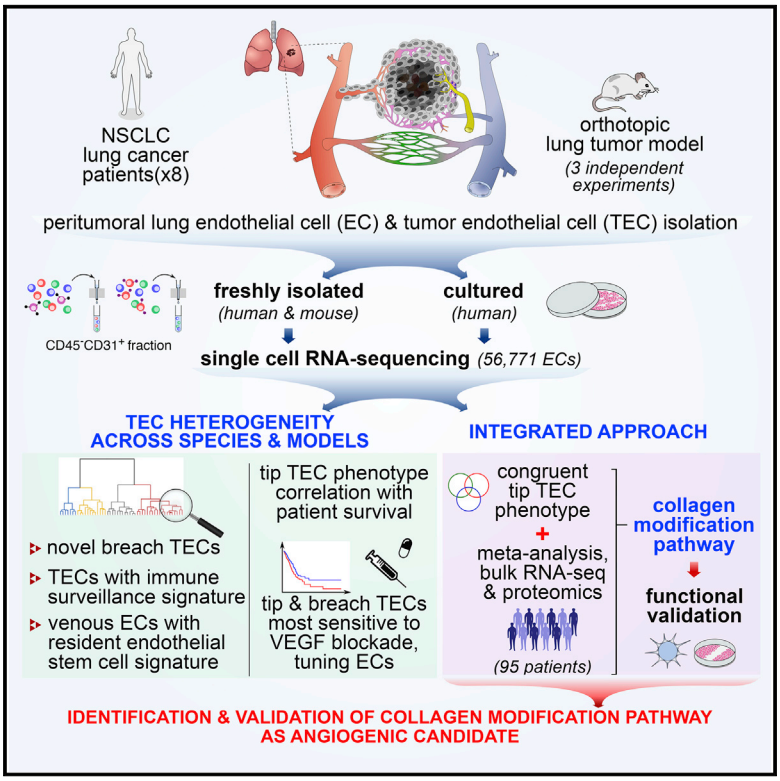
Article

Cancer Cell

An Integrated Gene Expression Landscape Profiling Approach to Identify Lung Tumor Endothelial Cell Heterogeneity and Angiogenic Candidates

Graphical Abstract



Highlights

- We single-cell RNA-sequenced 56,771 endothelial cells (ECs) from human, mouse, and cultured lung tumor models
- Tip ECs were resolved into migratory and basement-membrane remodeling phenotypes
- Capillary and venous ECs expressed immunoregulatory gene signatures
- Integrated analysis identified collagen modification as an angiogenic pathway

Authors

Jermaine Goveia, Katerina Rohlenova, Federico Taverna, ..., Diether Lambrechts, Yonglun Luo, Peter Carmeliet

Correspondence

lixr6@mail.sysu.edu.cn (X.L.), alun@biomed.au.dk (Y.L.), peter.carmeliet@kuleuven.vib.be (P.C.)

In Brief

Goveia et al. use single-cell RNA sequencing to provide an inventory of tumor endothelial cell (TEC) phenotypes from human and mouse non-small cell lung cancer and validate them functionally. Specific TEC phenotypes are associated with prognosis and response to anti-angiogenic therapy.

Data Resources

PXD014123
GSE77199

Goveia et al., 2020, Cancer Cell *37*, 21–36
January 13, 2020
https://doi.org/10.1016/j.ccell.2019.12.001

CellPress

Cancer Cell
Article

An Integrated Gene Expression Landscape Profiling Approach to Identify Lung Tumor Endothelial Cell Heterogeneity and Angiogenic Candidates

Jermaine Goveia,[1,29] Katerina Rohlenova,[1,29] Federico Taverna,[1,29] Lucas Treps,[1,29] Lena-Christin Conradi,[1,24] Andreas Pircher,[1,25] Vincent Geldhof,[1] Laura P.M.H. de Rooij,[1] Joanna Kalucka,[1,26] Liliana Sokol,[1] Melissa García-Caballero,[1] Yingfeng Zheng,[2] Junbin Qian,[3] Laure-Anne Teuwen,[1] Shawez Khan,[1] Bram Boeckx,[3] Els Wauters,[4] Herbert Decaluwé,[4,5] Paul De Leyn,[4,5] Johan Vansteenkiste,[4] Birgit Weynand,[6] Xavier Sagaert,[6] Erik Verbeken,[6] Albert Wolthuis,[7] Baki Topal,[7] Wouter Everaert,[8,9] Hanibal Bohnenberger,[10] Alexander Emmert,[11]

(Author list continued on next page)

[1]Laboratory of Angiogenesis and Vascular Metabolism, Department of Oncology and Leuven Cancer Institute (LKI), KU Leuven, VIB Center for Cancer Biology, Leuven 3000, Belgium
[2]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-Sen University, Guangzhou 510060, Guangdong, China
[3]Laboratory of Translational Genetics, Department of Oncology and Leuven Cancer Institute (LKI), KU Leuven, VIB Center for Cancer Biology, Leuven 3000, Belgium
[4]Respiratory Oncology Unit (Respiratory Medicine) and Leuven Lung Cancer Group, University Hospitals Leuven, Leuven 3000, Belgium
[5]Department of Thoracic Surgery, University Hospitals Leuven, Leuven 3000, Belgium
[6]Translational Cell & Tissue Research, Department of Imaging & Pathology, KU Leuven, Leuven 3000, Belgium
[7]Department of Abdominal Surgery, University Hospitals Leuven, Leuven 3000, Belgium
[8]Laboratory for Experimental Urology, Department of Development and Regeneration, KU Leuven, Leuven 3000, Belgium
[9]Department of Urology, University Hospitals Leuven, Leuven 3000, Belgium

(Affiliations continued on next page)

SUMMARY

Heterogeneity of lung tumor endothelial cell (TEC) phenotypes across patients, species (human/mouse), and models (*in vivo*/*in vitro*) remains poorly inventoried at the single-cell level. We single-cell RNA (scRNA)-sequenced 56,771 endothelial cells from human/mouse (peri)-tumoral lung and cultured human lung TECs, and detected 17 known and 16 previously unrecognized phenotypes, including TECs putatively regulating immune surveillance. We resolved the canonical tip TECs into a known migratory tip and a putative basement-membrane remodeling *breach* phenotype. Tip TEC signatures correlated with patient survival, and tip/breach TECs were most sensitive to vascular endothelial growth factor blockade. Only tip TECs were congruent across species/models and shared conserved markers. Integrated analysis of the scRNA-sequenced data with orthogonal multi-omics and meta-analysis data across different human tumors, validated by functional analysis, identified collagen modification as a candidate angiogenic pathway.

INTRODUCTION

During vessel sprouting in a retinal vascularization model (Blanco and Gerhardt, 2013), endothelial cells (ECs) exhibit heterogeneity: a navigating tip EC leads, while proliferating stalk ECs elongate the sprout. Tip and stalk ECs are not genetically predetermined fixed states, but dynamically interchangeable phenotypes (Blanco and Gerhardt, 2013). In tumors and eye diseases,

Significance

Therapeutic targeting of endothelial cells (ECs) to block tumor angiogenesis is clinically approved but suffers resistance. Here we surveyed EC phenotypes in human and murine lung cancer using single-cell RNA sequencing in combination with orthogonal bulk multi-omics approaches to provide a molecular atlas of tumor EC (TEC) phenotypes. Classical angiogenic tip and proliferating ECs comprised only a minority of TECs, and we identified distinct subpopulations expressing gene signatures related to basement-membrane breaching, immune cell recruitment, and semi-professional antigen presentation. Integrated analysis revealed previously overlooked tip cell markers, including transcription factors, matricellular proteins, and collagen cross-linking enzymes. Functional analysis confirmed novel angiogenic candidates and highlights the potential of our resource to revisit anti-angiogenic strategies.

Dena Panovska,[12] Frederik De Smet,[12] Frank J.T. Staal,[13] Rene J. Mclaughlin,[13] Francis Impens,[14,15] Vincenzo Lagani,[16,17] Stefan Vinckier,[1] Massimiliano Mazzone,[18] Luc Schoonjans,[1] Mieke Dewerchin,[1] Guy Eelen,[1] Tobias K. Karakach,[1,27,28] Huanming Yang,[19,20] Jian Wang,[19,20] Lars Bolund,[21,22] Lin Lin,[21,22] Bernard Thienpont,[23] Xuri Li,[2,30,*] Diether Lambrechts,[3] Yonglun Luo,[19,20,21,22,30,*] and Peter Carmeliet[1,2,30,31,*]

[10]Institute of Pathology, University Medical Center, Göttingen 37075, Germany
[11]Department of Thoracic and Cardiovascular Surgery, University Medical Center, Göttingen 37075, Germany
[12]Laboratory for Precision Cancer Medicine, Translational Cell & Tissue Research, Department of Imaging & Pathology, KU Leuven, Leuven 3000, Belgium
[13]Department of Immunology and Blood Transfusion, Leiden University Medical Center, Leiden 2300 RC, the Netherlands
[14]VIB Proteomics Core and VIB Center for Medical Biotechnology, Ghent 9000, Belgium
[15]Department of Biomolecular Medicine, Ghent University, Ghent 9000, Belgium
[16]Institute of Chemical Biology, Ilia State University, Tbilisi 0162, Georgia
[17]Gnosis Data Analysis PC, Heraklion GR-700 13, Greece
[18]Laboratory of Tumor Inflammation and Angiogenesis, Department of Oncology and Leuven Cancer Institute (LKI), KU Leuven, VIB Center for Cancer Biology, Leuven 3000, Belgium
[19]BGI-Shenzhen, Shenzhen 518083, China
[20]China National GeneBank, BGI-Shenzhen, Shenzhen 518120, China
[21]Department of Biomedicine, Aarhus University, Aarhus 8000, Denmark
[22]Lars Bolund Institute of Regenerative Medicine, BGI-Qingdao, Qingdao 266555, China
[23]Laboratory for Functional Epigenetics, Department of Human Genetics, KU Leuven, Leuven 3000, Belgium
[24]Present address: Clinic of General, Visceral and Pediatric Surgery, University Medical Center Göttingen, Göttingen 37075, Germany
[25]Present address: Department of Hematology and Oncology, Internal Medicine V, Medical University Innsbruck, Innsbruck 6020, Austria
[26]Present address: Aarhus Institute of Advanced Studies & Department of Biomedicine, Aarhus 8000, Denmark
[27]Present address: Bioinformatics Core Laboratory, Children's Hospital Research Institute of Manitoba, Winnipeg, MB, Canada
[28]Present address: Rady Faculty of Health Sciences, Department of Pediatrics and Child Health, Winnipeg, MB, Canada
[29]These authors contributed equally
[30]Senior author
[31]Lead Contact
*Correspondence: lixr6@mail.sysu.edu.cn (X.L.), alun@biomed.au.dk (Y.L.), peter.carmeliet@kuleuven.vib.be (P.C.)
https://doi.org/10.1016/j.ccell.2019.12.001

the vasculature is structurally disorganized and functionally abnormal (Carmeliet and Jain, 2011a), but the transcriptome heterogeneity of tumor ECs (TECs) at the single-cell level across patients, species (human versus mouse), and models (freshly isolated versus cultured), analyzed in a single study, has not been inventoried.

In the tumor angiogenesis field, some single-cell RNA (scRNA) sequencing (scRNA-seq) studies reported a descriptive list of previously known TEC phenotypes (Lambrechts et al., 2018; Zhao et al., 2018), but did not identify previously unknown functionally validated angiogenic candidates. It can even be questioned if a single scRNA-seq study of one tumor type is sufficiently powerful at all to prioritize angiogenic candidates. Indeed, tip-like TECs in 2 human colon carcinoma xenograft models in a single study already expressed different markers (Zhao et al., 2018). Here, we aimed to overcome species- and model-dependent variability and used an integrated approach, comprising, in addition to scRNA-seq, validation by orthogonal multi-omics approaches to identify conserved phenotypes and markers across patients, tumor types, species, and models, which are more likely to be of functional relevance.

RESULTS

Single-Cell Atlas of EC Phenotypes in Human Lung Cancer

Focusing on human non-small cell lung cancer (NSCLC), we profiled freshly isolated human tumor ECs (hTECs) and (paired from the same patient) human (peritumoral) non-tumor pulmonary ECs (hPNECs) from 1 large cell carcinoma, 4 squamous cell carcinomas, and 3 adenocarcinoma treatment-naive patients (Table S1). Single-cell suspensions, magnetic-activated cell sorting (MACS)-depleted for $CD45^+$ leukocytes and enriched for $CD31^+$ ECs (Figure 1A), were subjected to scRNA-seq using a 10× genomics-based single-tube protocol. After quality filtering for the number of detected genes and mitochondrial read counts, duplicates were assessed and unique transcripts were normalized for total read depth (STAR Methods; Table S2).

ECs were *in silico* selected (Figures S1A and S1B), and up to 12,323 hTEC and 8,929 hPNEC transcriptomes from 8 patients were pooled, batch-corrected, clustered, and visualized using *t*-distributed stochastic neighbor embedding (t-SNE) plots (Figures 1B, 1C, and S1C; Table S3). To ensure that batch correction did not remove relevant biological features, we also analyzed each sample separately. Clusters detected in batch-corrected data were largely similar to those obtained without batch correction (not shown); any relevant differences are indicated below. To assess cluster reproducibility (Tasic et al., 2018), we performed hierarchical clustering and bootstrap analysis (Figure S1D). For biologically relevant subclusters that were not resolved by bootstrapping (e.g., lymphatic ECs from tumor and peritumoral tissue), we ascertained that they were statistically separable using pairwise differential analysis (Innes and Bader, 2018) (Figure S1D, Table S4). Each of the 13 phenotypes is numbered as in Figure 1C (H1, H2, etc).

Clusters were biologically annotated based on the relative abundance of top-ranking marker genes in hPNECs and hTECs, and differences in marker gene expression levels shown in display items were statistically quantified to support the qualitative biological annotation (Figures 1D–1G and S1E; Tables S4

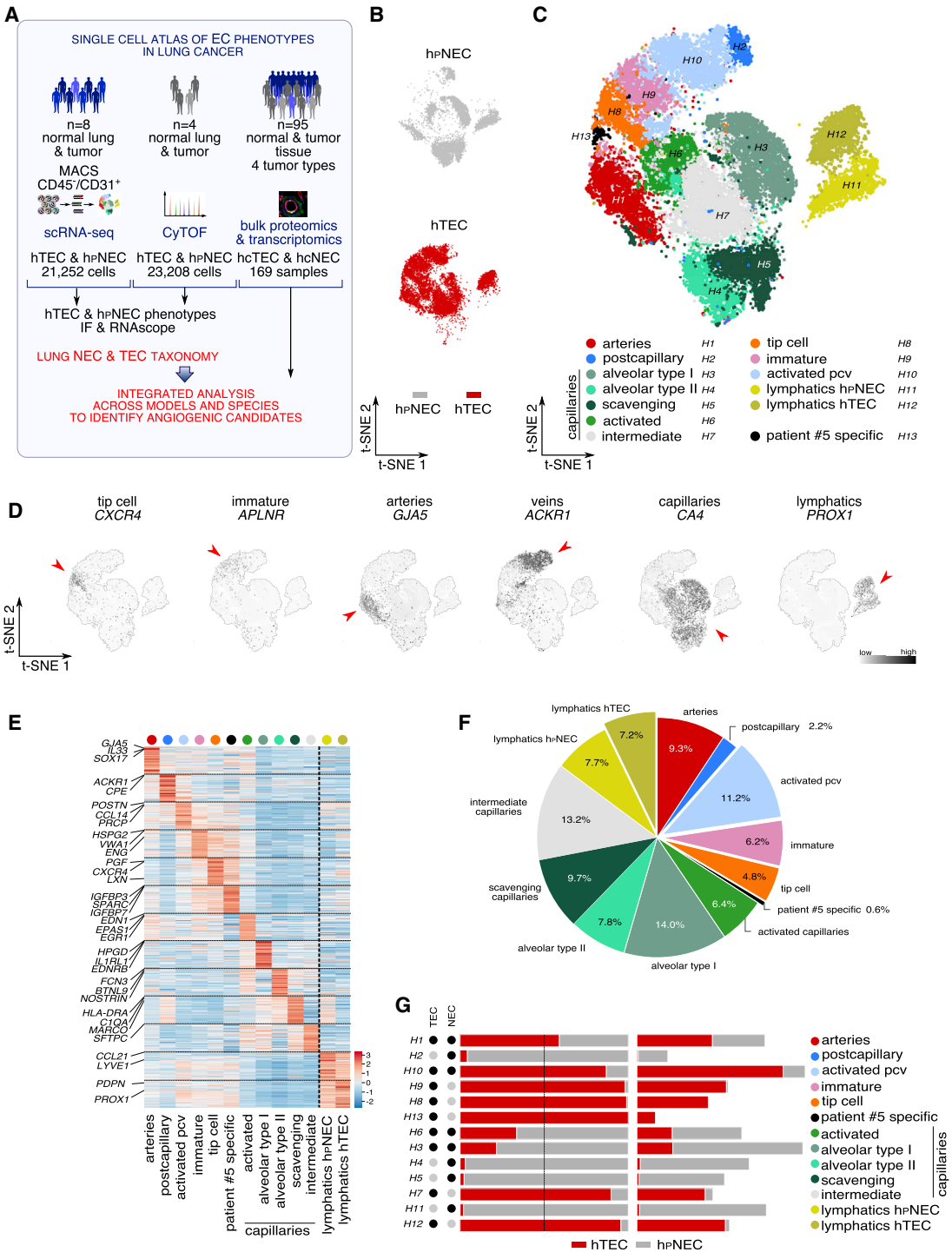


Figure 1. Construction of hTEC and hPNEC Taxonomy

(A) Study design. IF, immunofluorescence; NEC, normal EC; TEC, tumor EC.

(B) t-SNE plot color-coded for ECs from peritumoral non-malignant lung (hPNEC; gray) and tumor tissue (hTEC; red).

(C) t-SNE plot of hTEC and hPNEC transcriptomes, color- and number-coded for the 13 phenotypes identified by graph-based clustering.

(D) t-SNE plots, color-coded for expression of indicated marker genes (red arrowheads).

(E) Gene expression levels of top-ranking marker genes in different EC phenotypes. In this and all further heatmaps depicting marker genes, colors represent row-wise scaled gene expression with a mean of 0 and an SD of 1 (*Z* scores).

(F) Relative abundance of each phenotype in hTECs and hPNECs.

(legend continued on next page)

and S5). To obtain a more robust interpretation, we focused on clusters defined by gene set signatures rather than single genes (Table S4), and based cluster annotation on the expression of canonical gene signatures (e.g., arterial, capillary, venous, lymphatic, tip, and immature stalk phenotypes; listed in Table S5). To improve the generalizability of our analysis, we discuss all clusters at least once but focused mainly on clusters that were reproducible in multiple patients (Table S3). A detailed description of previously documented roles of the top-ranking markers, used to infer the putative biological role of each cluster, is provided in Tables S4 and S5.

