Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

PDCD10 loss is associated with inflammatory and barrier-related endothelial remodeling in cerebral cavernous malformation

Xue Xiao
Tsinghua University
Da Li
Capital Medical University
Hengwei Song
Capital Medical University
Yuxuan Ma
Capital Medical University
Jiancong Weng
galenweng92@gmail.com

China-Japan Friendship Hospital

Research Article

Keywords: Cerebral cavernous malformation, PDCD10, Endothelial heterogeneity, Single-cell RNA sequencing, Inflammation, Blood–brain barrier

Posted Date: September 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10922122/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

PDCD10-associated cerebral cavernous malformations have an aggressive clinical presentation, but the accompanying endothelial responses remain incompletely defined.

Methods

We reanalyzed public single-cell transcriptomes from neonatal endothelial-specific Pdcd10-knockout and wild-type mice, with two mice per genotype. Vascular affiliation and pathological states were evaluated alongside mouse-level pseudobulk differential expression, pathway scores, shared pseudotime and predicted ligand–receptor communication. Independent mouse datasets and a pooled human vascular-organoid dataset assessed reproducibility and model dependence.

Results

Among 31,353 singlets, 27,445 endothelial-lineage candidates were retained for state mapping. The mean fraction of endothelial-to-mesenchymal-transition-like cells increased descriptively from 21.42% to 28.57%, with the largest increase in capillary-assigned cells. Pseudobulk analysis identified 682 differentially expressed genes, with enrichment of tumor necrosis factor–nuclear factor κB signaling and hypoxia. Serpine1 and Pgf increased, whereas Mfsd2a and Cldn5 decreased. The shared trajectory was branched and did not establish a universal endothelial-to-fibroblast sequence. Additional brain endothelial datasets reproduced Serpine1, Pgf and Fn1 increases, whereas cultured endothelial cells and organoids showed incomplete replication.

Conclusions

PDCD10 loss was associated with inflammatory and barrier-related endothelial remodeling. These exploratory findings nominate targets for tissue validation but do not establish lineage conversion, pathway specificity for PDCD10 loss or a causal explanation for human hemorrhage.

Introduction

Cerebral cavernous malformations (CCMs) are common cerebrovascular lesions composed of dilated, thin-walled vascular channels that are prone to chronic leakage and intracranial hemorrhage. Patients may develop seizures, focal neurological deficits or hemorrhagic stroke, yet disease burden varies markedly across individuals and genotypes. Familial CCM is caused by loss-of-function variants in KRIT1 (CCM1), CCM2 or PDCD10 (CCM3). Among these, PDCD10-associated disease has particularly aggressive manifestations: affected patients frequently present during childhood, develop a greater

number of lesions and experience hemorrhagic events at younger ages. Experimental Pdcd10-loss models similarly exhibit rapid lesion formation and severe vascular abnormalities, providing experimental models of the vascular abnormalities associated with PDCD10 deficiency (Bergametti et al. 2005; Denier et al. 2006; Riant et al. 2013; Shenkar et al. 2015; Snellings et al. 2021).

Despite this clinical and experimental consistency, the mechanism underlying CCM3 severity remains unresolved. CCM1, CCM2 and CCM3 participate in a shared endothelial signaling complex, and disruption of each can perturb vascular quiescence, junctional integrity and MEKK3–KLF2/4 signaling. It is therefore unclear whether CCM3 loss activates a distinct pathogenic pathway or instead amplifies a common CCM programme across particularly susceptible endothelial compartments. Resolving this distinction is essential for understanding how genotype determines hemorrhagic severity and for identifying biomarkers suitable for validation in human CCM tissue. Other lesion-forming mechanisms include endothelial TLR4 signaling, clonal expansion, endothelial-to-mesenchymal transition (EndMT) and cooperation with somatic PIK3CA activation (Bravi et al. 2016; Maddaluno et al. 2013; Malinverno et al. 2019; Ren et al. 2021; Stockton et al. 2010; Tang et al. 2017; Weng et al. 2021; Z. Zhou et al. 2016).

Here, we examined how endothelial pathological states relate to inferred vascular compartments after Pdcd10 loss, which molecular programmes accompany these states, and which findings are reproduced in independent datasets. We combined a vascular-zonation-first single-cell reanalysis with sample-level pseudobulk comparisons, pathway scores, a shared transcriptional trajectory and predicted intercellular communication. Independent CCM3-loss datasets assessed the reproducibility of candidate responses, whereas an exploratory CCM1/2/3 organoid comparison examined their genotype and model dependence. This framework connects cell-state heterogeneity to inflammatory and barrier-related programmes while distinguishing directly observed expression changes from hypotheses about clinical severity.

Methods

Study design and datasets

The discovery analysis used raw mouse single-cell RNA-sequencing counts from GSE155788, comprising two postnatal day 8 wild-type and two endothelial-specific Pdcd10-knockout samples. Independent validation included untreated age-matched knockout and wild-type cells from GSE246598 and cerebellar endothelial bulk RNA-sequencing from three inducible Ccm3-knockout and three wild-type mice in GSE212018. Cross-genotype comparison used the supplied endothelial-cluster analysis of human CCM1-, CCM2- and CCM3-knockout blood-vessel organoids from GSE276497. Rapamycin-treated samples were excluded from the central validation. The biological sample, rather than the individual cell, was considered the replicate wherever replicate-level inference was possible (Globisch et al. 2022; Min et al. 2024; Orsenigo et al. 2020; Skowronek et al. 2025).

Dataset search and eligibility

During revision, GEO, the ArrayExpress collection within BioStudies, and ENA were searched on 9 September 2026 using disease names and PDCD10/CCM3, KRIT1/CCM1, CCM2 and MGC4607 terms. Eligible endothelial comparisons required defined gene loss and a matched control. Treatment studies were assessed for untreated or vehicle-matched contrasts; eligibility did not depend on effect direction or statistical significance. Repository mirrors, pooled libraries and non-endothelial specimens were distinguished from independent endothelial cohorts. The search records and sample-level assessments are provided in the Supplementary Table S10. The retrieval included 346 GEO series and 107 additional repository records before cross-repository deduplication. Thirty-eight GEO series underwent sample-level eligibility assessment, with the primary ArrayExpress record E-MTAB-13726 assessed separately. The register distinguishes eligible, contextual, excluded and unresolved records, and records completed versus outstanding analyses; not all eligible datasets were incorporated into the quantitative synthesis. GEO supplied the datasets in the completed quantitative synthesis; repository membership was not an inclusion criterion or a measure of evidence quality. Public raw data requiring reconstruction were distinguished from restricted-access data and from unresolved sample annotations. The retrieval and assessment totals are summarized in Online Resource 1; accession-level decisions and queries are supplied in Supplementary Table S10 (Online Resource 2).

Single-cell processing and endothelial annotation

Cells with fewer than 500 detected genes or at least 15% mitochondrial transcripts were excluded. Doublets were identified independently in each sample using scDblFinder. Singlets were log-normalized, integrated by reciprocal PCA and re-clustered using Seurat (Germain et al. 2022; Hao et al. 2024).

Endothelial-lineage candidates were distinguished from erythroid, perivascular, mixed and unresolved populations. A stringent endothelial subset was used for pseudobulk differential expression, whereas the broader endothelial-lineage set retained cells with partial endothelial-identity loss for vascular-origin and pathological-state analyses (Supplementary Fig. S1; Supplementary Table S1).

Vascular origin and pathological state were annotated as separate axes using different marker programmes; both annotations were derived from the same transcriptomic dataset. Cells retaining endothelial identity were annotated as arterial, arterial-capillary, capillary, venous-capillary or venous using coordinated vascular-zonation programmes. Cells with reduced endothelial identity were mapped to these vascular anchors by weighted nearest-neighbor voting, and assignment confidence was recorded (Supplementary Fig. S2); see Supplementary Table S2 (Vanlandewijck et al. 2018).

Pathological states were defined using coordinated endothelial-identity, mesenchymal, extracellular-matrix and proliferation programmes and classified as endothelial identity retained, cycling endothelial, early EndMT, intermediate EndMT, cycling EndMT-like or identity-lost/fibroblast-like candidate. EndMT-like labels indicate transcriptional states and do not establish lineage conversion. Angiogenesis was scored

as a functional programme rather than treated as an exclusive vascular segment (Supplementary Fig. S3; Supplementary Tables S2 and S3).

Differential expression and pathway analysis

Raw counts were aggregated by mouse across all endothelial cells or within endothelial subtypes. Knockout and wild-type samples were compared using edgeR robust quasi-likelihood models after gene-expression filtering and TMM normalization. Benjamini–Hochberg FDR below 0.05 defined differential expression (Supplementary Fig. S4); see Supplementary Table S4 (Chen et al. 2016).

Hallmark enrichment was performed with fgseaMultilevel using genes ranked by signed quasi-likelihood statistics. Predefined inflammatory, hypoxic, TGF-β, angiogenic, adhesion, extracellular-matrix and barrier programmes were localized at single-cell resolution using UCell. Mouse-level score summaries were used to avoid treating individual cells as independent biological replicates (Supplementary Fig. S4; Supplementary Tables S3 and S4) (Andreatta and Carmona 2021; Korotkevich et al. 2016; Liberzon et al. 2015).

