



Endothelial transcriptomic, epigenomic and proteomic data challenge the proposed role for TSAd in vascular permeability

James T. Brash¹ · Guillermo Diez-Pinel¹ · Luca Rinaldi² · Raphael F. P. Castellan¹ · Alessandro Fantin² · Christiana Ruhrberg¹

Received: 8 October 2024 / Accepted: 23 February 2025
© The Author(s) 2025

Abstract

The vascular endothelial growth factor VEGF drives excessive vascular permeability to cause tissue-damaging oedema in neovascular and inflammatory diseases across multiple organs. Several molecular pathways have been implicated in VEGF-induced hyperpermeability, including binding of the VEGF-activated tyrosine kinase receptor VEGFR2 by the T-cell specific adaptor (TSAd) to recruit a SRC family kinase to induce junction opening between vascular endothelial cells (ECs). Inconsistent with a universal role for TSAd in permeability signalling, immunostaining approaches previously reported TSAd only in dermal and kidney vasculature. To address this discrepancy, we have mined publicly available omics data for expression of TSAd and other permeability-relevant signal transducers in multiple organs affected by VEGF-induced vascular permeability. Unexpectedly, TSAd transcripts were largely absent from EC single cell RNAseq data, whereas transcripts for other permeability-relevant signal transducers were detected readily. TSAd transcripts were also lacking from half of the EC bulk RNAseq datasets examined, and in the remaining datasets appeared at low levels concordant with models of leaky transcription. Epigenomic EC data located the TSAd promoter to closed chromatin in ECs, and mass spectrometry-derived EC proteomes typically lacked TSAd. By suggesting that TSAd is not actively expressed in ECs, our findings imply that TSAd is likely not critical for linking VEGFR2 to downstream signal transducers for EC junction opening.

Keywords RNA-seq · Leaky gene expression · Active gene expression · VEGF · SH2D2A · TSAd · Endothelial cell · Vascular hyperpermeability

Introduction

The vascular endothelial growth factor VEGF (VEGFA) is one of the most widely studied proteins [1] due to its crucial roles in cardiovascular development, cancer, ischemic disease and neovascular eye disease (e.g., Refs. [2–5]). VEGF promotes both blood vessel growth and vascular permeability, and VEGF upregulation therefore drives excessive

vascular leakage and tissue-damaging oedema in neovascular and inflammatory diseases such as tumour ascites, diabetic retinopathy, age-related macular degeneration, stroke and myocardial infarction (e.g., Refs. [4, 6, 7]). Anti-VEGF therapies can be administered to attenuate pathological vascular leakage [5], but VEGF blockade is unsuitable when blood vessel growth should be stimulated, for example to promote the revascularisation of ischemic tissues. Accordingly, alternative therapeutic targets should be identified that are specific to the VEGF hyperpermeability signalling pathway without compromising beneficial VEGF-induced blood vessel growth in ischemic diseases. Much vascular biology research is therefore focussed on elucidating molecules that transduce VEGF signals in vascular endothelial cells (ECs) and could be targeted to modulate vascular permeability to fluids, solutes and larger molecules such as albumin [8–11].

To promote vascular hyperpermeability, VEGF binds the receptor tyrosine kinase VEGFR2 (also known as KDR or FLK1) and its co-receptor neuropilin 1 (NRP1) into a

James T. Brash and Guillermo Diez-Pinel have contributed equally.

✉ Alessandro Fantin
alessandro.fantin@unimi.it

✉ Christiana Ruhrberg
c.ruhrberg@ucl.ac.uk

¹ UCL Institute of Ophthalmology, University College London, 11-43 Bath Street, London EC1V 9EL, UK

² Department of Biosciences, University of Milan, Via G. Celoria 26, 20133 Milan, Italy

complex [12, 13]. In addition to NRP1, the proteoglycan syndecan 2 acts as a VEGFR2 co-receptor in the dermis and brain to segregate vascular permeability from angiogenic VEGF signalling [14]. Subsequent to VEGFR2 complex formation, SRC family kinase (SFK) signalling is activated, which sets in motion downstream pathways that target the junction protein vascular endothelial (VE-)cadherin (cadherin 5) in a process termed paracellular permeability [13, 15–19]. Agreeing with this model, mice globally deficient in either one of two SFKs, SRC or YES1, have an attenuated response to VEGF in vascular hyperpermeability assays [20], although another study reported that endothelial YES1 instead helps establish and maintain the vascular barrier [21].

The interaction between activated VEGFR2 and SRC has been proposed to be mediated by a third protein that binds the phosphorylated Y951 residue of VEGFR2 (Y949 in mice), the SH2 domain-containing T cell-specific adaptor protein (TSAd) that is encoded by the *SH2D2A* gene [22–26]. Supporting this model, VEGF-induced vascular leakage is reduced in mice expressing VEGFR2 with a Y949F mutation and in mice with a global TSAd deletion [24, 25]. However, TSAd was originally reported to be expressed in haemopoietic tissues and immune cells only [27], and an immunostaining-based survey of TSAd localisation in human tissues concluded that it was largely absent from blood vessels except in dermal vasculature [28]. Additionally, TSAd protein detection was reported in kidney vessels [22]. Nevertheless, endothelial-specific TSAd knockout mice remain to be examined for defects in vascular permeability to establish whether its role in VEGF-induced vascular permeability signalling is EC-autonomous.

In addition to the VEGFR2-TSAd-SRC axis, several other molecular pathways contribute to VEGF-induced hyperpermeability signalling through paracellular and/or transcellular routes, the former involving VE-cadherin-positive inter-endothelial junctions and the latter involving caveolae [9, 29, 30]. Amongst these pathways, signalling via the alternative VEGF receptor VEGFR1 and the caveolar protein caveolin 1 converges on the endothelial nitric oxide synthetase (eNOS), which nitrosylates beta-catenin and VE-cadherin, whereas the calcium-dependent AMPK/p38 MAP kinase pathway specifically enables paracellular permeability signalling downstream of VEGFR2 in the brain and retina, where they induce VE-cadherin rearrangements [9, 29–32]. Several intracellular kinases other than SRC also contribute to permeability signal transduction, including the ABL kinases ABL1 and ARG [13, 19] and the focal adhesion kinase FAK [33]. Vice versa, the vascular endothelial protein tyrosine phosphatase VE-PTP stabilises the endothelial barrier by targeting VE-cadherin phosphorylation and thus serves as a negative regulator of VEGF-induced hyperpermeability [34].

