissection of neural circuits by *Tol2* transposon-mediated Gal4 gene and enhancer trapping in zebrafish. *Proc. Natl. Acad. Sci. USA* **105**, 1255–1260 (2008).
46. Fukuhara, S. et al. Visualizing the cell-cycle progression of endothelial cells in zebrafish. *Dev. Biol.* **393**, 10–23 (2014).
47. Ando, K. et al. Clarification of mural cell coverage of vascular endothelial cells by live imaging of zebrafish. *Development* **143**, 1328–1339 (2016).
48. Kikuchi, K. et al. Retinoic acid production by endocardium and epicardium is an injury response essential for zebrafish heart regeneration. *Dev. Cell* **20**, 397–404 (2011).

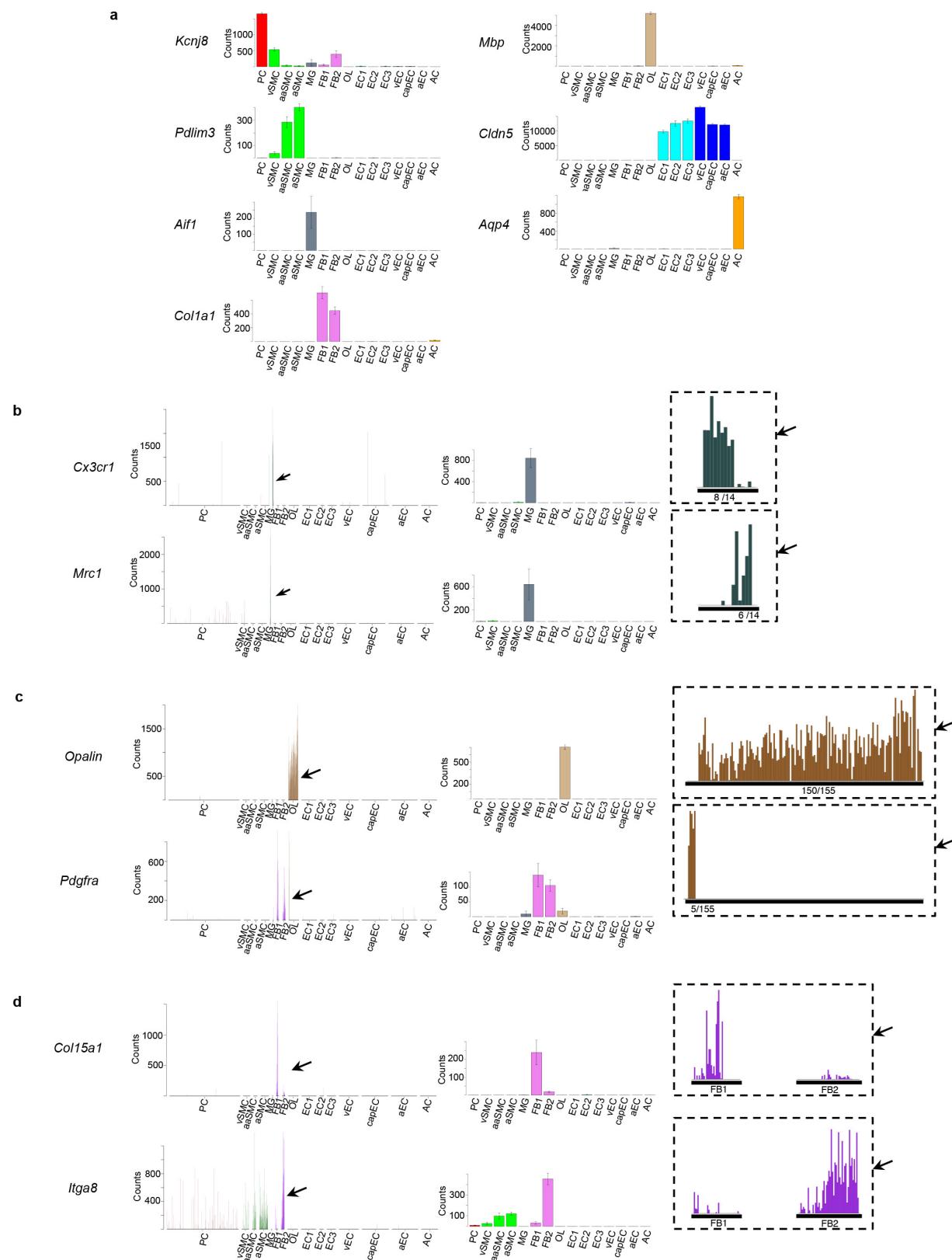
49. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
50. Ramsköld, D., Kavak, E. & Sandberg, R. How to analyze gene expression using RNA-sequencing data. *Methods Mol. Biol.* **802**, 259–274 (2012).
51. Storvall, H., Ramsköld, D. & Sandberg, R. Efficient and comprehensive representation of uniqueness for next-generation sequencing by minimum unique length analyses. *PLoS ONE* **8**, e53822 (2013).
52. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
53. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
54. Grün, D. et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525**, 251–255 (2015).
55. Forrest, A. R. et al. A promoter-level mammalian expression atlas. *Nature* **507**, 462–470 (2014).
56. Mackic, J. B. et al. Human blood-brain barrier receptors for Alzheimer's amyloid-beta 1–40. Asymmetrical binding, endocytosis, and transcytosis at the apical side of brain microvascular endothelial cell monolayer. *J. Clin. Invest.* **102**, 734–743 (1998).
57. Do, T. M. et al. Oatp1a4 and an L-thyroxine-sensitive transporter mediate the mouse blood-brain barrier transport of amyloid- $\beta$  peptide. *J. Alzheimers Dis.* **36**, 555–561 (2013).
58. Deane, R. et al. LRP/amyloid beta-peptide interaction mediates differential brain efflux of Abeta isoforms. *Neuron* **43**, 333–344 (2004).
59. Bell, R. D. et al. SRF and myocardin regulate LRP-mediated amyloid-beta clearance in brain vascular cells. *Nat. Cell Biol.* **11**, 143–153 (2009).
60. Xiong, H. et al. ABCG2 is upregulated in Alzheimer's brain with cerebral amyloid angiopathy and may act as a gatekeeper at the blood-brain barrier for Ab $\beta_{1–40}$  peptides. *J. Neurosci.* **29**, 5463–5475 (2009).
61. Agarwal, S. et al. Active efflux of Dasatinib from the brain limits efficacy against murine glioblastoma: broad implications for the clinical use of molecularly targeted agents. *Mol. Cancer Ther.* **11**, 2183–2192 (2012).
62. Lam, F. C. et al.  $\beta$ -Amyloid efflux mediated by p-glycoprotein. *J. Neurochem.* **76**, 1121–1128 (2001).
63. Corder, E. H. et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* **261**, 921–923 (1993).
64. Tarlungeanu, D. C. et al. Impaired amino acid transport at the blood brain barrier is a cause of autism spectrum disorder. *Cell* **167**, 1481–1494 (2016).
65. Dieterich, L. C. et al. Transcriptional profiling of human glioblastoma vessels indicates a key role of VEGF-A and TGF $\beta$ 2 in vascular abnormalization. *J. Pathol.* **228**, 378–390 (2012).



Extended Data Figure 1 | See next page for caption.

**Extended Data Figure 1 | Cell sorting, clustering and annotation and control for batch effects.** **a**, Representative FACS plots for cell sorts from the various transgenic reporter lines and C57Bl6 controls. For all sorts, the cells were first roughly selected with a broad gate based on forward scatter (FSC)/side scatter (SSC). This initial gate was then used as parent gate for cell selection based upon the relevant fluorescent signal of the different reporters. No additional gating was included, and the cut-off between positive and negative fluorescent signals was assessed using non-fluorescent C57Bl6 mice. **b**, BackSPIN analysis of cells obtained from *Cldn5*-GFP, *Pdgfrb*-GFP; *Cspg4*-DsRed and *Pdgfra*-H2BGFP mice. The BackSPIN algorithm splits the cells stepwise and also assigns genes uniquely to one branch with each split. At split level 1, all endothelial cells were assigned to the right branch, whereas all other cell types were assigned to the left branch. BackSPIN continues to split for as long as it recognizes expression heterogeneity among the genes assigned to that cluster leading to that some clusters continue to split dichotomously, whereas other clusters stop splitting at a certain level. Cluster 0, containing pericytes, stopped splitting at level 2. Any level can be arbitrarily chosen for cluster annotation; we chose level 6, at which 38 clusters had formed. **c**, The expression of the housekeeping gene *Actb* across the 38 clusters in **b** is illustrated as a bar plot, where each bar represents one cell (~3,500 in total). A gross annotation of the 38 clusters based on cell class-specific marker expression (see Fig. 1 for examples) is depicted below the bar plot. **d**, Abbreviations for the cell types. **e**, A reanalysis of cluster 0 from **b**, **c** by BackSPIN using all genes generated dichotomous splits over six levels,

showing that this cluster had stopped splitting in **b** because differentially expressed genes had been assigned to other branches. **f**, In the 64 clusters generated in **e**, expression of SMC markers, such as *Acta2*, was found in the right-most clusters. These clusters were subsequently identified as vSMC. **g**, The same tree as in **b**, with cluster 0 split in to PC and vSMC at level 7. Some clusters at level 6 appeared similar and were consolidated to a higher node, as indicated by the red boxes. **h**, The resulting simplified tree, which also shows the final cell type annotation that is used to display the data throughout the paper and at <http://betsholtzlab.org/VascularSingleCells/database.html>. **i**, An excerpt from the database illustrates the expression of *Actb* across the cell types. Cluster 1 at level 6 in **b**, **g** contained four cells displaying astrocyte markers; this cluster was considered too small and was removed from the dataset. Astrocytes were instead obtained from *Tagln-Cre;R26-stop-tdTomato* mice and added to the dataset for comparison. **j**, The position of the major cell classes in a t-SNE diagram, as annotated using marker expression. **k**, Representative marker distribution for the cell classes in t-SNE diagrams. **l**, **m**, Control for batch effects. Neither t-SNE nor SPIN showed any clustering bias related to mouse individual or experimental repetitions. **l**, The distribution of 1,385 mural cells colour-coded for their mouse origin in t-SNE (top) and SPIN diagrams (bottom: an arbitrarily chosen pan-mural cell marker *Myl9* is shown). **m**, The distribution of 1,100 endothelial cells in t-SNE (top) and SPIN (bottom: the zonated endothelial-specific transcript *Mfsd2a* is shown).