Arterial ECs (cluster H1) expressed genes involved in vascular integrity, homeostasis, and vasotonus (Figure 2A), while postcapillary venous ECs (H2) expressed genes implicated in leukocyte recruitment, tissue perfusion, and pulmonary blood pressure (Figure 2B). We identified a type I (H3) and type II (H4) alveolar capillary EC phenotype, characterized by differential expression of von Willebrand Factor (*VWF*, upregulated in type II compared with type I; fold change = 3.63, adjusted p value = 2.36×10^{-77}) and endomucin (*EMCN*, upregulated in type I compared with type II; fold change = 2.89, adjusted p value = 1.22×10^{-68}), which might be involved in, respectively, vasoregulation and anti-microbial defense (Figures 2C–2E). Compared with all other clusters, capillary ECs expressed a signature of genes involved in MHC-II-mediated antigen presentation and processing at higher levels than other phenotypes (Figure 2F), but the co-stimulatory molecules *CD80* and *CD86* were not detectable (not shown), suggesting a function as semi-professional antigen presentation cells (APCs). We also identified 2 novel capillary phenotypes that might be induced by tumor-derived cytokines: scavenging capillaries (H5) were by bootstrap analysis similar to type II capillary ECs, but upregulated scavenging receptors and genes associated with macrophages and antigen processing (Figures 2G, 2H, and S1D), while activated capillaries (H6) expressed EC activation markers (Figure 2I). An intermediate capillary EC phenotype (H7) resembled activated capillaries, but consisted mostly of cells derived from a single patient (Figure S1D; Table S3).

Traditional angiogenic phenotypes, such as tip and proliferating ECs (presumed targets of anti-angiogenic therapy [AAT]), were detectable only in hTECs (Figures 3A, S1E, and S2A; Table S3). Tip cells (H8), expressing gene signatures associated with EC migration, matrix remodeling, and vascular endothelial growth factor (VEGF) signaling (Figure 3B), comprised only a minority (<10%) of hTECs (Figure S1E; Table S3). Notably, tip hTECs showed highly restricted expression of the disease-specific molecule *PGF* (placental growth factor; PlGF [adjusted p value = 2.41×10^{-14}]) (Table S4). We also identified an immature hTEC phenotype (H9), which was similar to tip cells, but upregulated genes involved in the maturation of newly formed vessels, vessel barrier integrity, and Notch-signaling, possibly resembling stalk-like ECs (Figures 3C and S1D). Proliferating hTECs were only detected in a subset of patients, representing <1% of hTECs, and were no longer detected as a separate cluster after batch correction (Figure S2A). In hTECs, an activated postcapillary vein phenotype (H10) upregulated immunomodulatory factors and ribosomal proteins (Figure 3D), features of high endothelial venules (HEVs) in inflamed tissues (Girard and Springer, 1995). In contrast to the extensive phenotypic heterogeneity of blood vascular ECs, the gene expression signatures of tumor and peritumoral lymphatic ECs (LEC; H11, H12) were highly similar, and LEC subpopulations were not detected (Figures 3E and S1D; Table S4). An ambiguous phenotype, expressing markers of tip and arterial ECs (Figures 1C and S1D), predominantly consisted of cells from a single patient (patient 5, H13) and is not discussed.

To explore the generalizability of our NSCLC EC taxonomy, we used top-ranking marker genes to train a machine-learning algorithm (Kiselev et al., 2018) to automatically annotate 574 hPNECs and 638 hTECs, *in silico* selected from a 52,698-cell catalog from tumors of 5 treatment-naive NSCLC patients (Lambrechts et al., 2018). The majority of ECs were annotated with high confidence (similarity threshold >0.5) (Figure 3F), indicating that our taxonomy is an externally valid resource that comprises all major EC phenotypes detectable in 13 lung cancer patients.

We validated the taxonomy using orthogonal techniques. Immunostaining combined with quantitative RNAscope to count transcript numbers confirmed that *PGF* levels were upregulated in CD31^{+}/*CXCR4*^{high} tip hTECs, while triple immunostaining showed co-localization of CXCR4 and PlGF in TECs (Figures 3G and 3H; S2B). We also confirmed by immunostaining of (1) SELP the signature of activated postcapillary vein ECs (Figures S2C and S2D); (2) VWF and EFNB2, respectively, the venous and arterial phenotype (Figures S2E–S2H); and (3) the HEV-specific MECA-79 antigen the signature of the HEV phenotype (Figure S2I). RNAscope, combined with staining for the EC marker CD31, revealed that arterial (*EFNB2*) and venous (*ACKR1*) marker transcripts did not colocalize in the same hTECs (Figures S3A and S3B), while activated postcapillary vein hTECs expressed the marker CCL14 (Figure S3C). Further, transcripts of the scavenging capillary markers CD52 and CD68, and of the type I capillary markers EDNRB and IL1RL1 co-localized in the same hTECs (Figures S3D and S3E).

We used time-of-flight mass cytometry (CyTOF) to quantify protein levels of marker genes in single cancer, EC, and stromal cells, freshly isolated from NSCLC samples, using metal-conjugated antibodies against 26 preselected markers of these cells. Unbiased clustering and t-SNE visualization of the CyTOF single-cell data revealed separate clusters of ECs, stromal, and cancer cells (Figures 3I and S4A–S4D). Consistent with scRNA-seq, we detected type I (VWF^{low}) and type II (VWF^{high}) alveolar capillary ECs, postcapillary vein ECs (ACKR1^{high}/VWF^{high}), and lymphatic hPNEC and hTEC phenotypes (PROX1^{high}) (Figure S4E). A phenotype expressing the highest CXCR4 levels (tip cell marker)

(G) Relative contribution of hTECs (red) and hPNEC (gray) to each phenotype. Left: To correct for differences in absolute numbers of hTECs versus hPNECs, numbers in each phenotype were divided by the total number of hTECs and hPNECs, respectively. Black vertical line: equal relative contribution (50%). Right: contribution of hTECs and hPNECs in absolute numbers. Note the underrepresentation of capillary hTECs. Dots on the left indicate the relative absence (gray) or presence (black) of the phenotype in hTECs and hPNECs, where presence is defined as a relative contribution of >1% to the total pool of hTECs and hPNECs, respectively.

See also Figure S1 and Tables S1, S2, S3, S4, and S5.

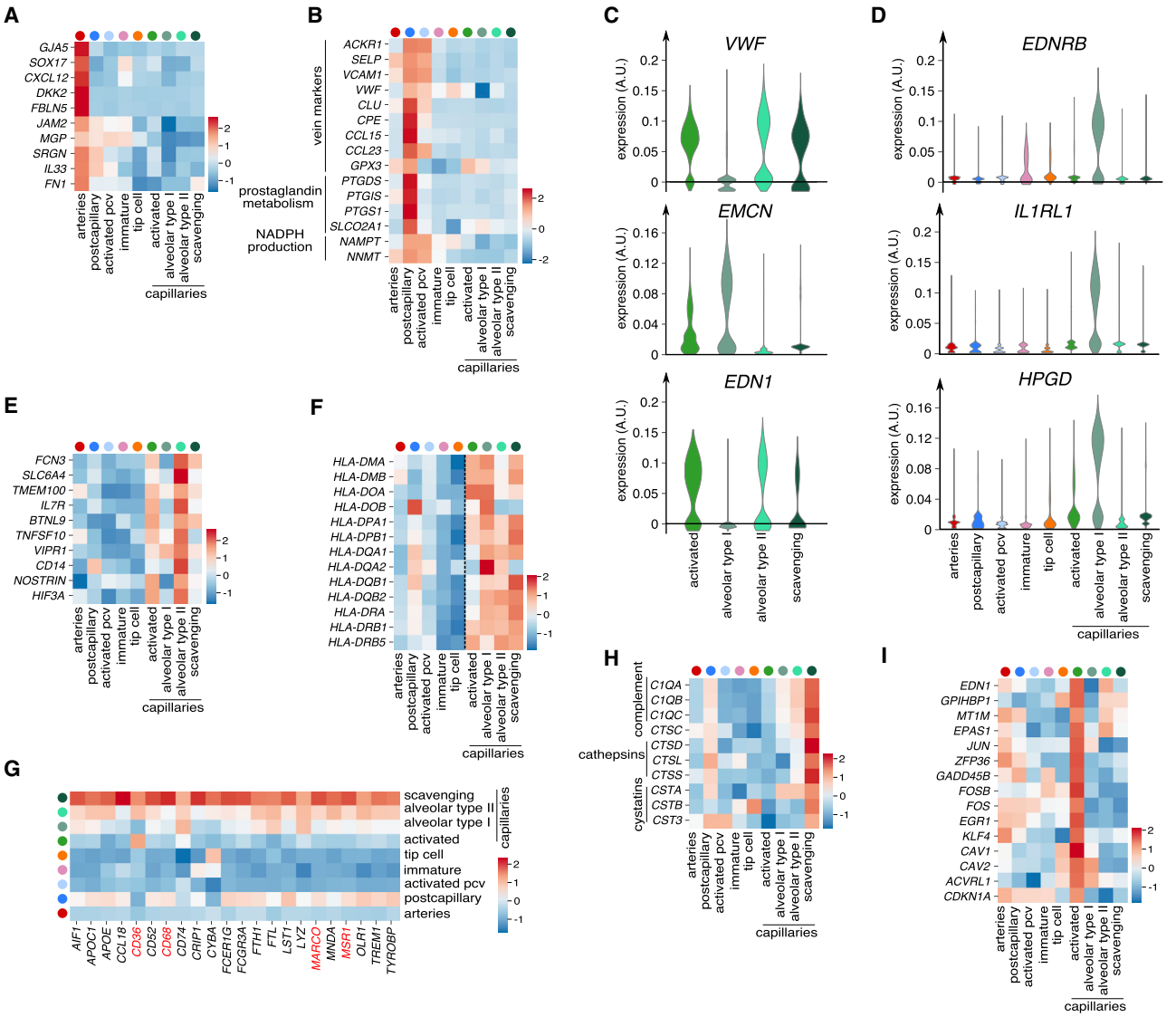


Figure 2. Quiescent and Activated Normal Pulmonary EC Phenotypes

(A and B) Expression levels of (A) arterial and (B) postcapillary venous EC markers.

(C and D) Expression levels of the indicated genes of (C) capillary EC phenotypes and (D) the type I alveolar capillary EC phenotype. A.U., arbitrary units.

(E–I) Expression levels of markers of (E) type II alveolar capillary ECs, (F) genes involved in MHC-II antigen presentation, (G) scavenging EC markers (scavenging receptors in red), (H) defense marker genes (complement, cathepsins and cystatins), and (I) genes associated with activated alveolar capillary ECs.

See also Table S4.

but downregulating the expression of capillary markers (CD36, CA4, HLA-II proteins) likely represented angiogenic hTECs. We also detected a poorly resolved cluster, putative activated capillaries that, compared with type I and type II capillaries, expressed reduced capillary markers (CA4, CD36, HLA-II), but increased levels of the vein TEC marker VWF and the tip hTEC marker CXCR4 (Figures 3J and S4E). CyTOF confirmed the increased levels of HLA and CD36 in capillary ECs and their downregulation in hTECs (Figures 3J and S4F). To validate the observation that the populations identified using scRNA-seq and CyTOF coincided, we performed hierarchical clustering using key surface markers (STAR Methods) (Figure 3K). LECs and capillary ECs clustered together, while vein ECs clustered with angiogenic ECs. Notably, multiscale bootstrap analysis revealed that scRNA-seq resolved type I and type II capillary ECs as most similar to their CyTOF counterparts, cross-validating the presence of 2 distinct capillary phenotypes at the mRNA and protein levels.

Possible Clinical Implications

Despite all patients having early-stage treatment-naive disease, we observed inter-patient heterogeneity in the relative fraction of the different EC phenotypes (Figure 4A; Table S3). In hPNECs, arterial, capillary, venous, and lymphatic ECs were detected at variable proportions, especially for the capillary hPNEC phenotypes. In hTECs, the number of capillary ECs was reduced (Table S3), a finding confirmed by immunohistochemistry (Figure 4B),

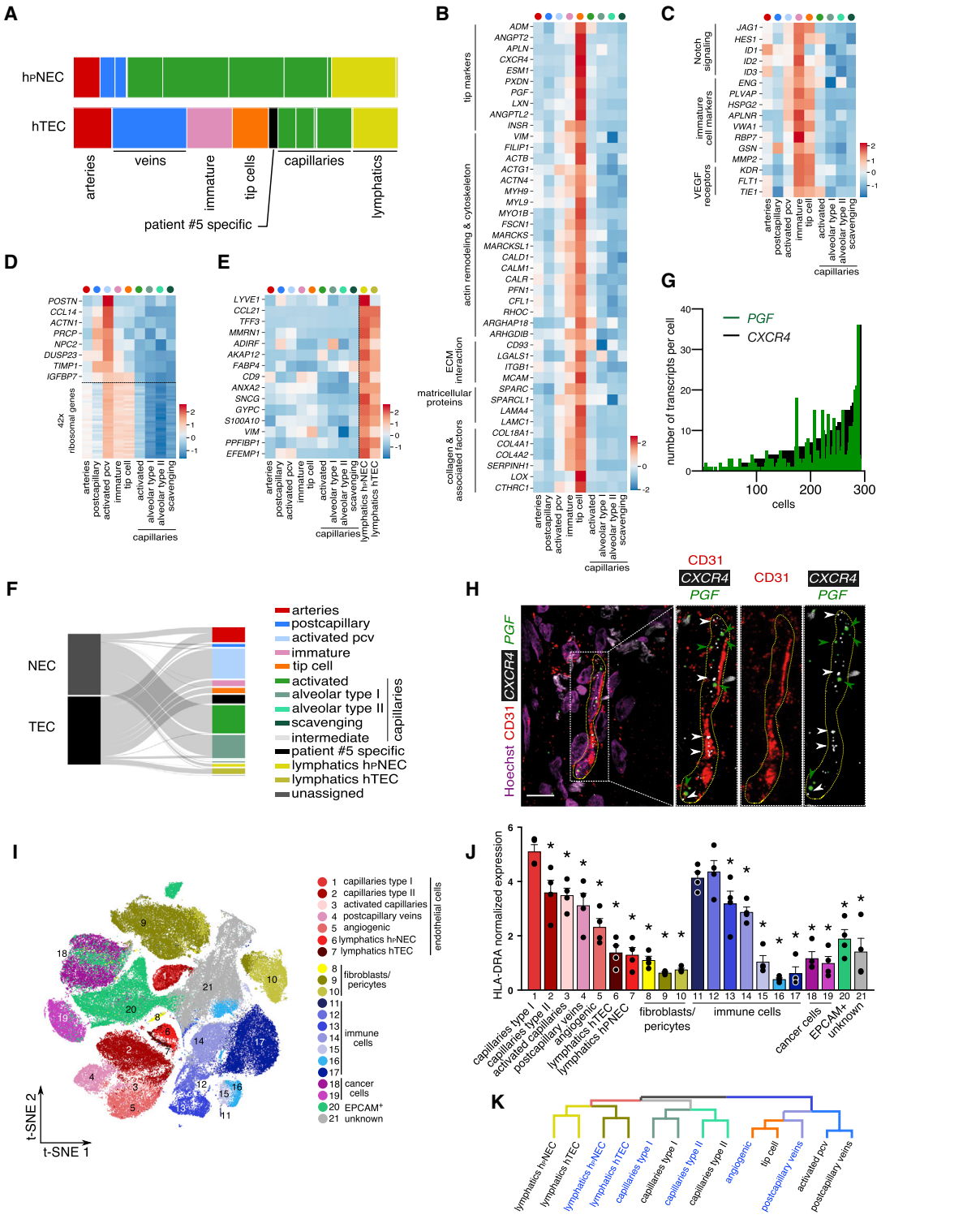


Figure 3. Pathological Tumor EC Phenotypes and Validation by CyTOF

(A) Relative abundance of EC phenotypes in hTECs and hPNECs, weighted by the number of cells per patient.

(B–E) Expression levels of marker genes of (B) tip, (C) immature, (D) activated postcapillary vein, and (E) lymphatic ECs.

(F) Assignment of hTEC and hPNEC phenotypes (identified in the taxonomy) to hTECs and hPNECs freshly isolated from an independent cohort of 5 NSCLC patients (publicly available data; Lambrechts et al., 2018).

(G) Number of *CXCR4* and *PGF* mRNA transcripts per cell in NSCLC tumor vessels as determined by RNAscope (n = 3 patients). *CXCR4* and *PGF* positively correlate (Spearman $r = 0.584$; two-sided p value <0.0001).

(legend continued on next page)

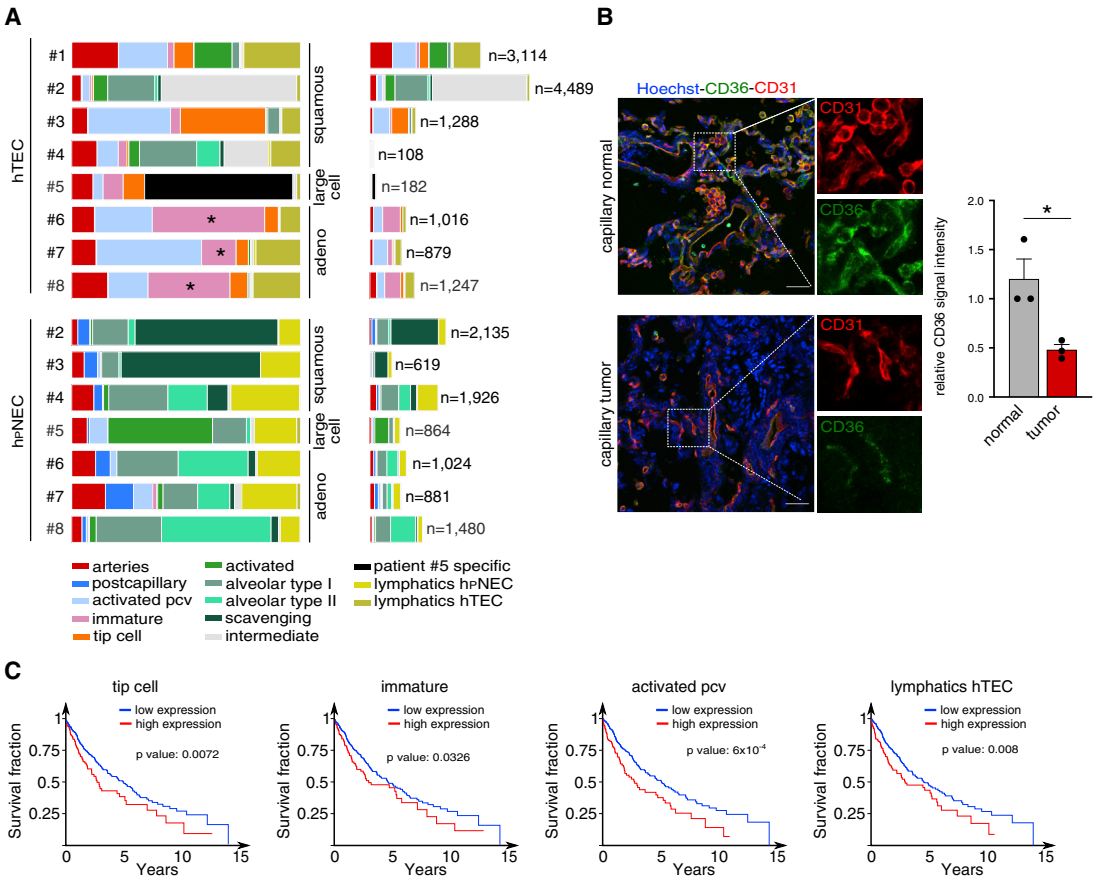


Figure 4. Translational and Possible Therapeutic Implications of the NSCLC EC Taxonomy

(A) EC phenotype composition in tumor and normal tissue from individual patients. Left: Relative contribution of each phenotype scaled to 100%. Right: Contribution of each phenotype in absolute numbers; total number of analyzed ECs on the right. Asterisks indicate the significantly increased fraction of immature TECs in adenocarcinoma compared with squamous cell carcinoma patients (n = 4 squamous and n = 3 adenocarcinoma patients, $p < 0.05$ by two-tailed unpaired t test).

