



Single Cell Transcriptomic Analysis Reveals Organ Specific Pericyte Markers and Identities

Seung-Han Baek¹, Enrico Maiorino², Hyunbum Kim¹, Kimberly Glass², Benjamin A. Raby^{1,2*} and Ke Yuan^{1*}

¹ Division of Pulmonary Medicine, Department of Pediatrics, Boston Children's Hospital and Harvard Medical School, Boston, MA, United States, ² Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States

OPEN ACCESS

Edited by:

Ngan F. Huang,
Stanford University, United States

Reviewed by:

Koji Ando,
Nippon Medical School, Japan
Walter Lee Murfee,
University of Florida, United States

*Correspondence:

Benjamin A. Raby
benjamin.raby@
childrens.harvard.edu
Ke Yuan
ke.yuan@childrens.harvard.edu

Specialty section:

This article was submitted to
General Cardiovascular Medicine,
a section of the journal
Frontiers in Cardiovascular Medicine

Received: 15 February 2022

Accepted: 21 April 2022

Published: 01 June 2022

Citation:

Baek S-H, Maiorino E, Kim H,
Glass K, Raby BA and Yuan K (2022)
Single Cell Transcriptomic Analysis
Reveals Organ Specific Pericyte
Markers and Identities.
Front. Cardiovasc. Med. 9:876591.
doi: 10.3389/fcvm.2022.876591

Pericytes are mesenchymal-derived mural cells that wrap around capillaries and directly contact endothelial cells. Present throughout the body, including the cardiovascular system, pericytes are proposed to have multipotent cell-like properties and are involved in numerous biological processes, including regulation of vascular development, maturation, permeability, and homeostasis. Despite their physiological importance, the functional heterogeneity, differentiation process, and pathological roles of pericytes are not yet clearly understood, in part due to the inability to reliably distinguish them from other mural cell populations. Our study focused on identifying pericyte-specific markers by analyzing single-cell RNA sequencing data from tissue-specific mouse pericyte populations generated by the Tabula Muris Senis. We identified the mural cell cluster in murine lung, heart, kidney, and bladder that expressed either of two known pericyte markers, *Cspg4* or *Pdgfrb*. We further defined pericytes as those cells that co-expressed both markers within this cluster. Single-cell differential expression gene analysis compared this subset with other clusters that identified potential pericyte marker candidates, including *Kcnk3* (in the lung); *Rgs4* (in the heart); *Myh11* and *Kcna5* (in the kidney); *Pcp4l1* (in the bladder); and *Higd1b* (in lung and heart). In addition, we identified novel markers of tissue-specific pericytes and signaling pathways that may be involved in maintaining their identity. Moreover, the identified markers were further validated in Human Lung Cell Atlas and human heart single-cell RNAseq databases. Intriguingly, we found that markers of heart and lung pericytes in mice were conserved in human heart and lung pericytes. In this study, we, for the first time, identified specific pericyte markers among lung, heart, kidney, and bladder and reveal differentially expressed genes and functional relationships between mural cells.

Keywords: pericytes, single cell RNA sequencing, heart, lung, tissue-specific

INTRODUCTION

Pericytes are mesenchymal-derived mural cells that cover capillary networks throughout the circulatory system. Partially surrounding capillary endothelial cells of precapillary arterioles and extending to post-capillary venules, pericytes are indispensable to the function of all organs by providing capillary structural support and facilitating capillary and endothelial cell homeostasis and

metabolism, contributing to vasoregulation of blood flow, capillary basement membrane synthesis, and initiation of angiogenesis (1–4). Pericytes are suggested to have multipotent cell-like properties. This plasticity has been implicated in the pathogenesis of many cardiovascular, pulmonary, and central nervous systems disorders and the metastasis and angiogenesis of diverse cancers (5–9).

Despite their physiological importance, a significant challenge in studying pericytes is the lack of a consensus definition to distinguish pericytes from other mural cell populations. Pericytes are recognized by their distinct morphology, consisting of an oval cell body and long extending processes that circumscribe vascular structures. Two proteins have been traditionally employed as pericyte markers: Chondroitin Sulfate Proteoglycan 4 (*Cspg4*) (also known as neural/glial antigen 2, NG2) and Platelet-Derived Growth Factor Receptor b (*Pdgfrb*), a cell surface tyrosine kinase receptor for members of the platelet-derived growth factor family (10, 11). However, these markers have generally failed to serve as unique pericyte markers. They are also expressed in other mesenchymal cell types, including smooth muscle (*Acta2* and *Tagln*) and fibroblasts (*Pdgfra*, *Lum*, and *Dcn*) (12, 13). Adding to this ambiguity, *Pdgfrb*+ or *Cspg4*+ cells expressing smooth muscle markers are inconsistently classified as either non-pericyte cells or a specific pericyte subset (e.g., *Acta2*-high expressing *Pdgfrb*+ pericytes in neonatal mouse lung) (14). Recent advances in single-cell RNA sequencing (scRNA-seq) have demonstrated substantial between-organ differences in pericytes gene expression, likely reflecting fundamental organ-specific differences in pericyte function. For example, compared with lung-derived pericytes, transcripts expressed in pericytes from the brain were enriched for genes associated with transmembrane transporter activity, consistent with the known role of pericytes in the blood-brain barrier function (15). Whether such heterogeneity in cellular phenotype and function extends to the vascular beds of other tissues remains largely unknown. The identification of unique tissue-specific pericyte markers has the potential to improve our understanding of pericyte diversity between and within the tissue and their impact on health and disease.

This study identified tissue-specific pericyte markers by analyzing murine and human scRNA-seq datasets generated in multiple tissues, including the lung, heart, kidney, and bladder. Using stringent criteria for defining pericyte clusters (*Cspg4*+/*Pdgfrb*+ dual positive), we identified several candidate markers whose expression was restricted to these “stringent pericytes,” several of which exhibited tissue-specific properties in both mice and humans. We demonstrated the ability of a subset to distinguish between lung and heart-derived human pericytes, illustrating their potential use in future studies.