Basing the search for novel therapeutics to treat VEGF-induced vascular leakage on molecular interactions between any of the above or other signal transduction molecules obviously requires that the signalling partners are co-expressed in ECs of the organ to be treated. This prerequisite is readily testable with computational approaches that interrogate transcriptomic data of ECs from multiple organs yet needs to consider that transcripts for individual genes may be under-represented in some types of datasets. For example, *SRC* transcripts are not readily detected in ECs of several publicly available sc-RNAseq datasets, but abundant in EC bulk RNAseq data [35]. Here, we have combined the analysis of publicly available sc-RNAseq, bulk RNAseq, epigenomic and proteomic datasets to compare the gene expression characteristics of signal transducers implicated in VEGF-induced vascular permeability signalling in ECs of several organs clinically affected by VEGF-driven oedema and in EC subtypes used for permeability studies in vitro. Our findings suggest that the *SH2D2A* gene encoding TSAd is not actively expressed in ECs of any of the organs examined, unlike all other signal transducers examined. This finding suggests that the current model of a VEGF-induced endothelial VEGFR2-TSAd-SRC complex needs to be revised to better support the search for novel therapeutics to reduce vascular hyperpermeability, and should be focussed on signal transducers expressed robustly in endothelial cells in vascular permeability-relevant organs.

Methods

Analysis of EC sc-RNAseq data

The human skin EC dataset was downloaded from the BIG Data Center (<https://bigd.big.ac.cn/>) [36]. The human trachea dataset was downloaded from the Human Cell Landscape project on GEO NCBI (<https://www.ncbi.nlm.nih.gov/geo/>) [37] and the EC cluster was selected according to expression of core endothelial markers such as *PECAM1*, as previously described [38]. We downloaded raw count data from the EC Atlas shiny app (http://endoteliomics.shinyapps.io/ec_atlas/) [39]. For the EC atlas and human datasets, we removed cells containing less than 200 feature counts. We downloaded *Tabula Muris* R objects containing gene expression data from cells isolated by fluorescence-activated cell sorting (FACS) or microfluidic droplet capture from http://figshare.com/articles/dataset/Robject_files_for_tissues_processed_by_Seurat/5821263; these Seurat objects had been pre-processed to remove cells containing less than 500 feature counts or less than 50,000 reads (FACS) and 10,000 UMIs (droplet), respectively [40]. EC clusters from brain, heart, lung and trachea obtained with the FACS isolation

method were selected as described above. Moreover, ECs, T cells or natural killer (NK) cells were merged into one Seurat object per isolation method. R v4.4.1 [41] and Seurat 5.1.0 [42] versions were used to explore sc-RNAseq data. Downstream analysis included data normalisation (“LogNormalize” method and scale factor of 10,000) and variable gene detection (“vst” selection method, returning 2000 features per dataset). Principal components (PC) analysis (PCA) was performed on variable genes, and the optimal number of PCs was chosen for each sample using the elbow plot. The selected PCs were used for Louvain graph-based clustering. Uniform manifold approximation and projection (UMAP) was chosen as a non-linear dimensionality reduction method, and cluster cell identity was assigned by manual annotation based on known marker genes. Each relevant gene was then examined using the *FeaturePlot* and *VlnPlot* functions.

Analysis of EC bulk RNAseq data

The Bulk-ECexplorer application for mining endothelial cell transcriptomic data (<https://www.ruhrberglab.shinyapps.io/BulkECexplorer>) compiles 240 high quality bulk RNAseq datasets from primary human ECs, human umbilical vein ECs (HUVECs) and human dermal microvascular ECs (HDMECs), and primary mouse ECs isolated from the lung, brain or retina [35]. We used the online interface of the Bulk-ECexplorer application to download graphs displaying expression data in TPM with or without a 1 TPM threshold, whereby the latter threshold is commonly used cut-off that was also validated to return results reflecting active gene expression in ECs [35]. We also used the online interface to download graphs returning validated predictions on active versus leaky gene expression, which are incorporated into the Bulk-ECexplorer application based on methods utilising Gaussian mixture models (GMM) and zTPM thresholds [35]. To evaluate gene expression correlations, *corr.test* from the ‘psych’ R package v2.4.3 was used to perform Pearson’s correlations (r) with Holm adjusted p-values, using zTPM values of genes with $\text{TPM} > 0$; *corrplot* from the ‘corrplot’ R package v0.92 and ‘ggplot2’ R package v3.5.0, were used to generate correlation plots. The relationship between the number of aligned reads and the probability of *SH2D2A* detection was assessed by logistic regression using the *glm* function in R v4.4.0. In our logistic regression model, the binary dependent variable represents the detection of *SH2D2A* (1 = detected, 0 = not detected) and the continuous predictor was the number of aligned reads. The final equation was $\text{logit}(\text{probability}) = \text{intercept} + \beta * \text{number of aligned reads} = -0.97 + 3.96\text{E-}08 * \text{number of aligned reads}$.

Analysis of bulk RNA-seq samples from VEGF-stimulated ECs

The European Nucleotide Archive (ENA; <https://www.ebi.ac.uk/ena/>) was queried with the term ‘VEGF endothelial’, which yielded 339 different projects. From these, we selected three bulk RNA-seq datasets that contained both control and VEGF-stimulated HUVECs. Raw reads from projects PRJNA431557 and PRJNA945485 were downloaded from GEO (<https://www.ncbi.nlm.nih.gov/geo/>; accession numbers GSE109625 and GSE227549, respectively). Raw reads from project PRJNA807293 were downloaded from ENA (<https://www.ebi.ac.uk/ena/>). To obtain count matrices and TPM matrices, the nf-core/rnaseq v3.14.0 pipeline (<https://zenodo.org/records/14537300>) was run using Nextflow v24.04.3 [43] to align reads to the human reference genome Homo_sapiens.GRCh38.dna.primary_assembly.fa.gz with the annotation file Homo_sapiens.GRCh38.113.gtf.gz (ENSMBL release 113). Differential expression analysis was carried out in R v4.4.1 using the DESeq2 v1.44.0 package [44].

Analysis of the Encyclopaedia of DNA Elements (ENCODE) DNase-seq database.

DNase I hypersensitive (HS) sites sequencing (DNase-seq) data from the ENCODE project were examined as meta-DNase tracks on the UCSC Genome Browser (<http://genome.ucsc.edu/>), as described [45]. The meta-DNase HS profile for ECs ($n = 15$) was comprised of HUVECs ($n = 1$), HDMECs ($n = 10$, including $n = 4$ blood vascular, $n = 2$ lymphatic vascular, $n = 4$ mixed blood and lymphatic vascular), lung ECs ($n = 4$, including $n = 2$ blood vascular and $n = 2$ lymphatic vascular). The meta-DNase HS profile for lymphoid samples ($n = 15$) was comprised of T cells ($n = 12$), natural killer cells ($n = 1$) and a B cell lymphoma line ($n = 2$). The meta-DNase HS profile for myeloid/erythroid samples ($n = 15$) was comprised of CD34+ erythroid ($n = 8$), myeloid ($n = 3$) and progenitor ($n = 4$) cells.