Additional endothelial comparisons

External evidence was organized by clinical relevance and experimental context, with replication assessed separately. Human lesion tissue provides the closest clinical context, but no genotype-resolved human hemorrhage comparison was available in the analyzed set. The main synthesis therefore emphasized mouse brain endothelial comparisons, followed by exploratory human vascular organoids; cultured endothelial and guide-perturbation models were evaluated separately in Supplementary Results. Human-derived cultures were not classified as patient-lesion evidence. This ordering did not assign numerical quality scores or determine eligibility from effect direction. Counts or estimated counts were analyzed with robust edgeR quasi-likelihood models, and normalized-expression datasets with platform-appropriate limma models. Unreplicated or pooled comparisons were descriptive (Supplementary Methods; Tables S9 and S10).

Trajectory and communication analyses

Monocle3 was fitted once to the combined wild-type and knockout endothelial-lineage object using the atlas UMAP coordinates. The shared starting-cell set comprised 404 WT cells selected from the original endothelial-identity-retained category by high endothelial-identity and low early/late mesenchymal scores; these mapped to 148 root principal nodes. The same root set defined pseudotime for both genotypes (Supplementary Table S5). For the retained shared graph, finite-pseudotime cells were assigned to 30 approximately equal-cell bins. Gene expression was summarized within mouse and bin, then averaged by genotype. The display connects these bin means without additional smoothing and

shows the observed mouse range where both mice contribute; it is a whole-graph summary and does not resolve branch-specific regulation (Supplementary Fig. S5) (Cao et al. 2019).

CellChat networks were inferred separately for each genotype using the mouse ligand–receptor database. Sample-by-state groups were downsampled to a maximum of 500 cells, and the main analysis was restricted to interactions involving pathological endothelial states using the retained grouped CellChat annotation, which retains separate endothelial and cycling-endothelial groups alongside pathological, mural and immune groups (Supplementary Fig. S6); see Supplementary Table S6 (Jin et al. 2021).

For GSE276497, the supplied endothelial-cluster comparisons were evaluated using a common threshold of absolute normalized average log fold change greater than 0.5 and adjusted P below 0.05.

Statistical interpretation

Cell-level analyses were used for annotation and localization rather than genotype-level inference. Composition, UCell, pseudotime and pooled-library comparisons were considered descriptive when biological replication was insufficient. CellChat interactions were interpreted as transcript-based predictions. Monocle3 pseudotime represents branched transcriptional ordering and does not demonstrate chronological progression or lineage conversion. Public mouse and organoid datasets were not interpreted as direct evidence of human hemorrhage.

Statistical design constraints

Sample sizes were fixed by the available public datasets; no prospective sample-size calculation was performed for this secondary analysis. No new animals or participants were allocated to groups. Genotype labels were available during reanalysis, which was not blinded. The small number of independent samples precluded reliable assessment of normality or variance homogeneity. RNA-sequencing counts were analyzed using negative-binomial quasi-likelihood models rather than tests assuming normally distributed cell-level measurements. Adjusted P values from the small discovery cohort should be interpreted together with effect sizes and external evidence; they do not overcome the limited biological replication.

Use of generative AI

OpenAI ChatGPT/Codex assisted with code development, figure preparation, and drafting and editing manuscript text. The authors remain responsible for the analyses, source verification, interpretation and final content.

Results

A vascular-origin atlas retains pathological endothelial cells

After quality control and doublet removal, the GSE155788 atlas contained 31,353 singlets, including 14,498 wild-type and 16,855 Pdcd10-knockout cells. Endothelial-focused reanalysis identified 25,270 high-confidence endothelial cells for pseudobulk analysis and 27,445 endothelial-lineage candidates for vascular-origin and pathological-state mapping (Supplementary Fig. S1). Sample-level quality-control, doublet, broad-atlas and endothelial-cluster outputs are provided in Supplementary Tables S1 and S2.

Pathological cells were retained during vascular-origin inference (Fig. 1a). The resulting atlas linked cells to arterial, arterial-capillary, capillary, venous-capillary and venous transcriptional anchors. Angiogenic activity was displayed separately as a continuous programme (Fig. 1b). Capillary and venous-capillary assignments accounted for much of the atlas, while the mean venous-origin fraction increased descriptively from 1.94% in wild type to 7.32% after Pdcd10 loss (Fig. 1c). These labels describe inferred transcriptomic affiliation and do not identify the lineage-traced origin of individual lesions (Supplementary Fig. S2; Supplementary Tables S2 and S3).

Pathological states increase descriptively in capillary and venous-capillary assignments

Six endothelial states were retained on the common atlas and supported by endothelial, matrix-associated and proliferation markers (Fig. 2a–c). The combined fraction of early EndMT-like, intermediate EndMT-like, cycling EndMT-like and fibroblast-like candidate cells increased from a mean of 21.42% in wild type to 28.57% in knockout mice. Endothelial-identity-retained cells decreased from 62.78% to 55.94% (Supplementary Fig. S3). These sample-level proportions are descriptive because each genotype comprised two mice. Functional definitions and annotation outputs are provided in Supplementary Tables S2 and S3.

The substantial EndMT-like fraction in wild-type P8 mice limits the specificity of these labels for established lesions. The annotated states may include developmental or activated endothelial programmes sharing features with partial EndMT. Their increased abundance after Pdcd10 loss therefore indicates an altered transcriptional-state distribution rather than direct evidence of endothelial conversion.

The largest increase occurred among capillary-origin cells, in which the pathological-state fraction rose from 22.70% to 35.00% (+ 12.3 percentage points). Venous-capillary cells increased from 22.50% to 29.37% (Fig. 2d). These descriptive differences nominate capillary and venous-capillary assignments for follow-up; they do not establish a genotype-by-vascular-compartment interaction.

The cycling EndMT-like population expressed proliferation genes including Mki67, Top2a, Cdk1, Ccnb1 and Ube2c together with transition-associated features. Proliferation was treated as a concurrent

programme, not a later stage. Fibroblast-like candidates were rare and had lower vascular-origin confidence (Fig. 2b), leaving their endothelial ancestry unresolved.

Inflammatory and barrier-related programmes accompany pathological states

All-endothelial sample-level pseudobulk analysis identified 682 differentially expressed genes after Pdcd10 loss, comprising 467 increased and 215 decreased genes. TNF signaling through NF-κB was the dominant enriched Hallmark programme (NES = 2.725, FDR = 2.62×10^{-23}), accompanied by hypoxia and other inflammatory and remodeling programmes (Supplementary Fig. S4; Supplementary Table S4). The separately defined activated-mesenchymal endothelial contrast also showed TNF–NF-κB enrichment (FDR = 3.8×10^{-19}; Fig. 3b). These statistics refer to different cell subsets and are reported separately.

In the all-endothelial comparison, NF-κB-associated genes including Cxcl2, Vcam1, Tnfaip3, Icam1 and Nfkbia increased. Pgf increased by 2.48 $\log_2$ fold and Serpine1, which encodes PAI-1, by 1.36 $\log_2$ fold. Mfsd2a and Cldn5 decreased. This pattern associates inflammatory and angiogenic transcription with altered hemostatic and matrix regulation and reduced barrier specialization. It does not establish that NF-κB directly regulates these candidate effectors.

At the state level, mouse-first summaries showed higher NF-κB and hypoxia scores in knockout cells across the identity-retained, early EndMT-like and intermediate EndMT-like categories (Fig. 3a,c,d). Barrier programme scores varied across states and mice and did not show a uniform within-state knockout reduction (Fig. 3e). Thus, reduced all-endothelial expression of individual barrier genes should not be equated with lower values of every composite barrier score in every state. Earlier functional-subtype contrasts are supplied separately in Supplementary Table S4.

A shared graph describes branched endothelial transcriptional variation

A common Monocle3 graph placed wild-type and knockout cells within the same transcriptional and pseudotime framework (Fig. 4). The graph extended from endothelial-identity-rich root cells towards several pathological branches rather than supporting a single obligatory sequence from non-EndMT through early, intermediate, cycling and fibroblast-like states (Supplementary Fig. S5).

Early and intermediate EndMT cells occupied related regions, whereas cycling EndMT-like cells accumulated on a distal proliferative branch. Identity-lost/fibroblast-like candidates occupied a separate branch and were not consistently positioned at the highest pseudotime. The trajectory therefore supports branched endothelial remodeling but does not demonstrate that the fibroblast-like population represents a universal late-EndMT endpoint.

Across the shared graph, Cldn5 expression declined with pseudotime, while Serpine1 and Pgf were generally higher and Mfsd2a lower in knockout cells (Fig. 4b–g). Nfkbia was also elevated across much of the ordering. Fn1 profiles overlapped and varied across bins, providing less consistent evidence of

genotype separation. Because branches are combined in this display, the curves describe graph-associated expression patterns rather than a single developmental sequence or the timing of pathway activation. Cell-level pseudotime and gene-expression summaries are provided in Supplementary Table S5.

Predicted communication differs across remodeling-associated pathways

Ligand–receptor inference identified genotype-associated differences in predicted communication involving pathological endothelial states (Fig. 5). The displayed pathway summary aggregates predicted probabilities across these interactions, while the selected ligand–receptor pairs include matrix, adhesion and angiogenic signals. These analyses nominate potential interactions for follow-up; they do not measure receptor activation or mouse-level differences in physical signaling. Supporting network outputs are provided in Supplementary Fig. S6 and Supplementary Table S6.