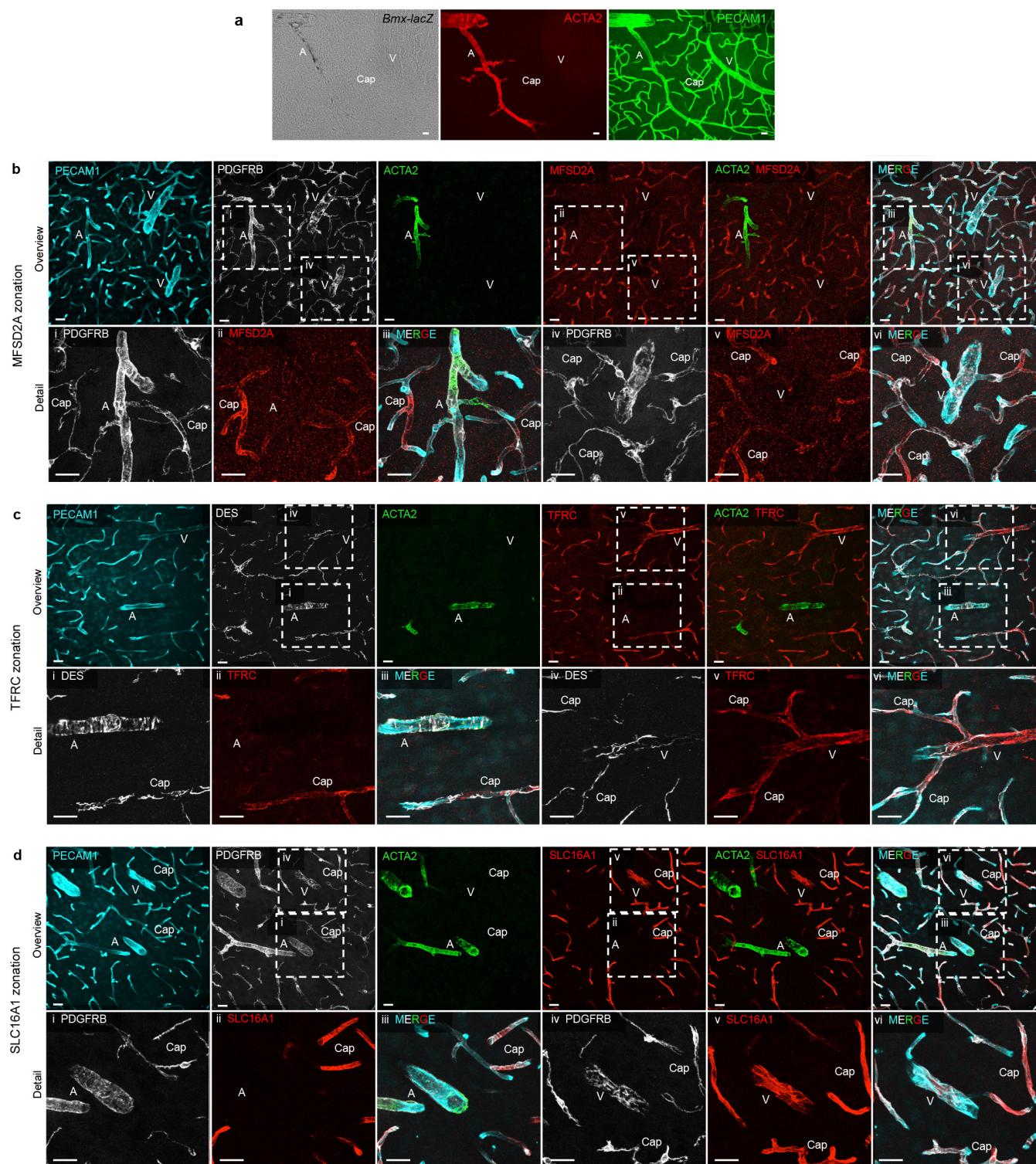


Extended Data Figure 2 | See next page for caption.

**Extended Data Figure 2 | Average expression and cell sub-types.**

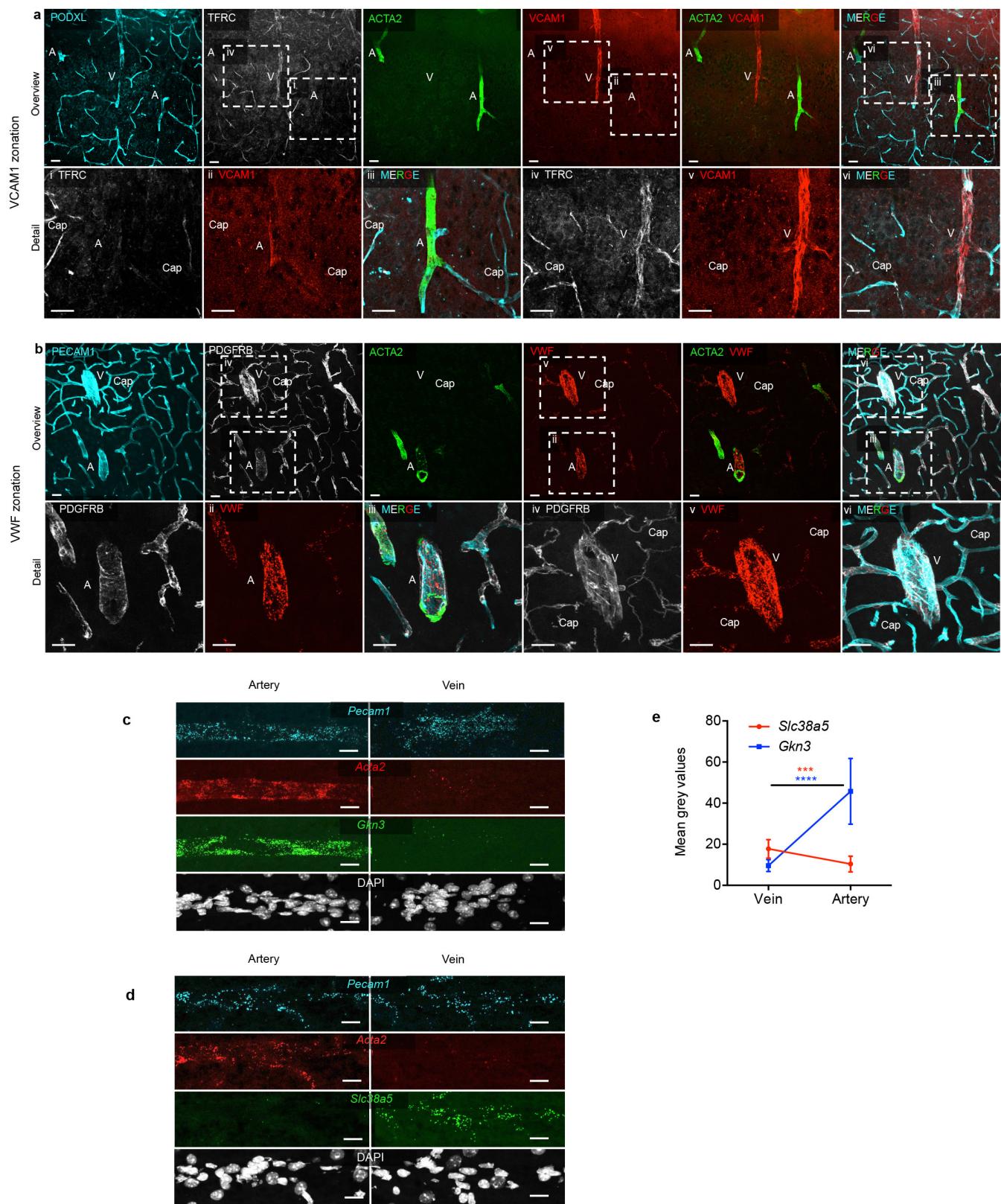
**a**, Average level of expression in the different cell types of the markers shown in Fig. 1b. A key to the cell type abbreviations is provided in Extended Data Fig. 1. Diagrams of the same type for all genes are available at <http://betsholtzlab.org/VascularSingleCells/database.html>. **b**, Microglial subtypes. The microglial markers *Cx3cr1* and *Mrc1* both mark the microglial (MG) cluster, as shown in the bar-plot diagrams (left) and in the average diagram (middle). Right: insets show magnified views of bar-plots indicated by black arrows. Of the 14 microglial cells present in the MG cluster, the eight left-most cells express *Cx3cr1* but not *Mrc1*, whereas the six right-most cells show the reverse pattern. Since *Cx3cr1* is a marker for interstitial microglia and *Mrc1* is a marker for perivascular microglia and macrophages, this pattern suggests that both these subtypes of microglia are present in the MG cluster, and that they are separated by

SPIN to opposite ends of the cluster. **c**, Oligodendrocyte (OL) subtypes. The marker for differentiating oligodendrocytes *Opalin* is expressed in the 150 right-most out of the 155 cells in the OL cluster, whereas the oligodendrocyte precursor cell (OPC) marker *Pdgfra* is expressed only in the five left-most cells (note that *Pdgfra* is also highly expressed in the FB1 and FB2 clusters). Since the OLs were sorted from an initial preparation of brain vascular fragments (in turn isolated using anti-PECAM1 coated magnetic beads) from *Pdgfra*-H2BGFP mice, the *Opalin*-positive cells appear physically associated with blood vessels. As for the microglial subtypes, SPIN separates the two oligodendrocyte lineage cell subtypes to opposite ends of the cluster. **d**, Subtypes of fibroblast-like cells. Numerous markers distinguish cells in the FB1 and FB2 clusters. Here we show the examples *Col15a1*, which is expressed preferentially in FB1 cells, and *Itga8*, which is expressed preferentially in FB2 cells.