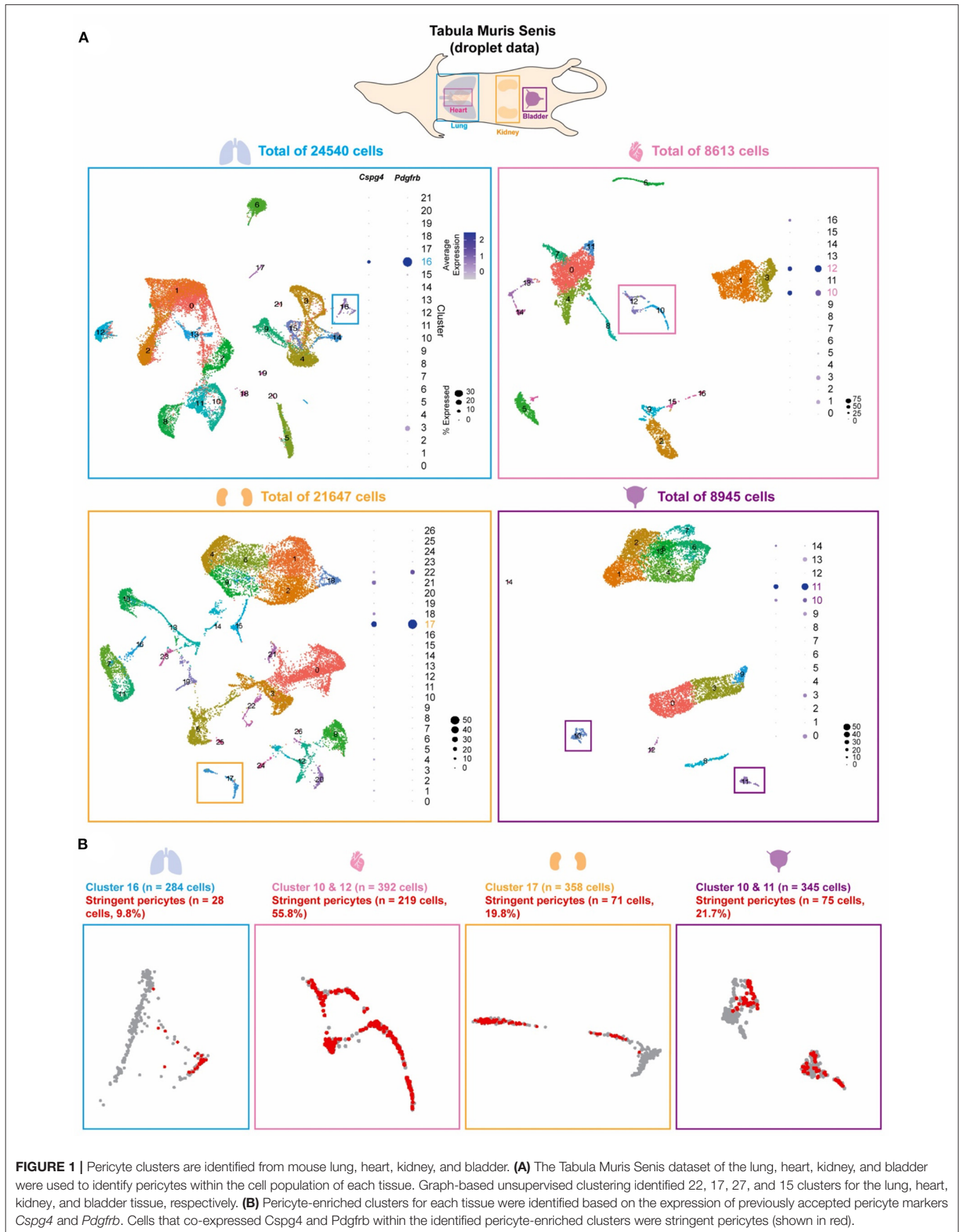
RESULTS

Identification of Pericyte Clusters From Mouse Lung, Heart, Kidney, and Bladder

The Tabula Muris Senis compendium is a single-cell survey across 23 C57BL/6JN mice tissues. From its website ([\[tabula-muris-senis.ds.czbiohub.org/all/droplet/\]\(https://tabula-muris-senis.ds.czbiohub.org/all/droplet/\)\), one of the two previously accepted pericyte markers, *Cspg4* or *Pdgfrb*, was positively expressed mostly in lung, heart, kidney, and bladder \(**Supplementary Figure 1**\). We therefore moved forward to analyze scRNA-seq data from lung \(\$n = 16\$ mice; 24,540 cells\), heart \(\$n = 11\$; 8,613\), kidney \(\$n = 16\$; 21,647\), and bladder \(\$n = 8\$; 8,945\) obtained from the Senis \(**Figure 1A**\) \(16\). Unsupervised clustering revealed substantial cellular heterogeneity within each tissue, with the number of distinct cell clusters ranging from 15 in the bladder to 27 in the kidney. To identify pericyte-enriched clusters, we first considered the individual and joint expression of two previously accepted pericyte markers, *Cspg4* and *Pdgfrb*. Consistent with prior observations \(10, 11, 14\), the preponderance of cells expressing either *Cspg4* and *Pdgfrb* localized to a limited number of clusters \(one each in lung and kidney, two in heart and bladder\). In contrast to all other clusters, which either consisted of few *Cspg4* or *Pdgfrb*-expressing cells or had low/undetectable expression of both markers, these clusters were tentatively deemed pericyte-enriched because of their large portions of cells expressing high levels of both *Cspg4* and *Pdgfrb* \(**Figure 1A**\). However, within these pericyte-enriched clusters, substantial heterogeneity in the expression of *Cspg4* and *Pdgfrb* was observed. Therefore, within these pericyte-enriched clusters, we defined a set of “stringent pericytes” by considering only those cells that expressed both *Cspg4* and *Pdgfrb* \(**Figure 1B**\). A total of 393 cells met these criteria \(28.4% of all cells\), including 28 in the lung, 219 in the heart, 71 in the kidney, and 75 in the bladder, with the proportion of stringent pericytes within each cluster varying widely by tissue \(highest in heart, 55.8%; lowest in the lung, 9.8%\). These refined subsets were considered tissue-specific stringent pericytes in all downstream analyses.](https://</p>
</div>
<div data-bbox=)

Identification of Tissue-Specific Pericyte Markers

We conducted a differential expression (DE) analysis between the stringent pericytes and all other cells based on the Wilcoxon rank-sum test to identify potential pericyte-specific markers. We identified transcripts that were enriched in the stringent pericytes in each tissue, prioritizing transcripts with high pericyte DE (adjusted p -value < 0.05) that were expressed (count > 0) in at least 80% of the stringent pericytes but <5% of all other cells. Based on these criteria, we identified 18, 4, 22, and 4 potential pericyte markers for the lung, heart, kidney, and bladder, respectively (**Figure 2, left**). No marker fulfilled these criteria in all four cell types. However, several markers were common to two or three tissues (**Figure 2, right**), including *Cox4i2* (in lung, kidney, and bladder); *Gucy1a3*, *Ndufa4i2*, and *Myl9* (in lung and kidney); *Gm13889* and *Notch3* (in heart and kidney); *Pcp4l1* (in kidney and bladder); and *Higd1b* (in lung and heart). Also, for each cluster, we determined the percentage of cells expressing the identified markers (**Figure 2, left and Supplementary Figure 6**). The pct3 represents the maximum percentage found across the clusters, excluding the pericyte enriched cluster(s). When considering markers with pct3 below 10%, we identified *Kcnk3* (in the lung); *Rgs4* (in the heart); *Myh11* and *Kcna5* (in the kidney); *Pcp4l1* (in the bladder); and *Higd1b*