Predicted VEGF, ANGPT and FN1 pathway differences were positive, whereas the summed collagen-associated difference was negative (Fig. 5a). Individual pairs could differ from their aggregate pathway pattern. This heterogeneity argues for testing selected ligand–receptor interactions in their cellular context rather than interpreting all matrix-associated communication as uniformly enhanced.

Independent brain endothelial datasets support selected candidate responses

Two independent mouse datasets were used to examine candidate response directions (Fig. 6). The replicate-level cerebellar endothelial dataset GSE212018 supported increased Serpine1 and Pgf and decreased Mfsd2a and Cldn5. The untreated, age-matched GSE246598 comparison provided weaker and gene-dependent directional support; it comprised pooled genotype libraries and was treated descriptively. The effect estimates are displayed in separate panels because the pooled-library mean log-expression differences are not the same estimand as replicate-level $\log_2$ fold changes (Supplementary Table S7).

Programme summaries were broadly concordant for inflammatory and barrier-related responses between discovery and bulk validation, with smaller changes in the pooled single-cell comparison (Fig. 6d). These summaries average the reported effects of module genes and are not formal pathway-enrichment tests or a meta-analysis. Their value is in assessing direction and model dependence, not demonstrating identical response magnitudes across datasets.

The expanded brain endothelial comparisons further supported selected inflammatory and matrix-associated responses. Serpine1, Pgf and Fn1 increased in GSE204979 and the normoxic GSE261224 comparison, with positive enrichment of TNF–NF-κB, hypoxia, TGF-β and coagulation programmes (Supplementary Fig. S7; Supplementary Table S9). The chronic GSE246373 cohort also supported increases in these three genes (Supplementary Table S9). Replication was incomplete across the full set

of brain endothelial and cultured models, including nonsignificant and opposite-direction effects. Detailed comparisons are reported in Supplementary Results and Table S9.

Organoid comparisons reveal shared and model-dependent CCM responses

The human vascular-organoid comparison provided exploratory genotype context (Fig. 7; Supplementary Table S8). CCM1, CCM2 and CCM3 loss shared selected endothelial responses, with differences in candidate expression. One pooled library per genotype precluded biological-replicate inference; the reported DEG counts cannot establish a severity ranking. The mouse candidate pattern was not fully reproduced in organoids, as detailed in Supplementary Results. These findings support model-dependent CCM responses rather than a universal CCM3-specific programme.

Discussion

This reanalysis links Pdcd10 loss to an increased burden of EndMT-like transcriptional states within capillary and venous-capillary assignments, accompanied by inflammatory and barrier-related expression changes. Its contribution is the joint description of vascular affiliation, pathological state and molecular response while retaining cells with reduced endothelial identity. Independent mouse data support several candidate directions, whereas the organoid comparison shows the limits of their transfer across models. Together, the results nominate a coordinated endothelial response for experimental testing, without resolving the cause of the more severe clinical phenotype associated with PDCD10 variants.

The original GSE155788 study established heterogeneity and susceptibility among venous-capillary and progenitor-like endothelial populations. Our analysis extends that framework by relating pathological-state burden to inferred vascular affiliation. The largest descriptive increase occurred among capillary-origin cells, with a parallel increase in venous-capillary cells. This supports involvement of these compartments, but origin inference becomes less reliable as vascular identity is lost. Weighted votes from transcriptomically similar anchors are not calibrated lineage probabilities, and forcing a label on an ambiguous cell cannot establish its anatomical ancestry (Orsenigo et al. 2020).

EndMT-like states require particular caution in this neonatal dataset. Wild-type cells already expressed parts of the programme, and shared inflammatory or extracellular-matrix genes can contribute to both state annotation and pathway scores. These analyses therefore provide complementary descriptions rather than fully independent proof of EndMT. The branched trajectory and the inconsistent pseudotime position of fibroblast-like candidates further argue against a universal endothelial-to-fibroblast sequence. Establishing conversion would require lineage information, spatial localization and coordinated loss of endothelial and gain of mesenchymal protein features (Bravi et al. 2016; Maddaluno et al. 2013).

TNF–NF-κB enrichment was a prominent feature of the discovery pseudobulk analysis and was supported by changes in several response genes. Hypoxia, TGF-β/EMT and hemostatic programmes

provided a broader context for the endothelial response. However, their co-occurrence does not establish an upstream hierarchy. NFKBIA and TNFAIP3 expression are compatible with an NF-κB-associated response but cannot substitute for measurements of pathway activation or functional dependence. Phospho-p65 and phospho-SMAD2/3 measurements, coupled with perturbation, would help distinguish associated transcription from a mechanism sustaining the pathological state (Z. Zhou et al. 2016). Prior work already linked Pdcd10 depletion to hypoxia, inflammatory responses and Serpine1 expression (Fusco et al. 2022). Moreover, endothelial Ikkb disruption did not prevent CCM and was associated with a trend toward increased lesions and immunothrombosis (Lai et al. 2022). Our transcriptional enrichment therefore neither establishes NF-κB as a uniformly pathogenic driver nor supports indiscriminate NF-κB inhibition.

The effector pattern provides a direct route to validation in human CCM3 tissue. SERPINE1/PAI-1 integrates inflammatory, hypoxic and TGF-β-associated responses and could promote persistence of fibrin-rich thrombo-inflammatory material by suppressing local fibrinolysis. PGF could amplify VEGFR1-dependent angiogenic and inflammatory signaling. MFSD2A loss could reduce suppression of caveolae-mediated transcytosis, thereby weakening blood–brain barrier function even when some junctional markers remain. CLDN5 loss supplies a complementary junctional readout. None of these observations alone proves hemorrhage. Their value lies in linking transcriptional-state changes to testable endothelial readouts; reproducibility varied across datasets and was incomplete in organoids (Andreone et al. 2017; Ben-Zvi et al. 2014; Carmeliet et al. 2001; Globisch et al. 2022).

The predicted communication network extends this model from cell-intrinsic state to tissue remodeling. Increased VEGF, ANGPT and FN1–integrin communication could link abnormal endothelial cells to neighbouring endothelial, mural and immune compartments. These axes are biologically plausible because they influence permeability, angiogenesis, matrix adhesion and inflammatory recruitment. However, transcript-based communication tools can be sensitive to cell abundance and cannot demonstrate receptor engagement. Multiplex immunofluorescence or spatial transcriptomics should therefore test whether ligand-positive pathological endothelial cells and receptor-positive target cells are co-localized within the same CCM3 lesions (H. J. Zhou et al. 2016).

Evidence was interpreted in relation to the biological setting it represents. Patient lesions are closest to the clinical question, whereas mouse brain endothelial comparisons provide the principal disease-model evidence here. Human vascular organoids offer a multicellular human-derived context, and cultured endothelial cells isolate selected cellular responses; neither is equivalent to a patient lesion. Organoid maturation, culture conditions and incomplete representation of the brain vascular environment may contribute to differences between models, but these explanations were not tested in this reanalysis. Incomplete and opposite-direction results therefore remain relevant limits on generalizability (Supplementary Results; Tables S8 and S9). Clinical relevance also does not establish causality: replicated perturbation and rescue experiments would still be needed to test the proposed mechanism. The available comparisons do not establish a universally stronger CCM3 programme or explain its clinical severity (Koskimäki et al. 2019; Skowronek et al. 2025).

The study has important limitations. The discovery dataset contains only two mice per genotype, so composition changes and pseudotime distributions require independent replication. The tissue is neonatal and actively developing; wild-type EndMT-like states may not resemble adult human lesion biology. Origin labels for identity-lost cells are inferred and become less certain as vascular markers disappear. Pseudotime is neither lineage tracing nor a measure of real time, and CellChat is hypothesis-generating. GSE246598 uses pooled condition-level libraries for the focused comparison, and GSE276497 contains one pooled library per organoid genotype. The same expression profiles contributed to both cell-state annotation and programme scoring, limiting their independence as evidence. Finally, none of the analysed datasets directly measures bleeding, and inter-model concordance cannot substitute for human tissue validation. The cross-repository search was retrospective and cannot guarantee coverage of all available data. Public raw-data reconstruction remained incomplete for E-MTAB-13726, and conflicting annotations prevented directional interpretation of GSE186523. These are reconstruction and annotation limitations, not documented permission restrictions; incomplete coverage may affect the synthesis.

A focused validation study would compare genotype-confirmed human CCM3 lesions with suitable non-lesional vessels and, where available, CCM1- and CCM2-associated lesions. Endothelial localization of PAI-1, PGF, phospho-p65, MFSD2A and CLDN5 could be evaluated with CD31 or ERG. Additional FN1, COL1A1 or α-SMA staining would require careful separation of endothelial and adjacent mural or stromal signals. Fibrin measurements could test the proposed hemostatic context. Patient-level quantification and direct genotype contrasts would be needed to connect these candidate features to disease severity.