(B) Representative micrograph of human peritumoral tissue (upper) and NSCLC tumor (lower) section immunostained for CD31 and CD36. Nuclei are stained with Hoechst. Images to the right are magnifications of the respective boxed areas. Scale bar, 50 μm. Right: Quantification of the CD36 signal intensity (mean ± SEM; n = 3 patients, $p < 0.05$ by two-tailed unpaired t test).

(C) Overall survival of 502 lung squamous cell carcinoma patients selected from the TCGA dataset and stratified by the indicated gene set expression signature scores.

See also Table S3.

although not to the same extent in each patient. Statistically, immature hTECs were more abundant in adenocarcinoma than squamous cell carcinoma patients (Table S3), although larger cohorts are required for meaningful interpretation (see below).

To correlate expression of EC phenotype-specific gene signatures with NSCLC patient survival, we took advantage of a large bulk RNA-seq dataset of 1,024 NSCLC patients of the Cancer Genome Atlas (TCGA). After identifying EC-specific marker gene sets for each phenotype utilizing a publicly available resource (Lambrechts et al., 2018) (STAR Methods), we used gene set variation analysis (GSVA) to score the enrichment of these signatures in each of the 1,024 NSCLC patients. Notably,

(H) RNAscope images of an NSCLC tumor vessel probed for *CXCR4* (white) and *PGF* (green) mRNA. ECs are immunostained for CD31 (red). Nuclei are stained with Hoechst (magenta). Dotted lines denote (part of) the vessel; boxed area is magnified on the right. Green and white arrowheads indicate *PGF* and *CXCR4* mRNA transcripts, respectively. Scale bar, 20 μm.

(I) t-SNE plot of CyTOF data, color-coded for stromal and cancer cell phenotypes. For visualization, all non-EC types have been randomly down-sampled to no more than the total number of ECs.

(J) Quantification of HLA-II expression by CyTOF in stromal cell subpopulations. Asterisks indicate clusters in which HLA-II expression is significantly down-regulated compared with type I capillary ECs (mean ± SEM; n = 4, $p < 0.05$ by one-way ANOVA with Dunnett's method).

(K) Dendrogram visualization of hierarchical clustering analysis of EC phenotypes identified by scRNA-seq and CyTOF analysis. Color differences in dendrogram indicate clusters that were resolved by multiscale bootstrap analysis. Clusters identified by CyTOF are indicated in blue.

See also Figures S2–S4 and Tables S3 and S4.

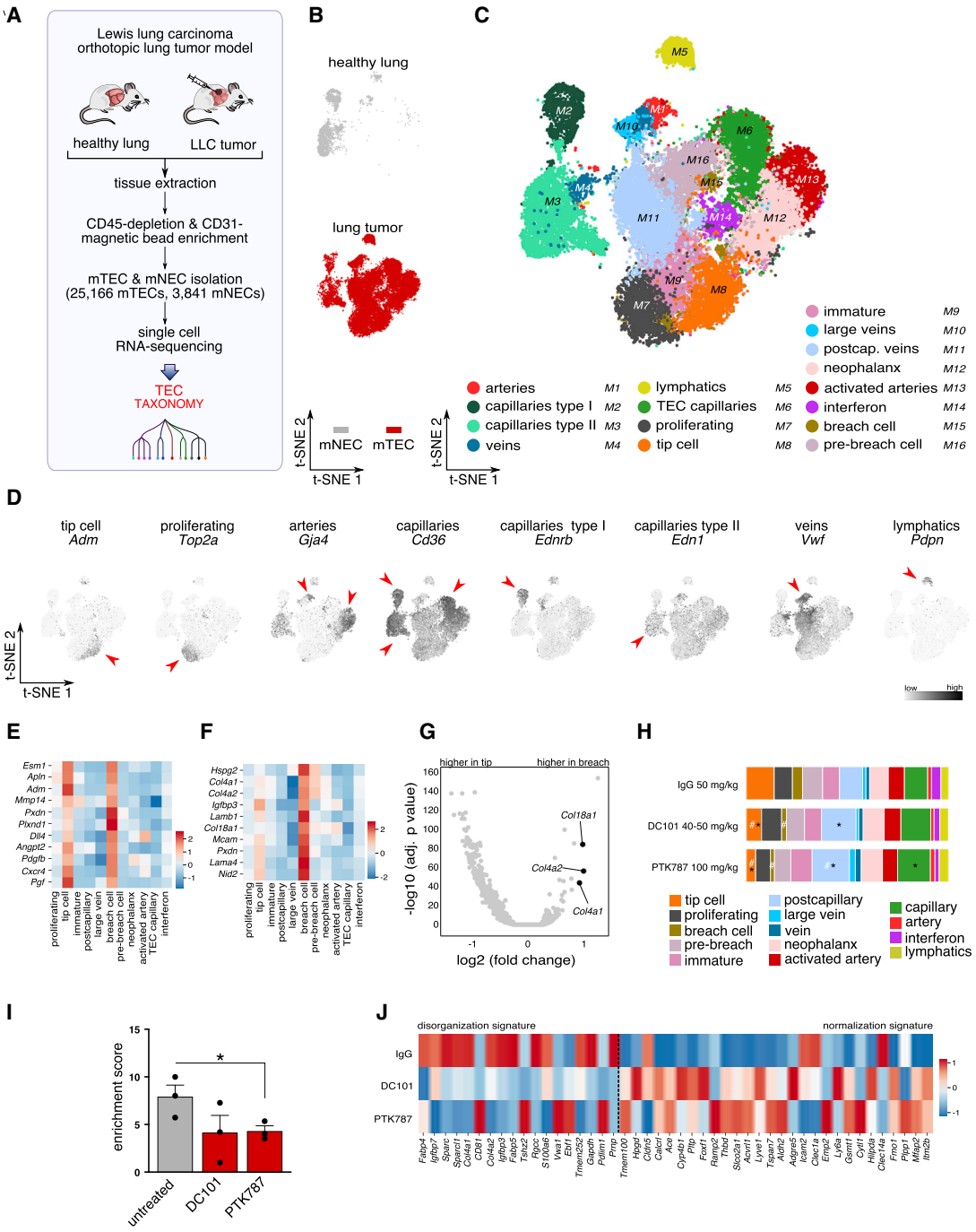


Figure 5. Construction of mTEC and mNEC Taxonomy

(A) Experimental design.

(B) t-SNE plot, color-coded for mNECs from healthy lung (gray) and mTECs from tumor-bearing mice (red).

(C) t-SNE plot, color-coded for the pulmonary mNEC and mTEC phenotypes.

(D) t-SNE plots of mNECs and mTECs, color-coded for the expression of indicated marker genes (red arrowheads).

(E and F) Expression levels of (E) tip and (F) breach mTEC marker genes.

(G) Differentially expressed genes in breach cells versus tip cells.

(H) Relative composition of mTEC phenotypes upon treatment with control immunoglobulin (Ig)G, DC101, or PTK787. *$p < 0.05$ by Dunnett's method; #$p < 0.05$ by two-tailed unpaired t test.

(I) Expression level of the 18-gene disorganization signature (mean ± SEM; n = 3, *$p < 0.05$ by one-tailed unpaired t test).

(legend continued on next page)

squamous (but not adenocarcinoma [not shown]) NSCLC patients, who expressed high levels of gene set signatures of angiogenic tip, immature, activated postcapillary, or lymphatic TECs, had shorter overall survival, presumably because these signatures reflect active angiogenesis and lymphatic spread (Figure 4C).

Lung Tumor EC Phenotypes in Other Species and Models

To compare lung TEC taxonomies across species and models for identification of common angiogenic TEC phenotypes and targets in an integrated congruency analysis, we used a similar strategy to construct 2 additional lung TEC taxonomies, one of 29,007 murine mNECs and mTECs micro-dissected from Lewis lung carcinoma (LLC) (another species), treated or not with anti-VEGF, and another of 6,512 human cultured hcTECs from human lung tumor (another model) (Tables S2, S3, and S4). We used the LLC model, since it relies on angiogenic vessel sprouting (the target of AAT), unlike other preclinical mouse NSCLC tumor models (Eldridge et al., 2016).

As in human lung, we identified mNEC-enriched phenotypes, expressing typical markers of arterial (M1), alveolar type I and II capillary (M2, M3), venous (M4) and lymphatic (M5) ECs (Figures 5A–5D and S5A–S5G; Tables S4 and S6). Type II alveolar capillary ECs expressed the highest levels of genes involved in MHC-II antigen presentation, processing, and loading (Figures S5B and S5H), but did not express the co-stimulatory genes *Cd80* and *Cd86* (not shown). As in NSCLC, the fraction of capillary ECs was underrepresented in tumors (Figure S5A; Table S3), and mTECs downregulated the typical capillary gene signature, including *MHC-II* expression (Figure S5H). mTECs also downregulated *Cd36* (a gene involved in fatty acid uptake), except in one capillary TEC phenotype (adjusted p value = 7.81×10^{-39}; M6) (Figure 5D). Unlike in human lung, the vein EC phenotype in murine lung expressed resident endothelial stem cell markers (*Cd200*, adjusted p value = 4.61×10^{-26}; *Bst1*, adjusted p value = 2.80×10^{-20}), previously identified in large pulmonary vessels (Wakabayashi et al., 2018) (Figure S5I).

Proliferating (M7) and tip (M8) mTECs were also detected (Figures 5C and 5E); proliferating mTECs were more abundant than in human lung tumors, consistent with the faster growth of murine lung tumors and a possible different type of tumor vascularization in NSCLC (Figures 5D and S5E). An immature mTEC phenotype (M9), lacking a strong transcriptome signature, was also recognized (Figure 5C). Consistent with human ECs, we observed high venous plasticity and detected TEC-enriched phenotypes including large (M10) and postcapillary (M11) vein ECs (Figure 5C). The postcapillary vein mTECs upregulated HEV markers (Figures S5J and S5K), confirmed by immunostaining (Figures S6A and S6B).

The remaining murine phenotypes were predominantly detected in mTECs (Figures 5C and S5F; Table S3). A neo-phalanx mTEC phenotype (M12) expressed capillary and arteriole markers, while activated artery mTECs (M13) upregulated neoarteriogenesis markers. Notably, only mTECs, not mNECs, exhibited a phenotype characterized by the expression of an interferon (IFN) response gene signature and chemokines, involved in immune cell recruitment and angiostasis (M14) (Figure S5C; Tables S3 and S4). We also identified a previously unrecognized EC phenotype, which we termed breach cells (M15), and their putative precursors, pre-breach cells (M16) (Figures 5E–5G). Breach cells upregulated expression of not only tip TEC markers, but also of genes previously involved in VEGF-induced podosome rosette-mediated basement-membrane (BM) and collagen remodeling (Tables S4 and S6) (Seano et al., 2014). Immunostaining confirmed the expression of signature markers of arterial, capillary, venous, and tip mNEC and mTEC phenotypes at the protein level (Figures S6C–S6J).

In cultured hcTECs, the *in vivo* artery, capillary, and vein EC phenotypes were no longer detectable (Figures S7A and S7B; Tables S2, S3, and S4). It was therefore surprising that the typical tip EC phenotype (C1), which is a plastic transient phenotype (Blanco and Gerhardt, 2013), was detectable in hcTECs. As expected for propagating ECs in culture, proliferating hcTECs (C2) were detected. Perhaps related to the culture conditions (presence of transforming growth factor β, an inducer of endothelial-to-fibroblast transition in serum), a phenotype (C3) exhibiting a signature of endothelial-to-mesenchymal transition (*TAGLN, SERPINE1, FN1, CD44*, adjusted p value < 5×10^{-4} for all genes) (Dejana et al., 2017) was identified. We also detected an intermediate transitioning phenotype (C4) that upregulated ribosomal genes, suggesting EC phenotypes on their way to adopting another phenotype (Figure S7B).

Effect of VEGF Blockade on TEC Phenotypes

Differential Sensitivity of TEC Phenotypes

Using the mTEC taxonomy, we explored if specific EC phenotypes were differentially sensitive to anti-VEGF AAT (VEGFR2 antibody [DC101], VEGFR tyrosine kinase inhibitor [PTK787]) (Figure 5H). These compounds were tested at doses that inhibit pathological angiogenesis and reduce EC numbers in tumors (Liu et al., 2005), and inhibited tumor growth (Figure S7C). Notably, control- and AAT-treated ECs were composed of the same clusters (Figure 5H; Table S3), suggesting that AAT did not alter the global transcriptome signatures of the different EC phenotypes. However, quantification of individual EC phenotypes revealed that tip and breach TECs were most sensitive, consistent with human colon carcinoma xenografts showing tip cell sensitivity to aflibercept (Zhao et al., 2018). Postcapillary vein and proliferating TECs were less sensitive to VEGF blockade, while capillary TECs were less sensitive to PTK787 treatment (Figure 5H). Whether this is due to the tumor switching from vessel sprouting to vessel cooption is unknown.

Molecular Signature of Tumor Vessel Disorganization and Effect of VEGF Blockade

Tumor vessels are structurally disorganized and functionally abnormal (Carmeliet and Jain, 2011b; Jain, 2005, 2014), but a detailed unbiased molecular footprint is lacking. Traditionally,

(J) Expression levels of the genes in the disorganization signature (18 genes on the left of the black dotted line in the control [IgG] condition [top row]) and more normal signature (28 genes on the right of the black dotted line in the control [IgG] condition [top row]), and effects of DC101 and PTK787 treatments on these signatures (middle and bottom rows).

See also Figures S5–S7 and Tables S2, S3, S4, and S6.

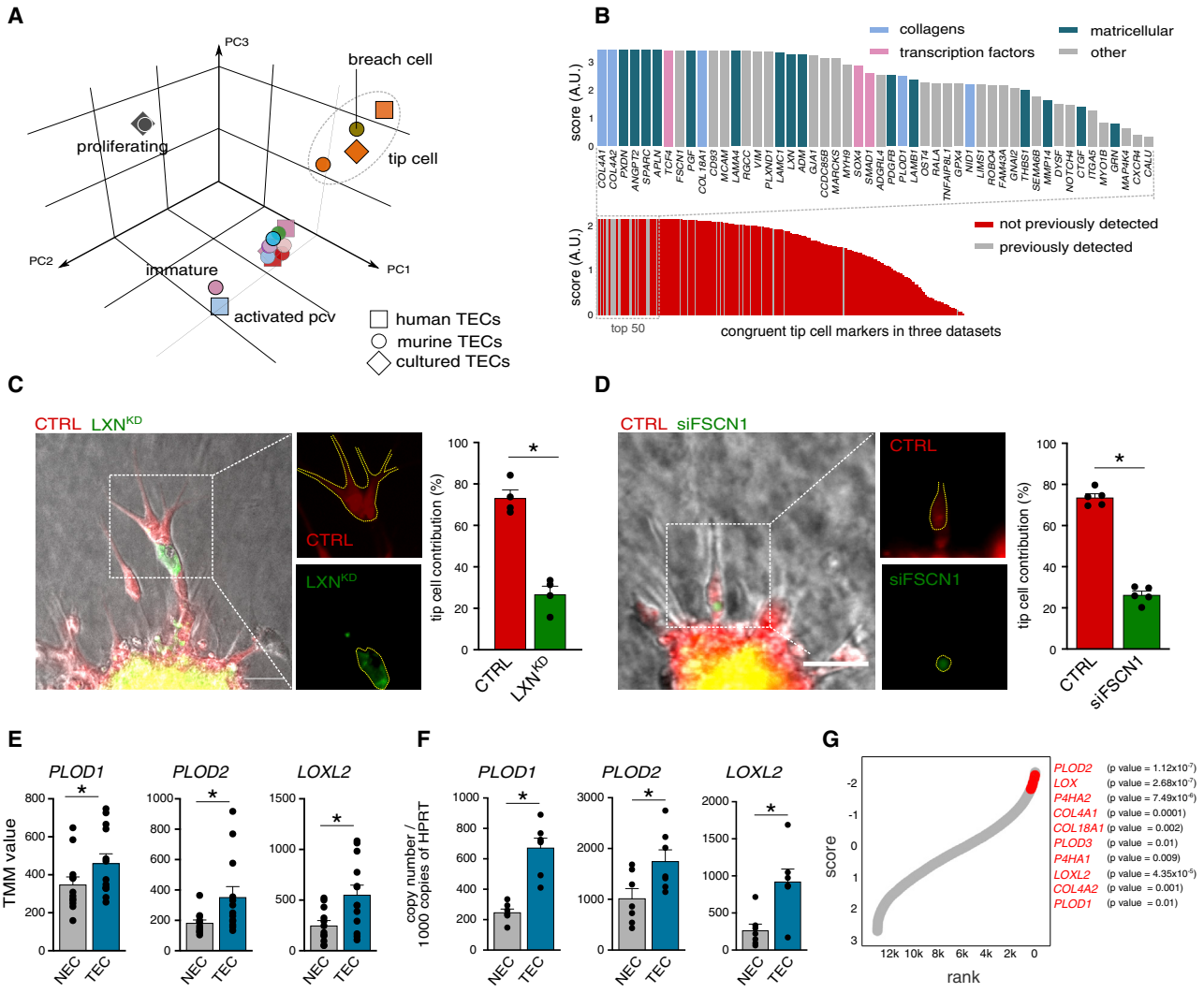


Figure 6. Congruency Analysis and Transcriptomics Meta-analysis

(A) PCA on the pairwise Jaccard similarity coefficients between the top 50 marker genes of freshly isolated human and murine, and cultured human TEC phenotypes.

(B) Bar plot visualization depicting the congruency score of conserved tip cell marker genes and their intersection with benchmark datasets (bottom panel) and the 50 top-ranking marker genes categorized by the indicated biological processes (top panel). A.U., arbitrary units.

(C and D) Mosaic EC spheroid competition of CTRL (red) and (C) LXN^{KD} (green, KD denotes small hairpin RNA [shRNA] knockdown) ECs or (D) siFSCN1 (green, si denotes siRNA silencing) ECs. Left: Representative micrographs of mosaic EC spheroid. Scale bar, 50 μm. Right: Quantification of the fraction of tip cells with the indicated genotype. Data are means ± SEM; n = 4 for LXN, n = 5 for FSCN1; *p < 0.05 by unpaired two-tailed t test.

(E and F) Expression levels of genes encoding collagen hydroxylation and cross-linking enzymes measured by (E) bulk RNA sequencing or (F) qRT-PCR. Data are means ± SEM; (E) n = 13 and (F) n = 7; *p < 0.05 by paired two-tailed t test. TMM, trimmed mean of *M* values.

(G) Gene expression meta-analysis of 6 publicly available NEC versus TEC datasets from 5 different tumor types. The S curve has 12,000 dots, representing genes that were detected in all 6 datasets. x axis: rank numbers from 1 to 12,000 (consistently overexpressed genes in TECs have a low rank number, consistently downregulated genes have a high rank number); y axis: the scaled meta-analysis score (consistently overexpressed genes in TECs have a low meta-analysis score, consistently downregulated genes have a high meta-analysis score). Genes involved in collagen modification or encoding collagens are shown as red dots and listed on the right (p < 0.01 by rank product analysis).

See also Figure S8 and Tables S3 and S4.