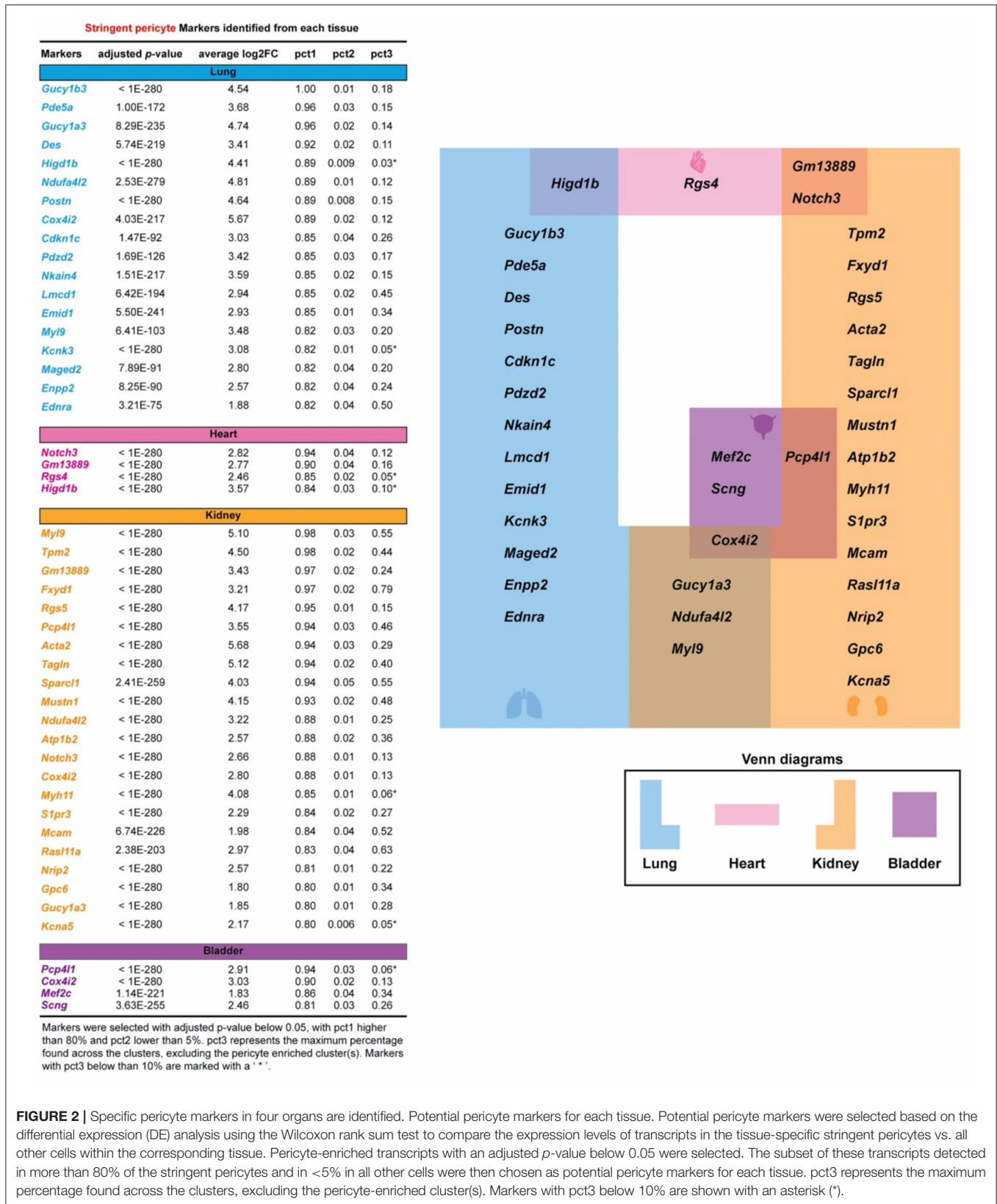


FIGURE 2 | Specific pericyte markers in four organs are identified. Potential pericyte markers for each tissue. Potential pericyte markers were selected based on the differential expression (DE) analysis using the Wilcoxon rank sum test to compare the expression levels of transcripts in the tissue-specific stringent pericytes vs. all other cells within the corresponding tissue. Pericyte-enriched transcripts with an adjusted *p*-value below 0.05 were selected. The subset of these transcripts detected in more than 80% of the stringent pericytes and in <5% in all other cells were then chosen as potential pericyte markers for each tissue. pct3 represents the maximum percentage found across the clusters, excluding the pericyte-enriched cluster(s). Markers with pct3 below 10% are shown with an asterisk (*).

(in lung and heart), which showed that they were very specific to pericyte-enriched clusters.

Pathway Enrichment Analysis of Pericytes in Each Tissue

To determine pathways associated with stringent pericytes across the lung, heart, kidney, and bladder, we conducted a pre-ranked Gene Set Enrichment Analysis (GSEA) (17) (Figure 3). First, genes were sorted based on their average $\log_2(\text{Fold Change})$ when comparing their expression in tissue-specific stringent pericytes to all cells in the corresponding tissue. GSEA was then performed on the resulting pre-ranked list using the Kyoto Encyclopedia of Genes and Genomes pathway database (18). We identified 7, 20, 36, and 8 pathways enriched (adjusted q -value < 0.05) in the stringent pericytes of the lung, heart, kidney, and bladder, respectively.

Vascular smooth muscle contraction dilated cardiomyopathy, and focal adhesion were top enriched pathways in the stringent pericytes of all four tissues, suggesting pathways involved in pericyte physiological and metabolic activity (Figure 3). This may be because pericytes share a similar lineage with smooth muscle cells (SMCs) and differentiate into SMCs under stress. Additionally, as broadly defined mural cells, their biological function can be identical to vital regulators of vessel structure support, permeability, and homeostasis (19, 20).

The other identified pathways in heart tissue were involved with neurodegenerative disorders and cardiac muscle contraction. Pericytes are one of the key regulators of the blood-brain barrier. Their involvement in the pathogenesis of neurological disorders may be associated with cardiovascular manifestations under the heart-brain axis (7, 21, 22). Pathways related to cancers are also enriched in the kidney-specific stringent pericytes. Pericytes contribute to the invasion and metastasis of cancer or tissue fibrosis through pericyte-fibroblast transition (PFT) mechanisms, which may be highlighted by the cell cycle, TGF β /Wnt/MAPK/Notch/Hedgehog pathways (19, 23). The urinary bladder is highly enriched with blood vessels, though the function of pericytes in this organ has been sorely under-investigated.

Expression of Pericyte Markers in Human Lung and Heart

To determine whether the potential pericyte markers identified from the mouse tissue were conserved in humans, we analyzed single-cell data from the Human Lung Cell Atlas and adult human heart single-cell data collected by Litvinukova et al. (24, 25) (Figure 4). The Human Lung Cell Atlas data consisted of scRNA-seq on 75,000 human lung cells and circulating blood collected from three individuals aged 75 (male), 46 (male), and 51 years (female) (24). The human heart data consisted of scRNA-seq on 486,134 human cells from all heart compartments and collected 14 individuals of both sexes with ages ranging from 40 to 75 years (25). To unify our analysis of these data with the Tabula Muris Senis data analysis, we selected only droplet-based scRNA-seq data processed using the 10 \times Genomics platform (10 \times 3'v2), which left us with three individuals-worth of lung

tissue, and seven individuals worth of heart tissue. The data across individuals were then integrated and processed using the Seurat pipeline, resulting in a final dataset consisting of 65,662 cells from the human lung tissue and 238,154 cells from the human heart tissue (26). Pericytes were found annotated based on *CSPG4*, *PDGFRB*, and *TRPC6* within the Human Lung Cell Atlas, while within the human heart data, pericytes were found annotated based on *RGS5*, *ABCC9*, and *KCNJ8*. We mapped the annotated pericytes onto the UMAP plot. We compared it with the expression density of the human orthologs of the potential pericyte markers identified in mouse lung and heart.

CSPG4 or *PDGFRB* was expressed in the human lung and heart (Figures 4A,B; Supplementary Figures 7, 8). We identified high expression of *GUCY1B1*, *HIGD1B*, *NDUFAL2*, *COX4I2*, and *KCNK3* in the annotated pericyte cluster from the Human Lung Cell Atlas. Except for *GUCY1A1*, *PDE5A*, *GUCY1B1*, and *DES*, all other markers were non-specific and expressed in other regions. The expression of *HIGD1B*, *NDUFAL2*, *COX4I2*, and *KCNK3* was more specific and almost fully overlapped with *CSPG4* or *PDGFRB* and mostly did not overlap with the expression of vascular smooth muscle cell (VSMC) markers *ACTA2*, *MYH11*, and *TAGLN* (Figure 4A and Supplementary Figure 7). For the human orthologs of the potential pericyte markers identified from the mouse heart tissue, we identified the expression of *NOTCH3* and *HIGD1B* to be relatively high and specific to the annotated pericytes compared with the other markers and relatively did not overlap with the VSMC markers.

Pericyte Localization in Human Heart Tissue by Spatial Transcriptomic Analysis

Most scRNA-seq approaches rely on dissociating cells from tissues, thereby losing the crucial spatial context of the cells' locations within the tissue. Techniques such as the image-based single-cell transcriptomics method, multiplexed RNA FISH, hybridized tissue sections to spatially barcoded microarrays, sequencing on intact tissues would simultaneously profile the expression of hundreds or thousands of genes within single cells whose spatial location is preserved (27). We, therefore, investigated the pericyte spatial localization within the human heart tissue. We compared it with regions of capillaries and arteries using the spatial transcriptomic data of human heart tissue from the Spatial Gene Expression Dataset by 10 \times Genomics Space Ranger (10 \times Genomics Space Ranger 1.1.0) (Figure 5).

Each spot in the dataset represented tissue regions where the transcriptomic reads were obtained. The dataset was projected onto a 2-dimensional map using UMAP, and unsupervised clustering of the spots was performed in which we identified 14 clusters (28, 29). In a sequential tissue section stained by hematoxylin and eosin (H&E), blood cells in dark pink color were visible within the hollow cavities where the same areas as vessels were mainly estimated to reside in. Integrating the transcriptomic and spatial data, the tissue section can be annotated as the same 14 clusters applied to the UMAP plot using the same color-coding (Figure 5A).