PDCD10 loss was associated with capillary and venous-capillary endothelial-state changes, inflammatory and hypoxic transcription, and reduced barrier specialization. Independent mouse datasets supported selected candidate responses, while pooled organoid comparisons revealed shared and model-dependent effects. These observations provide a framework for testing endothelial dysfunction in CCM, with the relationship to human lesion growth and hemorrhage remaining to be established.

Declarations

Code availability

Analysis scripts, source-data tables, software-session records and a provenance manifest are available from the corresponding author on request.

Ethics statement

This study analysed de-identified, publicly available mouse and human endothelial or organoid transcriptomic datasets and generated no new animal or human-subject data. Ethical approvals for the original experiments are reported in the source publications.

Funding

This work was supported by Youth Fund of the National Natural Science Foundation of China (Grant 82301470), Elite Medical Professionals Project of China-Japan Friendship Hospital (NO.ZRJY2025-GG10).

Author contributions

X.X. performed investigation, validation, methodology, data curation, and drafted the original manuscript. D.L. performed formal analysis, software implementation, data curation, visualization, investigation, and drafted the original manuscript. H.S. and Y.M. drafted the original manuscript. J.W. conceived and supervised the study, managed the project, acquired funding and resources, and critically reviewed and edited the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Supplementary information

Online Resource 1 (ESM_1.pdf) contains Supplementary Methods and Results, the retrieval and assessment summary, and Supplementary Figures S1–S7 with legends. Online Resource 2 (ESM_2.zip) contains Supplementary Tables S1–S10 and the complete-results CSV components of Table S9. Missing values indicate unavailable information, not zero.

References

1. Andreatta M, Carmona SJ (2021) UCell: Robust and scalable single-cell gene signature scoring. Comput Struct Biotechnol J 19:3796–3798. https://doi.org/10.1016/j.csbj.2021.06.043
2. Andreone BJ, Chow BW, Tata A, Lacoste B, Ben-Zvi A, Bullock K, Deik AA, Ginty DD, Clish CB, Gu C (2017) Blood-Brain Barrier Permeability Is Regulated by Lipid Transport-Dependent Suppression of Caveolae-Mediated Transcytosis. Neuron 94:581–594e5. https://doi.org/10.1016/j.neuron.2017.03.043
3. Ben-Zvi A, Lacoste B, Kur E, Andreone BJ, Mayshar Y, Yan H, Gu C (2014) Mfsd2a is critical for the formation and function of the blood–brain barrier. Nature 509:507–511. https://doi.org/10.1038/nature13324
4. Bergametti F, Denier C, Labauge P, Arnoult M, Boetto S, Clanet M, Coubes P, Echenne B, Ibrahim R, Irthum B, Jacquet G, Lonjon M, Moreau JJ, Neau JP, Parker F, Tremoulet M, Tournier-Lasserve E

(2005) Mutations within the Programmed Cell Death 10 Gene Cause Cerebral Cavernous Malformations. Am J Hum Genet 76:42–51. https://doi.org/10.1086/426952

5. Bravi L, Malinverno M, Pisati F, Rudini N, Cuttano R, Pallini R, Martini M, Larocca LM, Locatelli M, Levi V, Bertani GA, Dejana E, Lampugnani MG (2016) Endothelial Cells Lining Sporadic Cerebral Cavernous Malformation Cavernomas Undergo Endothelial-to-Mesenchymal Transition. Stroke 47:886–890. https://doi.org/10.1161/STROKEAHA.115.011867
6. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, Trapnell C, Shendure J (2019) The single-cell transcriptional landscape of mammalian organogenesis. Nature 566:496–502. https://doi.org/10.1038/s41586-019-0969-x
7. Carmeliet P, Moons L, Luttun A, Vincenti V, Compernolle V, De Mol M, Wu Y, Bono F, Devy L, Beck H, Scholz D, Acker T, DiPalma T, Dewerchin M, Noel A, Stalmans I, Barra A, Blacher S, Vandendriessche T, Ponten A, Eriksson U, Plate KH, Foidart JM, Schaper W, Charnock-Jones DS, Hicklin DJ, Herbert JM, Collen D, Persico MG (2001) Synergism between vascular endothelial growth factor and placental growth factor contributes to angiogenesis and plasma extravasation in pathological conditions. Nat Med 7:575–583. https://doi.org/10.1038/87904
8. Chen Y, Lun ATL, Smyth GK (2016) From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. https://doi.org/10.12688/f1000research.8987.2. F1000Research 5:1438
9. Denier C, Labauge P, Bergametti F, Marchelli F, Riant F, Arnoult M, Maciazek J, Vicaut E, Brunereau L, Tournier-Lasserve E (2006) Genotype–phenotype correlations in cerebral cavernous malformations patients. Ann Neurol 60:550–556. https://doi.org/10.1002/ana.20947
10. Frias-Anaya E, Gallego-Gutierrez H, Gongol B, Weinsheimer S, Lai CC, Orecchioni M, Sriram A, Bui CM, Nelsen B, Hale P, Pham A, Shenkar R, DeBiasse D, Lightle R, Girard R, Li Y, Srinath A, Daneman R, Nudleman E, Sun H, Guma M, Dubrac A, Mesarwi OA, Ley K, Kim H, Awad IA, Ginsberg MH, Lopez-Ramirez MA (2024) Mild Hypoxia Accelerates Cerebral Cavernous Malformation Disease Through CX3CR1-CX3CL1 Signaling. Arteriosclerosis, Thrombosis, and Vascular Biology 44:1246–1264. https://doi.org/10.1161/ATVBAHA.123.320367
11. Fusco C, Nardella G, Di Filippo L, Dejana E, Cacchiarelli D, Petracca A, Micale L, Malinverno M, Castori M (2022) Transcriptome Analysis Reveals Altered Expression of Genes Involved in Hypoxia, Inflammation and Immune Regulation in Pdcd10-Depleted Mouse Endothelial Cells. Genes 13:961. https://doi.org/10.3390/genes13060961
12. Germain PL, Lun A, Garcia Meixide C, Macnair W, Robinson MD (2022) Doublet identification in single-cell sequencing data using scDblFinder. https://doi.org/10.12688/f1000research.73600.2. F1000Research 10:979
13. Globisch MA, Onyeogaziri FC, Jauhiainen S, Yau ACY, Orsenigo F, Conze LL, Arce M, Corada M, Smith RO, Rorsman C, Sundell V, Fernando D, Daniel G, Mattsson O, Savander H, Wanders A, Rezai Jahromi B, Laakso A, Niemelä M, Dejana E, Magnusson PU (2022) Immunothrombosis and vascular

heterogeneity in cerebral cavernous malformation. Blood 140:2154–2169. https://doi.org/10.1182/blood.2021015350

14. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, Srivastava A, Molla G, Madad S, Fernandez-Granda C, Satija R (2024) Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol 42:293–304. https://doi.org/10.1038/s41587-023-01767-y
15. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, Myung P, Plikus MV, Nie Q (2021) Inference and analysis of cell-cell communication using CellChat. Nat Commun 12. https://doi.org/10.1038/s41467-021-21246-9
16. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A (2016) Fast gene set enrichment analysis. bioRxiv preprint. https://doi.org/10.1101/060012
17. Koskimäki J, Girard R, Li Y, Saadat L, Zeineddine HA, Lightle R, Moore T, Lyne S, Avner K, Shenkar R, Cao Y, Shi C, Polster SP, Zhang D, Carrión-Penagos J, Romanos S, Fonseca G, Lopez-Ramirez MA, Chapman EM, Popiel E, Tang AT, Akers A, Faber P, Andrade J, Ginsberg M, Derry WB, Kahn ML, Marchuk DA, Awad IA (2019) Comprehensive transcriptome analysis of cerebral cavernous malformation across multiple species and genotypes. JCI Insight 4:e126167. https://doi.org/10.1172/jci.insight.126167
18. Lai CC, Nelsen B, Frias-Anaya E, Gallego-Gutierrez H, Orecchioni M, Herrera V, Ortiz E, Sun H, Mesarwi OA, Ley K, Gongol B, Lopez-Ramirez MA (2022) Neuroinflammation Plays a Critical Role in Cerebral Cavernous Malformation Disease. Circul Res 131:909–925. https://doi.org/10.1161/CIRCRESAHA.122.321129
19. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P (2015) The Molecular Signatures Database Hallmark Gene Set Collection. Cell Syst 1:417–425. https://doi.org/10.1016/j.cels.2015.12.004
20. Maddaluno L, Rudini N, Cuttano R, Bravi L, Giampietro C, Corada M, Ferrarini L, Orsenigo F, Papa E, Boulday G, Tournier-Lasserve E, Chapon F, Richichi C, Retta SF, Lampugnani MG, Dejana E (2013) EndMT contributes to the onset and progression of cerebral cavernous malformations. Nature 498:492–496. https://doi.org/10.1038/nature12207
21. Malinverno M, Maderna C, Abu Taha A, Corada M, Orsenigo F, Valentino M, Pisati F, Fusco C, Graziano P, Giannotta M, Yu QC, Zeng YA, Lampugnani MG, Magnusson PU, Dejana E (2019) Endothelial cell clonal expansion in the development of cerebral cavernous malformations. Nat Commun 10:2761. https://doi.org/10.1038/s41467-019-10707-x
22. Min W, Qin L, Zhang H, López-Giráldez F, Jiang N, Kim Y, Mohan VK, Su M, Murray KN, Grutzendler J, Zhou JH (2024) mTORC1 Signaling in Brain Endothelial Progenitors Contributes to CCM Pathogenesis. Circul Res 135. https://doi.org/10.1161/CIRCRESAHA.123.324015
23. Orsenigo F, Conze LL, Jauhiainen S, Corada M, Lazzaroni F, Malinverno M, Sundell V, Cunha SI, Brännström J, Globisch MA, Maderna C, Lampugnani MG, Magnusson PU, Dejana E (2020) Mapping endothelial-cell diversity in cerebral cavernous malformations at single-cell resolution. eLife 9:e61413. https://doi.org/10.7554/eLife.61413