Analysis of proteome datasets

Summary data of proteomic datasets were retrieved from the ProteomeXchange database (<https://www.proteomexchange.org>) in August 2024. To identify relevant endothelial datasets, we queried the archive for the terms ‘HUVEC’, ‘human umbilical cord endothelial’ and ‘endothelial cells’. Our queries returned 139 projects for either *Homo sapiens* or *Mus musculus*, which were individually examined to determine their suitability for our analysis. We only retained published datasets for analysis that provided detection values for individual proteins in pre-processed lists, for example Log_2 of intensity Based Absolute Quantification

(iBAQ) or label-free quantitation (LFQ). We included only datasets derived from whole cell lysates or which were enriched in the cytoplasmic and plasma membrane fractions (not secretome, exosome or nuclear fractions), and which detected at least one of the 2 core EC markers KDR or CDH5. As we wished to examine the ‘basal’ proteome of ECs, we excluded datasets from cells that had been stimulated (for example, with a small molecule or by hypoxia) and/or had been genetically or functionally modified (for example, by gene deletion, protein overexpression or immortalisation). However, we retained datasets in these projects that were derived from control cells (for example, vehicle-stimulated or small interfering RNA control-transfected ECs). The following 15 datasets comprised of 52 individual samples were retained and mined for detection of individual proteins: whole cell proteomes for HUVECs (PXD009687, PXD003284, PXD020958, PXD000359, PXD034985 and PXD045899), heart ECs (PXD026673 and PXD045899), lung ECs (PXD029834, PXD046787 and PXD045899), retina ECs (PXD005972), skin ECs (PXD019909), kidney ECs (PXD045899) and brain ECs (PXD045899) as well as cytoplasmic/plasma membrane protein-enriched samples from HUVECs (PXD003406). Additionally, we examined the 3 VEGF-stimulated samples from human lung ECs that were included in PXD029834. Moreover, we mined whole cell proteomes for T cells (PXD015872) as well as cytoplasmic/plasma membrane protein-enriched samples from normal human dermal fibroblasts (NHDFs; PXD003412) and peripheral blood-derived mononuclear cells (PBMCs; PXD001416). The cytoplasmic/plasma membrane protein-enriched samples were part of a larger project that compared HUVEC, NHDF and PBMC proteomes [46, 47].

To determine the total number of proteins detected in each dataset, we considered all proteins expressed by at least one sample in that dataset; missing values, zeroes and imputed intensity values of 0.001 for an individual protein were considered as ‘not detected’. In datasets lacking gene names, we mapped protein names or Uniprot accessions to their corresponding gene names using the Uniprot human reference proteome (<https://www.uniprot.org/proteomes/UP000005640>, accessed in October 2024). To assess the cellular signature of each sample, samples were examined for core EC markers (in addition to KDR or CDH5: FLT1, NOS3, VWF, PLXND1, TIE1, TEK, or PLVAP), immune cell (IC) markers (PTPRC, CD3E, CD247, CD8A, ITGAM) and mesenchymal cell (MC) markers (ELN, PDGFRA, PDGFRB, CD248). The false discovery rate was q-value equal to 0.05. To assess cell type signatures of TSA-expressing HUVEC proteomes from PDX003406, we first identified the proteins exclusive to the TSA-expressing samples. Then, a hypergeometric test was carried out using the function *enricher* from the clusterProfiler v4.12.6 package to determine if any cell type signature gene sets from the

Human Molecular Signatures Database were significantly enriched in the HUVEC samples that detected TSA (file c8.all.v2023.2.Hs.symbols.gmt; <https://www.gsea-msigdb.org/gsea/msigdb>) [48]. For this, a list of all detected proteins across PDX003406’s EC samples was used as the background universe, and the Benjamini–Hochberg correction was applied to account for multiple testing.

Results

TSA (SH2D2A) transcripts are not readily detected in ECs by multi-organ sc-RNAseq analysis

We initially used publicly available sc-RNAseq data to evaluate the expression of hyperpermeability-implicated genes in ECs. As the dermis has been widely used to investigate the VEGF hyperpermeability pathway via the Miles assay (e.g., [13, 24, 49–51], we first examined the expression of relevant signalling molecules in sc-RNAseq data of dermal ECs. FACS-isolated PECAM1-positive ECs from human adult dermis [36] form a cluster of blood vascular ECs and a smaller cluster of lymphatic vascular ECs [35]. Transcripts for the main EC markers and permeability pathway components VEGFR2 (*KDR*) and VE-cadherin (*CDH5*) as well as the key signalling mediators caveolin 1 (*CAV1*) and eNOS (*NOS3*) were readily detected in ECs of both clusters (Fig. 1a, Tables 1 and 2). We also detected all other permeability-relevant genes examined, including transcripts for the alternative VEGF receptors NRP1 (*NRP1*) and VEGFR1 (*FLT1*) and numerous signalling transducers/regulators implicated in VEGF hyperpermeability signalling [9, 29, 32], such as the vascular endothelial protein tyrosine phosphatase VE-PTP (*PTPRB*), syndecan 2 (*SDC2*), β -catenin (*CTNNB1*), AMPKA (*PRKAA1*), p38 (*MAPK14*), ABL1 (*ABL1*), ARG (*ABL2*) and focal adhesion kinase (*PTK2*) (Tables 1 and 2). *YES1* and *SRC* detection rates in this human dermis dataset have previously been described [35]. Unexpectedly, this dataset contained no UMI counts assigned to *SH2D2A* in any of its 47,668 EC transcriptomes (Fig. 1a, Tables 1 and 2). As a negative control, we examined dermal ECs for expression of *CD4*, an endothelial-inactive gene that is a key marker for CD4+ T cells and NK cells [52–54], two cell types that express *SH2D2A* [27, 28, 55]. As expected, dermal EC transcriptomes did not include *CD4* transcripts (Fig. 1a, Tables 1 and 2).

Lack of *SH2D2A* detection in dermal ECs was unexpected, considering that a prior study reported TSA protein in dermal vasculature [27] and another study reporting that TSA knockout mice lack a VEGF response in the Miles assay [24]. As the trachea is well suited to investigate vascular hyperpermeability pathways [56] and has also been used to implicate TSA in permeability signalling [24], we