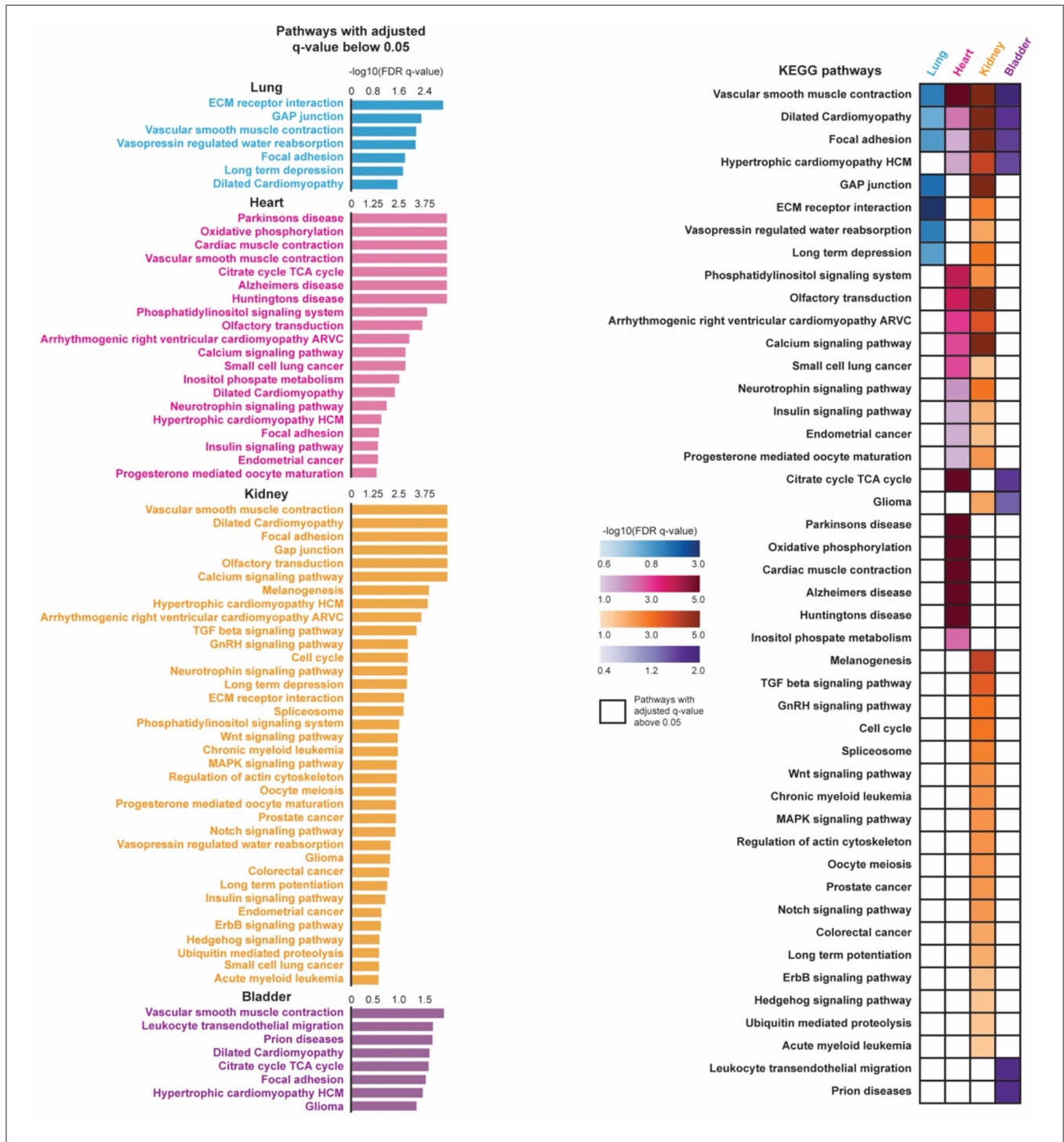
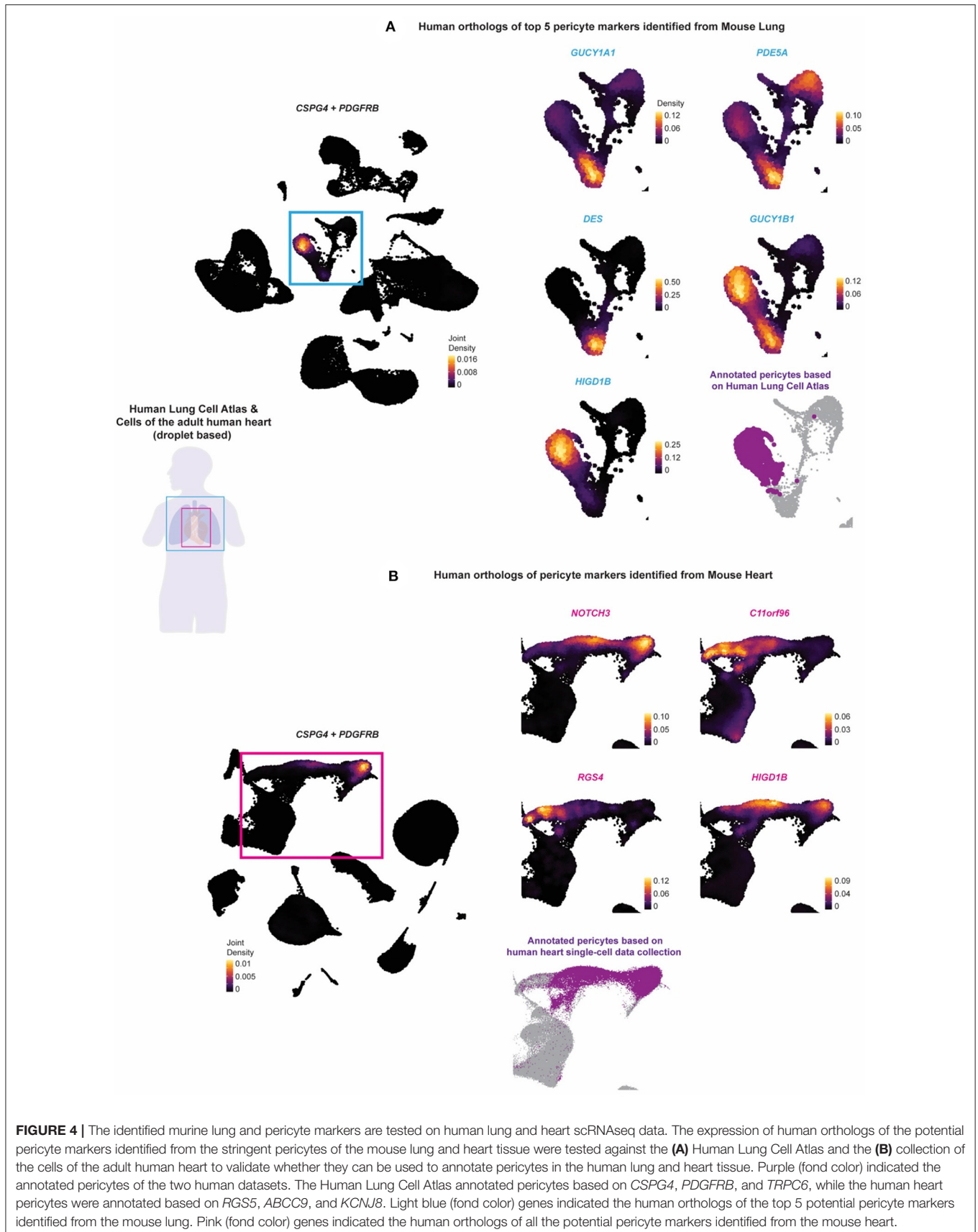
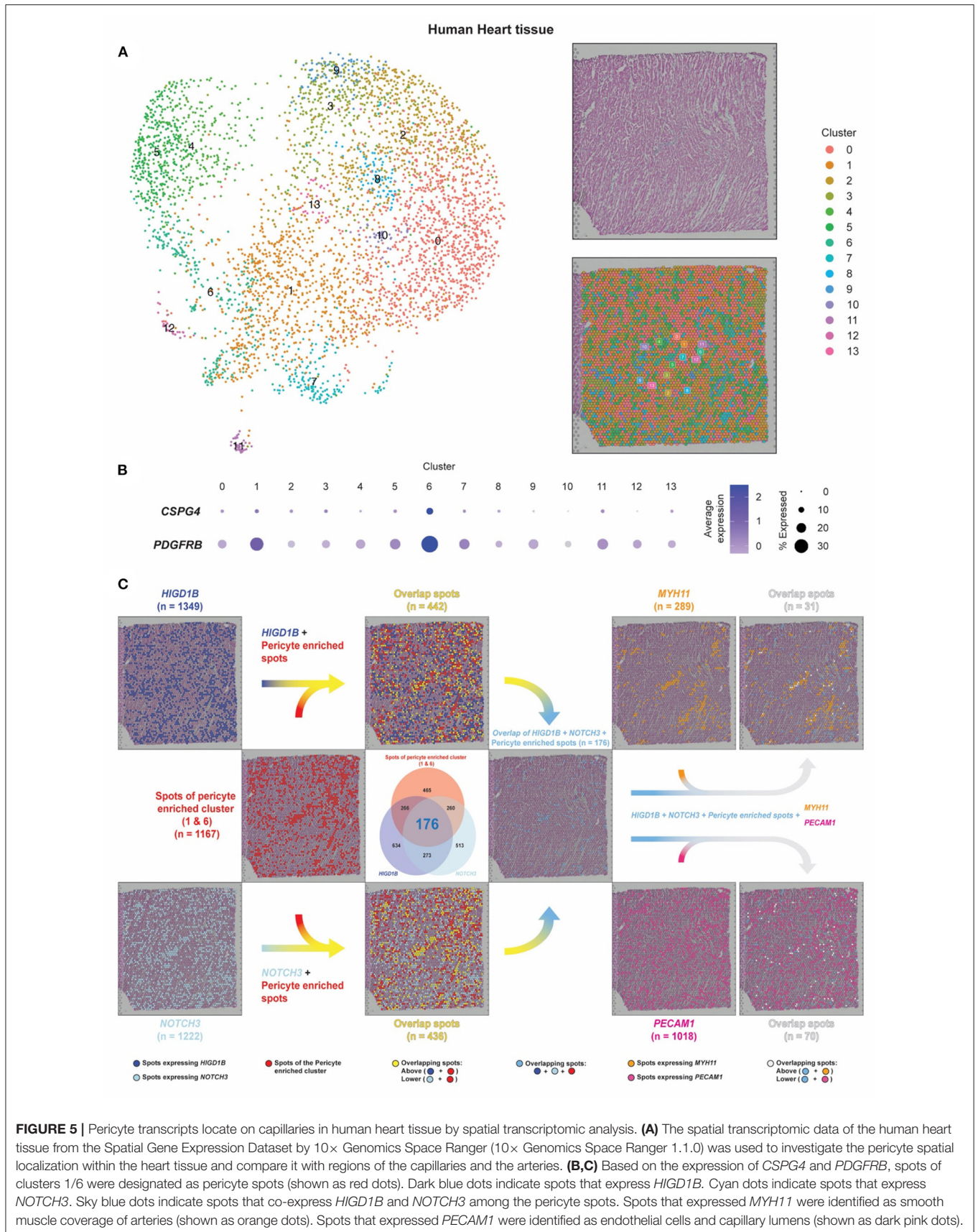