24. Ren AA, Snellings DA, Su YS, Hong CC, Castro M, Tang AT, Detter MR, Hobson N, Girard R, Romanos S, Lightle R, Moore T, Shenkar R, Benavides C, Beaman MM, Müller-Fielitz H, Chen M, Mericko P, Yang J, Sung DC, Lawton MT, Ruppert JM, Schwaninger M, Körbelin J, Potente M, Awad IA, Marchuk DA, Kahn ML (2021) PIK3CA and CCM mutations fuel cavernomas through a cancer-like mechanism. Nature 594:271–276. https://doi.org/10.1038/s41586-021-03562-8
25. Riant F, Bergametti F, Fournier HD, Chapon F, Michalak-Provost S, Cecillon M, Lejeune P, Hosseini H, Choe C, Orth M, Bernreuther C, Boulday G, Denier C, Labauge P, Tournier-Lasserve E (2013) CCM3 Mutations Are Associated with Early-Onset Cerebral Hemorrhage and Multiple Meningiomas. Mol Syndromol 4:165–172. https://doi.org/10.1159/000350042
26. Shenkar R, Shi C, Rebeiz T, Stockton RA, McDonald DA, Mikati AG, Zhang L, Austin C, Akers AL, Gallione CJ, Rorrer A, Gunel M, Min W, de Marcondes J, Lee C, Marchuk DA, Awad IA (2015) Exceptional aggressiveness of cerebral cavernous malformation disease associated with PDCD10 mutations. Genet Sci 17:188–196. https://doi.org/10.1038/gim.2014.97
27. Skowronek D, Pilz RA, Saenko VV, Mellinger L, Singer D, Ribback S, Weise A, Claaßen K, Büttner C, Brockmann EM, Hübner CA, Aung T, Haerteis S, Bekeschus S, Ekici AB, Felbor U, Rath M (2025) High-throughput differentiation of human blood vessel organoids reveals overlapping and distinct functions of the cerebral cavernous malformation proteins. Angiogenesis 28. https://doi.org/10.1007/s10456-025-09985-5
28. Snellings DA, Hong CC, Ren AA, Lopez-Ramirez MA, Girard R, Srinath A, Marchuk DA, Ginsberg MH, Awad IA, Kahn ML (2021) Cerebral Cavernous Malformation: From Mechanism to Therapy. Circul Res 129:195–215. https://doi.org/10.1161/CIRCRESAHA.121.318174
29. Stockton RA, Shenkar R, Awad IA, Ginsberg MH (2010) Cerebral cavernous malformations proteins inhibit Rho kinase to stabilize vascular integrity. J Exp Med 207:881–896. https://doi.org/10.1084/jem.20091258
30. Tang AT, Choi JP, Kotzin JJ, Yang Y, Hong CC, Hobson N, Girard R, Zeineddine HA, Lightle R, Moore T, Cao Y, Shenkar R, Chen M, Mericko P, Yang J, Li L, Tanes C, Kobuley D, Võsa U, Whitehead KJ, Li DY, Franke L, Hart B, Schwaninger M, Henao-Mejia J, Morrison L, Kim H, Awad IA, Zheng X, Kahn ML (2017) Endothelial TLR4 and the microbiome drive cerebral cavernous malformations. Nature 545:305–310. https://doi.org/10.1038/nature22075
31. Vanlandewijck M, He L, Mäe MA, Andrae J, Ando K, Del Gaudio F, Nahar K, Lebouvier T, Laviña B, Gouveia L, Sun Y, Raschperger E, Räsänen M, Zarb Y, Mochizuki N, Keller A, Lendahl U, Betsholtz C (2018) A molecular atlas of cell types and zonation in the brain vasculature. Nature 554:475–480. https://doi.org/10.1038/nature25739
32. Weng J, Yang Y, Song D, Huo R, Li H, Chen Y, Nam Y, Zhou Q, Jiao Y, Fu W, Yan Z, Wang J, Xu H, Di L, Li J, Wang S, Zhao J, Wang J, Cao Y (2021) Somatic MAP3K3 mutation defines a subclass of cerebral cavernous malformation. Am J Hum Genet 108:942–950. https://doi.org/10.1016/j.ajhg.2021.04.005

33. Zhou Z, Tang AT, Wong WY, Bamezai S, Goddard LM, Shenkar R, Zhou S, Yang J, Wright AC, Foley M, Arthur JSC, Whitehead KJ, Awad IA, Li DY, Zheng X, Kahn ML (2016) Cerebral cavernous malformations arise from endothelial gain of MEKK3–KLF2/4 signalling. Nature 532:122–126. https://doi.org/10.1038/nature17178
34. Zhou HJ, Qin L, Zhang H, Tang W, Ji W, He Y, Liang X, Wang Z, Yuan Q, Vortmeyer A, Toomre D, Fuh G, Yan M, Kluger MS, Wu D, Min W (2016) Endothelial exocytosis of angiopoietin-2 resulting from CCM3 deficiency contributes to cerebral cavernous malformation. Nat Med 22:1033–1042. https://doi.org/10.1038/nm.4169

Figures

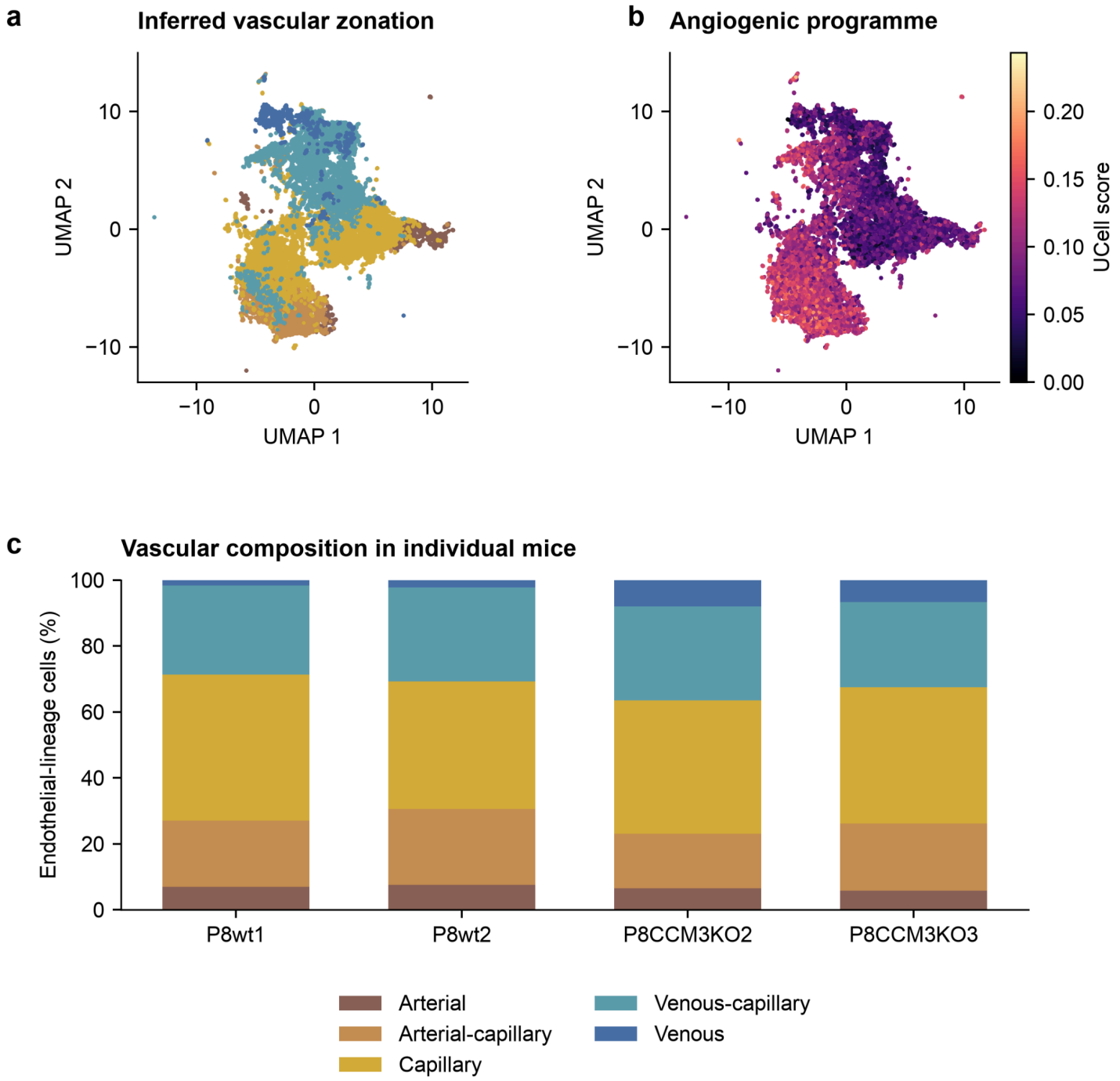


Figure 1

Vascular zonation and composition of endothelial-lineage candidates

a, Common UMAP of endothelial-lineage candidates colored by inferred vascular affiliation. b, Angiogenic programme scores on the same coordinates. c, Vascular-origin fractions in individual mice. No pathological cells were removed solely because of their EndMT-like label. Origin assignment is transcriptomic inference, not lineage tracing; n=2 mice per genotype. Supporting data: Supplementary Fig. S2; Supplementary Table S2