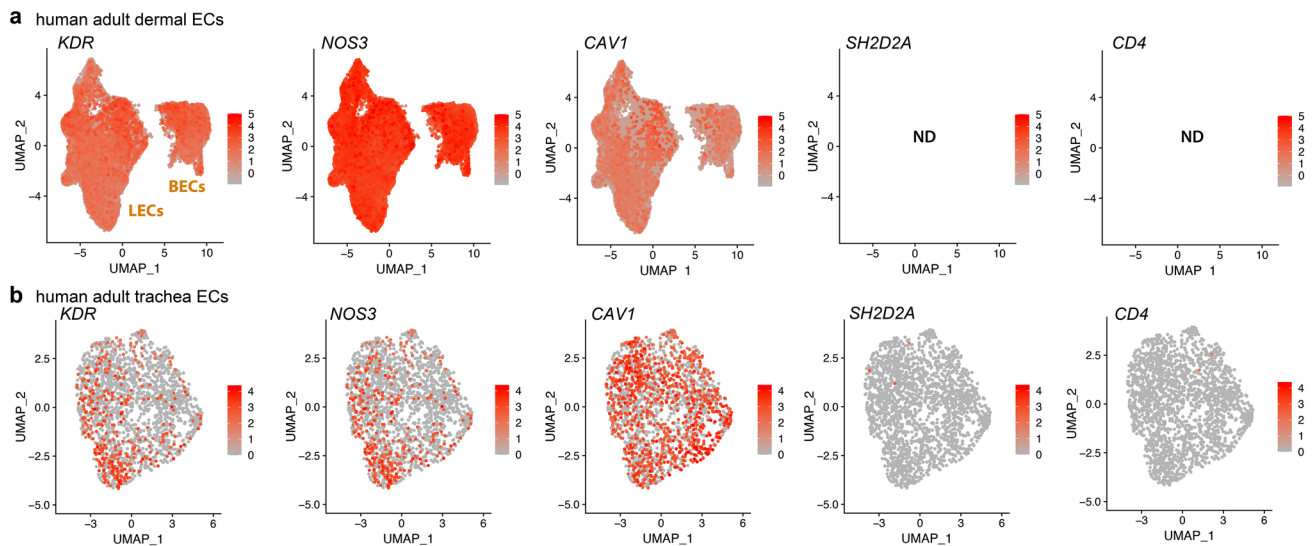


Fig. 1 *SH2D2A* is rarely detected in sc-RNAseq data from human skin or trachea ECs. sc-RNAseq data from human adult skin after isolation of PECAM1 ECs by FACS (a) and from human adult trachea after EC identification and selection (b). UMAP projections were generated to compare expression levels for *KDR*, *NOS3*, *CAV1*, *SH2D2A* and *CD4*; each data point represents the value for a single cell. ND, not detected; BECs, blood ECs; LECs, lymphatic ECs

Table 1 Average transcript levels for genes of the VEGF hyperpermeability pathway in ECs from sc-RNAseq datasets

dataset	PMID	species	method	organ	cell #	average normalised gene expression															
						<i>KDR</i>	<i>CDH5</i>	<i>NRP1</i>	<i>PTPRB</i>	<i>FLT1</i>	<i>CAV1</i>	<i>NOS3</i>	<i>CTNNB1</i>	<i>PRKAA1</i>	<i>MAPK14</i>	<i>ABL1</i>	<i>ABL2</i>	<i>PTK2</i>	<i>SDC2</i>	<i>SH2D2A</i>	<i>CD4</i>
PRJCA002692	33995668	Hs	droplet	skin	47668	1.10	1.70	1.58	2.00	1.52	13.82	0.18	1.14	0.50	0.34	0.11	0.31	0.59	0.02	ND	ND
GSE134355	32214235	Hs	microwell	trachea	2045	2.72	2.23	3.10	3.09	5.64	8.57	0.44	3.23	1.18	0.29	0.36	3.98	0.90	0.14	0.01	0.02
dataset	PMID	species	method	organ	cell #	<i>Kdr</i>	<i>Cdh5</i>	<i>Nrp1</i>	<i>Ptprb</i>	<i>Flt1</i>	<i>Cav1</i>	<i>Nos3</i>	<i>Ctnnb1</i>	<i>Prkaa1</i>	<i>Mapk14</i>	<i>Abl1</i>	<i>Abl2</i>	<i>Ptk2</i>	<i>Sdc2</i>	<i>Sh2d2a</i>	<i>Cd4</i>
EC atlas (E-MTAB-8077)	32059779	Mm	droplet	brain	3724	2.61	3.69	2.70	15.04	21.32	4.74	1.89	4.83	0.14	0.47	0.11	0.08	0.38	0.00	ND	ND
				lung	5152	3.83	11.20	8.80	11.44	3.63	15.88	0.27	1.76	0.14	0.45	0.16	0.07	0.61	0.09	ND	ND
				heart	4525	5.74	8.45	9.47	8.04	5.95	17.23	1.15	4.62	0.17	0.38	0.20	0.08	0.58	0.08	ND	ND
Tabula Muris (GSE109774)	30283141	Mm	SMART-Seq2	brain	733	4.06	1.68	3.51	9.35	24.90	5.69	2.23	8.36	0.13	0.40	0.10	0.21	0.31	0.02	0.00	0.00
				lung	702	4.64	5.22	9.00	12.91	3.14	25.52	0.45	2.94	0.17	0.32	0.27	0.16	0.65	0.07	0.00	0.00
				heart	1207	13.55	5.70	20.44	8.62	8.45	43.05	1.24	7.21	0.15	0.43	0.43	0.24	0.79	0.03	0.00	0.00
				trachea	115	10.49	4.42	23.28	10.40	22.63	19.61	1.62	6.23	0.11	0.34	0.17	0.77	0.63	0.07	0.00	0.00

Lowest value 25% percentile Highest value

Average transcript levels for genes of the VEGF hyperpermeability pathway and the non-EC, T cell marker *CD4* in ECs from the indicated organs in the indicated mouse and human sc-RNAseq datasets. EC selection was achieved either by FACS (E-MTAB-8077, PRJCA002692) or through clustering with Seurat (GSE109774, GSE134355). The table also lists the dataset name, accession number with Pubmed ID (PMID), species (Hs, *Homo sapiens* or Mm, *Mus musculus*), organ, method (droplet 10x, microwell or SMART-Seq2) and total EC number. Normalisation of gene expression was performed with Seurat. ND, not detected

next analysed adult sc-RNAseq data from ECs in the human trachea via the Human Cell Landscape compendium [37]. *YES1* and *SRC* detection rates have already been described for this dataset [35]. Transcripts for *KDR*, *CDH5*, *CAV1*, *NOS3* and all other examined genes implicated in the VEGF hyperpermeability were readily detected, whereas *SH2D2A* was detected in only 3/2045 trachea ECs with very low average transcript levels (Fig. 1b, c, Tables 1 and 2). This low detection level and frequency was again similar to that of

CD4, which was detected at low levels in 2/2045 trachea ECs (Fig. 1b, c, Tables 1 and 2).

Next, we analysed sc-RNAseq data from the mouse, which has been the main model organism employed to define the roles of molecules in the VEGF hyperpermeability pathway (e.g., Refs. [12, 13, 17, 24, 33, 49, 50, 57]). Specifically, we took advantage of two sc-RNAseq compendia that contain a large number of ECs each, the EC atlas [39] and *Tabula Muris* [40]. The EC atlas contains data from ECs isolated with a combined MACS and FACS pipeline from 11 adult