FIGURE 3 | Pathway enrichment analysis of pericytes in four organs are shown. Pre-ranked gene set enrichment analysis (GSEA) with the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database was conducted on the DE genes between the tissue-specific stringent pericytes and all other cells of its corresponding tissue. The DE genes were sorted based on their average log2(Fold Change) in a descending order to generate the pre-ranked gene list. Enriched pathways were identified for each tissue-specific stringent pericytes based on their adjusted q-value ($q < 0.05$). The identified enriched pathways of the tissue-specific stringent pericytes were compared across four tissue types.

Co-expression of *CSPG4* and *PDGFRB* was mainly found in Cluster 1/6, from which 1,167 spots were identified and designated as pericyte spots (Figures 5B,C, red dots). A total

of 289 spots that expressed *MYH11* were identified as smooth muscle coverage of arteries (orange dots), while 1,018 spots that expressed *PECAM1* were endothelial cells and capillary





lumens (dark pink dots). While capillary spots were randomly distributed, we observed the pattern and shape of arteries as continuous and vessel-like.

We further identified 1,349 spots expressing *HIGD1B* (dark blue dots) and 1,222 spots expressing *NOTCH3* (cyan dots), in which 449 spots co-expressed both *HIGD1B* and *NOTCH3* (Venn diagram, 176+273) (Figure 5C). After colocalization with Cluster 1/6, 37.87% overlapped spots were co-expressing *HIGD1B* (yellow dots on the upper-side flow), while 37.36% overlapped spots were co-expressing *NOTCH3* (yellow dots on the lower-sided flow). A total of 176 dots (6.8%) co-expressed *HIGD1B* and *NOTCH3* among the pericyte spots (sky blue dots on the middle right-sided flow), in concordance with their expected location independent from *MYH11* and *PECAM-1*. Although these spots were not cells (but rather aggregates of transcripts) and would have lower resolution than immunofluorescence staining, our spatial data revealed that pericytes were exclusively located on distal vessels, not on large vessels. This method also opened a new avenue to evaluate quantitative and spatially resolved maps of gene expression at a transcriptomic level no longer dependent on the accessibility of their corresponding antibodies.

DISCUSSION

Despite their functional importance and implications being one of the major components in capillary networks and indispensable in capillary homeostasis and function throughout the body in all organs, pericytes remain poorly characterized. Pericytes are also identified to play vital roles in vessel stabilization (30, 31), blood brain barrier formation (32), immune cell guidance (33), blood flow regulation (34), tissue repair, (9) and pathological scarring/fibrosis (35, 36). A significant challenge to studying pericytes is the lack of a consensus definition or cellular marker that can distinguish them from other mural cell populations. Additionally, genes and their enriched pathways that govern pericyte unique function in an organ-specific manner have not been demonstrated, primarily due to insufficient methodologies that can investigate a single pericyte cell type from various tissues in parallel.

With the advent of scRNA-seq technologies, our study investigated the transcriptomic landscape of pericytes that may be common to or differ across four tissues. Differentially expressed genes from *Cspg4/Pdgfrb* expressing compared to non-expressing cells revealed 18 lung-specific, 4 heart-specific, 22 kidney-specific, and 4 bladder-specific pericyte markers (Figure 2). The Venn diagram also suggested the commonly expressed pericyte transcriptome across four organs. *Gucy1a3*, the main receptor for nitric oxide, *Ndufa4l2*, which mediates activation of oxidative phosphorylation and produces ROS in the mitochondria; and *Myl9*, which regulates muscle contraction by modulating the ATPase activity of myosin heads, are all identified as potential pericyte markers in both the lung and kidney (37–39). Recently, *NDUFA4L2*, the human ortholog of *Ndufa4l2*, is involved in the vascular remodeling of smooth muscles in hypoxic pulmonary arterial hypertension (40). Although *Myl9* is well-expressed in muscles, it is also expressed in pericytes of the murine cortex (41). *Gm13889* and *Notch3* were identified as

potential pericyte markers in the heart and kidney. Although not much is known about *Gm13889*, *Notch3* is involved in neural development and brain pericyte proliferation (42). *Higd1b*, which belongs to the hypoxia-inducible gene 1 (HIG1) domain family and is involved in cell survival by maintaining mitochondrial integrity under hypoxia conditions, was identified as a potential pericyte marker in both the lung and heart (43). *Cox4i2*, which catalyzes the electron transfer from reduced cytochrome c to oxygen, is identified as a potential pericyte marker in the lung, kidney, and bladder (44). Human *COX4I2* is differentially expressed in pericytes and annotated as a pericyte marker in the Human Lung Cell Atlas (24). Our study also identified tissue-specific pericyte markers, which may potentially serve as promising targets for organ-specific delivery of therapeutics (e.g., AAV) that process efficient target specificity and reduce the risk of side effects. Based on the differentially expressed genes, we conducted a GSEA analysis to identify activated pathways in pericytes. Intriguingly, vascular smooth muscle contraction, dilated cardiomyopathy, and the focal adhesion pathways were suggested as commonly activated among all four tissues (Figure 3).

It was vital to ensure the results from mouse datasets could be translatable to human pericytes. Thus, the identified pericyte markers in mouse lung and heart were compared with the human lung and heart scRNA seq datasets. Overall, we found considerable overlapped gene expression between mouse and human tissues. Human orthologs of *Ndufa12*, *Cox4i2*, *Kcnk3*, and *Higd1b* were highly relevant and specific to the annotated lung pericyte clusters, while the human ortholog of *Notch3* and *Higd1b* was found in the heart pericyte cluster. Lastly, applying gene expression into context and delineating the spatial arrangement of cell types, we visualized the pericyte/EC/SMC distribution and further presented *Notch3* and *Higd1b* as pericyte markers and their spatial locations on a human heart tissue (Figure 5).