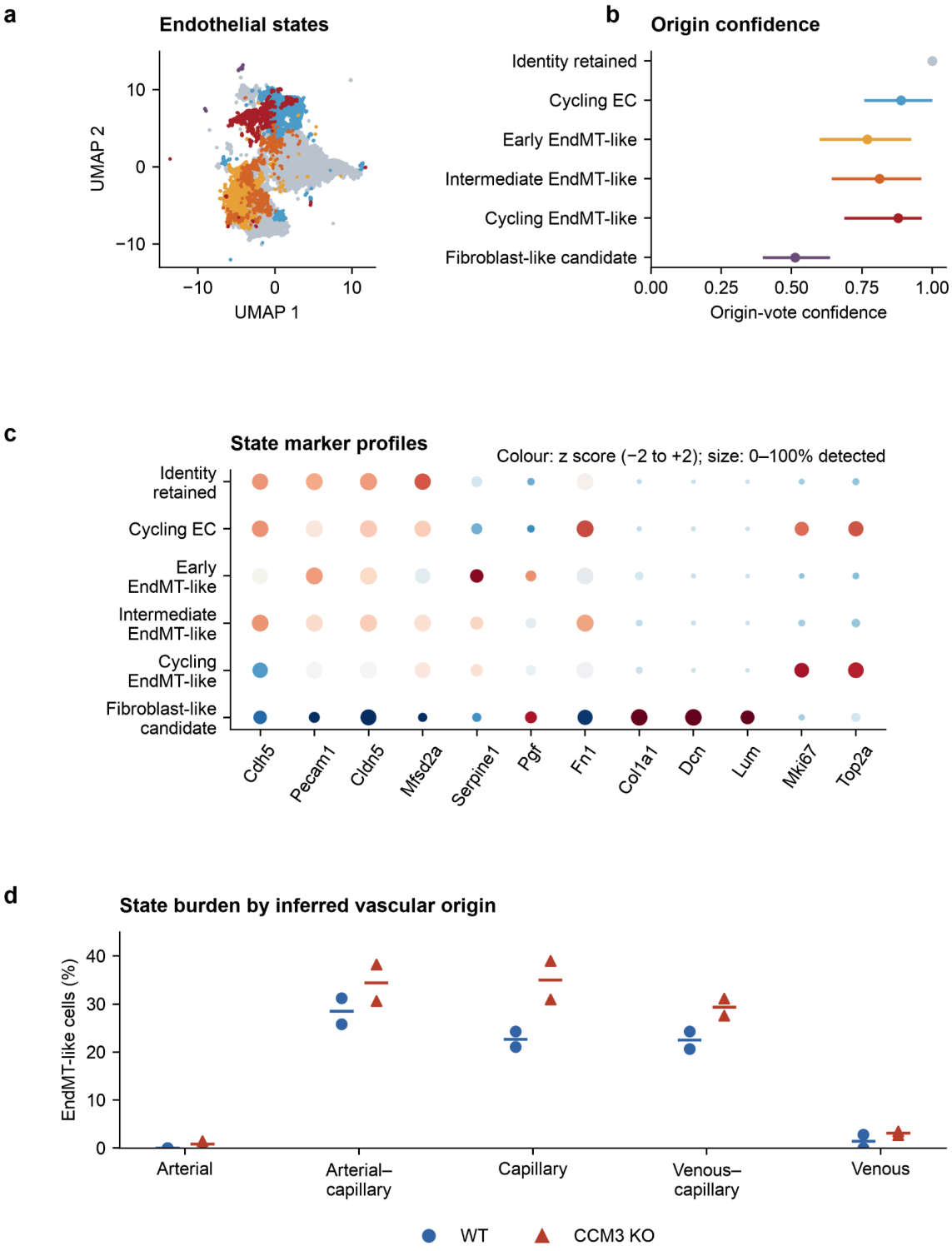


Figure 2

Pathological state annotation and vascular-origin-specific burden

a, The same atlas colored by the six retained endothelial states. b, Median origin-vote confidence with interquartile ranges; anchor values of 1 reflect assignment rules, not calibrated certainty. c, Marker profiles for the six states. Color indicates the z score of mean log-normalized expression across states;

point size increases with the fraction of cells expressing the gene. d, Combined early, intermediate, cycling EndMT-like and fibroblast-like candidate fractions within each inferred origin. Each point is one mouse; horizontal marks indicate genotype means. The small fibroblast-like population has uncertain lineage origin. The labels early and intermediate describe the annotation scheme, not verified chronological stages. Supporting data: Supplementary Fig. S3; Supplementary Table S2

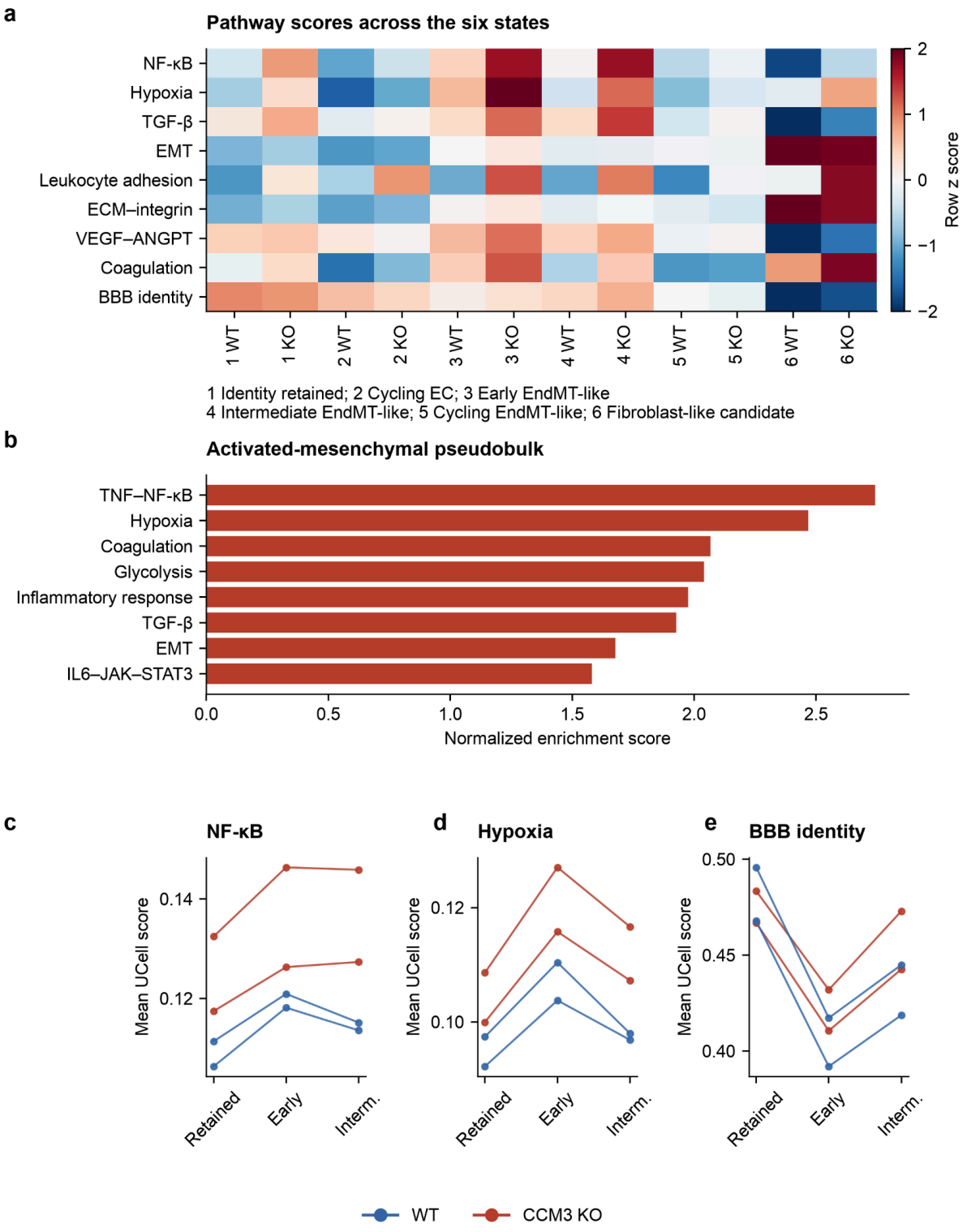


Figure 3

Inflammatory and barrier-related programme scores

a, UCell scores averaged within mouse and state, then within genotype, using the six-state assignments in Figure 2. Row z scores are for visualization. Numbers 1–6 identify the states in the order indicated below the heatmap. b, Hallmark enrichment in the retained activated-mesenchymal pseudobulk contrast, a different population definition from all-endothelial pseudobulk. Positive NES indicates enrichment in knockout. c–e, NF-κB, hypoxia and barrier programme scores in individual mice across identity-retained, early EndMT-like and intermediate EndMT-like states. Lines connect summaries from the same mouse and do not represent longitudinal observations. The state annotations and programme scores partly share transcriptomic information. Supporting data: Supplementary Fig. S4; Supplementary Tables S3 and S4