**Extended Data Figure 3 | Endothelial arteriovenous zonation of *Bmx*, *MFSD2A*, *TFRC* and *SLC16A1*.** All panels show microscopic fields from cerebral cortex with different labels for endothelial cells (cyan: PECAM1), mural cells including SMCs and pericytes (white: PDGFRB or DES), arterial SMCs (green: ACTA2), specific zoned proteins (red: MFSD2A, TFRC or SLC16A1), separately, or in the indicated combinations. A, artery or arteriole; V, vein or venule; Cap, capillary/ies. **a**, Validation of *Bmx* as an arterial marker in adult mouse brain using *Bmx-lacZ* mice. ACTA2 immunofluorescence distinguishes arteries from veins and capillaries. PECAM1 immunofluorescence depicts all vessels. **b–d**, The first row of each panel shows the same microscopic field at overview magnification for different immunofluorescence stainings. Hatched boxes delineate the areas shown at higher magnifications in the second row of each

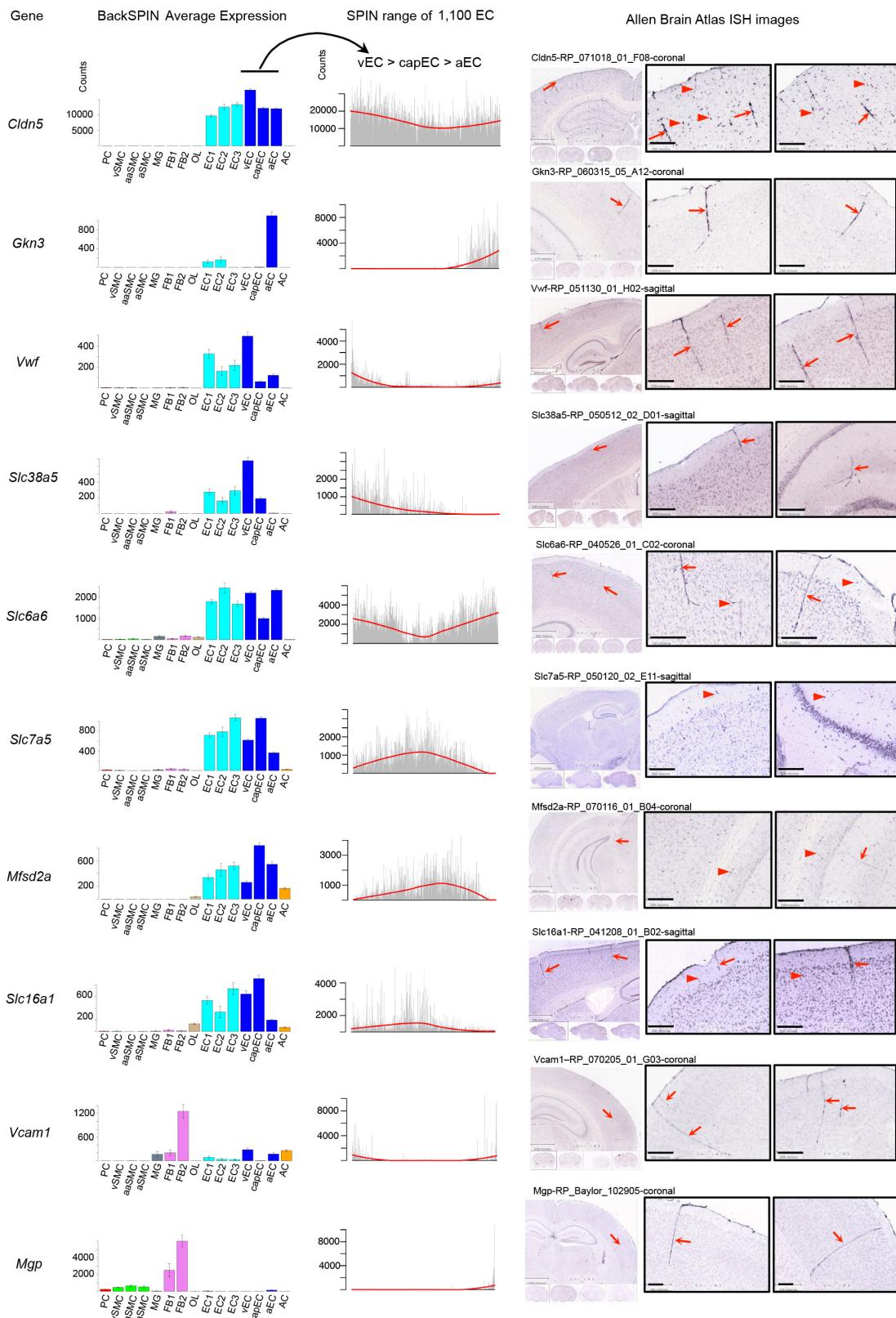
panel. The boxes are individually labelled i–vi. Arteries and arterioles are identified by ACTA2. Veins and venules are identified by vessel diameter in combination with the absence of ACTA2. Capillaries are identified by diameter and association with PDGFRB or DES-positive pericytes. An alternative choice of PDGFRB or DES labelling for mural cells was made to match the co-staining with other antibodies. **b**, Zonated expression of MFSD2A. Note the strong expression in capillaries and weak or absent staining in arteries/arterioles and veins. **c**, Zonated expression of TFRC. Note the strong expression in capillaries and veins/venules but absent staining in arteries/arterioles. **d**, Zonated expression of SLC16A1. Note the strong expression in capillaries and veins/venules but absent expression in arteries/arterioles. Data were reproduced in sections taken from three different mice. Scale bars, 20  $\mu$ m.



Extended Data Figure 4 | See next page for caption.

**Extended Data Figure 4 | Endothelial arteriovenous zonation of VCAM1, VWF, *Gkn3* and *Slc38a5*.** All panels show data from cerebral cortex. **a, b**, Immunofluorescence with different labels for endothelial cells (cyan, PECAM1 or PODXL), mural cells including SMC and pericytes (white, PDGFRB), arterial SMCs (green, ACTA2), and specific zoned proteins (red, VCAM1 or VWF), separately, or in the indicated combinations. The first row of each panel of **a, b** shows the same microscopic field at overview magnification for different immunofluorescent stainings. Hatched boxes delineate the areas shown at higher magnifications in the second row of each panel. The boxes are individually labelled i–vi. Arteries and arterioles (A) are identified by ACTA2. Veins and venules are identified by vessel diameter in combination with the absence of ACTA2. Capillaries are identified either by TFRC-positive staining and diameter, or by diameter and association with PDGFRB-positive pericytes. **a**, Zonated expression of VCAM1. Note the strong expression in veins, weak but clear expression in arteries/

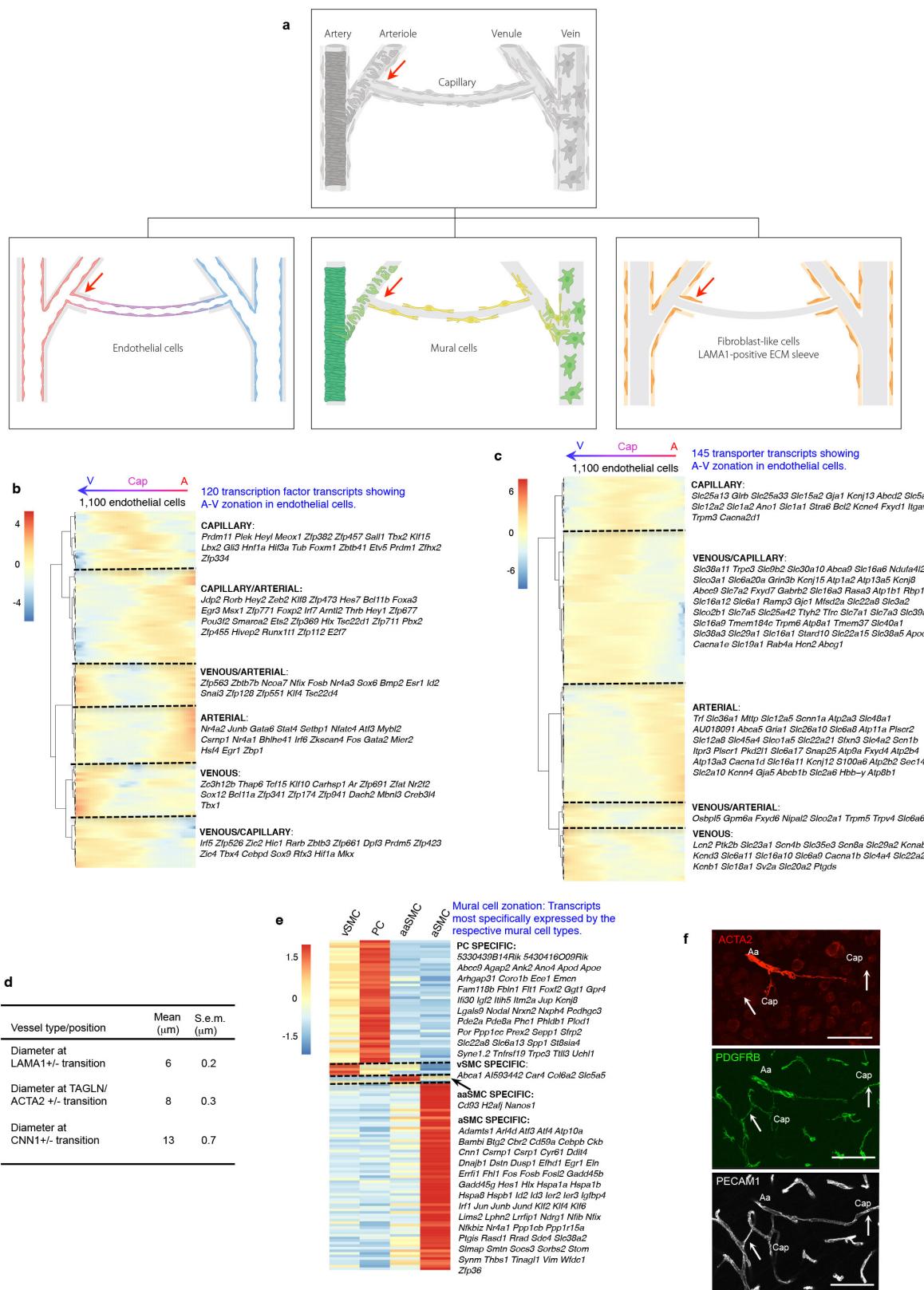
arterioles, and absent expression in capillaries. **b**, Zonated expression of VWF. Note the strong expression in arteries and veins, weak expression in arterioles and venules and weak expression in capillaries. **c, d**, RNA ISH for the indicated genes validates the zonal distribution of *Gkn3* to arteries and *Slc38a5* to veins. **e**, Quantification of the ISH results. The following number of mice and fields in total were analysed: *Slc38a5* ( $n = 3$ , 15 fields); *Gkn3* ( $n = 2$ , 20 fields). In Fiji software the ROI manager was used to mark and measure the mean grey values of the vascular segments of interest (vein, artery) and data were presented using GraphPad Prism 6 software. An unpaired *t*-test with Welsh correction was used to demonstrate statistical significance in gene expression between arteries and veins. *x*-axis shows the vascular segment, *y*-axis shows the mean grey values. Unpaired *t*-test with Welsh correction demonstrates statistically significant difference for both genes ( $****P < 0.0001$ ,  $***P < 0.001$ ). Data were reproduced in sections taken from three different mice. Scale bars, 20  $\mu\text{m}$ .