A significant challenge for studying pericyte pathobiology was lacking a consensus definition that helps distinguish pericytes from other mural cell populations, including smooth muscle cells, fibroblasts, myofibroblasts, and others. However, single-cell analysis of brain vasculature, particularly the distinguishment for brain pericytes and other mural cells, has been well-investigated (45–47). Murine brain pericytes share markers, including *Vtn*, *Higd1b*, *S1pr3*, *Mcam*, *Ifitm1*, *Baiap3*, and *Ehd3* with lung pericytes. Intriguingly, brain pericytes have substantial organotypic differences on markers such as *Anpep*, which is not expressed by lung pericytes (45). One of the top commonly altered pathways in all organ pericytes was associated with vascular smooth muscle contraction in our pathway analysis. This result was consistent with our most recent work that suggested NG2-expressing pericytes became smooth muscle like under hypoxia-induced pulmonary hypertension (48). Additionally, heart pericytes may serve as a progenitor for smooth muscle cells during embryonic heart development. Another lineage analysis suggests that kidney pericytes may differentiate into most smooth muscle actin expressing myofibroblasts during fibrosis (49). In the Human Lung Cell Atlas project, lung pericytes are identified using *COX4I2*, *TBX5*, and *KCNK3* and its potential implication in pulmonary hypertension is proposed (24). These abovementioned studies

laid some groundwork to elucidate the identification and mechanisms of pericytes; however, their refined biological role and potential for drug targeting still need more investigation.

Our work had several limitations. Only four datasets were included in our analysis due to minimum *Cspg4* and *Pdgfrb* expressions across all other organs. Therefore, larger sample sizes were needed to corroborate these results. These results should be further validated using the single-cell data generated by the FACS method. Our study did not include brain pericytes due to a lack of Droplet data on brain tissues. We also did not consider age as one of the variable factors. Pericytes may have different features that may be unique across developmental processes. Therefore, to identify diverse pericytes across developmental and differentiation processes, we needed to specify the pericyte population in each age group. Additionally, we did not look for variance between men and women, which would provide us with more insights into potential sex differences in pericyte biology and cardiovascular disease risk. There may be different pericyte subpopulations among the same organ between mice and humans regarding its heterogeneity. This subtype variation will be reflected at the transcriptomic level and needed more extensive analysis of co-expression of both pericyte markers and other mural cell markers. For instance, after pericytes were identified using *Pdgfrb* and *Cspg4*, sub-clustering these double-positive cells would provide a clearer understanding of subpopulation signature genes.

While the identified genes tested against the human lung and heart single-cell datasets have a high potential as pericyte unique markers, additional protein analysis and functional validation will be required to confirm each pericyte phenotype's putative role using the antibody RNA Scope staining. When going through the methods of several scRNAseq experiments used in this study, we noted that lung tissues underwent enzymatic digestion to achieve single-cell suspension before subjecting to scRNAseq analysis in the traditional dissociation method. The percentage of cells annotated as pericyte clusters seemed to be relatively small compared with the total cell population in most publicly available datasets. It was also possible that during dissociation of the tissues or FACS-based method, pericytes were not completely separated from endothelial cells, thus forming doublets and being excluded from the analysis. An optimized and enriched pericyte isolation method to preserve viability and subsequently pericyte scRNA-seq atlas will be valuable in addressing their unknown nature as indicated above. Lastly, pericyte spatial orientation alongside the vascular tree structure requires validation (for instance, by spatial scRNA-seq methods or Nanostring GeoMax) and their resulting data will help us to understand pericyte distribution or morphological change in the pathogenesis of diseases, especially clinically meaningful if we can identify them on patient biopsy samples during disease progression.

Our study provides important new insights into organ-specific pericyte localization using single-cell RNAseq from mouse and human databases and evaluating pericyte spatial transcriptomic orientation on human heart tissues. These datasets are effective in identifying tissue-specific markers and in understanding biological processes and active pathways in other diverse cell types (16, 24, 25). The single cell and spatial data revealed that pericyte transcripts were exclusively located on distal vessels,

which was consistent with pericyte immunostaining results (1, 48). This method also revealed the relationship between single cell and their spatial transcriptomics data and thus could spatially map any cell types of interest across the tissue.

In conclusion, our work provides innovative and insightful knowledge on facilitating the identification of pericyte phenotypes across different tissues. Future studies with a wider range of tissues considering developmental processes and focusing on the trajectory of pericyte differentiation will provide us with unique markers for organ-specific pericytes, which would aid us in understanding and characterizing the role of pericyte plays in diverse biological processes and pathogenesis of related diseases.

METHODS

Data Availability

In this manuscript, we analyzed microfluidic droplet based scRNA-seq data processed using the 10× Genomics platform (10 × 3'v2) of the (1) Tabula Muris Senis, (2) Human Lung Cell Atlas, and the (3) Cells of the human heart that were all obtained from the Cellxgene collections (<https://cellxgene.cziscience.com>) (16, 24, 25). The specific pre-processed Seurat objects for these datasets we obtained are as follows:

1. The lung, heart, kidney, and bladder tissue of the Tabula Muris Senis dataset: (<https://cellxgene.cziscience.com/collections/0b9d8a04-bb9d-44da-aa27-705bb65b54eb>).
2. The Human Lung Cell Atlas dataset: (<https://cellxgene.cziscience.com/collections/5d445965-6f1a-4b68-ba3a-b8f765155d3a>).
3. The dataset of the cells of the adult human heart: (<https://cellxgene.cziscience.com/collections/b52eb423-5d0d-4645-b217-e1c6d38b2e72>).

In addition, spatial transcriptomics for the human heart was obtained from the Spatial Gene Expression Dataset by 10x Genomics Space Ranger (10x Genomics Space Ranger 1.1.0) at: (<https://www.10xgenomics.com/resources/datasets/human-heart-1-standard-1-1-0>).

Pre-processing Single-Cell RNA Sequencing Data

The obtained scRNA-seq data were analyzed based on the pipeline of the Seurat R package (26). The dataset of each subject and tissue was identified and processed through the Seurat pipeline in which the feature counts were log-normalized and subsequently scaled. The top 2,000 variable features were then selected based on the Variance Stabilizing Transformation method for integration (50). Within each tissue, the datasets of the according subjects were then integrated using the IntegrateData function of Seurat. The annotated cell types within the Seurat objects of the human lung dataset and the human heart datasets were used to identify pericytes.

Dimensionality Reduction and Clustering

For each integrated scRNA-seq dataset, principal component (PC) analysis was conducted on the highly variable genes to reduce the dimensionality. Based on the percentage of variance

explained by the top PCs, we selected the top 45 PCs for the mouse kidney, the top 30 PCs for the mouse bladder, and the top 35 PCs for the remaining (lung and heart of both the mouse and human) datasets to conduct the downstream analysis. Based on the selected PCs, for each dataset, cells were projected into a 2-dimensional space using the Uniform Manifold Approximation and Projection (UMAP) algorithm (28). We conducted an unsupervised clustering of single cells using the FindCluster function in Seurat, setting the resolution measure to 0.5, which constructs a shared nearest neighbor graph by identifying the k -nearest neighbors for each cell in the PC space and weights the edges between the cells based on the number of the k -nearest neighbors or their Jaccard similarity (29). The densely connected cells are then considered as clusters in which the modularity is optimized using the Louvain community detection method (26, 29).

Differential Gene Expression Analysis

Differential expression analysis was conducted by comparing the expression of stringent pericytes (cells that co-expressed *Cspg4* and *Pdgfrb* within the identified pericyte-enriched cluster) with all other cells according to the tissue for the Tabula Muris Senis dataset using a Wilcoxon rank sum test (26). Genes with an adjusted p -value below 0.05 that are expressed in more than 80% of the stringent pericytes and in <5% of all other cells were selected as potential pericyte markers according to the tissue.