AAT is considered to *prune* angiogenic ECs, although VEGF blockade can also normalize the abnormal tumor vasculature (“tumor vessel normalization”) (Jain, 2005). We explored if VEGF blockade tuned ECs by inducing more subtle gene expression changes. We therefore first constructed “normal” and “tumor vessel disorganization” gene signatures by comparing angiogenic versus non-angiogenic phenotypes in a pairwise differential analysis to identify genes that were significantly (adjusted p value <0.01) differentially expressed from an effective fold change threshold of ±1.5. This revealed 18 and 28 genes that were significantly higher expressed in disorganized (among which tip, breach, and immature EC marker genes)

and normal (including capillary marker genes) phenotypes, respectively (Figures 5I, 5J, and S7D and Table S4).

GSVA using the 18 disorganization gene signature revealed that tumor vessel disorganization was partially reversed by VEGF blockade (Figures 5I and 5J). Gene set enrichment analysis further showed that both VEGFR inhibitors reduced the expression of tip EC gene signatures (glycolysis; De Bock et al., 2013; BM remodeling), while increasing the expression of signatures associated with mature homeostatic functions (antigen presentation, barrier, and blood vessel development) (Figures S7E and S7F). Thus, VEGF blockade not only pruned but also tuned TECs to promote a more quiescent and mature tumor vasculature with homeostatic functions, suggesting partial molecular tumor vessel normalization.

Identification of Conserved TEC Phenotypes and Tip Cell Markers

Anti-VEGF therapy efficacy is limited by resistance due to upregulation of alternative pro-angiogenic signals (Ebos and Kerbel, 2011). To identify alternative angiogenic candidates, we hypothesized that TEC phenotypes, conserved across species and models, and congruently overexpressed genes might be stronger and more robust angiogenic candidates. We thus performed an integrated combined single-cell transcriptomics, bulk multiomics, and transcriptomics meta-analysis to identify such candidates (Figure S8A).

We used the scRNA-seq data to assess similarity of angiogenic EC phenotypes across species and models using pairwise Jaccard similarity coefficients of marker gene sets and applied principal component analysis (PCA) for visualization (Figure 6A and S8A). The *in vivo* artery, capillary, and vein phenotypes were lost in hcTECs *in vitro*; in contrast, the tip and breach EC transcriptome signature of *in vivo* hTECs and mTECs was conserved in cultured hcTECs. ECs with a proliferation signature were only detected in cultured and murine TECs, and in hTECs from some NSCLC patients (see above). Freshly isolated human vein and murine immature TECs, and intermediate hcTECs highly expressed ribosomal genes, suggesting plastic phenotypes, transitioning to other angiogenic EC phenotypes. All other phenotypes were model- or species-specific. Since tip TECs expressed a similar marker gene signature across all species and models tested, we focused on identifying congruent tip TEC markers.

To construct a ranked list of tip TEC markers, we first selected 296 conserved tip cell-enriched genes that were consistently highest expressed in tip TECs across species and models. Second, for each conserved gene, we defined a rank score by calculating the product of the marker gene rank in each of the 3 datasets (in this list, genes that are consistently among the most upregulated genes in tip cells rank the highest) (Figure 6B). For subsequent analyses, we focused on the top-ranking 50 enriched markers (adjusted p value $< 5 \times 10^{-3}$). Validating our approach, we confirmed the inclusion of tip EC markers (*ANGPT2*, *APLN*, *FSCN1*, *PGF*, *PLXND1*, *ADM*, *PDGFB*, *CXCR4*, others), previously detected in tip ECs but not necessarily further characterized at the expression pattern and functional levels (del Toro et al., 2010; Strasser et al., 2010; Zhao et al., 2018). Half of the 50 top-ranking genes were not previously described as tip TEC markers (not detected in transcriptomics studies of ECs in tumor or physiological angiogenesis). Congruent tip TEC markers were associated with the migratory tip EC phenotype, including laminins (*LAMA4*, *LAMC1*, *LAMB1*), matricellular proteins (*SPARC*, *LXN*), cytoskeleton-associated genes (*VIM*, *MARCKS*, *MYH9*, *MYO1B*), and cell adhesion molecules (*CD93*, *MCAM*, *ITGA5*). We also identified novel tip TEC-enriched transcription factors (*TCF4*, *SOX4*, *SMAD1*).

The disorganized vascular architecture in tumors impedes/precludes topographical identification of tip TECs: Figure S6J shows heterogeneous expression of a tip cell marker, but its position at the forefront of the tumor vessel sprout and its expression in morphological tip cells cannot be unambiguously identified. We, therefore, used the postnatal retinal angiogenesis model to validate the expression of LXN (Latexin) and FSCN1 (Fascin) in tip ECs at the vascular front (Figure S8B). To document their functional role, we examined if silencing these tip cell markers affected tip cell competitivity (De Bock et al., 2013). We silenced LXN or FSCN1 expression in human umbilical vein ECs (HUVECs; >70%–80% silencing efficiency; Figure S8C) and generated mosaic spheroids containing a 1:1 mixture of control wild-type HUVECs (red) and HUVECs silenced for LXN or FSCN1 (green). Silenced cells were less often at the tip position, confirming the tip cell role of these markers (Figures 6C and 6D).

Integrated Analysis Highlights Collagen Modification as an Angiogenic Pathway

Genes encoding collagens (*COL4A1*, *COL4A2*, *COL18A1*) ranked in the top 10 most congruent tip cell markers, while other collagen-modifying (cross-linking) enzymes (*PXDN*, *PLOD1*) ranked in the top 50 (Figure 6B). Bulk RNA-seq analysis of 13 independent NSCLC patients (Figure S8A; Table S3) and confirmatory RT-PCR analysis showed that the expression of genes involved in collagen cross-linking (*LOXL2*, fold change = 2.5; $p = 5.15 \times 10^{-3}$) and hydroxylation (*PLOD1-3*, fold change 1.24–1.74; $p < 0.05$ for all genes) was higher in TECs than NECs (Figures 6E and 6F). Procollagen-lysine, 2-oxoglutarate 5-dioxygenase (PLOD) isoenzymes (known as lysyl hydroxylases) intracellularly hydroxylate lysine residues in collagen, while lysyl oxidase (LOX), or its homologue LOXL2, extracellularly catalyze the first step in collagen crosslink formation (Gilkes et al., 2014). To explore upregulation of collagen-modifying enzymes in TECs from other tumor types, we performed a meta-analysis of 6 publicly available NEC versus TEC bulk transcriptomics datasets, freshly isolated from patients with 5 different tumor types (Table S4). Also in this analysis, transcripts encoding collagen-modifying enzymes were enriched ($p < 0.05$ for all genes) and ranked among the top 1%–5% most consistently upregulated genes in TECs (Figure 6G).

To explore upregulation of collagen-modifying enzymes in TECs at the protein level, we performed proteomics analysis on an in-house-generated cohort of 144 prospectively collected TEC and NEC samples from lung, kidney, and colorectal (CRC) tumors and colorectal liver metastasis (CRCLM) (Figure S8A; Table S3). In TECs from NSCLC patients (n = 27 patients), LOXL2 was the highest upregulated protein (>4-fold; $p = 6.89 \times 10^{-7}$), while PLOD1 and PLOD2 were upregulated 1.7-fold ($p = 5.52 \times 10^{-5}$) and 1.9-fold ($p = 4.39 \times 10^{-4}$), respectively (Figure 7A). Meta-analysis across all 4 tumor types identified 288

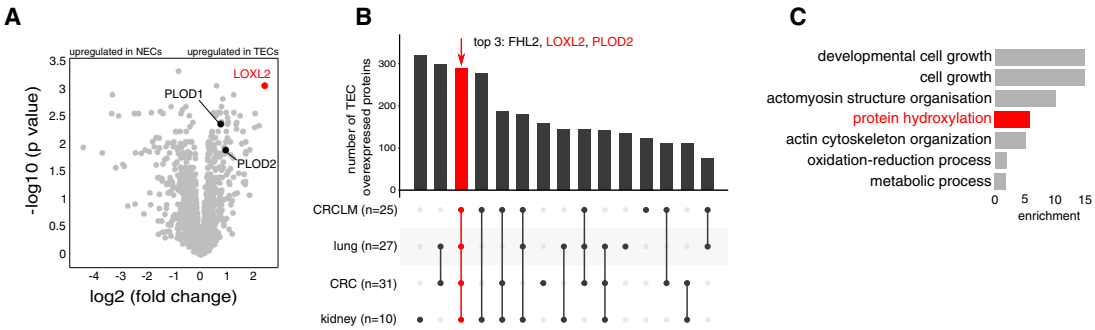


Figure 7. Proteomics and Gene Ontology Analysis

(A) Volcano plot of differentially expressed proteins in TECs compared with NECs.

(B) Upset plot visualization of the results of a cultured hcNEC versus hcTEC differential protein expression meta-analysis of 144 bulk proteomics samples (n = 93 patients), showing the number of proteins that were higher expressed in TECs than NECs isolated from the indicated tumor type. The bar graph represents the number of proteins, detected in the tumor type(s) indicated by the dot plot below; 288 proteins were consistently higher expressed in TECs than NECs in all 4 tumor types (red bar graph and intersection). LOXL2 and PLOD2 were ranked in the top 3.

(C) Bar graph representation of significantly enriched gene ontology terms ($p < 10^{-3}$). See also Figure S8 and Tables S3, S4, and S7.

consistently higher expressed proteins, including upregulated LOX ($p = 7.88 \times 10^{-3}$), LOXL2 ($p = 1.21 \times 10^{-6}$), PLOD1 ($p = 7.19 \times 10^{-3}$) and PLOD2 ($p = 4.52 \times 10^{-5}$) (Figure 7B; Table S4). Gene ontology (GO) analysis on the subset of consistently higher expressed proteins revealed an enrichment of glycolytic enzymes ($p = 4.98 \times 10^{-4}$), growth ($p = 4.80 \times 10^{-4}$) and actin cytoskeleton organization ($p = 1.16 \times 10^{-4}$), as well as protein hydroxylation and cross-linking ($p = 7.55 \times 10^{-4}$) (Figures 7C and S8D; Table S7). We consistently identified PLOD1, PLOD2, and LOXL2 and protein hydroxylation as highly ranked targets, and thus assessed if these genes functionally regulate vessel sprouting.

In agreement with the above GO analysis, migration, proliferation, and vessel sprouting were higher in TECs than NECs (Figures 8A–8C). Silencing of PLOD1, PLOD2, or LOXL2 in HUVECs (>70%–80% silencing efficiency; Figure S8E) reduced EC migration (Figure 8D) and impaired vessel sprouting in HUVEC spheroids (Figures 8E and 8F). Minoxidil, a pharmacological PLOD blocker (Shao et al., 2018), reduced EC migration *in vitro* (Figure 8G). Administration of minoxidil or the lysyl oxidase inhibitor beta-aminopropionitrile (BAPN) (Rodriguez et al., 2010) inhibited corneal angiogenesis *in vivo* (Figure 8H). Overall, functional validation revealed that targeting genes identified by integrated analysis inhibits vessel sprouting.

DISCUSSION

Previous bulk analysis-based transcriptomics studies revealed only a limited number of EC phenotypes (Coppiello et al., 2015; Marcu et al., 2018; Nolan et al., 2013; Sabbagh et al., 2018). Highly powered single-cell transcriptomics studies now allow characterization of EC phenotypes in more detail and enable revision of the traditional outlook on the tumor endothelium. Our data suggest a possible role for peritumoral (scavenging) capillary NECs and TECs with a transcriptome signature of HEVs or semi-professional APCs in tumor immune surveillance. Notably, TECs represent the first-line defense contact for immune cells in the tumor micro-environment. We also identified ECs that expressed markers of tip cells and an EC phenotype, implicated in the initiation of vessel sprouting by creating an opening in the basement membrane to assist invasive tip cells to sprout (Seano et al., 2014), which we coined breach ECs. We additionally identified vein ECs with a resident endothelial stem cell signature in mice, which might contribute to lung tumor vascularization, although this remains to be further studied.

Extending previous findings (Zhao et al., 2018), our data show that a VEGFR2-selective inhibitor not only affects more sensitively tip but also breach mTECs, while having smaller effects on vein and capillary mTECs. In addition, this study provides an unbiased molecular characterization of disorganized TECs and shows that VEGF blockade induces partial molecular tumor vessel normalization, including the induction of the activated postcapillary vein phenotype (resembling HEVs) and other more quiescent NEC phenotypes. This raises the question if AAT, in addition to pruning the small number of angiogenic and proliferating TECs, might benefit from tuning the more abundant TEC phenotypes contributing to tumor vessel disorganization. A concern is that tip TECs in human lung tumors make up only <10% of all TECs, raising the question if targeting such a small TEC subpopulation suffices to inhibit tumor angiogenesis and if the paucity of angiogenic tip and proliferating TECs in human lung tumors contributes to the insufficient efficacy and resistance to VEGF blockade AAT. Regardless, our study provides initial evidence for a correlation of angiogenic signatures with NSCLC patient survival.

Our integrated scRNA-seq and multi-omics approach showed that only the tip TEC phenotype was conserved across species and models. This was surprising, given that the tip EC phenotype is not a genetically predetermined state (Blanco and Gerhardt, 2013), but may be due to the presence of VEGF in the culture medium, an inducer of the tip EC phenotype (Siemerink et al., 2012). The identification of the tip TEC phenotype in culture raises the opportunity to study tip TEC targets *in vitro*, in contrast to arterial, venous, capillary, and other *in vivo* TEC phenotypes that are lost in culture.

To overcome inter-patient heterogeneity and contextual differences in marker gene signatures due to differences in species (human versus mouse), tumor type (NSCLC versus LLC),

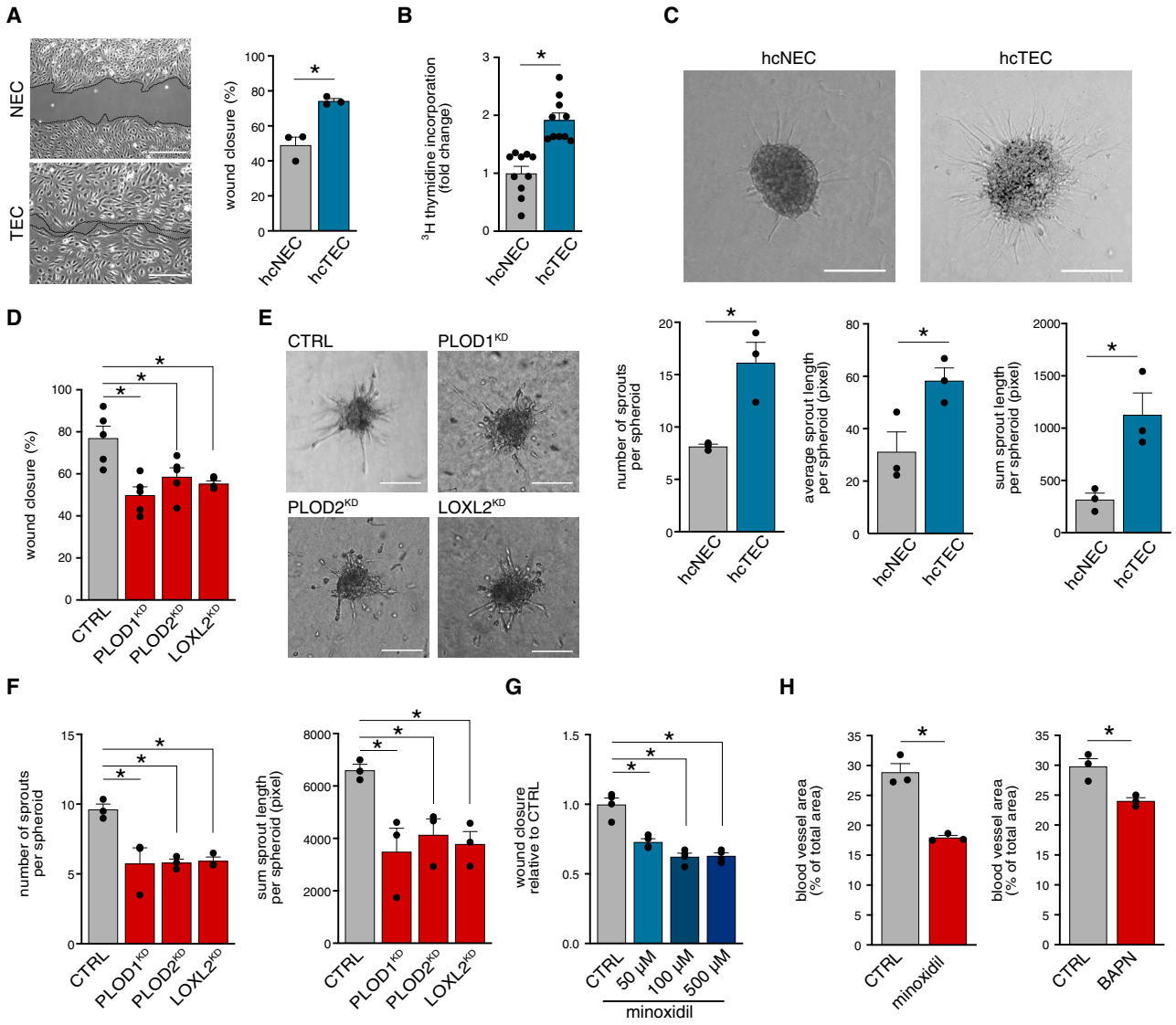


Figure 8. Inhibition of Collagen Hydroxylation/Cross-linking Impairs Vessel Sprouting

(A) Micrographs (left) and quantification (right) of hcNEC and hcTEC migration in scratch wound assays (mean ± SEM; n = 3; *p < 0.05, unpaired two-tailed t test). Scale bar, 300 μm.

(B) ^{3}H-thymidine incorporation assay (mean ± SEM, n = 10, *p < 0.05, unpaired two-tailed t test).

(C) Bright field photographs (top) and morphometric quantification (bottom) of hcNEC and hcTEC spheroid sprouting (mean ± SEM; n = 3; *p < 0.05 versus NECs, unpaired two-tailed t test). Scale bars, 100 μm.

(D) Scratch wound migration assay with control and PLOD1, PLOD2, or LOXL2 silenced (KD) ECs (mean ± SEM, n = 5, *p < 0.05, one-way ANOVA with Dunnett's test).

(E and F) (E) Bright field photographs and (F) morphometric quantification of spheroid sprouting using control and PLOD1, PLOD2, or LOXL2 silenced (KD) ECs. Scale bars, 100 μm. Data are means ± SEM; n = 3; *p < 0.05 versus control, one-way ANOVA with Dunnett's test.

(G) Scratch wound migration assay of ECs without or with pharmacological inhibition of PLODs by minoxidil (mean ± SEM, n = 3, *p < 0.05).

(H) Corneal angiogenesis upon corneal cauterization-induced injury in mice treated with vehicle (CTRL), minoxidil, or beta-aminopropionitrile (BAPN). Data are means ± SEM; n = 3 independent experiments, each of 6 mice per group; *p < 0.05; unpaired, two-tailed t test.

See also Figure S8.

and model (freshly isolated versus cultured TECs), we used scRNA-seq in combination with orthogonal multi-omics approaches to identify multiple genes involved in posttranslational collagen modification, illustrating that the discovery of these genes as angiogenic candidates was not accidental. We confirmed LOXL2 as known angiogenic candidate (Baker et al., 2013; Gilkes et al., 2014; Osawa et al., 2013; Zaffryar-Eilot et al., 2013), and identified and functionally validated procollagen cross-linking PLOD isoenzymes. Collagens have been implicated in tumor angiogenesis, albeit contextually (Fang et al., 2014; Sottile, 2004), but PLOD isoenzymes have not been substantially and consistently involved in angiogenesis yet. Hence,

an integrated approach, based on orthogonal multi-omics single-cell and bulk methods, may offer improved chances to identify conserved TEC markers that are biologically and translationally relevant.