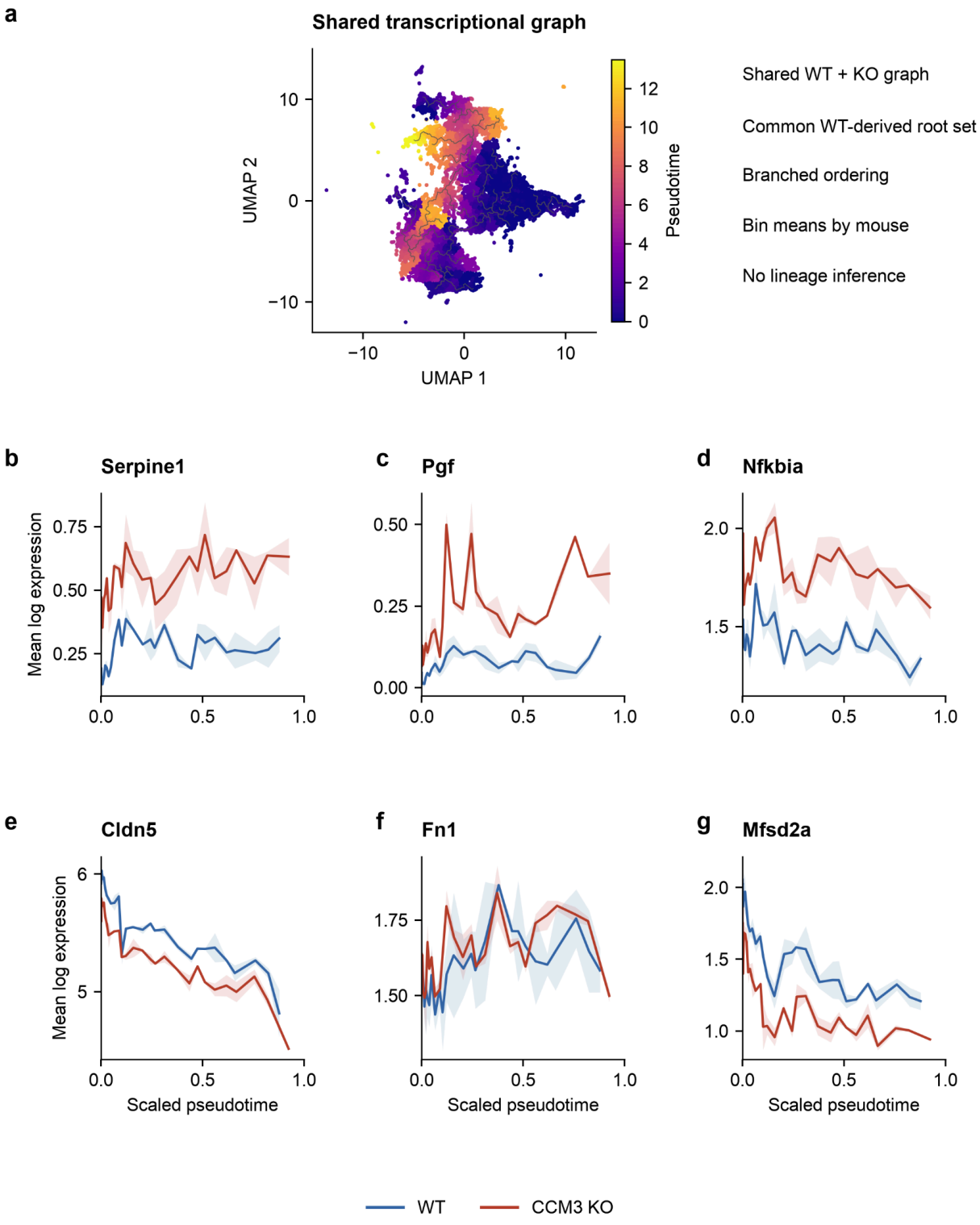

a
Shared transcriptional graph
UMAP 2
UMAP 1
Pseudotime
Shared WT + KO graph
Common WT-derived root set
Branched ordering
Bin means by mouse
No lineage inference
b
Serpine1
c
Pgf
d
Nfkbia
e
Cldn5
f
Fn1
g
Mfsd2a
Mean log expression
Scaled pseudotime
WT
CCM3 KO

Figure 4

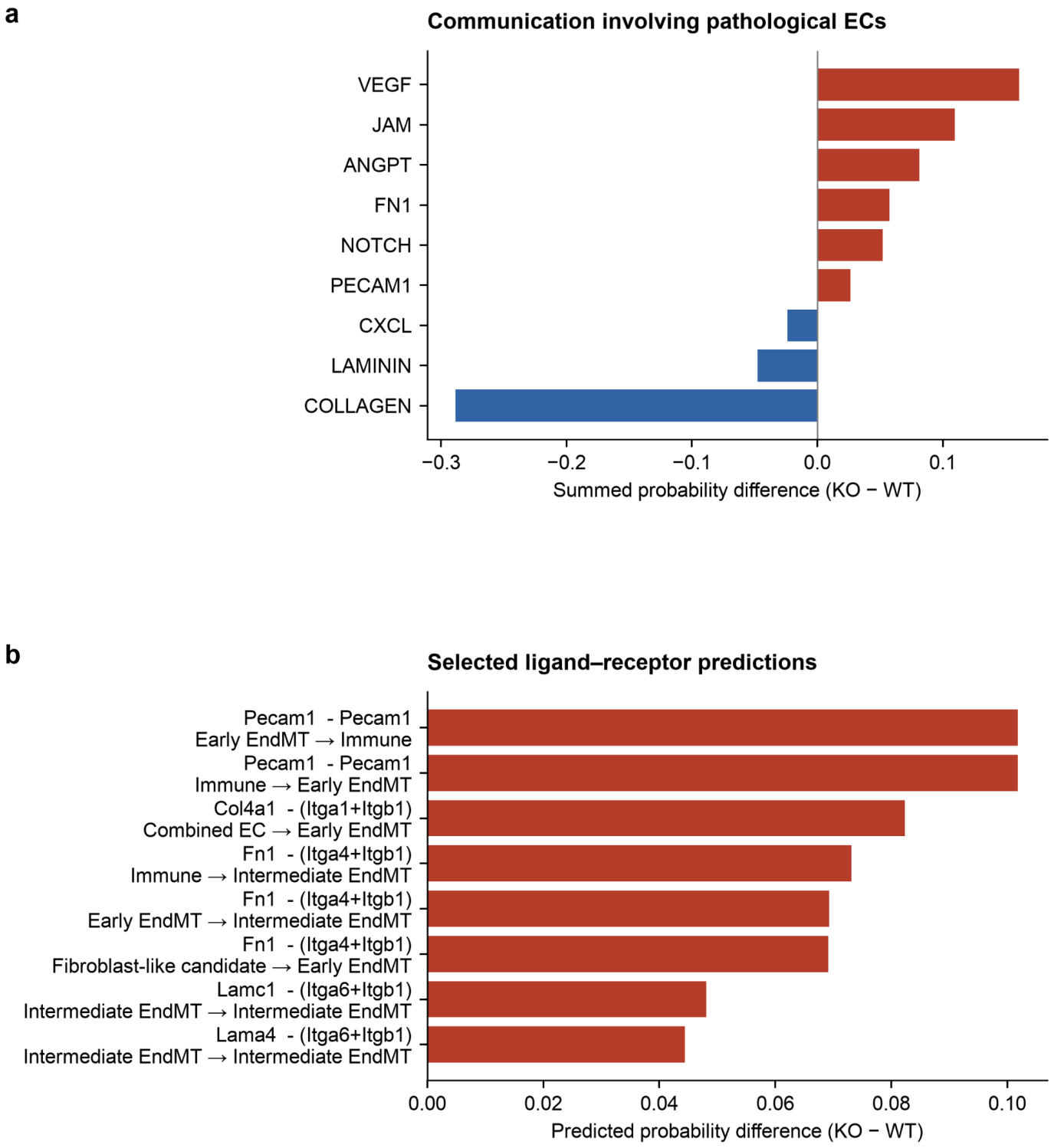


Figure 5

Gene expression across a shared endothelial transcriptional graph

a, The retained Monocle3 principal graph with finite-pseudotime cells on the atlas coordinates. WT and knockout were fitted jointly with a common WT-derived root set. b–g, Serpine1, Pgf, Nfkbia, Cldn5, Fn1 and Mfsd2a expression across 30 approximately equal-cell bins of the same graph. Expression was

averaged within mouse and bin, then by genotype; bins required at least 10 cells within a mouse. Lines connect bin means without additional smoothing. The curves combine graph branches and do not show a single selected disease path, chronological progression or endothelial lineage conversion. Shading shows the observed range across mice where both mice contribute; it is not a confidence interval. Supporting data: Supplementary Fig. S5; Supplementary Table S5