Gene Set Enrichment Analysis

The Gene Set Enrichment Analysis (GSEA) (17) was used to interpret the gene expression data using the software available at <https://www.gsea-msigdb.org/gsea/index.jsp>. A pre-ranked GSEA determines whether genes within an *a priori* defined specific gene tend to occur toward the top or the bottom of a pre-ranked gene list, which then identifies whether the corresponding gene set shows statistically significant differences between two biological states. We generated pre-ranked gene lists by sorting the DE genes based on their log₂ fold change from the DE analysis in a descending order. GSEA was conducted using the resulting pre-ranked gene lists and compared with the gene sets defined by the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway gene sets. A permutation test estimates the statistical significance of the enrichment score. GSEA generates a version of the dataset with phenotype labels randomly scrambled, produces the corresponding ranked list, and then recomputes the enrichment score of the gene set for this permuted dataset. Pathways with an adjusted q -value below 0.05 in this analysis were considered statistically significant.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the the Cellxgene collections (<https://cellxgene.cziscience.com>). The lung, heart, kidney, and bladder tissue of the Tabula Muris Senis dataset: (<https://cellxgene.cziscience.com/collections/0b9d8a04-bb9d-44da-aa27-705bb65b54eb>). The Human Lung Cell Atlas dataset: ([https://cellxgene.cziscience.com/collections/5d445965-](https://cellxgene.cziscience.com/collections/5d445965-6f1a-4b68-ba3a-b8f765155d3a)

[6f1a-4b68-ba3a-b8f765155d3a](https://cellxgene.cziscience.com/collections/6f1a-4b68-ba3a-b8f765155d3a)). The dataset of the cells of the adult human heart: (<https://cellxgene.cziscience.com/collections/b52eb423-5d0d-4645-b217-e1c6d38b2e72>).

ETHICS STATEMENT

The study was approved by the Institutional Biosafety Committee of Boston Children's Hospital.

AUTHOR CONTRIBUTIONS

S-HB and KY planned and performed experiments. S-HB was responsible for data analysis. S-HB, EM, HK, KG, BR, and KY provided intellectual input. BR and KY conceived the study. S-HB, BR, and KY wrote the manuscript. The order of the authors was decided to depend on the combined contribution of intellectual input and experiments performed. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by NIH/NHLBI 5R01HL150106-02, ATS/PHA Aldrighetti Research Award for Young Investigators and Pulmonary Hypertension Accelerated Bayer Awards (to KY).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcvm.2022.876591/full#supplementary-material>

Supplementary Figure 1 | UMAP plots depict expressions of *Cspg4* and *Pdgfrb* across different organs of the Tabula Muris Senis dataset. UMAP showing expression levels of *Cspg4* (top) and *Pdgfrb* (bottom) across organs of the Tabula Muris Senis dataset.

Supplementary Figure 2 | UMAP plots depict expressions of *Cspg4*, *Pdgfrb*, and vascular smooth muscle cell (VSMC) markers in the mouse lung. UMAP plots showed expression levels of *Cspg4*, *Pdgfrb*, and VSMC markers (*Acta2*, *Myh11*, *Tagln*) from the Tabula Muris Senis lung droplet dataset.

Supplementary Figure 3 | UMAP plots depict expression of *Cspg4*, *Pdgfrb*, and vascular smooth muscle cell (VSMC) markers in the mouse heart. UMAP plots showed expression levels of *Cspg4*, *Pdgfrb*, and VSMC markers (*Acta2*, *Myh11*, *Tagln*) from the Tabula Muris Senis heart droplet dataset.

Supplementary Figure 4 | UMAP plots depict expression of *Cspg4*, *Pdgfrb*, and vascular smooth muscle cell (VSMC) markers in mouse kidney. UMAP plots showed expression levels of *Cspg4*, *Pdgfrb*, and VSMC markers (*Acta2*, *Myh11*, *Tagln*) from the Tabula Muris Senis kidney droplet dataset.

Supplementary Figure 5 | UMAP plots depict expression of *Cspg4*, *Pdgfrb*, and vascular smooth muscle cell (VSMC) markers in the mouse bladder. UMAP plots showed expression levels of *Cspg4*, *Pdgfrb*, and VSMC markers (*Acta2*, *Myh11*, *Tagln*) from the Tabula Muris Senis bladder droplet dataset.

Supplementary Figure 6 | Heatmap plot depicts expression levels of potential pericyte markers. Heatmap plot showed expression levels of identified tissue-specific pericyte markers from four tissues.

Supplementary Figure 7 | Expression levels of murine pericyte markers are identified in the human lung single cell RNAseq dataset. Density plot of the expression of human orthologs of potential pericyte markers identified from the mouse lung were further identified on the collection of the single cells of the Human Lung Cell Atlas.

Supplementary Figure 8 | Expression levels of murine pericyte markers are identified in the human heart single-cell RNAseq dataset. Density plot of the expression of human orthologs of potential pericyte markers identified from the mouse heart were further identified on the collection of the single cells of the adult human heart.