This study provides a public resource for data exploration, available at https://www.vibcancer.be/software-tools/lungTumor_ECTax. While our study shows how this integrated approach can identify highly relevant angiogenic targets, we acknowledge limitations of our work. First, the inferred biological role for each EC phenotype is putative and requires functional validation to probe their biological role. Second, larger numbers of patients must be analyzed to probe inter-patient heterogeneity and identify all possible TEC phenotypes at deep resolution. Third, it should not be surprising that mouse and human TEC taxonomies do not completely overlap, given that tumors in both species grow differently, and patients exhibit larger genetic and environmental heterogeneity than mice.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Patients
  - Mice
  - Cell Lines and Primary Cells
- METHOD DETAILS
  - Human Tumor and Peritumoral Normal Endothelial Cell Isolation
  - Primary Endothelial Cell Culture
  - Mouse Model of Lewis Lung Carcinoma & Tumor and Normal EC Isolation
  - Knock Down Strategy
  - *In Vitro* Functional Assays
  - *In Vivo* Functional Assays in Corneal Angiogenesis
  - Immunohistochemistry
  - RNA Isolation and Quantitative RT-PCR
  - Single-cell Droplet-based RNA Sequencing
  - Single-Cell Transcriptomics Analysis
  - RNAscope *In Situ* Hybridization and Quantification
  - Time-of-Flight Mass Cytometry (CyTOF)
  - CyTOF Data Analysis
  - Integration of scRNA-seq and CyTOF Data
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND CODE AVAILABILITY
  - Data Resources
  - Software

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.ccell.2019.12.001.

ACKNOWLEDGMENTS

We acknowledge K. Peeters, T. Van Brussel, and R. Schepers for assistance. J.G., K.R., L.T., and J.K. are supported by the 'Fonds Wetenschappelijk Onderzoek' (FWO); L.-C.C. by the Else Kröner-Fresenius and Fritz Thyssen Stiftung (10.16.2.017MN); A.P. by the Austrian Science Fund (J3730-B26); L.-A.T. by Antwerp University; V.G. by Strategisch Basisonderzoek FWO (SB-FWO); S.K. by Kom op Tegen Kanker (Stand up to Cancer, Flemish Cancer Society); M.M. by Methusalem funding (Flemish Government); L.B. and Y.L. by Sanming Project of Medicine in Shenzhen (SZSM201612074); L.L. by Lundbeck Foundation (R219–2016-1375) and DFF Sapere Aude Starting grant (8048-00072A); Y.L. by BGI-Research, Danish Research Council for Independent Research (DFF–1337–00128), Sapere Aude Young Research Talent Prize (DFF-1335–00763A), and Aarhus University Strategic Grant (AU-iCRISPR); X.L. by State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center (Sun Yat-sen University) and National Natural Science Foundation of China (81670855), Key Program of Guangzhou Scientific Research Plan (3030901006074); and P.C. by VIB TechWatch, Methusalem funding, FWO, Foundation against Cancer (2016-078), Kom op tegen Kanker (Stand up to Cancer, Flemish Cancer Society), ERC Proof of Concept (ERC-713758), and Advanced ERC Research Grant (EU-ERC743074).

AUTHOR CONTRIBUTIONS

J.G., K.R., F.T., and L.T. designed and analyzed experiments. J.G., L.T., L.-C.C., A.P., J.K., and M.G-C. set up the EC isolation protocols. L.-C.C., E.W., H.D, P.DL., J.V., B.W., X.S., E.V., A.W., B. Topal, W.E., H.B., and A.E. provided human tumor samples. K.R., L.T., L.-C.C., V.G., J.K., L. Sokol, J.Q, and L.-A.T. prepared the scRNA-seq samples. L.L., B. Thienpont, D.L., and Y.L. did 10× single-cell sequencing. S.K., B.B., and T.K. processed scRNA-seq data. L.dR., J.K., D.P., F.D.S., F.J.T.S., R.J.M., and Y.Z. did CyTOF. L.T., J.K., and S.V. did immunohistochemistry. F.I. did proteomics analysis. V.L. provided advice on the statistical methodology. F.D.S., M.M., G.E., M.D., L. Schoonjans, H.Y., J.W., L.B., L.L., B. Thienpont, X.L., D.L., and Y.L. provided advice and discussed the results. J.G., K.R., and P.C. wrote the manuscript. P.C. conceptualized the study. All authors discussed the results and the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 24, 2019
Revised: August 30, 2019
Accepted: December 6, 2019
Published: January 13, 2020

REFERENCES

Angelidis, I., Simon, L.M., Fernandez, I.E., Strunz, M., Mayr, C.H., Greiffo, F.R., Tsitsiridis, G., Ansari, M., Graf, E., Strom, T.M., et al. (2019). An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. Nat. Commun. *10*, 963.

Baker, A.M., Bird, D., Welti, J.C., Gourlaouen, M., Lang, G., Murray, G.I., Reynolds, A.R., Cox, T.R., and Erler, J.T. (2013). Lysyl oxidase plays a critical role in endothelial cell stimulation to drive tumor angiogenesis. Cancer Res. *73*, 583–594.

Blanco, R., and Gerhardt, H. (2013). VEGF and Notch in tip and stalk cell selection. Cold Spring Harb. Perspect. Med. *3*, a006569.

Bruning, U., Morales-Rodriguez, F., Kalucka, J., Goveia, J., Taverna, F., Queiroz, K.C.S., Dubois, C., Cantelmo, A.R., Chen, R., Loroch, S., et al. (2018). Impairment of angiogenesis by fatty acid synthase inhibition involves mTOR malonylation. Cell Metab. *28*, 866–880.e15.

Cancer Genome Atlas Research Network, Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., and Stuart, J.M. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. Nat. Genet. *45*, 1113–1120.

Cantelmo, A.R., Conradi, L.C., Brajic, A., Goveia, J., Kalucka, J., Pircher, A., Chaturvedi, P., Hol, J., Thienpont, B., Teuwen, L.A., et al. (2016). Inhibition of the glycolytic activator PFKFB3 in endothelium induces tumor vessel

normalization, impairs metastasis, and improves chemotherapy. Cancer Cell *30*, 968–985.

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496–502.

Carmeliet, P., and Jain, R.K. (2011a). Molecular mechanisms and clinical applications of angiogenesis. Nature *473*, 298–307.

Carmeliet, P., and Jain, R.K. (2011b). Principles and mechanisms of vessel normalization for cancer and other angiogenic diseases. Nat. Rev. Drug Discov. *10*, 417–427.

Colaprico, A., Silva, T.C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T.S., Malta, T.M., Pagnotta, S.M., Castiglioni, I., et al. (2016). TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. *44*, e71.

Conway, J.R., Lex, A., and Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics *33*, 2938–2940.

Coppiello, G., Collantes, M., Sirerol-Piquer, M.S., Vandenwijngaert, S., Schoors, S., Swinnen, M., Vandersmissen, I., Herijgers, P., Topal, B., van Loon, J., et al. (2015). Meox2/Tcf15 heterodimers program the heart capillary endothelium for cardiac fatty acid uptake. Circulation *131*, 815–826.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

De Bock, K., Georgiadou, M., Schoors, S., Kuchnio, A., Wong, B.W., Cantelmo, A.R., Quaegebeur, A., Ghesquière, B., Cauwenberghs, S., Eelen, G., et al. (2013). Role of PFKFB3-driven glycolysis in vessel sprouting. Cell *154*, 651–663.

Dejana, E., Hirschi, K.K., and Simons, M. (2017). The molecular basis of endothelial cell plasticity. Nat. Commun. *8*, 14361.

del Toro, R., Prahst, C., Mathivet, T., Siegfried, G., Kaminker, J.S., Larrivee, B., Breant, C., Duarte, A., Takakura, N., Fukamizu, A., et al. (2010). Identification and functional analysis of endothelial tip cell-enriched genes. Blood *116*, 4025–4033.

Ebos, J.M., and Kerbel, R.S. (2011). Antiangiogenic therapy: impact on invasion, disease progression, and metastasis. Nat. Rev. Clin. Oncol. *8*, 210–221.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics *10*, 48.

Eldridge, L., Moldobaeva, A., Zhong, Q., Jenkins, J., Snyder, M., Brown, R.H., Mitzner, W., and Wagner, E.M. (2016). Bronchial artery angiogenesis drives lung tumor growth. Cancer Res. *76*, 5962–5969.

Fang, M., Yuan, J., Peng, C., and Li, Y. (2014). Collagen as a double-edged sword in tumor progression. Tumour Biol. *35*, 2871–2882.

García-Caballero, M., Zecchin, A., Souffreau, J., Truong, A.-C.K., Teuwen, L.-A., Vermaelen, W., Martín-Pérez, R., de Zeeuw, P., Bouché, A., Vinckier, S., et al. (2019). Role and therapeutic potential of dietary ketone bodies in lymph vessel growth. Nat. Metab. *1*, 666–675.

Gilkes, D.M., Semenza, G.L., and Wirtz, D. (2014). Hypoxia and the extracellular matrix: drivers of tumour metastasis. Nat. Rev. Cancer *14*, 430–439.

Giordani, L., He, G.J., Negroni, E., Sakai, H., Law, J.Y.C., Siu, M.M., Wan, R., Corneau, A., Tajbakhsh, S., Cheung, T.H., and Le Grand, F. (2019). High-dimensional single-cell cartography reveals novel skeletal muscle-resident cell populations. Mol. Cell *74*, 609–621.e6.

Girard, J.P., and Springer, T.A. (1995). High endothelial venules (HEVs): specialized endothelium for lymphocyte migration. Immunol. Today *16*, 449–457.

Haghverdi, L., Lun, A.T.L., Morgan, M.D., and Marioni, J.C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat. Biotechnol. *36*, 421–427.

Hanzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics *14*, 7.

Heskes, T., Eisinga, R., and Breitling, R. (2014). A fast algorithm for determining bounds and accurate approximate p-values of the rank product statistic for replicate experiments. BMC Bioinformatics *15*, 367.

Hong, F., Breitling, R., McEntee, C.W., Wittner, B.S., Nemhauser, J.L., and Chory, J. (2006). RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. Bioinformatics *22*, 2825–2827.

Innes, B.T., and Bader, G.D. (2018). scClustViz - single-cell RNAseq cluster assessment and visualization. F1000Res *7*, ISCB Comm J-1522.

Jain, R.K. (2005). Normalization of tumor vasculature: an emerging concept in antiangiogenic therapy. Science *307*, 58–62.

Jain, R.K. (2014). Antiangiogenesis strategies revisited: from starving tumors to alleviating hypoxia. Cancer Cell *26*, 605–622.

Khan, S., Taverna, F., Rohlenova, K., Treps, L., Geldhof, V., de Rooij, L., Sokol, L., Pircher, A., Conradi, L.C., Kalucka, J., et al. (2019). EndoDB: a database of endothelial cell transcriptomics data. Nucleic Acids Res. *47*, D736–D744.

Kiselev, V.Y., Yiu, A., and Hemberg, M. (2018). scmap: projection of single-cell RNA-seq data across data sets. Nat. Methods *15*, 359–362.

Lambrechts, D., Wauters, E., Boeckx, B., Aibar, S., Nittner, D., Burton, O., Bassez, A., Decaluwe, H., Pircher, A., Van den Eynde, K., et al. (2018). Phenotype molding of stromal cells in the lung tumor microenvironment. Nat. Med. *24*, 1277–1289.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Levandowsky, M., and Winter, D. (1971). Distance between sets. Nature *234*, 34.

Li, J.C.A. (2003). Modeling survival data: extending the Cox model. Sociol. Method Res. *32*, 117–120.

Liu, Y., Poon, R.T., Li, Q., Kok, T.W., Lau, C., and Fan, S.T. (2005). Both anti-angiogenesis- and angiogenesis-independent effects are responsible for hepatocellular carcinoma growth arrest by tyrosine kinase inhibitor PTK787/ZK222584. Cancer Res. *65*, 3691–3699.

Marcu, R., Choi, Y.J., Xue, J., Fortin, C.L., Wang, Y., Nagao, R.J., Xu, J., MacDonald, J.W., Bammler, T.K., Murry, C.E., et al. (2018). Human organ-specific endothelial cell heterogeneity. iScience *4*, 20–35.

Nolan, D.J., Ginsberg, M., Israely, E., Palikuqi, B., Poulos, M.G., James, D., Ding, B.S., Schachterle, W., Liu, Y., Rosenwaks, Z., et al. (2013). Molecular signatures of tissue-specific microvascular endothelial cell heterogeneity in organ maintenance and regeneration. Dev. Cell *26*, 204–219.

Osawa, T., Ohga, N., Akiyama, K., Hida, Y., Kitayama, K., Kawamoto, T., Yamamoto, K., Maishi, N., Kondoh, M., Onodera, Y., et al. (2013). Lysyl oxidase secreted by tumour endothelial cells promotes angiogenesis and metastasis. Br. J. Cancer *109*, 2237–2247.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Rodriguez, H.M., Vaysberg, M., Mikels, A., McCauley, S., Velayo, A.C., Garcia, C., and Smith, V. (2010). Modulation of lysyl oxidase-like 2 enzymatic activity by an allosteric antibody inhibitor. J. Biol. Chem. *285*, 20964–20974.

Sabbagh, M.F., Heng, J.S., Luo, C., Castanon, R.G., Nery, J.R., Rattner, A., Goff, L.A., Ecker, J.R., and Nathans, J. (2018). Transcriptional and epigenomic landscapes of CNS and non-CNS vascular endothelial cells. eLife *7*, e36187.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schmidt, T., Kharabi Masouleh, B., Loges, S., Cauwenberghs, S., Fraisl, P., Maes, C., Jonckx, B., De Keersmaecker, K., Kleppe, M., Tjwa, M., et al.

(2011). Loss or inhibition of stromal-derived PlGF prolongs survival of mice with imatinib-resistant Bcr-Abl1(+) leukemia. Cancer Cell *19*, 740–753.

Schoors, S., Bruning, U., Missiaen, R., Queiroz, K.C., Borgers, G., Elia, I., Zecchin, A., Cantelmo, A.R., Christen, S., Goveia, J., et al. (2015). Fatty acid carbon is essential for dNTP synthesis in endothelial cells. Nature *520*, 192–197.

Seano, G., Chiaverina, G., Gagliardi, P.A., di Blasio, L., Puliafito, A., Bouvard, C., Sessa, R., Tarone, G., Sorokin, L., Helley, D., et al. (2014). Endothelial podosome rosettes regulate vascular branching in tumour angiogenesis. Nat. Cell Biol. *16*, 931–941, 931-938.

Shao, S., Zhang, X., Duan, L., Fang, H., Rao, S., Liu, W., Guo, B., and Zhang, X. (2018). Lysyl hydroxylase inhibition by minoxidil blocks collagen deposition and prevents pulmonary fibrosis via TGF-beta(1)/Smad3 signaling pathway. Med. Sci. Monit. *24*, 8592–8601.

Siemerink, M.J., Klaassen, I., Vogels, I.M., Griffioen, A.W., Van Noorden, C.J., and Schlingemann, R.O. (2012). CD34 marks angiogenic tip cells in human vascular endothelial cell cultures. Angiogenesis *15*, 151–163.

Smillie, C.S., Biton, M., Ordovas-Montanes, J., Sullivan, K.M., Burgin, G., Graham, D.B., Herbst, R.H., Rogel, N., Slyper, M., Waldman, J., et al. (2019). Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. Cell *178*, 714–730.e22.

Sottile, J. (2004). Regulation of angiogenesis by extracellular matrix. Biochim. Biophys. Acta *1654*, 13–22.

Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. U S A *100*, 9440–9445.

Strasser, G.A., Kaminker, J.S., and Tessier-Lavigne, M. (2010). Microarray analysis of retinal endothelial tip cells identifies CXCR4 as a mediator of tip cell morphology and branching. Blood *115*, 5102–5110.

Suzuki, R., and Shimodaira, H. (2006). Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics *22*, 1540–1542.

Sweeney, P., Karashima, T., Kim, S.J., Kedar, D., Mian, B., Huang, S., Baker, C., Fan, Z., Hicklin, D.J., Pettaway, C.A., and Dinney, C.P. (2002). Anti-vascular endothelial growth factor receptor 2 antibody reduces tumorigenicity and metastasis in orthotopic prostate cancer xenografts via induction of endothelial cell apoptosis and reduction of endothelial cell matrix metalloproteinase type 9 production. Clin. Cancer Res. *8*, 2714–2724.

Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature *563*, 72–78.

van den Brink, S.C., Sage, F., Vertesy, A., Spanjaard, B., Peterson-Maduro, J., Baron, C.S., Robin, C., and van Oudenaarden, A. (2017). Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. Nat. Methods *14*, 935–936.

Wakabayashi, T., Naito, H., Suehiro, J.I., Lin, Y., Kawaji, H., Iba, T., Kouno, T., Ishikawa-Kato, S., Furuno, M., Takara, K., et al. (2018). CD157 marks tissue-resident endothelial stem cells with homeostatic and regenerative properties. Cell Stem Cell *22*, 384–397.e6.

Young, M.D., Mitchell, T.J., Vieira Braga, F.A., Tran, M.G.B., Stewart, B.J., Ferdinand, J.R., Collord, G., Botting, R.A., Popescu, D.M., et al. (2018). Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. Science *361*, 594–599.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287.

Zaffryar-Eilot, S., Marshall, D., Voloshin, T., Bar-Zion, A., Spangler, R., Kessler, O., Ghermazien, H., Brekhman, V., Suss-Toby, E., Adam, D., et al. (2013). Lysyl oxidase-like-2 promotes tumour angiogenesis and is a potential therapeutic target in angiogenic tumours. Carcinogenesis *34*, 2370–2379.