Table 2 Prevalence of expression of genes in the VEGF hyperpermeability pathway in ECs from sc-RNAseq data

dataset	PMID	species	method	organ	cell #	% of ECs with detectable transcript levels															
						<i>KDR</i>	<i>CDH5</i>	<i>NRP1</i>	<i>PTPRB</i>	<i>FLT1</i>	<i>CAV1</i>	<i>NOS3</i>	<i>CTNNB1</i>	<i>PRKAA1</i>	<i>MAPK14</i>	<i>ABL1</i>	<i>ABL2</i>	<i>PTK2</i>	<i>SDC2</i>	<i>SH2D2A</i>	<i>CD4</i>
PRJCA002692	33995668	Hs	droplet	skin	47668	27.73	40.18	34.95	41.69	33.43	85.95	6.26	29.60	15.73	11.06	3.81	9.24	17.53	0.74	ND	ND
GSE134355	32214235	Hs	microwell	trachea	2045	18.14	16.48	22.74	23.47	33.94	45.18	3.81	23.81	9.19	2.49	3.08	27.24	7.04	1.37	0.15	0.15
dataset	PMID	species	method	organ	cell #	<i>Kdr</i>	<i>Cdh5</i>	<i>Nrp1</i>	<i>Ptprb</i>	<i>Flt1</i>	<i>Cav1</i>	<i>Nos3</i>	<i>Ctnnb1</i>	<i>Prkaa1</i>	<i>Mapk14</i>	<i>Abl1</i>	<i>Abl2</i>	<i>Ptk2</i>	<i>Sdc2</i>	<i>Sh2d2a</i>	<i>Cd4</i>
EC atlas (E-MTAB-8077)	32059779	Mm	droplet	brain	3724	46.80	60.20	44.20	94.63	96.13	64.47	40.12	68.82	4.03	12.65	3.14	2.63	10.50	0.00	ND	ND
				lung	5152	48.35	84.74	72.38	85.38	53.49	91.52	6.66	32.94	3.80	12.01	4.41	2.04	14.11	1.88	ND	ND
				heart	4525	55.82	71.25	70.59	71.62	59.80	86.17	21.72	53.81	3.69	7.71	4.31	1.99	11.47	1.94	ND	ND
Tabula Muris (GSE109774)	30283141	Mm	SMART-Seq2	brain	733	73.26	82.67	75.58	95.09	99.45	77.22	75.85	95.09	11.60	25.78	10.23	15.28	28.51	4.91	0.41	0.82
				lung	702	56.55	94.16	80.06	83.19	67.24	97.44	20.51	62.68	10.11	19.37	11.97	10.11	26.64	6.70	0.85	1.28
				heart	1207	85.58	93.21	92.87	81.36	86.58	95.28	49.88	78.62	8.53	15.99	13.01	9.94	31.40	5.63	0.99	1.16
				trachea	115	93.04	98.26	92.17	95.65	93.91	93.91	74.78	92.17	10.43	20.87	13.91	33.04	37.39	6.96	0.00	1.74

Percentage of ECs with detectable transcript levels for the indicated genes in the VEGF hyperpermeability pathway and the non-EC, T cell marker *CD4* in ECs from the indicated organs in the indicated mouse and human sc-RNAseq datasets. EC selection was achieved either by FACS (E-MTAB-8077, PRJCA002692) or through clustering with Seurat (GSE109774, GSE134355). The table also lists the dataset name, accession number with Pubmed ID (PMID), species (Hs, *Homo sapiens* or Mm, *Mus musculus*), organ, method (droplet 10x, microwell or SMART-Seq2) and total EC number. ND, not detected

organs, whereas *Tabula Muris* contains data from ECs isolated with a FACS pipeline or droplet method from 10 adult organs. Notably, both compendia included the brain, lung and heart as examples of organs in which VEGF-induced vascular hyperpermeability has been observed.

In both the EC atlas and *Tabula Muris*, transcripts for *Kdr* and *Cdh5* were readily detected in the brain, lung and heart EC populations (Tables 1 and 2; *Yes1* and *Src* detection rates have been described [35]). ECs from the brain, lung and heart in both compendia also contained abundant transcripts for *Cav1*, *Nrp1*, *Flt1*, *Ctnnb1* and *Ptprb* (Tables 1 and 2). Transcripts for other known signal transducers in the VEGF hyperpermeability pathway were also detectable, although less frequently and at more moderate levels, including *Nos3*, *Sdc2*, *Ptk2*, *Prkaa1*, *Mapk14*, *Abl1* and *Abl2* (Tables 1 and 2). By contrast, no UMI counts assigned to *Sh2d2a* were identified with the EC atlas in brain, lung and heart ECs, and both its detection level and detection rate were very low with *Tabula Muris* in these organs (Tables 1 and 2; *Tabula Muris* detection rate: brain ECs 3/733, lung ECs 6/702, heart ECs 12/1207). *Tabula Muris* also contains data for the mouse trachea, which was studied to put forward a role for TSA in the VEGF hyperpermeability pathway [24]; however, *Sh2d2a* transcripts were also lacking from trachea ECs (Tables 1 and 2).

Next, we extended analysis of *Sh2d2a* transcripts to all ECs across all organs in the EC atlas; however, *Sh2d2a* transcripts were again not detected. Further, we analysed the entire pool of FACS-isolated ECs across all 10 *Tabula Muris* organs alongside pooled T cell and NK cells, which are known to express *Sh2d2a* (Fig. 2a). Moreover, we performed a similar, parallel analysis of the entire pool of ECs isolated with the alternative droplet method from 9 adult organs (Fig. 2b). We again found that *Sh2d2a* transcripts

were rarely detected in ECs, i.e., in only 1% of FACS-captured ECs (39/4155) and in only 0.5% of droplet-captured ECs (18/3725) of *Tabula Muris*. Moreover, ECs with detectable *Sh2d2a* were as rare as ECs with detectable *Cd4* (Fig. 2a, b), and average *Sh2d2a* transcript levels in ECs were as low as those for *Cd4* (Fig. 2a, b). Low *Sh2d2a* detection rates and low transcript levels in ECs could not be attributed to technical difficulties in identifying *Sh2d2a* transcripts by sc-RNAseq, because *Sh2d2a* transcripts were readily detected in NK cells and T cells by both the FACS and droplet method in *Tabula Muris* (Fig. 2a, b). Vice versa, and as expected, the EC marker *Kdr* was not detected in NK cells or T cells (Fig. 2a, b). These observations on mouse and human organ ECs can be confirmed by directly exploring the publicly available interactive webtools of the original sc-RNAseq datasets (EC atlas: https://endotheliomics.shinyapps.io/ec_atlas/; *Tabula Muris*: <https://tabula-muris.sf.czbiohub.org/visualizations>; Human Cell Landscape: <https://db.cngb.org/HCL/landscape.html>).

Taken together, the above transcriptomic analyses suggest that TSA transcripts are either not expressed in ECs or expressed at levels that are not readily detectable by sc-RNAseq. We therefore complemented sc-RNAseq analysis with bulk RNAseq analysis.

TSA (*SH2D2A*) transcripts are absent or detected at low levels in EC bulk RNAseq data

Bulk RNAseq has an excellent capacity for detecting lowly expressed genes in homogenous cell populations [58]. For example, *SRC* is not easily detectable in sc-RNAseq datasets from mouse and human organ ECs but readily detected across most of the 240 mouse and human endothelial bulk RNAseq datasets included in the Bulk-ECexplorer

compendium [35]. Relevant for the present investigation, the Bulk-ECexplorer datasets are derived from EC subtypes that have been used to characterise the VEGFR2-TSAd-SRC pathway, namely human dermal microvascular ECs (HDMECs) [13, 24], human umbilical vein ECs (HUVECs) [22, 59], mouse lung ECs [13, 25] and mouse retina ECs [26]. The Bulk-ECexplorer also contains datasets from ECs of the mouse brain, an organ in which VEGF-induced vascular leakage is a significant concern [60]. Moreover, transcripts for the permeability-relevant EC markers *KDR* and *CDH5* are defining features of the Bulk-ECexplorer datasets. To analyse the gene expression characteristics of permeability-relevant genes in ECs, we therefore took advantage of the Bulk-ECexplorer application, hosted at <https://ruhrberglab.shinyapps.io/BulkECexplorer> [35].