**Extended Data Figure 5** | See next page for caption.

**Extended Data Figure 5 | Validation of endothelial arteriovenous zoned gene expression using Allen Brain Atlas.** Allen Brain Atlas data for nine zonated endothelial transcripts (*Gkn3*, *Vwf*, *Slc38a5*, *Slc6a6*, *Slc7a5*, *Mfsd2a*, *Slc16a1*, *Vcam1* and *Mgp*) and one highly expressed endothelial transcript that does not show statistically significant zonation (*Cldn5*). The left panel of diagrams shows the expression distribution across all vascular cell types clustered by BackSPIN and depicted by average expression for each cluster. These diagrams are excerpts from <http://betsholtzlab.org/VascularSingleCells/database.html>, where bar plots can also be accessed and viewed at high magnification. A key to the cell type abbreviations is provided in Extended Data Fig. 1. The middle panel of diagrams shows bar-plots for 1,100 endothelial cells (the clusters indicated by the black arrow) organized in one dimension by SPIN. The grey bars show the expression level in individual cells and the red curve (loess function) depicts the expression pattern across the range. The arteriovenous zonation indicated above the middle panel of diagrams is

suggested by the data shown in Fig. 2 and Extended Data Figs 3, 4: venous endothelial cells to the left, capillary in the middle and arterial cells to the right. The right panel of micrographs shows excerpts from the Allen Brain Atlas, one overview and two at higher magnification (black frames). Genes with preferential expression in arteries (for example, *Gkn3*, *Mgp*), veins (for example, *Slc38a5*) or both (for example, *Vwf*, *Vcam1* and *Slc6a6*) show distinct signals in large penetrating as well as in deep vascular structures. Genes that are highly expressed in capillaries (*Mfsd2a*, *Slc16a1*, and *Cldn5*) show distinct labelling of small vessel structures at all locations. The strong expression of *Vcam1* and *Mgp* in fibroblast-like cells may contribute to the labelling of large vessels, although immunofluorescence shows a preferential localization of *Vcam1* to endothelial cells (Fig. 3). Red arrows point to labelled profiles consistent with arteries/arterioles or veins. Red arrowheads point to profiles consistent with microvessels. Allen Brain Atlas experiment codes are provided above each overview image. Scale bars, 200  $\mu$ m.

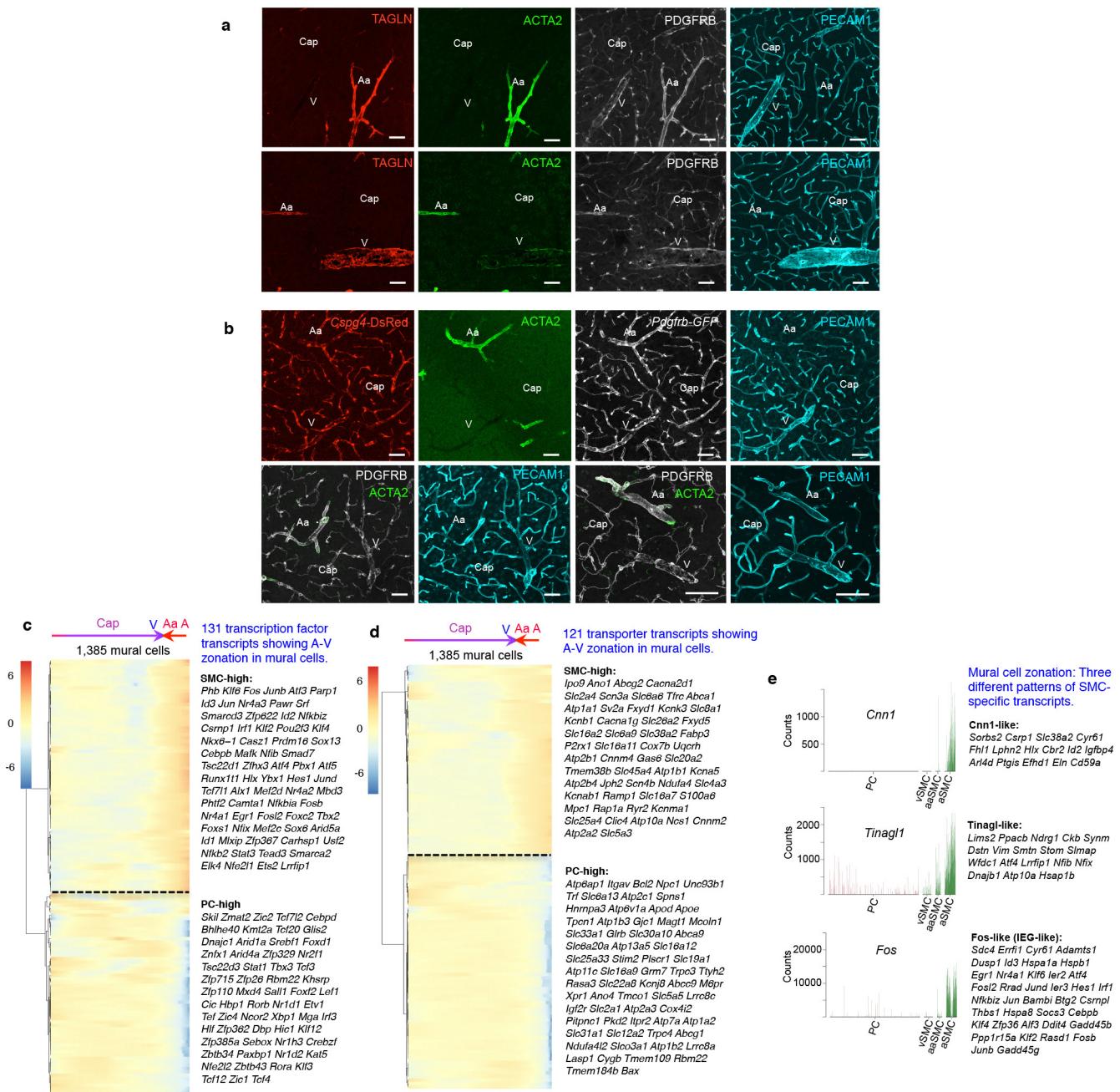


Extended Data Figure 6 | See next page for caption.

**Extended Data Figure 6 | Zonation principles, endothelial transcription factors and transporters, and genes specific to mural cell subtypes.**

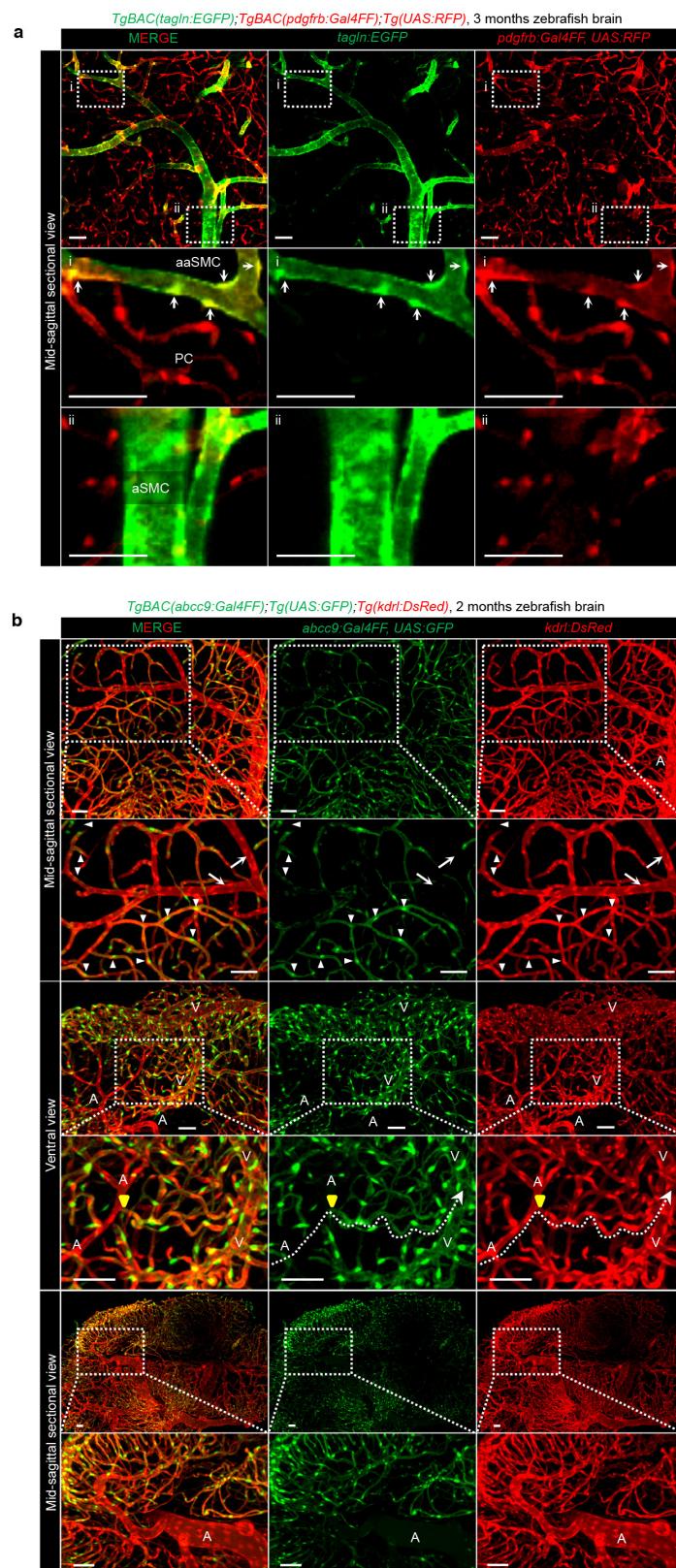
**a**, Schematic illustration of zonation principles and distribution of cell types. Endothelial arteriovenous zonation: red–blue; mural arteriovenous zonation: green–yellow; fibroblast-like cells and LAMA1 sleeve: brown. Arrow, arteriolar–capillary boundary. **b, c**, Heat map of relative expression (red, high; blue, low) of 120 transcription factor (**b**) and 145 transporter (**c**) transcripts across the SPIN range of 1,100 endothelial cells. Venous (V), capillary (Cap) and arterial (A) position in the range is depicted at the top. The identities of the transcripts distributing according to six or five major patterns of zonation, respectively, are provided to the right of the