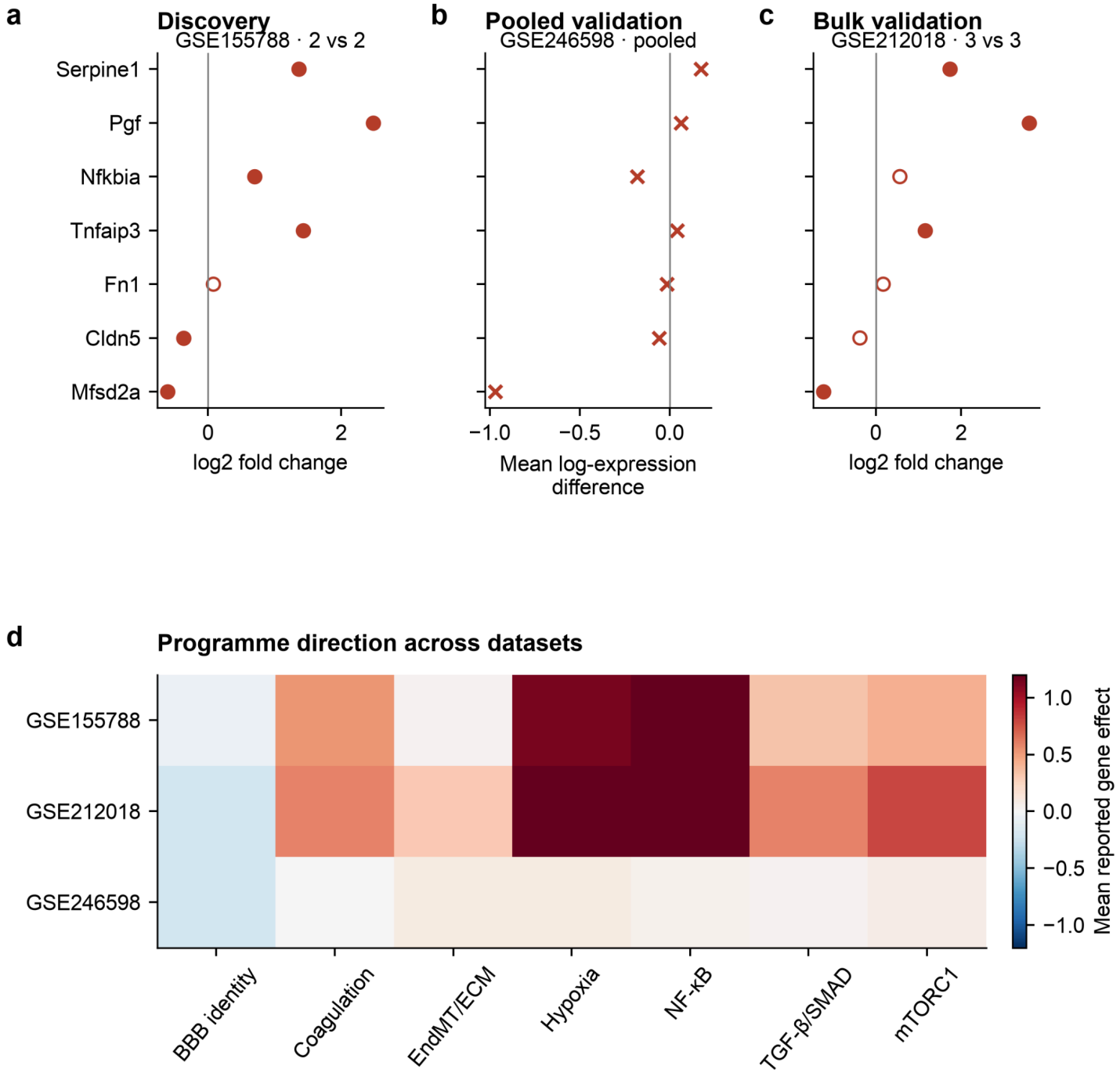


Figure 6

Independent mouse datasets assess candidate response directions

a–c, Candidate-gene effect estimates in GSE155788, GSE246598 and GSE212018. Filled circles indicate FDR<0.05 and open circles indicate nonsignificant replicate-level estimates. Crosses denote descriptive pooled-library estimates without replicate-level inference. The GSE246598 axis displays mean log-normalized-expression differences, separately from the replicate-level $\log_2$ fold changes. d, Mean reported gene effects for predefined programmes, shown to compare direction rather than effect magnitude or formal enrichment significance. Module coverage and complete outputs are provided in Supplementary Table S7

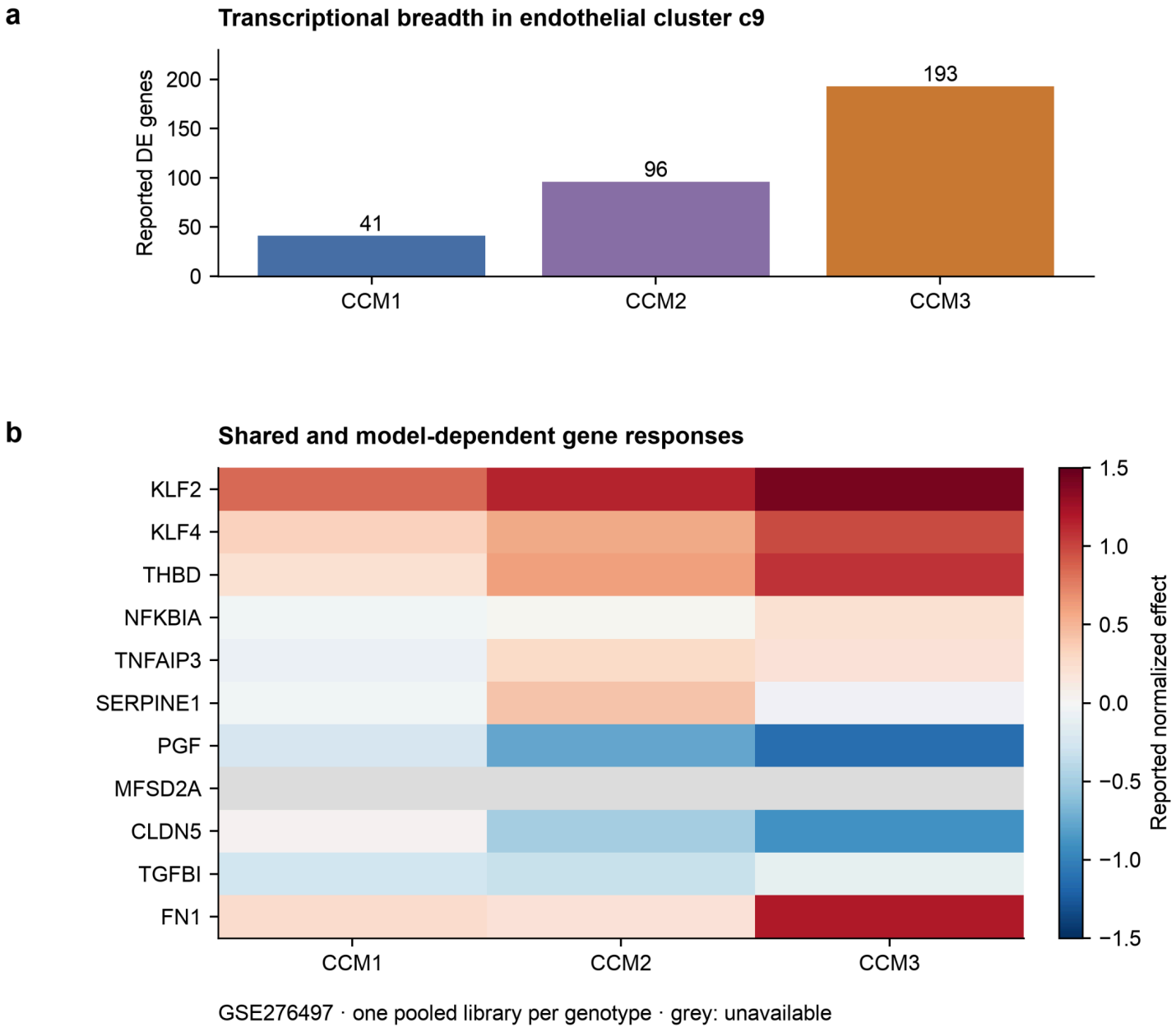


Figure 7

Exploratory cross-genotype comparison in human vascular organoids

a, Reported numbers of differentially expressed genes in organoid endothelial cluster c9 after CCM1, CCM2 or CCM3 loss, using the supplied common threshold. b, Selected KO-versus-WT normalized gene effects from the supplied comparison; grey indicates unavailable values. One pooled library was available per genotype. Counts and heatmap effects are descriptive and do not establish genotype-level differences in severity. PGF and SERPINE1 do not reproduce the same pattern as the mouse discovery data. Statistical stars from cell-level comparisons are omitted to avoid implying biological-replicate support. Supporting data: Supplementary Table S8

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.pdf
- ESM2.zip
- GraphicalAbstract.tif