Zhao, Q., Eichten, A., Parveen, A., Adler, C., Huang, Y., Wang, W., Ding, Y., Adler, A., Nevins, T., Ni, M., et al. (2018). Single-cell transcriptome analyses reveal endothelial cell heterogeneity in tumors and changes following antiangiogenic treatment. Cancer Res. *78*, 2370–2382.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-human CD31 (clone JC70A)	Dako	Cat# M0823; RRID: AB_2114471
Rabbit anti-human PlGF	Abcam	Cat# ab9542; RRID: AB_307330
Rabbit anti-human CXCR4	Novus Biologicals	Cat# NLS1380; RRID: AB_2292628
Rabbit anti-human SELP	Sigma-Aldrich	Cat# HPA005990; RRID: AB_1078479
Rabbit anti-human VWF	Dako	Cat# A0082; RRID: AB_2315602
Rabbit anti-human EFNB2	Sigma-Aldrich	Cat# HPA008999; RRID: AB_1078721
Rat anti-human peripheral node addressin (clone MECA-79)	Novus Biologicals	Cat# NB100-77673, RRID: AB_1085374
Rat anti-mouse CD31	BD Biosciences	Cat# 550274; RRID: AB_393571
Goat anti-mouse CD105	R&D	Cat# AF1320; RRID: AB_354735
Rabbit anti-mouse VWF	Abcam	Cat# ab6994; RRID: AB_305689
Goat anti-mouse CAR4	R&D	Cat# AF2414; RRID: AB_2070332
Rabbit anti-mouse LXN	Abcam	Cat# ab154744
Mouse anti-SMA-FITC	Sigma-Aldrich	Cat# F3777; RRID: AB_476977
Rabbit anti-human CD36	Abcam	Cat# ab133625; RRID: AB_2716564
Rabbit anti-mouse FSCN1	Abcam	Cat# ab126772; RRID: AB_11143231
Donkey anti-rat Biotinylated	Jackson Immuno Research	Cat# 712-065-153; RRID: AB_2315779
Donkey anti-mouse Biotinylated	Jackson Immuno Research	Cat# 715-065-151; RRID: AB_2340785
Donkey anti-goat Biotinylated	Jackson Immuno Research	Cat# 705-065-003; RRID: AB_2340396
Donkey anti-Rabbit Alexa 647	Thermo Fisher Scientific	Cat# A-31573; RRID: AB_2536183
Donkey anti-Rabbit Alexa 568	Thermo Fisher Scientific	Cat# A10042; RRID: AB_2534017
Donkey anti-Rabbit Alexa 488	Thermo Fisher Scientific	Cat# A-21206; RRID: AB_2535792
Donkey anti-Rat Alexa 488	Thermo Fisher Scientific	Cat# A-21208; RRID: AB_2535794
ACKR1 153Eu	R&D	Cat# MAB4139; RRID: AB_2088599
CA4 166Er	R&D	Cat# AF2414; RRID: AB_2070332
CD274 175Lu	Fluidigm	Cat# 3175017B; RRID: AB_2687638
CD36 152Sm	Fluidigm	Cat# 3152007B; RRID: AB_2802106
CDH5 155Gd	Fluidigm	Cat# 3155010A
COL1A1 165Ho	R&D	Cat# AF6220; RRID: AB_10891543
CXCR4 173Yb	Fluidigm	Cat# 3173001B
EDNRB 163Dy	Novus	Cat# NBP2-16334
ENO1 172Yb	Abgent	Cat# AT1908a; RRID: AB_1552127
EPCAM 141Pr	Fluidigm	Cat# 3141006B; RRID: AB_2687653
HLA-DRA 150Nd	Fluidigm	Cat# 3150028B
IL33 158Gd	Abcam	Cat# ab118503; RRID: AB_10899254
KDR 154Sm	Fluidigm	Cat# 3154009B
LXN 143Nd	Abcam	Cat# ab154744
MKI67 164Dy	Thermo Fisher Scientific	Cat# MA5-14520, RRID: AB_10979488
NRP1 169Tm	Fluidigm	Cat# 3169018B
PDGFRB 171Yb	R&D	Cat# AF385; RRID: AB_355339
PECAM1 145Nd	Fluidigm	Cat# 3145004B; RRID: AB_2737262
PGF 149Sm	R&D	Cat# AF-264-PB; RRID: AB_2161055
POSTN 151Eu	Abcam	Cat# ab14041; RRID: AB_2299859

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PROX1 176Yb	Abcam	Cat# ab101851; RRID: AB_10712211
PTPRC 156Gd	Fluidigm	Cat# 3156010; RRID: AB_2714154
S1PR1 161Dy	Abcam	Cat# ab72806; RRID: AB_1268719
SOX17 168Er	Novus	Cat# NBP2-24568
SPARCL1 144Nd	Abcam	Cat# ab107533; RRID: AB_10863538
VWF 174Yb	Dako	Cat# A0082; RRID: AB_2315602
Prolong Gold Antifade Mountant	Thermo Fisher Scientific	Cat # P36934; RRID: SCR_015961
HRP-conjugated anti-mouse IgG	Cell Signaling Technology	Cat# 7076; RRID: AB_330924
HRP-conjugated anti-rabbit IgG	Cell Signaling Technology	Cat# 7074; RRID: AB_2099233
DC101 (neutralizing anti-VEGFR2 antibody), Bio X Cell	Bio-Connect	Cat# BP0060; RRID: AB_1107766
1C8 (control antibody)	(Schmidt et al., 2011)	
Biological Samples		
Human umbilical vein endothelial cells (HUVECs), primary; Approved by Ethics Committee Research UZ/KU Leuven; approval number S57123	This paper	N/A
Human patient NSCLC tumor specimen (approved by the Medical Ethics Committee UZ/KU Leuven under the protocol S57123)	UZ Leuven/This paper	N/A
LLC orthotopic tumor samples (study approved by the Animal Ethics Committee of the KU Leuven under protocol number P012/2014 and P156/2015)	This paper	N/A
Chemicals, Peptides, and Recombinant Proteins		
[^{3}H]-thymidine	Perkin Elmer	NET355L005MC
Antibiotic-antimycotic	Thermo Fisher Scientific	15240062
Bovine serum albumin	Sigma-Aldrich	A2058
Bovine serum albumin (UltraPure BSA)	Thermo Fisher Scientific	AM2616
CellTracker Deep Red	Thermo Fisher Scientific	C34565
CellTrace CFSE dye	Thermo Fisher Scientific	C34554
Collagen type I	Merck Millipore	08-115
Collagenase I	Thermo Fisher Scientific	GIBCO™ 17100017
Collagenase II	Thermo Fisher Scientific	GIBCO™ 17101015
Collagenase type 4	Worthington Biochemical	LS004188
Dispase	Thermo Fisher Scientific	171055-041
DMEM	Thermo Fisher Scientific	11965-084
RPMI	Thermo Fisher Scientific	21875-034
DMSO	Sigma-Aldrich	D2438
DNase I	Sigma-Aldrich	D4527
ECGS (endothelial cell medium growth supplement mix)	Bio-Connect	PromoCell C-39216
ECGS/H (endothelial cell growth supplement/heparin)	Bio-Connect	PromoCell C-30140
EDTA	VWR Chemicals	20302.293
EGM2 (Endothelial growth medium)	Bio-Connect	PromoCell C-22011
Fc receptor block TruStain FcX	BioLegend	422301
Fetal bovine serum (FBS)	Merck - Biochrom	S 0415
Gelatin from bovine skin	Sigma-Aldrich	G9391
Glutamine	Thermo Fisher Scientific	25030149
Glycerol	Merck Millipore	1.04091.1000
Heparin (bullet of ECGS/H)	Bio-Connect	PromoCell C-30140
Hoechst 33258	Sigma-Aldrich	B2261
Isolectin B4	Thermo Fisher Scientific	I21411

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Ketamine 100 mg/mL	Eurovet Animal Health B.V.	NIMATEK
Matrigel, growth factor reduced	Corning	356230
Medium 199, HEPES	Thermo Fisher Scientific	22340020
MEM NEAA	Thermo Fisher Scientific	11140035
MEM vitamin solution	Thermo Fisher Scientific	11120052
Opti-MEM	Thermo Fisher Scientific	11058021
Methylcellulose 4000 cP	Sigma-Aldrich	M0512
Minoxidil	Sigma-Aldrich	M4145
beta-aminopropionitrile	Sigma-Aldrich	A3134
Lipofectamine RNAiMAX	Thermo Fisher Scientific	13778150
Paraformaldehyde	Merck Millipore	8.18715.1000
Penicillin/Streptomycin	Thermo Fisher Scientific	15140122
Penicillin/Streptomycin/Glutamine	Thermo Fisher Scientific	10378016
Phosphate buffered saline (PBS)	Thermo Fisher Scientific	14190094
PTK787 (VEGFR tyrosine kinase inhibitor)	Novartis Pharma	N/A
Sodium dodecyl sulfate	Acros Organics	230425000
Sodium pyruvate	Thermo Fisher Scientific	11360070
TCA (trichloroacetic acid)	Sigma-Aldrich	T6399
Trypsin-EDTA (0.25%)	Thermo Fisher Scientific	25200056
Tween 80	Sigma-Aldrich	P1754
Unicaïne 0.4%	Théa Pharma	So1ha02
Xylazine	VMD	XYL-M 2%
Ketamine 100 mg/mL	Eurovet Animal Health B.V.	NIMATEK
Critical Commercial Assays		
CD31 MicroBeads, mouse	Miltenyi Biotec	130-097-418
CD31 MicroBead Kit, human	Miltenyi Biotec	130-091-935
CD45 MicroBeads, mouse	Miltenyi Biotec	130-052-301
CD45 MicroBeads, human	Miltenyi Biotec	130-045-801
Chromium Single Cell 3' Library, Gel Bead & Multiplex Kit and Chip Kit, v2	10x Genomics	PN-120237
Chromium Single Cell A Chip Kit	10x Genomics	PN-120236
Chromium i7 Multiplex Kit	10x Genomics	PN-120262
CyTOF Barcode Perm Buffer	Fluidigm	201057
CyTOF Cell-ID Cisplatin	Fluidigm	201064
CyTOF Cell-ID Intercalator-Ir	Fluidigm	201192A
CyTOF Maxpar Cell Staining Buffer	Fluidigm	201068
CyTOF Palladium barcodes	Fluidigm	201060
CyTOF MaxPar X8 antibody labeling kits	Fluidigm	
CyTOF EQ Four Element Calibration beads	Fluidigm	201078
Cytotoxicity Detection Kit	Roche Applied Science	11644793001
Fix/Perm buffer (eBiosciences intracellular fixation & permeabilization buffer set)	Thermo Fisher Scientific	88-8824-00
iScript cDNA synthesis kit	Bio-Rad	1708891
Pierce ECL Western Blotting Substrate	Thermo Fisher Scientific	32106
PureLink RNA Mini Kit	Thermo Fisher Scientific	12183018A
RNAscope Multiplex Fluorescent v2 Assay	ACDBio	323110
TSA Cyanine 3 (Cy3) System	Perkin Elmer	NEL704A001KT
TSA Cyanine 5 (Cy5) System	Perkin Elmer	NEL705A001KT
TSA Fluorescein System	Perkin Elmer	NEL701A001KT

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
scRNA-sequencing raw and analyzed data human NSCLC, freshly isolated and cultured	This paper	ArrayExpress: E-MTAB-6308
scRNA-sequencing raw and analyzed data LLC	This paper	ArrayExpress: E-MTAB-7458
Bulk RNA-sequencing raw and analyzed data	This paper	ArrayExpress: E-MTAB-8031
Proteomics	This paper	PRIDE: PXD014123
NSCLC scRNA-seq dataset	Lambrechts et al., 2018	ArrayExpress: E-MTAB-6149
Bulk RNA-seq dataset – renal cell carcinoma and colorectal cancer	EndoDB	GEO: GSE77199
Bulk RNA-seq dataset – bladder cancer	EndoDB	ArrayExpress: E-GEOD-41614
Bulk RNA-seq dataset – hepatocellular carcinoma	EndoDB	ArrayExpress: E-GEOD-51401
Experimental Models: Cell Lines		
Human umbilical vein endothelial cells (HUVECs), primary; Approved by Ethics Committee Research UZ/KU Leuven; approval number S57123	This paper	N/A
293T cells	ATCC	RRID: CVCL_0063
Lewis Lung Carcinoma cells	ATCC	RRID: CVCL_4358
Experimental Models: Organisms/Strains		
C57BL6/J mice	KU Leuven animal facility	N/A
C57BL6/J mice	Charles River	
Oligonucleotides		
qRT-PCR primers	Table S8	
shRNAs	Table S8	
siRNAs	Table S8	
RNAscope probe Hs-ACKR1	ACDBio	525131-C2
RNAscope probe Hs-CCL14	ACDBio	490401-C1
RNAscope probe Hs-CXCR4	ACDBio	310511-C2
RNAscope probe Hs-EFNB2	ACDBio	430651-C1
RNAscope probe Hs-PGF	ACDBio	311771-C1
RNAscope 3-Plex positive control probe-Hs	ACDBio	320861
RNAscope 3-Plex negative control probe	ACDBio	320871
RNAscope probe Hs-CD52	ACDBio	519381
RNAscope probe Hs-CD68	ACDBio	560591-C2
RNAscope probe Hs-IL1RL1	ACDBio	437481-C2
RNAscope probe Hs-EDNRB	ACDBio	528301
Recombinant DNA		
pLKO-shRNA2 vector	Sigma-Aldrich	N/A
pLVX-shRNA2 vector	Clontech	PT4052-5
pLKO.1-puro non-mammalian shRNA control	Sigma-Aldrich	SHC002V
Software and Algorithms		
R version 3.4.4 (2018-03-15) system: x86_64, mingw32 ui: RStudio (1.1.456) language: (EN) collate: English United States.1252	CRAN (R 3.4.4)	
asinh (base R)	CRAN (R 3.4.4)	
base R (version 3.4.4)	local	

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cell Ranger	10x Genomics	(tenx, RRID: SCR_01695)
clusterProfiler; version 3.6.0	Bioconductor	(clusterProfiler, RRID: SCR_016884)
edgeR; version *3.20.9*	Bioconductor	(edgeR, RRID: SCR_012802)
flashPCA; version 2	https://github.com/gabraham/flashpca	
Gorilla, version April 27, 2019	http://cbl-gorilla.cs.technion.ac.il/	(Gorilla, RRID: SCR_006848)
GSVA; version *1.26.0*	Bioconductor	
heatmaply; version *0.15.2*	CRAN (R 3.4.4)	
limma; version *3.34.9*	Bioconductor	(LIMMA, RRID: SCR_010943)
Pathview; version *1.18.2*	Bioconductor	
premessa R package	https://github.com/ParkerICI/premessa	
pvclust; version 2.0.0	CRAN (R 3.4.4)	
qvalue 2.10.0	Bioconductor	(Qvalue, RRID: SCR_001073)
Rtsne; version *0.13*	CRAN (R 3.4.4)	
scmap; version 1.1.5	Bioconductor	
scran; version *1.6.9*	Bioconductor	(scran, RRID: SCR_016944)
Seurat; version *2.3.2*	CRAN (R 3.4.4)	(Seurat, RRID: SCR_016341)
Survival; version 2.41.3	CRAN (R 3.4.4)	
TCGAbiolinks	Bioconductor	
UpSetR	CRAN (R 3.4.4)	
GraphPad Prism8, version 8.1.1		(GraphPad Prism, RRID: SCR_002798)
Fiji/ImageJ, 1.52n	https://fiji.sc	RRID: SCR_002285
FlowJo (version 8.8.6)	FlowJo, https://www.flowjo.com	(FlowJo, RRID: SCR_008520)
Other		
scRNA-seq interactive resource	This paper, https://www.vibcancer.be/software-tools/lungTumor_ECTax	

LEAD CONTACT AND MATERIALS AVAILABILITY

Correspondence and requests for materials should be addressed to the Lead Contact, Peter Carmeliet (peter.carmeliet@kuleuven.vib.be).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Patients

This study was approved by the local ethics committee (Medical Ethics Committee UZ/KU Leuven) under the protocol S57123. Only treatment-naïve patients were included for the analysis. scRNA-seq and CyTOF were performed on primary, non-metastatic non-small-cell lung carcinoma (NSCLC) tumors, resected from patients who underwent lung lobe resection with curative intent. For bulk RNA-seq and proteomics analysis, TEC and NEC samples were prospectively (2015-2019) collected from patients with lung, kidney, colorectal cancer, and colorectal cancer liver metastasis tumors. No statistical method was used to predetermine sample size. Details on the gender, age and clinical data of the patients included are shown in Table S1. Sample numbers per dataset are listed in Table S3.

Mice

Experiments were performed in 7 to 10 week-old immunocompetent C57BL6/J mice (male for tumor experiments; female for corneal cauterization-induced injury assay) or postnatal P5 neonates (retinal vasculature analysis) (obtained from the KU Leuven animal facility or purchased from Charles Rivers). They had not been involved in other, previous procedures. Animals were maintained in individually ventilated cages in a room with controlled temperature and humidity under a 12 hr light/12 hr dark cycle and with food and drink *ad libitum*. Animals were closely followed-up by the animal caretakers and the experimenters, with regular inspection by a veterinarian, as per the standard health and animal welfare procedures of the local animal facility. No statistical method was used to predetermine sample size. Animal housing and all experimental procedures were approved by the Institutional Animal Ethics Committee of the KU Leuven (Belgium) under protocol number P012/2014 and P156/2015.

Cell Lines and Primary Cells

Lewis Lung Carcinoma (LLC) Cells

LLC cells were purchased from ATCC and cultured in RPMI (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS) (Merck-Biochrom), and 100 IU/ml penicillin and 100 μg/ml streptomycin (Thermo Fisher Scientific).

293T Cells

293T cells were purchased from ATCC and cultured in DMEM supplemented with 10% fetal bovine serum (FBS) (Merck-Biochrom), 2 mM L-glutamine (Thermo Fisher Scientific) and 100 IU/ml penicillin and 100 μg/ml streptomycin (Thermo Fisher Scientific).

Primary Human Umbilical Vein Endothelial Cells (HUVECs)

HUVECs were freshly isolated from umbilical cords obtained from multiple donors (with approval from the Medical Ethical Committee KU Leuven / UZ Leuven and informed consent obtained from all subjects) as previously described (Schoors et al., 2015), and cultures were maintained as described in Method Details.

Primary Human Tumor and Peritumoral Endothelial Cells

Human tumor (hTEC) and peritumoral (hPNEC) ECs from biopsies of an NSCLC patient (Table S1) were isolated and cultured as described in Method Details. All cultures were maintained at 37°C and 5% CO_2, and were regularly tested for mycoplasma.

METHOD DETAILS

Human Tumor and Peritumoral Normal Endothelial Cell Isolation

Following surgical resection, samples from tumor and adjacent tissue (as far away from the tumor border as possible) were taken and transported to the research facility. Upon arrival, samples were rinsed with PBS, minced into pieces smaller than 1 mm^3, and transferred to 5 ml digestion medium consisting of Dulbecco's Modified Eagle medium (4,500 mg/l glucose), supplemented with penicillin/streptomycin (100 U/ml) (Thermo Fisher Scientific), sodium pyruvate (1 mM), MEM NEAAs (1x), MEM vitamin solution (1x), glutamine (2 mM), heparin (10 U/ml) (PromoCell), ECGF (bullet kit of EGM2; PromoCell) and containing 0.1% collagenase I, 0.1% collagenase II (Thermo Fisher Scientific), DNase I (75 U/ml) and dispase (2.5 U/ml). Samples were incubated for 30 min at 37°C with manual shaking every 5 min. Next, 10 ml of cold PBS containing 0.1% bovine serum albumin (BSA) (Sigma-Aldrich) were added and the samples were filtered using a 100 μm strainer (Corning). Following centrifugation (1,200 rpm) at room temperature (RT) for 5 min, the supernatant was decanted and discarded. Single-cell suspensions were enriched for ECs using the MACS system (Miltenyi Biotec) by $CD45^+$ cell depletion and subsequent enrichment for $CD31^+$ cells according to the manufacturer's instructions. For further processing for scRNA-seq, the samples were resuspended in PBS containing 0.4% UltraPure BSA (50 mg/ml; Thermo Fisher Scientific) and filtered over 40 μm cell strainers on ice. The number of cells and fraction of live cells in the suspension was counted using an automated cell counter (Luna, Logos Biosystems). The volume of suspension containing the required number of live cells was used for scRNA-seq as described below. Throughout the dissociation procedure, cells were maintained on ice whenever possible.

Primary Endothelial Cell Culture

Human Umbilical Vein Endothelial Cell Isolation and Culture

Human umbilical vein endothelial cells (HUVECs) were maintained on 0.1% gelatin-coated dishes in M199 medium (1 mg/ml D-glucose) (Thermo Fisher Scientific) supplemented with 20% fetal bovine serum (FBS) (Merck-Biochrom), 2 mM L-glutamine (Thermo Fisher Scientific), Endothelial Cell Growth Supplement (ECGS)/ Heparin (PromoCell), 100 IU/ml penicillin and 100 μg/ml streptomycin (Thermo Fisher Scientific) or in endothelial cell basal medium (EGM2) (PromoCell) supplemented with endothelial cell growth medium supplement pack (PromoCell). In all experiments, HUVECs were always used as single-donor cultures and were used between passage (p) 1 and 4.