heat maps. **d**, Morphometric analysis of vessel diameter at the transitions between cells positive (+) or negative (−) for immunofluorescent labelling of the indicated proteins. The following number of mice and vessel profiles were analysed for each value: diameter at LAMA1 +/− transition, 5 mice, 49 vessels; diameter at TAGLN/ACTA2 +/− transition, 3 mice, 66 vessels; diameter at CNN1+− transition, 3 mice, 34 vessels. **e**, Heat map of relative expression (red, high; blue, low) of the transcripts most specific for each subtype of mural cells ordered according to anatomical axis (vSMC, PC, aaSMC and aSMC). Transcript identities for each group are provided to the right of the heat map. **f**, Immunofluorescence shows abrupt transition from arteriolar SMCs to pericytes (arrows). Scale bars, 50 μm.



**Extended Data Figure 7 | Zonation of mural cell markers in mouse brain.** **a, b**, Immunofluorescence with different labels for endothelial cells (cyan, PECAM1), and mural cells including SMCs and pericytes (white, PDGFRB), arterial SMC (green, ACTA2), specific zoned proteins/reporters (red, TAGLN or CSPG4-DsRed) or in the indicated combinations. All images are from cerebral cortex. These panels show protein distribution to vessel types that match the zonal distribution of transcripts to mural cell subtypes, as shown in Fig. 3a. Data were reproduced in sections taken from three different mice. Scale bars, 50  $\mu$ m. **c, d**, Heat map of relative expression (red, high; blue, low) of 131 transcription factor (**c**) and 121 transporter (**d**) transcripts across the SPIN

range of 1,385 mural cells. The anatomical positions (capillary (Cap), venous (V), arteriolar (Aa) and arterial (A)) are shown at the top. Note the presence of two zones and their non-alignment with the anatomical axis (punctuated zonation where arterioles and veins meet). The identity of the transcripts distributing according to the two major patterns of zonation (SMC-high and PC-high) is provided to the right. **e**, Examples of three arterial SMC transcript distribution patterns exemplified by *Cnn1*, *Tinagl1* and *Fos* and other transcripts with similar distribution. The transcripts with *Fos*-like distribution include products of numerous immediate early genes (IEGs), including *Egr1*, *Fosb*, *Jun* and *Junb*.

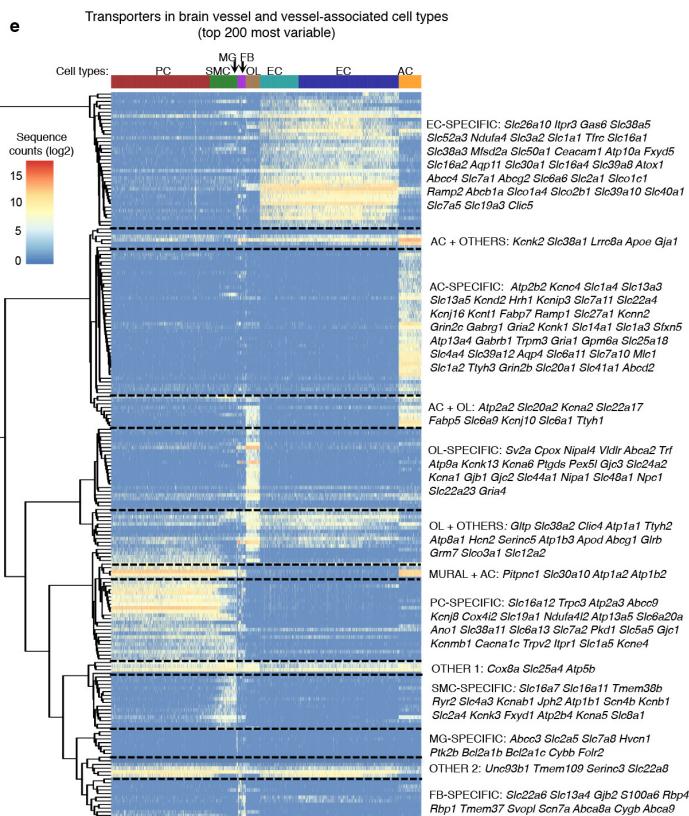
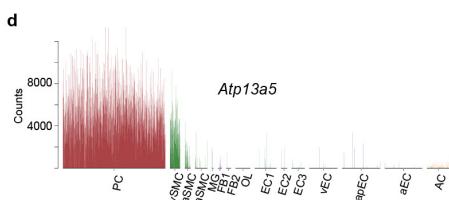
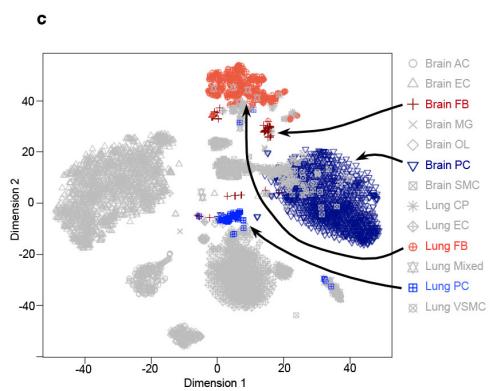
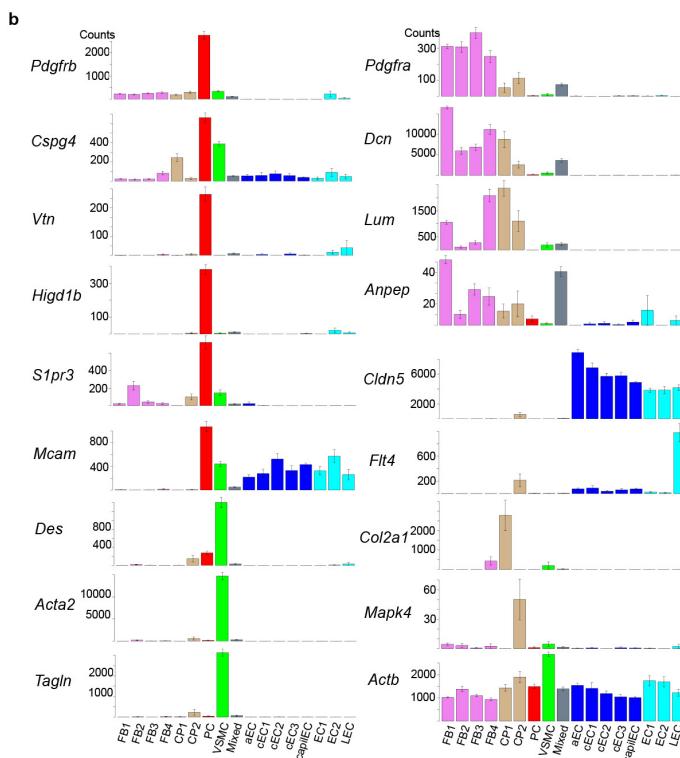
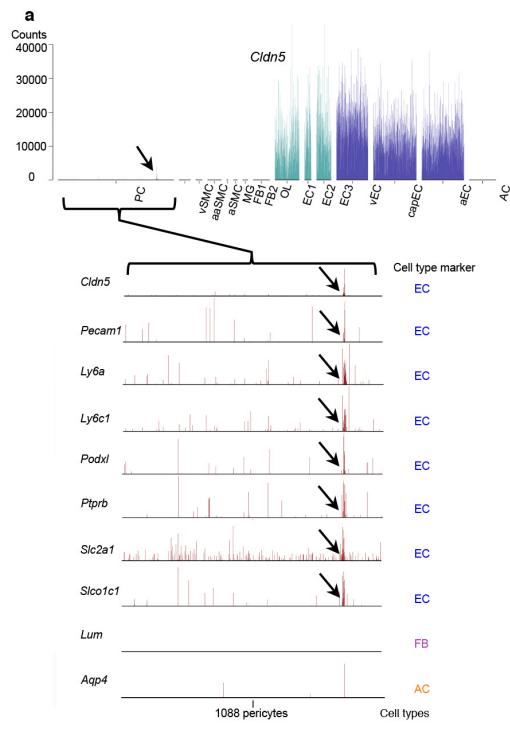


Extended Data Figure 8 | See next page for caption.