Similar to transcripts for the vascular barrier regulators *YES1* and *SRC* [35], transcripts for other permeability-relevant genes were detected by the Bulk-ECexplorer at high frequency (Figs. 3a, S1A, Supplemental Table 1). Thus, transcripts for *CAVI* and *NOS3* were detected in all 240 EC RNAseq datasets of the Bulk-ECexplorer (Fig. 3a), and also for all permeability signalling regulators examined, including *NRP1*, *FLT1*, *PTPRB*, *SDC2*, *PTK2*, *CTNBB1*, *PRKAA1*, *MAPK14*, *ABL1* and *ABL2* (Supplemental Fig. 1). By contrast, *SH2D2A* transcripts were not detected in 45% of the 240 EC datasets in the Bulk-ECexplorer (Fig. 3a; Supplemental Table 1). *SH2D2A* detection in the other 55% of the EC datasets was mostly in HUVECs, the patho-physiologically least relevant EC subtype for vascular permeability studies amongst the five EC subtypes included in the Bulk-ECexplorer (e.g., *SH2D2A* detection in 27% of HDMEC versus 70% of HUVEC datasets; Fig. 3a). Moreover, *SH2D2A* transcripts, when present, were typically detected at very low levels, within a TPM range that was orders of magnitude smaller than the TPM range for *CAVI*, *NOS3* or other permeability-relevant signal transducers (Figs. 3c; S1B). We therefore applied the 1 TPM threshold as a commonly used cut off for selecting genes suitable for downstream analysis, and thus found that only 13% of the 240 EC datasets in the Bulk-ECexplorer scored positive for *SH2D2A*, and these datasets were again almost exclusively derived from HUVECs (Fig. 3b).

Insufficient read depth can sometimes explain the absence of a transcript in bulk RNAseq data [61]. In agreement, *SH2D2A* detection in the 240 datasets of the Bulk-ECexplorer positively correlated with higher read depth (Supplemental Fig. S2). Nevertheless, genes detectable only at high read depth are naturally expressed at low levels [62]. Further, we considered that expression levels below 1 TPM in only a subset of EC samples is also observed for numerous genes not expected to be functional within ECs, for example, several ocular genes (*LENEP*, *CRYBB2*) and sex cell-specific genes (*DDX4*, *GDF9*, *YBX2*, *SPACA4*) [35]. We therefore

compared the expression characteristics of *SH2D2A* to those of the T cell marker *CD4*, which is not expected to be expressed in ECs and is only rarely detected in EC scRNAseq data (see Fig. 2). *CD4* was detected in nearly half of the EC datasets included in the Bulk-ECexplorer (48% detection frequency), but only 11% of EC datasets expressed *CD4* at levels above 1 TPM (Fig. 3b, data resolved per EC type). When *CD4* expression was detected in ECs, it was within a TPM range orders of magnitude smaller than the TPM range for *CAVI*, *NOS3* or other permeability-relevant signal transducers, as observed for *SH2D2A* (Figs. 3c, S1B). Therefore, TSAd transcripts were either not detected in EC bulk RNAseq datasets or were detected at very low levels, akin to an EC-inactive gene.

TSAd (*SH2D2A*) transcript levels in ECs are not upregulated in response to VEGF

To assess whether *SH2D2A* expression was increased upon VEGF stimulation, we performed differential expression analyses of three published bulk RNAseq data, which we identified by searching the European Nucleotide Archive (ENA; <https://www.ebi.ac.uk/ena/>) for the term ‘VEGF endothelial’. We identified three HUVEC datasets that contained both control and VEGF-stimulated samples. Amongst these, two different projects [63, 64] stimulated HUVECs cultured in 2D with VEGF and analysed them at time points from 1 to 12, or at 24 h post treatment, respectively. Among the VEGF-induced hyperpermeability pathway members, *FLT1* appeared upregulated from 1 to 24 h and *CTNBB1* from 1 to 12 h after VEGF exposure (Supplemental Table S2). *ABL2* and *YES1* appeared transiently upregulated 1 h after VEGF treatment, and *CAVI*, *KDR* and *NRP1* transiently 4 h after VEGF treatment (Supplemental Table S2). In another project, HUVECs were grown as spheroids to model sprouting angiogenesis [65]. *FLT1*, *CTNBB1* and *NRP1* were upregulated 18 h after VEGF exposure (n=5 per group). By contrast, transcript levels of the other permeability pathway genes did not appear to be significantly enriched in any of the datasets (Supplemental Table S2). As observed for HUVEC datasets in the Bulk-ECexplorer, all permeability pathway genes examined were detected robustly across all three datasets, with exception of *SH2D2A*, whose TPM values ranged from 0 to 1 TPM in 14/14 control samples and in 12/14 VEGF-stimulated samples, and was detected at 1 or 2 TPM in the remaining two samples (Supplemental Table S2).

Low TSAd (*SH2D2A*) levels in ECs are predicted to arise from leaky gene expression

Studies across several different cell types have found that lowly expressed (LE) transcripts can arise from leaky

transcription, a phenomenon in which highly expressed (HE) genes impart a ‘transcriptional ripple effect’ on nearby genes that lack active chromatin markers [62, 66–69]. Notably, such LE genes do not usually give rise to detectable protein, and their transcripts are therefore thought to be biologically irrelevant [70–72]. Like other cell types, ECs contain a mixture of LE and HE transcripts whose combined expression values assume a bimodal distribution [35]. To predict whether transcripts arise from actively and leakily expressed genes, a two component Gaussian mixture model (GMM) can be fitted to such gene expression data and the probability of genes belonging to either the LE or HE distribution calculated, as initially done to examine the T cell transcriptome [73, 74], and subsequently applied to each one of the 198 bimodally distributed EC datasets included in the Bulk-ECexplorer [35].

Similar to *SRC* and *YES1* [35], *CAVI* and *NOS3* were classified by the GMM approach as actively expressed in the vast majority of eligible (i.e., bimodal) Bulk-ECexplorer datasets in which their transcripts were detected (> 98%), and similar results were obtained for *NRP1*, *FLT1*, *PTPRB*, *SDC2*, *PTK2*, *CTNBB1*, *PRKAA1*, *MAPK14*, *ABL1* and *ABL2* (Figs. 4, S3 and Supplemental Table 1). Amongst the permeability regulators examined, two exceptions were *SDC2* and *SH2D2A*. *SDC2* was classified as actively expressed in 84% of datasets in which it was detected (164 out of 196), whereas *SH2D2A* was classified as actively expressed in only 8% of datasets in which it was detected (9 out of 114). When resolved by EC subtype, *SDC2* was predicted to be actively expressed in most HUVEC and lung EC datasets; furthermore, it was never classified as leakily expressed in datasets from brain, retina or dermal ECs, the organs in which *SDC2* was reported to regulate vascular permeability signalling [14]. By contrast, *SH2D2A* was classified as actively expressed more often in HUVECs than in EC subtypes of vascular permeability-relevant organs (Fig. 4). Notably, *CD4* was classified as actively expressed at a frequency similar to *SH2D2A* (14% of 103 samples in which transcripts were detected; Fig. 4). GMM classification therefore predicted that all signalling molecules examined were expressed at functional levels in ECs of vascular permeability-relevant organs, except TSAd, which is classified as leakily expressed when transcript is detectable at all, similar to the EC inactive *CD4* gene.