Human Tumor and Peritumoral Endothelial Cell Culture

Human hTEC and hPNECs, isolated as described above, were resuspended in 6 ml of a 1:1 mix of M199 (containing 10% FBS and sodium pyruvate, MEM NEAAs, MEM vitamin solution, glutamine and heparin as above) and EGM2 medium, further supplemented with antibiotic/antimycotic (2x) (Thermo Fisher Scientific), and the single cell suspension was plated out in 3 wells of a 6-well plate pre-coated with 0.1% gelatin. The next day, the medium was changed to EGM2 supplemented with antibiotic/antimycotic and thereafter changed every other day. When reaching confluency and upon detection of EC cell colonies (patches with cobblestone appearance), ECs were purified using anti-CD31 coated magnetic beads (MACS Technology, Miltenyi Biotec). The resulting MACS-purified ECs were further cultured in EGM2 medium and used at passage 2-5 for bulk RNA-sequencing and proteomics, and passage 8 for scRNA-seq. Treatment of EC cultures with the PLOD inhibitor minoxidil (Sigma-Aldrich) was done at 100 and 500 μM for 24 hr in fully supplemented EGM2 medium.

Mouse Model of Lewis Lung Carcinoma & Tumor and Normal EC Isolation

LLC Model

The syngeneic LLC cells (ATCC) were injected orthotopically into the parenchyma of the right lung through the rib cage using a 30G needle (1×10^6 cells in 70 μl matrigel (BD Biosciences)). The animals were sacrificed on day 10 of the experiment and lungs were immediately processed for TEC and NEC isolation. Since both lungs become colonized by cancer cells in this orthotopic LLC model, even upon injection of cancer cells into a single lung, and since cancer cells in a tumor can affect peritumoral ECs, we isolated ECs from healthy non-tumor-bearing mice as controls (which we termed mNECs).

AAT Treatment

Mice were treated with rat anti-mouse VEGFR2 antibody DC101 (Bio-Connect) (40-50 mg/kg i.p., 3x per week), the VEGFR tyrosine kinase inhibitor PTK787 (Novartis Pharma) (100 mg/kg by daily gavage) or with, respectively, irrelevant control antibody 1C8 (Schmidt et al., 2011) (50 mg/kg i.p., 3x per week) or 5% DMSO and 1% Tween 80 in water (daily gavage) starting from day 3 after LLC injection (Liu et al., 2005; Sweeney et al., 2002). Mice were randomly allocated to the treatment condition.

EC Isolation

Mice were anesthetized and once the withdrawal reflex was absent in the pelvic limbs, they were perfused with 5 ml of PBS at a perfusion rate of 1 ml/min. For isolation of TECs, lung tumors were micro-dissected from lung tissue. For isolation of NECs, both lungs of control (non-tumor bearing) mice were dissected. The dissected tissue was placed into a C-tube (Miltenyi Biotec) containing 5 ml digestion buffer (0.1% collagenase II (Thermo Fisher Scientific), 0.25% collagenase type 4 (Worthington Biochemical), DNase I (150 U/ml) in DMEM supplemented with 1x sodium pyruvate, 1x MEM NEAAs, ECGF/Heparin, antibiotic/antimycotic (2x) and 1% penicillin/streptomycin (Thermo Fisher Scientific)) and processed with the m_lung_01 program of gentleMACS (Miltenyi Biotec). The samples were then incubated in a water bath at 37°C for 30 min with manual shaking every 5 min. At the end of digestion, the samples were processed with the m_lung_02 program of gentleMACS, and the reaction was stopped by adding 10 ml of an isolation buffer containing 1x PBS, 0.1% BSA, 5 mM EDTA. Subsequently, the cell suspension was filtered through a 100 μm cell strainer (Corning) on ice. Finally, ECs were enriched by magnetic bead sorting using MACS system (Miltenyi Biotec) at RT. Immune cells were depleted using selection with CD45 MicroBeads (Miltenyi Biotec), followed by positive selection of ECs with CD31 MicroBeads (Miltenyi Biotec), according to the manufacturer's procedures. On average the isolation time was 3 hr, the viability between 60-90%, and cells were kept on ice whenever possible.

Knock Down Strategy

To silence the expression of *PLOD1*, *PLOD2*, *LOXL2*, gene-specific oligonucleotides (see Key Resources Table and Table S8) were cloned into the pLKO-shRNA2 vector (Clontech). To silence the expression of *LXN*, gene-specific oligonucleotides were cloned into the pLVX-shRNA2 vector (Clontech). All constructs were sequence verified. Lentiviral particles were produced in 293T cells as previously described (Cantelmo et al., 2016). For lentiviral transduction, a MOI of 20 was used. Transductions were performed on day 0 in the evening, cells were refed with fresh medium on day 1 in the morning and experiments were performed from day 3 or 4 onwards. Knockdown efficiency was monitored for each experiment either at the mRNA (qRT-PCR) or protein level or both and compared to expression in cells transduced with a negative control shRNA (see Key Resources Table and Table S8). To silence the expression of *FSCN1* with siRNA (Table S8) we used the Lipofectamine RNAiMAX (Thermo Fisher Scientific) as per the manufacturer's instructions. Cells were seeded at day 0 in the morning (with 200 K/well using a 6-well plate format) and transfection was performed the same day in the evening with 25 pmol siRNA per well, diluted in Opti-MEM. On day 1 in the morning, medium was refreshed, and in the afternoon cells were collected for use in the mosaic spheroid capillary sprouting assay.

In Vitro Functional Assays

Proliferation

EC proliferation was quantified by incubating cells for 2 hr with 1 μCi/ml [^{3}H]-thymidine (Perkin Elmer). Thereafter, cells were fixed with 100% ethanol for 15 min at 4°C, precipitated with 10% TCA and lysed with 0.1 N NaOH. The amount of [^{3}H]-thymidine incorporated into DNA was measured by scintillation counting.

Scratch Wound Migration Assay

A scratch wound was applied on confluent EC monolayers using a 200 μl tip, 24 hr after seeding (100,000 cells per well in 24-well plates). After scratch wounding and photography (T0), the cultures were further incubated in fully supplemented EGM2 medium for 18 hr and photographed again (T18). Migration was measured with the *fiji/ImageJ* software package (Schindelin et al., 2012) and is expressed as % wound closure (gap area at T0 minus gap area at T18 in % of gap area at T0).

Spheroid Capillary Sprouting Assay

ECs were incubated overnight in hanging drops in EGM2 medium containing methylcellulose (20 vol% of a 1.2% solution of methylcellulose 4000 cP) (Sigma-Aldrich) to form spheroids. Spheroids were then embedded in collagen gel and cultured for 20 hr to induce sprouting as described previously (De Bock et al., 2013). Cultures were fixed with 4% paraformaldehyde (PFA) at RT and imaged under bright field using a Motic AE 31 microscope (Motic Electric Group Co Ltd.) or Leica DMI6000 microscope (Leica Microsystems). Analysis of the number of sprouts per spheroid and the total sprout length (cumulative length of primary sprouts and branches per spheroid) was done on phase contrast images using the *fiji/ImageJ* imaging software package.

Mosaic Spheroid Capillary Sprouting Assay

Control and silenced ECs (lentiviral or siRNA) were collected as described in the knock down strategy section, and fluorescently labeled with intracellular dyes. Control ECs were stained with the CellTracker Deep Red (Thermo Fisher Scientific) and *FSCN1* silenced ECs with the CellTrace CFSE dye (Thermo Fisher Scientific). *LXN* silenced cells did not need labeling as the shRNA viral vector contains a ZsGreen1 fluorescent reporter. The staining protocol was done according to the manufacturer's guideline. Briefly, suspensions containing 125,000 control and silenced ECs were placed in a separate tube, spun down and the media changed to EGM2 without growth factors and supplement during the incubation time with the dye (10 μM for CellTracker Deep Red, 50 μM of the CellTrace CFSE), 30 min, room temperature, protected from the light. Full EGM2 was then added to the staining mix to stop the reaction, the cells centrifuged and the supernatant removed. Finally, control (Deep Red) and silenced (CFSE or ZsGreen1) cells

were mixed at equal ratio (1:1, 250,000 cells in total) and used for spheroid formation as described in the section above. Using a Leica DMI6000 microscope (Leica Microsystems) at least 10 spheroids were acquired per replicate and per condition. Using the *fiji/ImageJ* imaging software package, the contribution to the tip position of each sprout was quantified and is represented as the percentage of green (CFSE or ZsGreen1) or red (Deep Red) stained ECs occupying the tip position.

In Vivo Functional Assays in Corneal Angiogenesis

Corneal angiogenesis was induced by thermal cauterization as previously described (García-Caballero et al., 2019). After anaesthetizing 8-week-old female mice with an intraperitoneal injection of ketamine (100 mg/kg body weight) and xylazine (10 mg/kg body weight), a local anesthetic (Unicaine 0.4%; Théa Pharma) was applied to the eye and the central cornea was thermally cauterized using an ophthalmic cautery (Optemp II V; Alcon Surgical). Mice were daily treated with ophthalmic drops of Minoxidil (5% w/w), 3-beta-aminopropionitrile (800 mg/kg body weight), or the vehicle for 6 days, initiated the day after induction of the corneal injury. Mice were euthanized 7 days after cauterization, eyes were removed and corneas were dissected. Whole-mounted corneas were fixed in 70% ethanol for 1 hr at room temperature and blocked with 3% BSA-3% Gloria milk (Nestlé) for 1 hr. Corneas were incubated overnight with a rat anti-mouse CD31 antibody (BD Biosciences), subsequently incubated with AlexaFluor 546-conjugated goat anti-rat secondary antibodies (Molecular Probes) for 2 hr, and flat-mounted on a microscopy slide with ProLong Gold antifade reagent (Thermo Fisher Scientific). Corneas were imaged using a Leica DMI6000 microscope (Leica Microsystems) and the blood vessel area was quantified ($CD31^+$ area as a percentage of the total corneal area).

Immunohistochemistry

Lung Tissue

Formalin-fixed paraffin-embedded murine or human lung tumor tissue sections were subjected to immunohistochemistry. For a full list of primary and secondary antibodies, see Key Resources Table. Briefly, after incubation overnight at 4°C with the primary antibodies, sections were incubated with the appropriate secondary antibodies followed by amplification with the proper tyramide signal amplification system (Perkin Elmer) (for CD31 staining), or with Alexa Fluor-conjugated secondary antibodies (for all marker stainings, except MECA-79 and CXCR4/PLGF, no amplification was used to avoid loss of differential signal intensity between vessel/EC types), also overnight at 4°C. Nuclei were counterstained with Hoechst 33342 (Sigma-Aldrich) and slides were mounted using ProLong Gold Antifade Mountant (Thermo Fisher Scientific). Imaging was performed using a Zeiss AxioScan Z1 at 20x magnification, or by confocal imaging using a Zeiss LSM 780 confocal microscope (Carl Zeiss) at 100x magnification (alpha Plan-Apochromat 100x/1.46 Oil DIC M27).

Neonatal Retinas

Retinas were isolated at P5 as previously described (De Bock et al., 2013) and fixed in 2% paraformaldehyde for 2 hr. Briefly, after overnight permeabilization of the retinas, the samples were incubated overnight at 4°C with the primary antibody (LXN Abcam 1/100; Fascin Abcam, 1/100). After incubation with the primary antibodies, retinas were incubated overnight at 4°C with the appropriate secondary antibodies and IsolectinB4-Alexa 488 (Thermo Fisher Scientific, 1/100). Stained retinas were flatmounted and imaged with a Zeiss LSM 780 confocal microscope (Carl Zeiss) at 20x magnification (Fluar 20x/0.75).

RNA Isolation and Quantitative RT-PCR

RNA was collected and purified with the PureLink RNA Mini Kit (Thermo Fisher Scientific) and converted to cDNA using the iScript cDNA synthesis kit (Bio-Rad). RNA expression analysis was performed by Taqman quantitative RT-PCR (Thermo Fisher Scientific) as described using premade primer sets (see Key Resources Table and Table S8). For comparison of gene expression between conditions, mRNA levels (normalized to the housekeeping gene HPRT) were expressed relative to control conditions.

Single-cell Droplet-based RNA Sequencing

The single cell suspensions of freshly isolated TECs and NECs (MACS-bead enriched to 20-30% as described above) or cultured TECs were converted to barcoded scRNA-seq libraries using the Chromium Single Cell 3' Library, Gel Bead & Multiplex Kit and Chip Kit (10x Genomics), aiming for 6,000 cells per library. Samples were processed using kits pertaining to V2 barcoding chemistry of 10x Genomics. Single samples were always processed in a single well of a PCR plate, allowing all cells from a sample to be treated with the same master mix and in the same reaction vessel. For each experiment, all samples (TEC and NEC) were processed in parallel in the same thermal cycler. Libraries were sequenced on an Illumina HiSeq4000, and mapped to the human genome (build GRCh38) or to the mouse genome (build mm10) using CellRanger software (10x Genomics, version 2.1.1).

Single-Cell Transcriptomics Analysis

Data from MACS-enriched TEC and NEC samples were aggregated using *Cell Ranger* and data from the raw unfiltered matrix was further processed using R (version 3.4.4 - *Someone to Lean On*).

Quality Control and Data Normalization

For the human freshly-isolated ECs, the following quality control steps were applied: (i) genes expressed by <50 cells or with a row mean of <0.01 were not considered; (ii) cells that had either fewer than 201 (low quality cells) or over 6,000 expressed genes (possible doublets), or over 5% of unique molecular identifiers (UMIs) derived from the mitochondrial genome were removed. The data of the remaining 100,512 cells were natural-log transformed using log1p and normalized using the *Seurat* package (Satija et al., 2015).

Quality control and normalization of murine and cultured ECs were performed using the same approach. Dataset specific cut-off values and parameter settings are listed in Table S8.

In Silico *EC Selection*

After auto-scaling, the normalized data were first summarized by principal component analysis (PCA) using the *flashPCA* package, and the first 15 PCAs were visualized using t-Distributed Stochastic Neighbor Embedding (t-SNE, *Rtsne* package) with a perplexity value of 200 and a learning rate of 100. Graph-based clustering was performed to group cells according to their gene expression profiles as implemented in *Seurat* (see Table S8 for parameter settings). Cell clusters were annotated based on canonical markers, including *PECAM1* and *CDH5* (ECs), *PROX1* (lymphatic ECs), *COL1A1* (fibroblasts), *PTPRC* (leukocytes), and *PDGFRB* (pericytes) to discriminate ECs from contaminating cells. All down-stream analyses were performed on ECs only.

Batch Effect Correction

ECs from freshly-isolated human and murine samples were prospectively collected, resulting in several batches of sequencing data. We first analyzed each patient and / or batch separately, and removed clusters expressing ambiguous marker genes (i.e. representing low quality cells, red blood cells, possible doublets, etc.), as performed in other scRNA-seq studies (Cao et al., 2019; Smillie et al., 2019; Young et al., 2018). We then used a recently developed algorithm, *mnnCorrect* available from the *scran* package (Haghverdi et al., 2018), for batch effect correction of: (i) all human samples that were used to reconstruct the NSCLC EC taxonomy; and (ii) all murine mNECs and mTECs. The optimal neighborhood size, k, was empirically defined as 50 after optimizing over a range of 10-300 for both human and murine datasets. Importantly, differential analysis and GSVA analyses were based on the original expression data, using additive linear models to model the batch of origin (for murine samples) or patient (for human samples) as a blocking factor (*limma* package) (Haghverdi et al., 2018; Smillie et al., 2019).

Feature Selection and Dimensionality Reduction

After *in silico* EC selection and batch correction for joint analysis, we identified genes with high variability using the Seurat *FindVariableGenes* function. This function calculates the mean expression and dispersion for each gene, then groups genes into bins (of size 20) by their mean expression and identifies any gene for which the z-score calculated from the dispersion exceeds a pre-defined cut-off. For most experiments, we used a cut-off of z=0.25 and mean expression in the range 0.00125 to 8, all other parameters were default (see Table S8 for parameter settings for each analysis). The normalized data were auto-scaled and principal component analysis was performed on variable genes, followed by t-SNE to construct a two-dimensional representation of the data. This representation was only used to visualize the data.

EC Cluster Identification

To estimate the number of distinct phenotypes in batch corrected data, we color-coded t-SNE plots for each of the ~14,000 detected genes using an in-house developed R/Shiny-based web tool and identified clusters of cells with discriminating gene expression patterns in all datasets. To unbiasedly group ECs, we performed PCA on highly variable genes, and used graph-based clustering as implemented in the *FindClusters* function of the *Seurat* package (Satija et al., 2015). Cluster results were visualized using t-SNE to verify that all visually identified clusters were captured and not under-partitioned. Over-partitioned clusters that represent the same biological phenotype were merged into a single cluster.

We adapted previously described methods to build taxonomic trees (Innes and Bader, 2018; Tasic et al., 2018). We first compiled the top 10 marker genes of each cluster (see below) into a marker list. We calculated the mean of these marker genes and applied hierarchical clustering with Euclidean distance and average linkage. The confidence of each branch of the tree was estimated by the bootstrap resampling approach from the R-package *pvclust* (Suzuki and Shimodaira, 2006). Second, to ensure that biologically relevant branches that could not be directly resolved by bootstrapping with a confidence score of >0.4 were statistically separable (e.g. human tumor and peritumoral LECs, postcapillary veins, and tip and immature cells), we performed pair-wise differential analysis and confirmed that these clusters had at least ten genes that exceed a 0.2 $\log_2$fold enrichment with a FDR corrected p value <0.05 (Table S3). Details of clustering parameters are provided in Table S8.

Pair-Wise Differential Analysis

Differential expression analysis between two specific clusters was performed using *limma* (Ritchie et al., 2015). As described previously (Haghverdi et al., 2018), we parameterized the design matrix such that each batch/patient–cluster combination formed a separate group in a one-way layout, by using the clusters derived from the batch-corrected data (described above). For each gene, we used this design to (a) fit a linear model to the normalized uncorrected log expression values, (b) perform an empirical Bayes shrinkage to stabilize the sample variances, and (c) compute a moderated t test by comparing the two clusters across all batches. Differentially expressed genes were defined as those for which the $\log_2$ fold change exceeds 0.2 at a false discovery rate of 5% (Table S4).

Marker Gene Analysis

We used a two-step approach to obtain ranked marker gene lists for each cluster. As a first criterion, marker genes for a given cluster should have the highest expression in that cluster compared to all other clusters and are therefore uniquely assigned to one cluster. Second, we ranked marker genes using a product-based meta-analysis (Hong et al., 2006). Briefly, we performed pair-wise differential analysis of all clusters against all other clusters separately and ranked the results of each pair-wise comparison by $\log_2$ fold change. The most upregulated genes received the lowest rank number (top ranking marker genes) and the most downregulated genes received the highest rank number. For each cluster, we combined the rank numbers for all genes in all pair-wise comparisons by calculating their product to obtain a final list of ranked marker genes for each cluster. To assess statistical significance, we used a recently developed algorithm to determine accurate approximate p values for each marker gene based on the rank product statistic

(Heskes et al., 2014), and obtained Benjamini-Hochberg adjusted p values using the R package *qvalue* (Storey and Tibshirani, 2003) (Table S4). For murine samples, marker gene analysis was performed using gene expression in untreated samples only, to avoid treatment-induced effects.