**Extended Data Figure 8 | Zonation of mural cell markers in zebrafish brain.**

**a**, The zonated distribution of *tagln* and *pdgfrb* as shown in the three-month-old zebrafish brain using transgenic reporter fish. GFP is expressed from the *tagln* promoter and RFP from the UAS promoter activated by *pdgfrb* promoter-driven Gal4FF. Similar to mice, *tagln* is expressed in a/aaSMCs (white arrows, inset i) whereas *pdgfrb* marks mural cells along all vessel types, including capillary pericytes (PC). The lower panels show insets at high magnification. Note the lower expression of *pdgfrb* in the aSMCs of the larger arteries (inset ii). **b**, GFP expression from a UAS promoter activated by Gal4FF expressed from the *abcc9* promoter in forebrain mural cells of a two-month-old zebrafish. Endothelial cells are labelled by DsRed expression from the *kdr1* (*Vegfr2*) promoter. Note that,

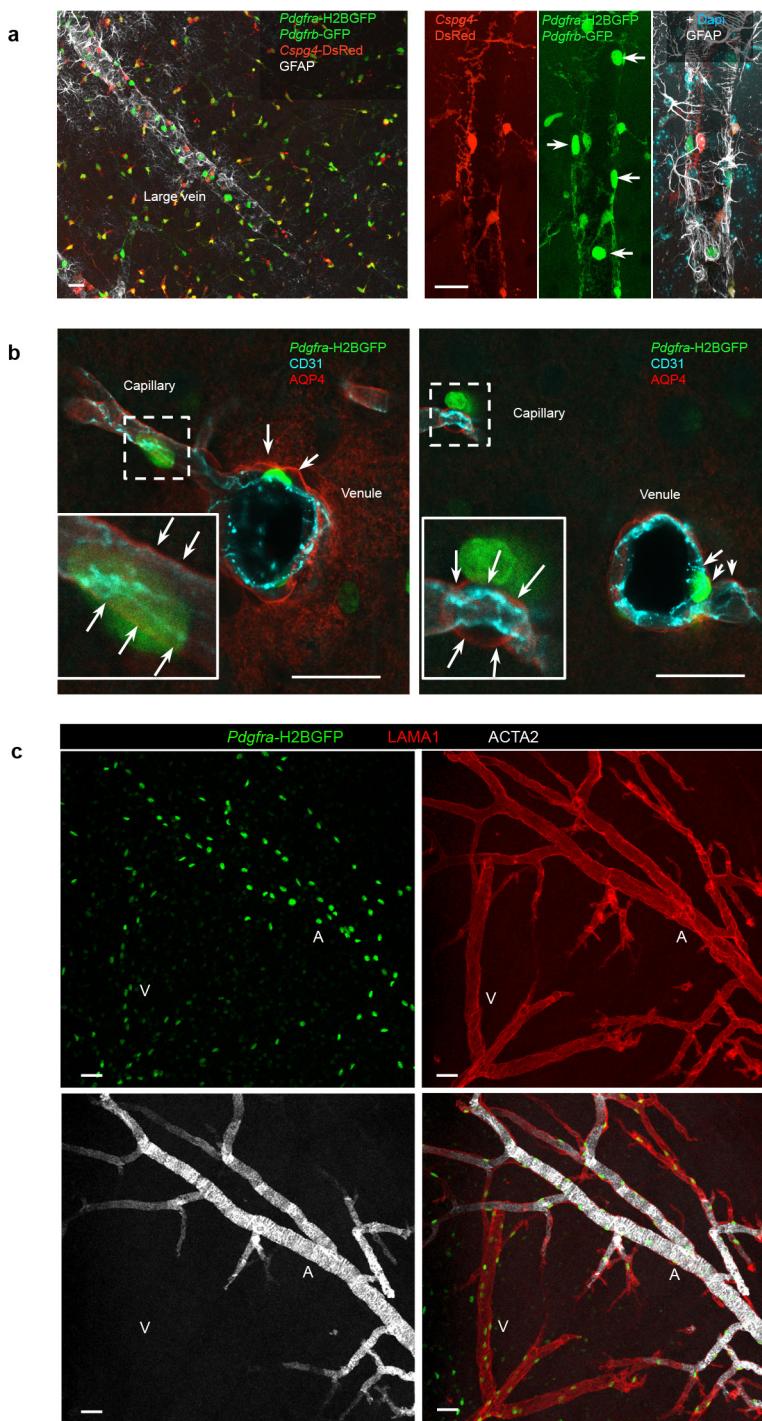
similar to the expression pattern of *Abcc9* in mice, fish *abcc9* expression is localized mainly to capillary pericytes and venous mural cells, but not to arteries or arterioles (A). Top, a midsagittal view shows GFP-positive pericytes (white arrowheads) on capillary branches, whereas the SMC of the feeding arteriole as well as the upstream artery (A, white arrows) are GFP-negative. Middle, in a ventral view of the capillary network branching off an arteriole (A), numerous GFP-positive pericytes are seen, while the arteriole is negative. An arteriole–capillary transition is indicated by a yellow arrowhead. Dashed line indicates a vascular continuum stretching from the arteriole (A) across capillary branches to a venule (V). Bottom, another midsagittal view illustrates that arterial SMCs (A) lack *abcc9* expression. Scale bars, 50  $\mu$ m.



Extended Data Figure 9 | See next page for caption.

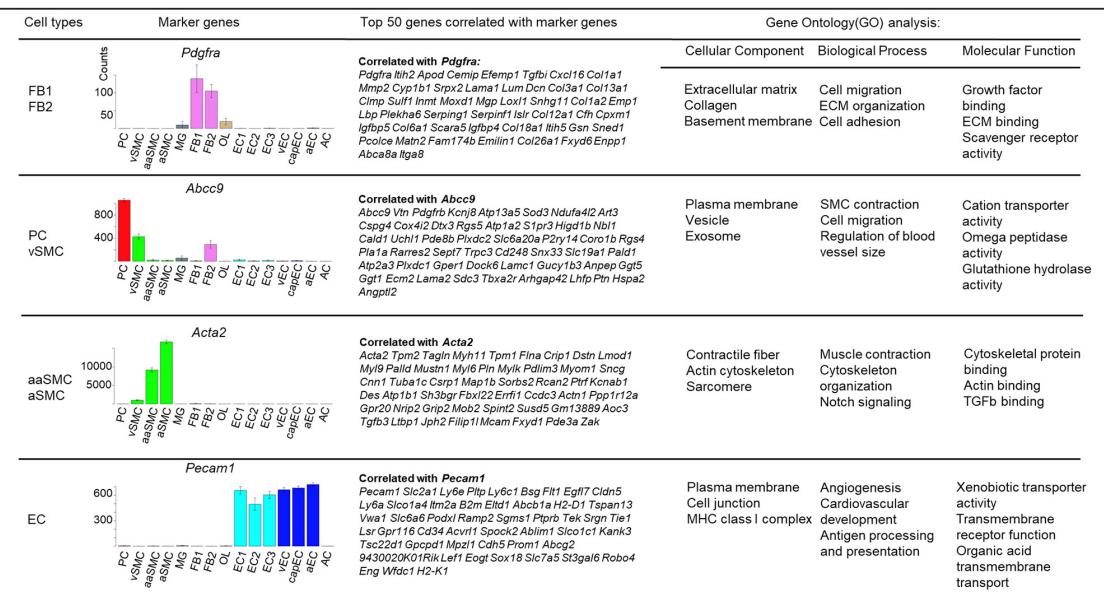
**Extended Data Figure 9 | Pericyte definition, organotypicity and transporter transcript expression.** **a**, Bar plot showing the expression of the highly abundant endothelial marker *Cldn5* across the BackSPIN clusters; note the abundant and specific expression in EC clusters. The arrow points to a small group of cells in the PC cluster showing *Cldn5* expression. The BackSPIN range of 1,088 pericytes is shown at larger magnification in the lower panels. An additional seven high abundant endothelial markers (*Pecam1*, *Ly6a*, *Ly6c1*, *Podxl*, *Ptprb*, *Slc2a1* and *Slco1c1*) all show a peak of expression in the same roughly 10 cells, suggesting endothelial contamination of these pericytes. The high abundance markers for fibroblast-like cells (*Lum*) and astrocytes (*Aqp4*) do not reveal signs of contamination of the pericytes by these cell types. **b**, Annotation of lung vascular cell BackSPIN clusters and identification of lung pericytes. The PC cluster was identified according to the expression of known and/or shared markers with brain pericytes, combined with the absence of markers for vascular smooth muscle cells (VSMC), fibroblasts

(FB), chondrocytes/perichondrium (CP), or endothelial cells (EC). The mixed cluster contains markers for epithelial cells together with markers for fibroblasts and probably represents contaminated cells. Among the endothelial cells (EC), several clusters were identified, including arterial (a), capillary (c/capil) and lymphatic (LEC). **c**, t-SNE analysis of brain and lung cells shows separation of brain and lung pericytes, as well as of brain and lung fibroblast-like cells, into distinct clusters suggesting organotypic differences for both cell classes. **d**, Bar plot of expression of the transporter transcript *Atp13a5* in brain shows that it is highly specific for pericytes and vSMCs. **e**, Heat map illustrating the expression (red, high; blue, low) of the 200 most variable transporter transcripts across the brain vascular and vessel-associated cell types. Hierarchical clustering divides the patterns of transporter distribution into approximately 13 groups (separated by hatched lines). The identity of the transcripts in each group is provided to the right of the heat map. Note the abundance of transporters that are pericyte-specific.

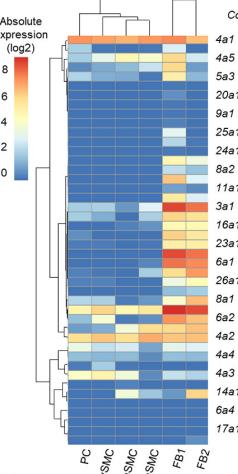
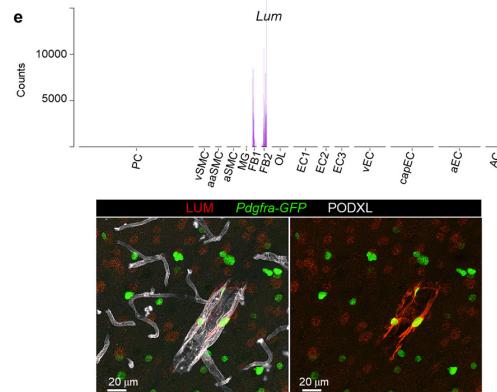
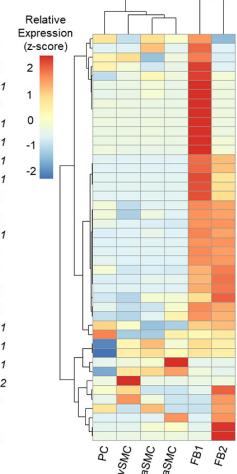


**Extended Data Figure 10 | Localization of perivascular fibroblast-like cells.** **a**, *Pdgfra*-H2BGFP-positive cells in veins. Right panels show that the H2BGFP-positive nuclei are distinct from *Cspg4*-DsRed- and *Pdgfrb*-GFP-positive mural cells and GFAP-positive astrocytes. White arrows indicate *Pdgfra*-H2BGFP-positive nuclei. **b**, *Pdgfra*-H2BGFP-positive nuclei associated with venules are located between the endothelial cells and AQP4-positive astrocyte end-feet. Capillary-associated *Pdgfra*-H2BGFP-

positive nuclei are located outside the astrocyte end-feet (insets); these cells probably represent oligodendrocyte lineage cells. **c**, The distribution of LAMA1 correlates with the distribution of vessel-associated *Pdgfra*-H2BGFP-positive cells. Image in **c** is from striatum and images in **a** and **b** are from cerebral cortex. Data were reproduced in sections taken from three different mice. Scale bars, 20  $\mu$ m.