We next used the zTPM tool in the Bulk-ECexplorer as an alternative method to predict whether a gene is actively or leakily expressed [35]. This tool applies a zTPM transformation to each of 220 datasets in the Bulk-ECexplorer without unimodal expression distribution before determining whether a queried gene exceeds the predicted zTPM threshold for actively expressed genes [35]; this threshold was previously determined based on HUVEC epigenomic and RNAseq data from the Encyclopaedia of DNA Elements

(ENCODE) project [62]. Using this method, we found that VEGF hyperpermeability-relevant genes examined exceeded the threshold for actively expressed genes in the vast majority of datasets that contained their transcripts (> 92%), with exception of *SH2D2A*, which exceeded the threshold in only 18% of the 220 datasets, and these were again mostly derived from HUVECs (Figs. 5, S4, Supplemental Table 1). *CD4* transcript levels similarly exceeded the threshold in only 11% of the 220 datasets. zTPM-based classification therefore agrees with GMM-based classification to predict that all signalling molecules examined are expressed at functional levels in ECs of vascular permeability-relevant organs, except TSAd, which instead is expressed like a leaky gene, similar to the EC inactive gene *CD4*.

Taken together, *SH2D2A*’s gene expression characteristics in ECs were similar to those of a typical non-EC gene and therefore inconsistent with the prediction of functional gene transcription in ECs.

***SH2D2A*’s genomic position suggests susceptibility to a transcriptional ripple effect**

In immune cells, LE genes often have HE genomic neighbours, with the distance of HE genes to the nearest LE gene typically smaller than the distance of HE genes to the nearest not expressed (NE) gene [67]. We therefore asked whether low *SH2D2A* transcript levels in ECs correlated with the gene’s localisation near HE genes that might confer a transcriptional ripple effect on *SH2D2A*. Thus, we first examined gene expression characteristics in *SH2D2A*’s genomic neighbourhood and then correlated *SH2D2A*’s zTPM levels with those of its neighbouring genes across all 220 Bulk-ECexplorer datasets suitable for this analysis.

The organisation and orientation of the genes surrounding *SH2D2A* are conserved between human chromosome 1 and mouse chromosome 3 (data discoverable at <https://genome.ucsc.edu/>). Specifically, *SH2D2A* resides between an upstream gene cluster (*ISG20L2*, *RRNAD1*, *MRPL24*, *HDGF* and *PRCC*) and a downstream gene cluster (*NTRK1* and *INSRR*), both in the human genome and in the mouse genome (not shown). The read distribution in this genomic neighbourhood was mapped for a representative HUVEC sample that contained *SH2D2A* transcripts, which raised the possibility that *SH2D2A* resided at the border between an HE and NE gene cluster in ECs (Fig. 6a). To corroborate this finding across additional HUVEC datasets and further EC subtypes, we used the Bulk-ECexplorer to examine the expression characteristics of transcripts in the *SH2D2A* genomic neighbourhood. Thus, the upstream genes *ISG20L2*, *RRNAD1* (also known as *METTL25B*), *MRPL24*, *HDGF* and *PRCC* were consistently reported as HE genes expressed above the 1 TPM threshold (> 99% of EC datasets, Figs. 6b, Supplemental Fig. S5, Supplemental Table 1). The

GMM and zTPM methods also classified these upstream genes consistently as actively expressed (> 99% of EC datasets; Figs. 6c, d; Supplemental Fig. S5, Supplemental Table 1). The downstream *PEAR1* gene (Fig. 6a) that resides within 100 kb of the *SH2D2A* transcription start site (TSS) was also classified as HE and actively expressed in > 99% of EC datasets (Fig. 6b–d; Supplemental Table 1). By contrast, the immediate downstream neighbours of *SH2D2A*, *NTRK1* encoding the neurotrophin receptor TRK-A, and *INSRR* encoding the urogenital insulin-related receptor, were only rarely detected in the Bulk-ECexplorer datasets (Fig. 6a; Supplemental Figure S5, Supplemental Table 1), thus identifying them as NE genes in ECs. When datasets contained some transcripts for these genes, the GMM and zTPM methods classified *NTRK1* as not actively expressed and *INSRR* mostly as not actively expressed in all EC subtypes (Fig. 6b–d; Supplemental Figure S5, Supplemental Table 1). The presence of *INSRR* in brain EC datasets may reflect neural cell contamination of primary ECs (see Discussion in [35]). The genomic neighbourhood of *SH2D2A* therefore resembles that of LE genes in ECs, in analogy to a prior study, in which LE genes were found to reside between HE and NE gene clusters in immune cells [67].

We next correlated the expression levels of *SH2D2A* with those of its neighbouring HE genes across 220 human and mouse Bulk-ECexplorer datasets, again using the standardised zTPM unit to exclude that gene expression might appear correlated only because of global variations in transcript levels between individual datasets. We found that *SH2D2A* expression significantly and positively correlated with the expression of all but one of its HE upstream neighbours across the 220 datasets (*SH2D2A* correlation with *ISG20L2* $r=0.53$, *MRPL24* $r=0.67$, *HDGF* $r=0.59$, *PRCC* $r=0.54$; Fig. 6e). A less strong, but nevertheless significant, correlation was also observed between the expression of *SH2D2A* and its downstream HE neighbour *PEAR1* ($r=0.45$; Fig. 6e). In particular, linear regression analysis revealed that variation in *MRPL24* expression could account for 44.9% of the variation in *SH2D2A* expression amongst datasets in the Bulk-ECexplorer (Fig. 6e, f). By contrast, *SH2D2A* expression did not positively correlate with the expression of two other VEGF-induced vascular permeability signalling-relevant HE genes in unlinked genomic loci, *CDH5* (chromosome 16; $r=0.10$) and *KDR* (chromosome 4; $r=-0.38$) (Fig. 6e). Taken together, low level *SH2D2A* expression correlates with genomic proximity to neighbouring, actively expressed genes in ECs, as previously described for leakily expressed genes in other cell types.