Cluster Quantification

The representation of clusters across patients, tumor type, and tumor and peritumoral tissue was quantified using *limma,* using patient as a covariate and using a p value cutoff of < 0.05 to call differences in fractional composition.

Cluster Annotation

We annotated clusters based on literature-curated marker genes of canonical artery, capillary, vein, lymphatic, proliferating, high endothelial venule and tip ECs. In case of an entirely unknown phenotype or previously unrecognized sublineages of a canonical phenotype, which could not be annotated based on canonical marker genes or gene sets, we used a three-step approach to identify a putative biological function. First, we searched through the top 50 ranking list of markers for a coherent set of genes involved in similar biological processes. Second, if we identified a putative signature (e.g. antigen presentation, scavenging, basement remodeling, etc.), we determined whether other genes associated with such a signature were also highest expressed in this phenotype. Third, we integrated insights from additional analysis (e.g. cross-model comparison, color-coded tSNEs, heatmap analysis) into our assessment. For the murine HEV signature, we used Gene set variation analysis (GSVA) as implemented in the *GSVA* R-package (version 1.26.0) to convert the gene-by-cell matrix into a HEV gene set-by-cell matrix using default settings (Hanzelmann et al., 2013). Enrichment scores for the HEV signature were highest in postcapillary venous mTECs.

Cells that could not be unambiguously assigned to a biologically meaningful phenotype might represent low quality cells or doublets and were excluded from the analysis. In human ECs, we excluded clusters of lowly sequenced cells that expressed recently identified markers of ambient RNA contamination (Angelidis et al., 2019). In murine ECs, a small fraction of tumor sample-derived ECs had the same phenotype (gene expression signature) as mNECs (Figure S5E) and were considered peritumoral ECs, given that micro-dissected tumor samples unavoidably contained some peritumoral ECs; these cells were similar to, and clustered with mNECs and were not further characterized as a separate population.

Evaluation of Dissociation Artifacts

We performed gene set variation analysis to determine which individual cells strongly expressed a recently published dissociation gene signature, consisting mainly of immediate-early and other stress response genes (van den Brink et al., 2017). In one murine NEC sample, we identified a cluster consisting of cells that were strongly enriched for this signature but did not express any other marker genes that could biologically explain a stress response. Cells in this cluster were therefore removed from all analyses.

Heatmap Analysis

All heatmaps are based on cluster-averaged gene expression to account for cell-to-cell transcriptomic stochastics. Data was autoscaled for visualization. Heatmaps were produced using the *heatmaply* package (version 0.15.2). The datamatrix for each heatmap can be downloaded from the accompanying web tool (see Data Resources below).

Cross Dataset Validation of the Human EC Taxonomy

We used the *scmapCluster* algorithm as implemented in the *scmap* package (version 1.1.5) to project normalized and quality filtered EC data (Kiselev et al., 2018). *scmapCluster* carries out a search by cluster, in which each cluster is represented by its centroid, and measures the similarity between a new cell, c, and each cluster centroid or cell. We used the top ten marker genes for each EC cluster in the reference taxonomy and a similarity threshold of 0.5, all other parameters were default. The projection was visualized using a Sankey plot.

TCGA Analysis

Raw count gene expression data was obtained from the lung squamous cell carcinoma and lung adenocarcinoma NSCLC cohorts catalogued in The Cancer Genome Atlas (TCGA, accessed through *TCGAbiolinks* package) (Cancer Genome Atlas Research Network et al., 2013; Colaprico et al., 2016), and normalized using trimmed mean of M-values (TMM) available from the *edgeR* package (Robinson et al., 2010). Marker gene sets of each phenotype in human TECs and NECs were filtered for EC-enriched genes utilizing a publicly available resource (Lambrechts et al., 2018) (EC-enriched genes were defined as all genes that were highest expressed in human lung TECs compared to all other cell types described in this resource). GSVA scores were only calculated for EC-enriched marker gene sets with a minimum of five detected genes, all other parameters were default. Analysis was performed on the output of GSVA to test gene expression signatures. Patients were stratified in high (top 25%) and low (bottom 75%) expression groups, and survival analysis was performed via the Kaplan-Meier estimator, using the log-rank test to determine significance (Li, 2003).

Jaccard Similarity Analysis

To assess conservation of cell phenotypes in lung ECs, we calculated similarity of marker gene sets using pair-wise Jaccard similarity coefficients for all clusters consisting of at least 50% TECs against all other clusters. The Jaccard similarity coefficient is defined as the size of the intersection divided by the size of the union of sets:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A| \cap |B|}$$

Where J is the Jaccard index and A and B are two sets of marker genes (Levandowsky and Winter, 1971).

Congruent Tip Cell Marker Analysis

To identify congruent tip cell markers across freshly-isolated human and murine and cultured TECs, we first selected all genes that were highest expressed in tip and breach cells in all three datasets. Second, we ranked these conserved genes via a rank-product meta-analysis, by calculating the product of the rank numbers of each gene in each of the three datasets (Heskes et al., 2014; Hong et al., 2006), and corrected (Benjamini-Hochberg) the associated p values using the R package *qvalue* (Storey and Tibshirani, 2003). The results were visualized using bar plots, depicting the top 50 most congruent tip cell marker genes with an adjusted p value < 0.005.

Gene Set Enrichment Analysis

We used gene set enrichment analysis (GSEA) as implemented in the *clusterProfiler* package (version 3.6.0) to compare gene expression signatures between treated and untreated mTEC samples (Yu et al., 2012). Gene set analysis was performed using a set of vascular related gene sets selected from the Molecular Signatures Database (MSigDB version 5.2 downloaded from http://bioinf.wehi.edu.au/software/MSigDB/), a collection of expert annotated gene sets. GSEA scores were calculated for sets with a minimum of 10 detected genes, all other parameters were default.

Meta-Analysis of Transcriptomics Data

We performed a differential gene expression meta-analysis using previously published human tumor ECs (TEC) versus normal ECs (NEC) transcriptomics. Briefly, we screened the EndoDB database for relevant studies and identified three studies comprising six distinct TEC versus NEC datasets of five tumor types (Khan et al., 2019). We performed pair-wise, TEC versus NEC, differential expression analysis for each dataset independently as described previously (Bruning et al., 2018; Cantelmo et al., 2016). Differentially expressed genes and their false discovery rate (FDR) corrected p values were identified by the *limma* package (Ritchie et al., 2015). We ranked genes in each dataset by fold change; genes upregulated in TECs received the lowest rank, downregulated genes the highest rank. We combined the rank numbers for all genes using a product-based meta-analysis approach (Heskes et al., 2014; Hong et al., 2006), and calculated Benjamini-Hochberg adjusted p values using the R package *qvalue* (Storey and Tibshirani, 2003). For visualization purposes only, we plotted the scaled median ranks.

RNA-Sequencing Analysis

RNA was extracted using TRIzol (Thermo Fisher Scientific). Starting from 1 μg total RNA, poly-adenylated fragments were isolated, reverse transcribed and converted into indexed sequencing libraries using the KAPA stranded mRNA-seq kit (Sopachem). The first 50 bases of these libraries were sequenced on a HiSeq 2500 system (Illumina). The raw sequenced reads were mapped to the human reference transcriptome and genome (build GRCh38) using the Bowtie TopHat pipeline (Langmead and Salzberg, 2012). Mapped reads were assigned to ensemble gene IDs by HTSeq. Differentially expressed genes and their false discovery rate (FDR) corrected p values were identified by the *Limma* package (Ritchie et al., 2015) using the patient factor as a covariate.

Proteomics Sample Preparation

Human TEC and NEC were isolated from patients and cultured until passage 1, at which point 37,500 cells were washed twice with ice cold PBS and scraped in 300 μl of a 50:30:20 (methanol: acetonitrile: 20 mM Tris, pH 9.3) extraction buffer. Samples were then centrifuged for 5 min at 15,000 rcf at 4°C and the supernatant was discarded. The resulting protein pellet was re-dissolved in 50 μl 8 M urea, 20 mM HEPES pH 8.0 and proteins were reduced by addition of 15 mM DTT and incubation for 30 min at 55°C. Proteins were then alkylated by addition of 30 mM iodoacetamide and incubation for 15 min at RT in the dark. The samples were diluted by addition of 20 mM HEPES pH 8.0 to reach a final urea concentration of 4 M; then, the proteins were digested with 0.125 μg endoLysC (Wako 1/250, w/w) for 4 hr at RT. Next, all samples were further diluted by addition of 20 mM HEPES pH 8.0 to a final urea concentration of 2 M and the proteins were digested with 0.125 μg trypsin (Promega) (1/100, w/w) overnight at 37°C. The resulting peptides were then purified on OMIX C18 tips (Agilent), dried completely by vacuum drying and stored at -20°C until LC-MS/MS analysis.

Proteomics LC-MS/MS and Data Analysis

Peptides from each sample were re-dissolved in 20 μl loading solvent A (0.1% TFA in water/acetonitrile (98 : 2, v/v)) of which 10 μl was injected for LC-MS/MS analysis on an Ultimate 3000 RSLCnano system (Thermo Fisher Scientific) in-line connected to a Q Exactive HF mass spectrometer (Thermo Fisher Scientific) equipped with a Nanospray Flex Ion source (Thermo Fisher Scientific). Trapping was performed at 10 μl/min for 4 min in solvent A on a home-made 100 μm internal diameter (I.D.) × 20 mm trapping column (5 μm beads, C18 Reprosil-HD, Dr Maisch) and peptides were separated on a reverse-phase column (made in-house, 75 μm I.D. × 400 mm, 1.9 μm beads C18 Reprosil-HD, Dr Maisch). The peptides were eluted by a non-linear increase from 2% to 56% MS solvent B (0.1% FA in water/acetonitrile (2 : 8, v/v)) over 140 min at a constant flow rate of 250 nl/min. The column temperature was kept constant at 50°C (CoControl 3.3.05, Sonation).

The mass spectrometer was operated in data-dependent mode, automatically switching between MS and MS/MS acquisition for the 16 most abundant ion peaks per MS spectrum. Full-scan MS spectra (375 to 1500 m/z) were acquired at a resolution of 60,000 in the Orbitrap analyzer after accumulation to a target value of 3,000,000. The 16 most intense ions above a threshold value of 13,000 were isolated (window of 1.5 Th) for fragmentation at a normalized collision energy of 28% after filling the trap at a target value of 100,000 for maximum 80 ms. MS/MS spectra (200 to 2,000 m/z) were acquired at a resolution of 15,000 in the orbitrap analyzer. The S-lens RF level was set at 55, and we excluded precursor ions with single and unassigned charge states from fragmentation selection.

Data analysis was performed with MaxQuant (version 1.6.0.16) (Cox and Mann, 2008) using the Andromeda search engine with default search settings, including a false discovery rate set at 1% on peptide spectrum match (PSM), peptide and protein level. The spectra were searched against the human proteins in the UniProt/Swiss-Prot database (database release version of September

2017 containing 20,237 human protein entries, downloaded from www.uniprot.org). The mass tolerances for precursor and fragment ions were set to 4.5 and 20 ppm, respectively, during the main search. Enzyme specificity was set as carboxy-terminal to arginine and lysine (trypsin), also allowing cleavage at arginine/lysine–proline bonds with a maximum of two missed cleavages. Carbamidomethylation of cysteine residues was set as a fixed modification and variable modifications were set to oxidation of methionine (to sulfoxides) and acetylation of protein amino-termini. Proteins were quantified by the MaxLFQ algorithm integrated in the MaxQuant software. Only proteins with at least one unique or razor peptide were retained for identification, while a minimum ratio count of two was required for quantification. Matching between runs was enabled, with a matching time window of 2 min and an alignment time window of 20 min.

Further data analysis was performed in R after loading the proteinGroups results files from MaxQuant. Proteins with less than 20% valid values were removed, and missing values were imputed using minimum values. Differentially expressed proteins and their false discovery rate (FDR) corrected p values were identified by the *Limma* package (Ritchie et al., 2015), non-corrected p values are reported for targeted analysis of collagen modification enzymes. We used the *UpsetR* package to visualize proteins consistently higher expressed across tumor types (Conway et al., 2017), and calculated rank products statistics to assess significance (Heskes et al., 2014) (Table S4). The subset of 288 proteins that were higher expressed in all four tumor types were further ranked using Fisher's combined probability test (as implemented in the *metap* R-package) to obtain combined p values across datasets. This ranked list was used for gene ontology analysis using the *Gorilla* web tool using default settings (Eden et al., 2009).

RNAscope *In Situ* Hybridization and Quantification

Formalin-fixed paraffin-embedded human NSCLC tissue blocks were sectioned and subjected to RNAscope *in situ* hybridization using the RNAscope Multiplex Fluorescent v2 assay (ACDBio) according to the manufacturer's instructions (Pretreatment and RNAscope Multiplex Fluorescent v2 Assay according to protocol 323100-USM). Briefly, after deparaffinization, the slides were incubated with hydrogen peroxide for 10 min at RT. After several washing steps, manual target retrieval was performed followed by incubation with Protease Plus before proceeding to the RNAscope Multiplex Fluorescent v2 protocol. Hybridization was performed with the RNAscope probes Hs-PGF, Hs-CXCR4, Hs-ACKR1, Hs-EFNB2, Hs-CCL14, Hs-CD52, Hs-CD68, Hs-IL1RL1, Hs-EDNRB and the RNAscope 3-plex Positive (low expression Polr2a, medium expression PPIB, and high expression UBC) and Negative Control Probes. A negative control probe targeting a bacterial gene was used to assess background. Slides were then processed according to the RNAscope Multiplex Fluorescent v2 protocol (Hybridization, Amplification, and Signal Development), prior to immunofluorescent staining for CD31. Images were acquired using a Zeiss LSM 780 confocal microscope (Carl Zeiss). For quantification, the *in situ* hybridization images were first converted to a thresholded image and numbers of particles per $CD31^+$ cell were counted using a circularity of 0.00 – 1.00, and a size pixel between 0 – ∞. Doublets and triplets were split by considering signal over the average size of a dot. Results are depicted as number of dots per cell as estimated by counting nuclei (Hoechst signal). Depicted data are representative tumor sections of n=3 NSCLC patients.

Time-of-Flight Mass Cytometry (CyTOF)

Custom conjugated antibodies were generated using MaxPar X8 antibody labeling kits as per the manufacturer's instructions (Fluidigm). Freshly isolated TEC and NEC single cell suspensions were washed twice in PBS followed by incubation with 0.5 mM Cell-ID Cisplatin (Fluidigm) for 3 min at RT. Reactions were quenched with 5 ml Cell Staining Buffer (Fluidigm), washed and fixed in 2% paraformaldehyde for 20 min at RT. Cells were then washed and permeabilized with freshly prepared 1X Barcode Perm Buffer in PBS (Fluidigm), followed by incubation with Palladium barcodes (Fluidigm) for 20 min at RT. Samples were subsequently washed with Cell Staining Buffer, resuspended with Fc receptor block for 5 min (TruStain FcX, BioLegend) and stained with a cocktail of surface staining antibodies for 1 hr at RT under constant rotation. Following washes with Cell Staining Buffer and PBS, samples were fixed for 30 min in Fix/Perm buffer (Thermo Fischer Scientific). After 2 washes with freshly prepared 1X Perm Buffer (Thermo Fischer Scientific), cells were stained with a cocktail of intracellular staining antibodies for 1 hr at RT under constant rotation. Cells were then washed twice in 1X Perm Buffer and once in PBS, followed by incubation with 1.6% paraformaldehyde and 0.5 μM Intercalator-Ir (Fluidigm) for 1 hr at RT. After multiple washes with Cell Staining Buffer, PBS and finally ultrapure water, EQ Four Element Calibration beads (Fluidigm) were added 1:10 to each sample followed by analysis on a Helios instrument (Fluidigm) at an event rate of 150–200 cells per second. CyTOF was performed on human NCSLC tissue samples in four patients independent from those included in the scRNA-seq analysis (Table S1).

CyTOF Data Analysis

Mass cytometry data were randomized using the Fluidigm acquisition software (version 6.7.1014) and normalized using four element calibration beads. Bead-normalized samples were debarcoded using the *premessa* R package (version 0.2.2) using default settings. The viable fraction (DNA^+, $Cisplatin^-$) of debarcoded single cells was then selected using FlowJo software (version 8.8.6), and viable cells were sequentially gated on $CD45^-$/$PDGFRB^-$/$COL1A1^-$/$ACTA2^-$ /$PECAM1^+$/$CDH5^+$ endothelial cells. The data was arcsine-normalized using the *asinh* function (*base R*), and visualized using t-SNE with a perplexity value of 35 and a learning rate of 100. The data were clustered using the graph-based clustering algorithm as implemented in *Seurat* (version) using the first 15 principal components, ten neighbors and a resolution of 1. The differential expression of marker genes in endothelial clusters was assessed using ANOVA with Dunnett's post-hoc test to determine which differences between clusters were significant.

Integration of scRNA-seq and CyTOF Data

We adapted a recently described approach to assess the similarity between the scRNA-seq and the CyTOF derived populations (Giordani et al., 2019), and obtained average gene expression for each scRNA-seq cluster and average protein expression values for each CyTOF cluster. We then scaled both the scRNA-seq and CyTOF derived expression matrix using the *scale()* function available from *base R* and merged them. Finally we used the *heatmaply* package to obtain a heatmap where both markers and clusters were rearranged based on hierarchical clustering with Euclidean distance and average linkage using CA4, CD36, HLA-DP isoforms, HLA-DR isoforms, ACKR1, CXCR4, VWF and PROX1 as markers. The similarity of each branch of the tree was estimated by the bootstrap resampling approach from the R-package *pvclust* (Suzuki and Shimodaira, 2006), branches that could not be resolved by bootstrapping with a confidence score of >0.4 were considered similar.

QUANTIFICATION AND STATISTICAL ANALYSIS

RNAscope data are representative tumor sections of n=3 NSCLC patients. Spearman correlation and two-sided statistical significance were calculated in GraphPad Prism8. A p value <0.05 was considered significant. Statistical significance of differences between Kaplan Meier survival curves was calculated using a Log Rank (Mantel Cox) test as implemented in the *Survival* package. A p value <0.05 was considered significant. A t test was used for all pair-wise comparisons, ANOVA with the appropriate posthoc test for multiple group comparisons.

DATA AND CODE AVAILABILITY

Data Resources

The accession number for all raw sequencing data are available in ArrayExpress under accession numbers ArrayExpress: E-MTAB-6308 (human scRNA-seq data), ArrayExpress: E-MTAB-7458 (murine scRNA-seq data), ArrayExpress: E-MTAB-8031 (human bulk RNA-seq). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PRIDE: PXD014123. Publicly available single cell transcriptome data from NSCLC was derived from ArrayExpress: E-MTAB-6149. Processed publicly available endothelial transcriptomics data was derived from the EndoDB, raw data can be obtained from GEO (accession number GEO: GSE77199) or ArrayExpress (accession numbers ArrayExpress: E-GEOD-41614, E-GEOD-51401). To ensure data accessibility to non-bioinformaticians, reproducibility, and resource value, we made our scRNA-seq data available for further exploration via an interactive webtool at https://www.vibcancer.be/software-tools/lungTumor_ECTax. With this tool users can interactively visualize gene expression and clustering on t-SNE, search marker genes for all subclusters, and export gene expression data.

Software

All software is freely or commercially available and is listed in the STAR Methods description and Key Resources Table.