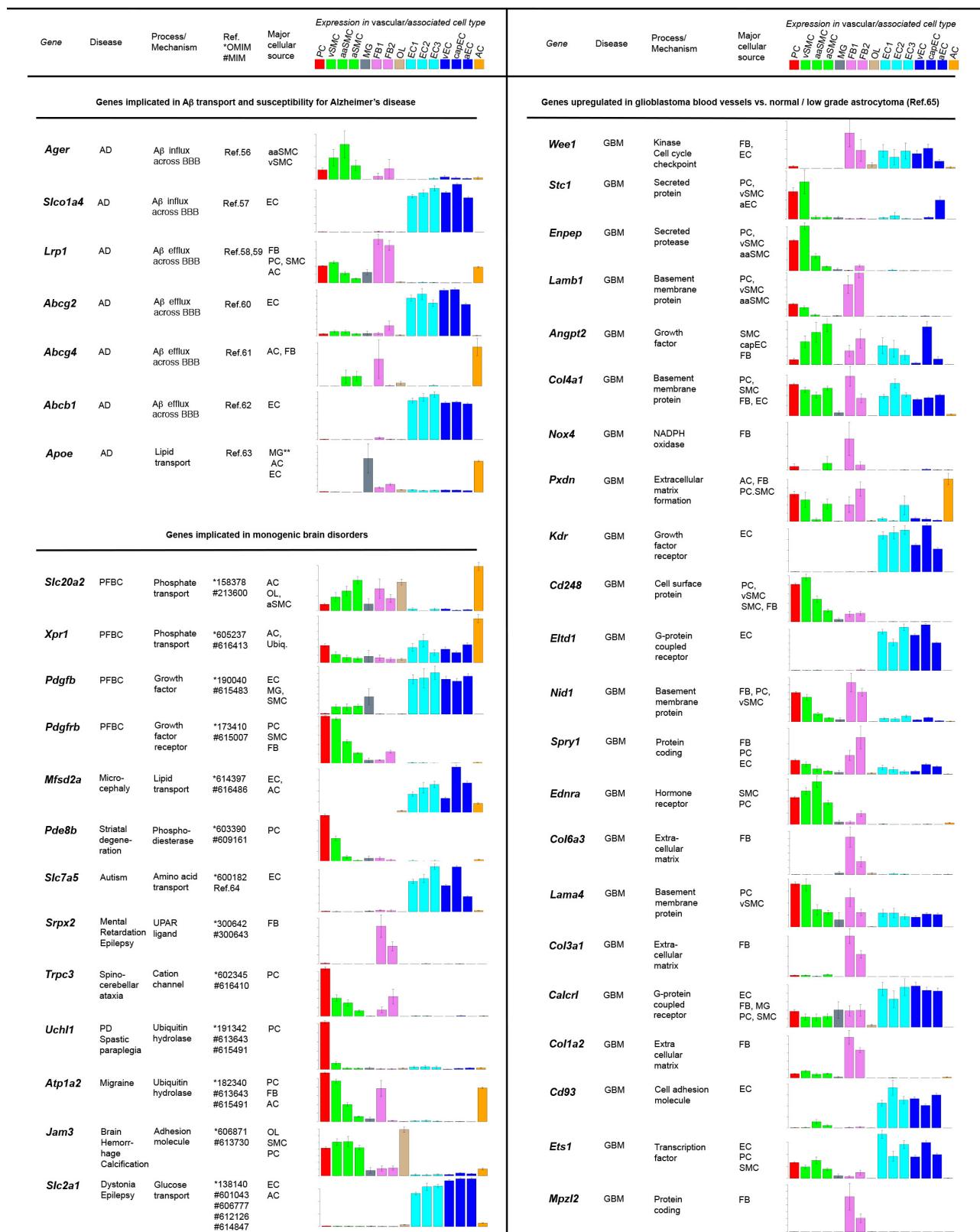
**a****b**

Component	ECM		Innate immunity	Lung fibroblast expression	Component	ECM		Innate immunity	Lung fibroblast expression
	Enzyme/inhibitor	Binding				Enzyme/inhibitor	Binding		
<i>Pdgfra</i>			X		<i>Lbp</i>	X	X	X	
<i>Ith2</i>	X	X	X		<i>Plekha6</i>		X		
<i>Apod</i>					<i>Serping1</i>	X	X		
<i>Cemip</i>	X		X		<i>Islr</i>				
<i>Etemp1</i>	X		X		<i>Col12a1</i>	X		X	
<i>Tgfb1</i>	X		X		<i>Cfn</i>		X		
<i>Cxcl16</i>			X		<i>Cpxm1</i>				
<i>Col1a1</i>	X		X		<i>Igfbp5</i>		X		
<i>Mmp2</i>	X		X		<i>Col6a1</i>	X			
<i>Cyp12a1</i>			X		<i>Scara5</i>				
<i>Sirpx2</i>			X		<i>Igfbp4</i>				
<i>Lama1</i>	X				<i>Col18a1</i>	X			
<i>Lum</i>	X		X		<i>Itih5</i>	X	X		
<i>Dcn</i>	X		X		<i>Gsn</i>				
<i>Col3a1</i>	X		X		<i>Sned1</i>	X			
<i>Col13a1</i>	X		X		<i>Pcolce</i>	X			
<i>Cimp</i>					<i>Matn2</i>	X			
<i>Sulf1</i>	X		X		<i>Fam174b</i>				
<i>Inmt</i>			X		<i>Emilin1</i>	X			
<i>Moxd1</i>					<i>Col26a1</i>	X			
<i>Mgp</i>	X		X		<i>Fxyd6</i>				
<i>Loxl1</i>	X		X		<i>Enpp1</i>	X			
<i>Snhg11</i>					<i>Abca8a</i>				
<i>Col1a2</i>	X		X		<i>Itgae8</i>	X			
<i>Emp1</i>	X		X						

**c****d**

**Extended Data Figure 11 | Characterization of fibroblast-like cells in the brain.** **a**, GO analysis of the 50 most specific transcripts for the following four cell categories: fibroblast-like cells (FB1, FB2), pericytes and venous SMCs (PC, vSMC), arteriolar and arterial SMCs (aaSMC, aSMC) and endothelial cells (EC). The transcripts were selected on the basis of their distribution correlation with *Pdgfra*, *Abcc9*, *Acta2* and *Pecam1*, respectively. Right panel shows a selection of significant terms for each cell class in the three mentioned categories. **b**, Further categorization of the 50 top markers for fibroblast-like cells reveals that most of them are associated with ECM components, turnover, or function. A few are related to innate immunity function. Nearly all markers are also expressed by lung fibroblasts.

fibroblasts. **c, d**, Expression heat map (blue, low; red, high) of collagen transcripts in mural cells and fibroblast-like cells. **c**, Absolute expression. **d**, Relative expression. Note the abundance of collagen transcripts in the fibroblast-like cell clusters (FB1 and FB2) and the difference between them, indicating FB subtypes. **e**, Bar plot distribution of *Lum* shows strong specificity for the FB clusters. Immunofluorescence for LUM shows its specific localization to blood vessels larger than capillaries and a correlation with the distribution of *Pdgfra*-H2BGFP-positive perivascular cells. **f**, Bar plot distribution of *Dcn* shows strong specificity for the FB clusters. Allen Brain Atlas images confirm expression in large penetrating brain vessels (red arrows).



Extended Data Figure 12 | See next page for caption.

**Extended Data Figure 12 | Vascular cell type-specific expression of genes linked to brain diseases.** Brain vascular scRNA-seq data provide insight into brain disease mechanisms. Three types of disease mechanisms are illustrated. For genes implicated in the transport of Alzheimer's amyloid- $\beta$  (A $\beta$ ) peptide across the BBB, it is noteworthy that several different cell types of the neurovascular unit are implicated in both A $\beta$  influx and efflux mechanisms. Surprisingly, *Lrp1* shows its highest expression in perivascular fibroblast-like cells (FB) and pericytes (PC) or vSMCs, but no detectable expression in endothelial cells (EC). \*OMIM and #MIM numbers are provided for the monogenic brain disorders. We list the four genes known to be mutated in primary familial brain calcification (PFBC) to illustrate the different cellular expression of these genes, suggesting that changes in cell–cell communication in the vascular wall plays a role in the pathogenesis of this disease. It is noteworthy that several of the genes mutated in other listed monogenic brain disorders

are commonly expressed in the cell types molecularly defined through the present study, such as *Srpx2* (FB) and *Uchl1* (PC). Other transcripts, including *Slc2a1* and *Mfsd2a*, show the expected expression in EC. Unexpectedly, *Jam3*, which is thought to be expressed mainly in EC, has its major vascular expression in mural cells and oligodendrocytes (OL). Of genes overexpressed in glioblastoma grade IV (GBM) vessels, it is striking that most have their normal expression site in FB and PC or vSMC. This may suggest that GBM vasculature has an altered cellular composition, including a higher abundance of mural and perivascular cells, possibly reflecting a fibrotic reaction originating from FB. The diagrams display relative levels of expression. High-resolution quantitative versions of the diagrams and single cell bar plots can be viewed at <http://betsholtzlab.org/VascularSingleCells/database.html>. \*\*A detailed analysis of the expression of the microglial expression of *Apoe* shows that it is predominantly expressed in perivascular brain macrophages.

## Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### ► Experimental design

#### 1. Sample size

Describe how sample size was determined.

In this study, we aimed to sequence as many vascular single cells as possible and we achieved 3186 brain vascular cells and 1504 lung vascular cells. To our knowledge, this is the largest study ever performed in the vascular research community. No statistical methods were used to predetermine sample size.

#### 2. Data exclusions

Describe any data exclusions.

We have used the following pre-established cell exclusion criteria: cells with total aligned reads less than 100,000 were removed, and also outlier cells with maximum spearman's correlation with other cells less than 0.3 were filtered out.