REFERENCES

- Yuan K, Orcholski ME, Panaroni C, Shuffle EM, Huang NF, Jiang X, et al. Activation of the Wnt/planar cell polarity pathway is required for pericyte recruitment during pulmonary angiogenesis. *Am J Pathol.* (2015) 185:69–84. doi: 10.1016/j.ajpath.2014.09.013
- Yuan K, Shamskhou EA, Orcholski ME, Nathan A, Reddy S, Honda H, et al. Loss of endothelium-derived Wnt5a is associated with reduced pericyte recruitment and small vessel loss in pulmonary arterial hypertension. *Circulation.* (2019) 139:1710–24. doi: 10.1161/CIRCULATIONAHA.118.037642
- Yuan K, Shao NY, Hennigs JK, Discipulo M, Orcholski ME, Shamskhou E, et al. Increased pyruvate dehydrogenase kinase 4 expression in lung pericytes is associated with reduced endothelial-pericyte interactions and small vessel loss in pulmonary arterial hypertension. *Am J Pathol.* (2016) 186:2500–14. doi: 10.1016/j.ajpath.2016.05.016
- Armulik A, Genove G, Betsholtz C. Pericytes: developmental, physiological, and pathological perspectives, problems, and promises. *Dev Cell.* (2011) 21:193–215. doi: 10.1016/j.devcel.2011.07.001
- Yamazaki T, Mukoyama YS. Tissue specific origin, development, and pathological perspectives of pericytes. *Front Cardiovasc Med.* (2018) 5:78. doi: 10.3389/fcvm.2018.00078
- Hung C, Linn G, Chow YH, Kobayashi A, Mittelsteadt K, Altemeier WA, et al. Role of lung pericytes and resident fibroblasts in the pathogenesis of pulmonary fibrosis. *Am J Respir Crit Care Med.* (2013) 188:820–30. doi: 10.1164/rccm.201212-2297OC
- Lendahl U, Nilsson P, Betsholtz C. Emerging links between cerebrovascular and neurodegenerative diseases—a special role for pericytes. *EMBO Rep.* (2019) 20:e48070. doi: 10.15252/embr.201948070
- Sun R, Kong X, Qiu X, Huang C, Wong PP. The emerging roles of pericytes in modulating tumor microenvironment. *Front Cell Dev Biol.* (2021) 9:676342. doi: 10.3389/fcell.2021.676342
- Crisan M, Yap S, Casteilla L, Chen CW, Corselli M, Park TS, et al. A perivascular origin for mesenchymal stem cells in multiple human organs. *Cell Stem Cell.* (2008) 3:301–13. doi: 10.1016/j.stem.2008.07.003
- Schiffer D, Mellai M, Boldorini R, Bisogno I, Grifoni S, Corona C, et al. The significance of chondroitin sulfate proteoglycan 4 (CSPG4) in human gliomas. *Int J Mol Sci.* (2018) 19:2724. doi: 10.3390/ijms19092724
- Winkler EA, Bell RD, Zlokovic BV. Pericyte-specific expression of PDGF beta receptor in mouse models with normal and deficient PDGF beta receptor signaling. *Mol Neurodegener.* (2010) 5:32. doi: 10.1186/1750-1326-5-32
- Hughes S, Chan-Ling T. Characterization of smooth muscle cell and pericyte differentiation in the rat retina *in vivo*. *Invest Ophthalmol Vis Sci.* (2004) 45:2795–806. doi: 10.1167/iovs.03-1312
- Wang CH, Lee YS, Lin SJ, Mei HF, Lin SY, Liu MH, et al. Surface markers of heterogeneous peripheral blood-derived smooth muscle progenitor cells. *Arterioscler Thromb Vasc Biol.* (2012) 32:1875–83. doi: 10.1161/ATVBAHA.112.245852
- Guo M, Du Y, Gokey JJ, Ray S, Bell SM, Adam M, et al. Single cell RNA analysis identifies cellular heterogeneity and adaptive responses of the lung at birth. *Nat Commun.* (2019) 10:37. doi: 10.1038/s41467-018-07770-1
- Brown LS, Foster CG, Courtney JM, King NE, Howells DW, Sutherland BA. Pericytes and neurovascular function in the healthy and diseased brain. *Front Cell Neurosci.* (2019) 13:282. doi: 10.3389/fncel.2019.00282
- Tabula Muris C. A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. *Nature.* (2020) 583:590–5. doi: 10.1038/s41586-020-2496-1
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* (2000) 28:27–30. doi: 10.1093/nar/28.1.27
- Bergers G, Song S. The role of pericytes in blood-vessel formation and maintenance. *Neuro Oncol.* (2005) 7:452–64. doi: 10.1215/S1152851705000232
- Kumar A, D'Souza SS, Moskvina OV, Toh H, Wang B, Zhang J, et al. Specification and diversification of pericytes and smooth muscle cells from mesenchymoangioblasts. *Cell Rep.* (2017) 19:1902–16. doi: 10.1016/j.celrep.2017.05.019
- Sagare AP, Bell RD, Zhao Z, Ma Q, Winkler EA, Ramanathan A, et al. Pericyte loss influences Alzheimer-like neurodegeneration in mice. *Nat Commun.* (2013) 4:2932. doi: 10.1038/ncomms3932
- Laredo F, Plebanski J, Tedeschi A. Pericytes: problems and promises for CNS repair. *Front Cell Neurosci.* (2019) 13:546. doi: 10.3389/fncel.2019.00546
- Hosaka K, Yang Y, Seki T, Fischer C, Dubey O, Fredlund E, et al. Pericyte-fibroblast transition promotes tumor growth and metastasis. *Proc Natl Acad Sci USA.* (2016) 113:E5618–27. doi: 10.1073/pnas.1608384113
- Travaglini KJ, Nabhan AN, Penland L, Sinha R, Gillich A, Sit RV, et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature.* (2020) 587:619–25. doi: 10.1038/s41586-020-2922-4
- Litvinukova M, Talavera-Lopez C, Maatz H, Reichart D, Worth CL, Lindberg EL, et al. Cells of the adult human heart. *Nature.* (2020) 588:466–72. doi: 10.1038/s41586-020-2797-4
- Hao Y, Hao S, Andersen-Nissen E, Mauck WM III, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell.* (2021) 184:3573–87.e29. doi: 10.1016/j.cell.2021.04.048
- Lohoff T, Ghazanfar S, Missarova A, Kouloua N, Pierson N, Griffiths JA, et al. Integration of spatial and single-cell transcriptomic data elucidates mouse organogenesis. *Nat Biotechnol.* (2021) 40:74–85. doi: 10.1038/s41587-021-01006-2
- Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol.* (2018) 37. doi: 10.1038/nbt.4314
- Levine JH, Simonds EF, Bendall SC, Davis KL, Amir el AD, Tadmor MD, et al. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell.* (2015) 162:184–97. doi: 10.1016/j.cell.2015.05.047
- Lindahl P, Johansson BR, Leveen P, Betsholtz C. Pericyte loss and microaneurysm formation in PDGF-B-deficient mice. *Science.* (1997) 277:242–5. doi: 10.1126/science.277.5323.242
- Yuan K, Agarwal S, Chakraborty A, Condon DF, Patel H, Zhang S, et al. Lung pericytes in pulmonary vascular physiology and pathophysiology. *Compr Physiol.* (2021) 11:2227–47. doi: 10.1002/cphy.c200027
- Armulik A, Genove G, Mae M, Nisancioglu MH, Wallgard E, Niaudet C, et al. Pericytes regulate the blood-brain barrier. *Nature.* (2010) 468:557–61. doi: 10.1038/nature09522
- Proebstl D, Voisin MB, Woodfin A, Whiteford J, D'Acquisto F, Jones GE, et al. Pericytes support neutrophil subendothelial cell crawling and breaching of venular walls *in vivo*. *J Exp Med.* (2012) 209:1219–34. doi: 10.1084/jem.20111622
- Hartmann DA, Coelho-Santos V, Shih AY. Pericyte control of blood flow across microvascular zones in the central nervous system. *Annu Rev Physiol.* (2022) 84:331–54. doi: 10.1146/annurev-physiol-061121-040127
- Chang FC, Chou YH, Chen YT, Lin SL. Novel insights into pericyte-myofibroblast transition and therapeutic targets in renal fibrosis. *J Formos Med Assoc.* (2012) 111:589–98. doi: 10.1016/j.jfma.2012.09.008
- Sava P, Ramanathan A, Dobronyi A, Peng X, Sun H, Ledesma-Mendoza A, et al. Human pericytes adopt myofibroblast properties in the microenvironment of the IPF lung. *JCI Insight.* (2017) 2:e96352. doi: 10.1172/jci.insight.96352

37. Kessler T, Wolf B, Eriksson N, Kofink D, Mahmoodi BK, Rai H, et al. Association of the coronary artery disease risk gene GUCY1A3 with ischaemic events after coronary intervention. *Cardiovasc Res.* (2019) 115:1512–8. doi: 10.1093/cvr/cvz015
38. Kumar CC, Mohan SR, Zavodny PJ, Narula SK, Leibowitz PJ. Characterization and differential expression of human vascular smooth muscle myosin light chain 2 isoform in nonmuscle cells. *Biochemistry.* (1989) 28:4027–35. doi: 10.1021/bi00435a059
39. Vinogradov AD, Grivennikova VG. Oxidation of NADH and ROS production by respiratory complex I. *Biochim Biophys Acta.* (2016) 1857:863–71. doi: 10.1016/j.bbabi.2015.11.004
40. Liu Y, Nie X, Zhu J, Wang T, Li Y, Wang Q, et al. NDUFA4L2 in smooth muscle promotes vascular remodeling in hypoxic pulmonary arterial hypertension. *J Cell Mol Med.* (2021) 25:1221–37. doi: 10.1111/jcmm.16193
41. Zeisel A, Munoz-Manchado AB, Codeluppi S, Lonnerberg P, La Manno G, Jureus A, et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science.* (2015) 347:1138–42. doi: 10.1126/science.aaa1934
42. Wang Y, Pan L, Moens CB, Appel B. Notch3 establishes brain vascular integrity by regulating pericyte number. *Development.* (2014) 141:307–17. doi: 10.1242/dev.096107
43. Pang Y, Zhu Z, Wen Z, Lu J, Lin H, Tang M, et al. HIGD1B inhibits hypoxia-induced mitochondrial fragmentation by regulating OPA1 cleavage in cardiomyocytes. *Mol Med Rep.* (2021) 24:549. doi: 10.3892/mmr.2021.12188
44. Huttemann M, Lee I, Liu J, Grossman LI. Transcription of mammalian cytochrome c oxidase subunit IV-2 is controlled by a novel conserved oxygen responsive element. *FEBS J.* (2007) 274:5737–48. doi: 10.1111/j.1742-4658.2007.06093.x
45. Vanlandewijck M, He L, Mae MA, Andrae J, Ando K, Del Gaudio F, et al. A molecular atlas of cell types and zonation in the brain vasculature. *Nature.* (2018) 554:475–80. doi: 10.1038/nature25739
46. Gastfriend BD, Foreman KL, Katt ME, Palecek SP, Shusta EV. Integrative analysis of the human brain mural cell transcriptome. *J Cereb Blood Flow Metab.* (2021) 41:3052–68. doi: 10.1177/0271678X211013700
47. He L, Vanlandewijck M, Mae MA, Andrae J, Ando K, Del Gaudio F, et al. Single-cell RNA sequencing of mouse brain and lung vascular and vessel-associated cell types. *Sci Data.* (2018) 5:180160. doi: 10.1038/sdata.2018.160
48. Yuan K, Liu Y, Zhang Y, Nathan A, Tian W, Yu J, et al. Mural cell SDF1 signaling is associated with the pathogenesis of pulmonary arterial hypertension. *Am J Respir Cell Mol Biol.* (2020) 62:747–59. doi: 10.1165/rcmb.2019-0401OC
49. Humphreys BD, Lin SL, Kobayashi A, Hudson TE, Nowlin BT, Bonventre JV, et al. Fate tracing reveals the pericyte and not epithelial origin of myofibroblasts in kidney fibrosis. *Am J Pathol.* (2010) 176:85–97. doi: 10.2353/ajpath.2010.090517
50. Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Res.* (2008) 36:e11. doi: 10.1093/nar/gkm1075

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Baek, Maiorino, Kim, Glass, Raby and Yuan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.