#### 3. Replication

Describe whether the experimental findings were reliably reproduced.

The gene expression profiles we reported are observed from dozens to hundreds of individual cells. All experimental findings were reliably reproduced

#### 4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Mice were selected for primary brain/lung cell isolation based on their genotype (i.e. fluorescent reporter expression) and age (10-19 weeks old, adult). All other criteria were not considered and as such, randomized.

#### 5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Selection of a specific mouse with the appropriate genotype was completely random. Clustering of the single cell data was performed using the BackSPIN algorithm, which is working in a completely unbiased fashion

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

## 6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- The test results (e.g.  $P$  values) given as exact values whenever possible and with confidence intervals noted
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

## ► Software

Policy information about [availability of computer code](#)

### 7. Software

Describe the software used to analyze the data in this study.

For analyse the data in this study, we used all academical softwares. For clustering the cells, we used BackSPIN software (<https://github.com/linnarsson-lab/BackSPIN>). To compare the gene expression difference between different clusters, we used R edgeR software (R version 3.3.2, edgeR version 3.16.4). The 2D tSNE visualization was performed using R tsne package (version 0.1-3). The heat map visualization was performed using R pheatmap packages (version 1.0.8). FACS plots were analyzed with FlowJo v10.1 and FACSDiva v8.0.2. Image acquisition and adjustment was done with Leica Application Suite 3.2.1.9702, Fiji v1.51s and PhotoShop CC 2015, while the former two softwares were also used for quantifications. The GO analysis was performed using GOSTAT packages (version 2.40.0) in R software, with the associated annotation packages GO.db (version 3.4.0) and org.Mm.eg.db (version 3.4.0).

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). [Nature Methods guidance for providing algorithms and software for publication](#) provides further information on this topic.

## ► Materials and reagents

Policy information about [availability of materials](#)

### 8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No unique materials were used

## 9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

Antibodies were used in this study for immunofluorescence only, as validation of the transcriptomic changes we found in the single cell database. The study does not include any claims based on antibody staining only. Specificity of primary antibodies was verified in two ways: 1) Negative controls were performed using only secondary antibody. 2) Antibody staining was verified to occur only in the correct anatomical location and cell type (based on single cell transcriptomics both from our own database and the Zeisel et al (Science, 2015) database (PMID: 25700174 DOI: 10.1126/science.aaa1934) . To assess this, we have used the following reporter mice to visualize components of the brain vasculature:

- All endothelial specific antibodies were routinely used in conjunction with each other and the endothelial specific Cldn5-GFP reporter mouse: CD31, Glut1, ColIV, TFRC, MFSD2a, ERG, VCAM1, Podocalyxin, VWF, AQP4 (astrocyte end-feet localized on the outer layer of vessels)
- All mural cell specific antibodies were routinely used in conjunction with each other and the mural cell specific PDGFRb-GFP reporter mouse: CD13, Acta2-Cy3, Acta2-Alexa fluor 647, Acta2-FITC, PDGFRb, TAGLN, CNN1, DES
- All perivascular fibroblast-like cell specific antibodies were used in conjunction with each other and the PDGFRa-H2BGFP reporter mouse: LAMA1, LUM, CD13
- GFAP staining (Z0334) was in agreement with the known morphology of astrocytes.

Specific use of all antibodies is detailed below:

- PECAM1 (CD31), dilution 1:250 R&D Systems cat: AF3628, lot not available as this antibody is heavily used with several lot numbers simultaneously. We have seen no variation between lot numbers in more than 5 years of use.
- CD13 (anpep), dilution 1:100, Bio-Rad, cat: MCA2183EL, lot: 1113, clone R3-63.
- GLUT1, dilution 1:200, Merck Millipore, cat: 07-1401, lot: 2630745.
- ColIV, dilution 1:100, Bio-Rad, cat: 2150-1470, lot: 161115
- ACTA2-Cy3, dilution 1:200, Sigma, cat: C6198, lot: 037M4783V, clone 1A4.
- ACTA2-Alexa Fluor 647, dilution 1:200, Santa Cruz, cat: sc-32251, lot: E1916, clone 1A4.
- ACTA2-FITC, dilution 1:200, Sigma, cat: F3777, lot: 105M4838V, 027M4784V, clone 1A4
- TFRC, dilution 1:200, Novus Biologicals; Thermo Fischer Scientific, cat: NB100-64979; 13-6800, lot: 1607, RB232679, clone H68.4
- LAMA1, dilution 1:800 was a kind gift from Prof. Lydia Sorokin. Specificity of the antibody has been assessed previously on a KO mouse for LAMA1 in Edwards et al, JBC 2010: PMID: 20048158, DOI: 10.1074/jbc.M109.069575.
- PDGFRb, dilution 1:100, eBioscience, cat: 14-1402, lot: E04060-1632, clone APB5
- TAGLN (SM22), dilution 1:100, Abcam, cat: ab14106, lot: GR3178029-1
- CNN1, dilution 1:100, Abcam, cat: ab46794, lot: GR144432-38
- AQP4, dilution 1:100, Millipore. cat: ab2218, lot: 2827164
- MFSD2a, dilution 1:100 was a kind gift from Prof. David Silver. Specificity of the antibody has been assessed previously on a KO mouse for MFSD2a in Nguyen et al, Nature 2014: PMID: 24828044 DOI: 10.1038/nature13241
- DES, dilution 1:100, Abcam, cat: ab15200, lot: GR94416-1
- ERG, dilution 1:200, Abcam, cat: ab92513, lot: GR219881-25
- VCAM1, dilution 1:100, Merck, cat: CBL1300, lot: 2836934
- VWF, dilution 1:100, Dako, cat: A00822, lot: 20044707
- SLC16a1, dilution 1:100, Acris, cat: TA321556, lot: D814AA0111
- PODXL, dilution 1:500, R&D Systems, cat: AF1556, lot: JPC0114021, JPC0115031
- LUM, dilution 1:100, Abcam, cat: ab98067, lot: GR86656-1
- GFAP, dilution 1:100, Dako, cat: Z0334, lot: 20019135

The information above can also be found in Supplementary Table 4.

For all antibodies (except LAMA1 and MFSD2a), additional information on specificity and species cross-reactivity, with links to key publications can be found on the manufacturer's website.

## 10. Eukaryotic cell lines

- a. State the source of each eukaryotic cell line used.
- b. Describe the method of cell line authentication used.
- c. Report whether the cell lines were tested for mycoplasma contamination.
- d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

No eukaryotic cell lines were used

No eukaryotic cell lines were used

No eukaryotic cell lines were used

No commonly misidentified cell lines were used

## ► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

### 11. Description of research animals

- Provide details on animals and/or animal-derived materials used in the study.

The following transgenic mouse strains were used: C57Bl6 (The Jackson laboratory, C57BL6/J), Cspg4-dsRed (The Jackson laboratory, Tg(Cspg4-dsRed.T1)1Akik/J, Pdgfrb-eGFP (Gensat.org. Tg(Pdgfrb-eGFP)JN169Gsat/Mmucd), Pdgfra-H2B-eGFP (Pdgfratm11(EGFP)Sor, gift from Prof. Philippe Soriano), Cldn5-eGFP (Tg(Clnd5-GFP)Cbet/U, described previously), SM22-Cre (Tg(Tagln-cre)1Her/J and R26-stop-tdTomato (B6;129S6-Gt(ROSA)26Sortm14(CAG-tdTomato)Hze). All animals were backcrossed on a C57BL6/J genetic background. Adult mice of either sex, with age ranging between 11 and 19 weeks were used for all experiments.

Policy information about [studies involving human research participants](#)

### 12. Description of human research participants

- Describe the covariate-relevant population characteristics of the human research participants.

The study did not involve human research participants

## Flow Cytometry Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

### ► Data presentation

For all flow cytometry data, confirm that:

- 1. The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- 2. The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- 3. All plots are contour plots with outliers or pseudocolor plots.
- 4. A numerical value for number of cells or percentage (with statistics) is provided.

### ► Methodological details

5. Describe the sample preparation.  

Detailed description of sample preparation can be found in Methods section, with links to the relevant Protocols Exchange submissions.
6. Identify the instrument used for data collection.  

Becton Dickinson FACSAria III
7. Describe the software used to collect and analyze the flow cytometry data.  

FlowJo v10.1r1
8. Describe the abundance of the relevant cell populations within post-sort fractions.  

Purity of samples is assessed from the single cell transcriptomes as described
9. Describe the gating strategy used.  

Cells were selected using forward scatter area/side scatter area (FSC-A/SSC-A) with a very large margin for possible cell sizes and complexities, to avoid any bias in cell selection. Next, fluorescent events were selected based on the parent FSC-A/SSC-A gate. GFP was excited with a 488 nm laser, and emission was detected through a 530/30 filter, while dsRed or TdTomato was exited with a 561 laser and emission detected through a 582/15 filter. A C57BL6 mouse was used as a negative control to set the threshold for fluorescent event selection. A figure describing the gating strategy is shown in Extended Data Figure 1a

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

# CORRECTIONS & AMENDMENTS

---

---

## CORRECTION

<https://doi.org/10.1038/s41586-018-0232-x>

## Author Correction: A molecular atlas of cell types and zonation in the brain vasculature

Michael Vanlandewijck, Liqun He, Maarja Andalousi Mäe, Johanna Andrae, Koji Ando, Francesca Del Gaudio, Khayrun Nahar, Thibaud Lebouvier, Bárbara Laviña, Leonor Gouveia, Ying Sun, Elisabeth Raschperger, Markus Räsänen, Yvette Zarb, Naoki Mochizuki, Annika Keller, Urban Lendahl & Christer Betsholtz

Correction to: *Nature* <https://doi.org/10.1038/nature25739>, published online 14 February 2018.

In Fig. 1b of this Article, ‘*Csf1r*’ was misspelt ‘*Csfr1*’. In addition, in Extended Data Fig. 11b, owing to an error during figure formatting, the genes listed in the first column shifted down three rows below the first gene on the list, causing a mismatch between the gene names and their characteristics. These errors have been corrected online, and the original Extended Data Fig. 11b is provided as Supplementary Information to this Amendment.

**Supplementary Information** is available for this Amendment at <https://doi.org/10.1038/s41586-018-0232-x>.