The *SH2D2A* promoter resides within closed chromatin in ECs

As chromatin accessibility to transcription factors is an inherent feature of active genomic elements, the promoters of actively expressed genes are accessible also to DNase, and DNase I hypersensitivity sites (DNase HS) can be mapped across a genome, as reported for the human genome in the ENCODE project [75]. In phase 3 of the ENCODE project, a total of 733 human samples were analysed by DNase-seq [45]. Further, the DNase-seq signal profile for 15 samples each for ECs, lymphoid and myeloid/erythroid cells was averaged to provide a DNase HS meta-profile that could be directly compared for accessible chromatin sites in these cell types [45]. By re-analysing these published meta-DNaseq HS data, we found that a strong DNase HS signal encompassed the *SH2D2A* promoter and transcription start site (TSS) in immune cells, as expected, whereas no obvious DNase HS signal was detected at the *SH2D2A* promoter/TSS in ECs, including in HUVECs, HDMECs and lung ECs (Fig. 7). These data show that the *SH2D2A* promoter resides within open chromatin in immune cells but within closed chromatin in ECs. *SH2D2A* therefore has the chromatin hallmarks of a gene that is not actively expressed in ECs, consistent with the gene expression characteristics we have described above.

Mass spectrometry studies do not report TSAd (*SH2D2A*) protein in ECs except in rare HUVEC samples with an immune signature

We next examined which vascular permeability regulators were present in publicly available EC proteomes. Having surveyed the ProteomeXchange database (<https://www.proteomexchange.org>) for suitable EC mass spectrometry datasets (see methods), we retrieved 12 datasets with 52 independent proteomes derived from whole cell lysates or the cytoplasmic/transmembrane fractions of HUVECs as well as whole cell lysates from primary ECs across 7 organs. Since each proteomic dataset was obtained with quantitative approaches that differed in terms of estimation, sensitivity and specificity, we reported protein expression in a binary outcome of ‘detected’ or ‘not detected’.

As *SH2D2A* had been classified as actively expressed in only a minor proportion of EC bulk RNAseq datasets that were mostly derived from HUVECs (see above), we initially examined six proteome datasets from HUVEC whole cell lysates with a range of recorded proteins that spanned from 8210 to 4207 [76–82]. The detection of 6/9 pre-defined core EC markers, including *KDR* and/or *CDH5* (see methods), confirmed cell identity (Supplemental Table 3). Notably, TSAd and *SDC2* were not present in the list of reported proteins in any of the 18 samples contained

Table 3 Detection of proteins in the VEGF hyperpermeability pathway

dataset (PMID)	species	cell type	proteins	sample	protein detection in whole cell lysates (D, detected; ND, not detected)																signature						
					KDR	CDH5	NRP1	PTPRB	FLT1	CAV1	NOS3	SRC	YES1	CTNNB1	PRKAA1	MAPK14	ABL1	ABL2	PTK2	SDC2		SH2D2A					
PXD045899 (39143368)	Hs	HUVECs	8210	1/3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC					
				2/3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC			
				3/3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC	
PXD034985 (35887356)	Hs	HUVECs	8198	1/3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC			
				2/3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC	
				3/3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC
PXD009687 (30983154)	Hs	HUVECs	6876	1/1	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC		
PXD000359 (23979707)	Hs	HUVECs	6678	1/3	D	D	D	D	ND	D	D	D	D	D	D	D	D	ND	ND	D	ND	ND	EC				
				2/3	D	D	D	D	ND	D	D	D	D	D	D	D	D	D	D	ND	ND	D	ND	ND	EC		
				3/3	D	D	D	D	ND	D	D	D	D	D	D	D	D	D	D	D	ND	ND	D	ND	ND	EC	
PXD003284 (27246970)	Hs	HUVECs	4591	1/1	D	D	D	D	D	D	D	D	D	ND	D	D	ND	ND	D	ND	ND	EC					
PXD020958 (36971339)	Hs	HUVECs	4207	1/7	ND	D	D	D	ND	D	D	ND	D	D	ND	ND	ND	ND	D	ND	ND	EC					
				2/7	D	D	D	D	ND	D	D	ND	D	D	ND	D	ND	ND	D	ND	ND	EC					
				3/7	D	D	D	D	ND	D	D	ND	D	D	ND	D	ND	ND	D	ND	ND	EC					
				4/7	ND	D	D	D	ND	D	D	ND	D	D	ND	ND	ND	ND	D	ND	ND	EC					
				5/7	ND	D	D	D	ND	D	D	ND	D	D	ND	ND	ND	ND	D	ND	ND	EC					
				6/7	ND	D	D	D	ND	D	D	ND	D	D	ND	ND	ND	ND	D	ND	ND	EC					
				7/7	D	D	D	D	ND	D	D	ND	D	D	D	ND	ND	ND	D	ND	ND	EC					
PXD045899 (39143368)	Hs	heart ECs	7935	1/2	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC			
				2/2	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC
PXD026673 (35614104)	Mm	heart ECs	4694	1/3	ND	D	ND	D	D	D	ND	ND	D	D	D	D	D	D	D	D	D	D	D	ND	EC/MC		
				2/3	ND	D	ND	D	D	D	ND	ND	D	D	D	D	D	D	D	D	D	D	D	D	ND	EC/MC	
				3/3	ND	D	ND	D	D	D	ND	ND	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	EC/MC
PXD045899 (39143368)	Hs	lung ECs	8163	1/2	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC		
				2/2	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND
PXD029834 (37252964)	Hs	lung ECs	5765	1/3	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	D	D	D	ND	ND	EC			
				2/3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	D	D	D	ND	ND	EC		
				3/3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	D	D	D	ND	ND	EC	
PXD046787 (38047313)	Hs	lung ECs	3322	1/6	D	D	D	D	D	D	D	ND	D	D	D	D	D	ND	ND	D	ND	ND	EC				
				2/6	D	D	D	D	ND	D	D	ND	D	D	D	D	D	D	ND	ND	D	ND	ND	EC			
				3/6	D	D	D	D	D	D	D	ND	D	D	D	D	D	D	ND	ND	D	ND	ND	EC			
				4/6	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	D	ND	ND	EC			
				5/6	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	D	ND	ND	EC			
				6/6	D	D	D	D	D	D	D	D	ND	D	D	D	D	D	ND	ND	D	ND	ND	EC			
PXD045899 (39143368)	Hs	brain ECs	8217	1/3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC		
				2/3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC
				3/3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND
PXD005972 (29559410)	Hs	retina ECs	5614	1/5	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	D	ND	ND	EC			
				2/5	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	D	ND	ND	EC	
				3/5	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC
				4/5	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC
				5/5	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	D	ND	ND	EC		
PXD045899 (39143368)	Hs	kidney ECs	8165	1/2	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC		
				2/2	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	ND	EC
PXD019909 (33154365)	Hs	skin ECs	9553	1/1	ND	D	D	D	ND	D	D	D	D	D	D	D	D	D	D	D	D	D	D	ND	EC/IC/MC		
PXD021272 (33383959)	Hs	T cells	4971	1/2	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	D	D	D	ND	ND	ND	ND	D	D	EC			
				2/2	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	D	D	D	ND	ND	ND	ND	D	D	EC		

The table summarizes whether the indicated proteins were detected (D) or not detected (ND) in 16 publicly available human and mouse datasets obtained by mass-spectrometry of whole cell lysates. Only samples from healthy, untreated conditions were scored; HUVEC data were derived from cultured cells. The table also lists the ProteomeXchange identifiers and Pubmed ID (dataset), species (Hs, *Homo sapiens* or Mm, *Mus musculus*), cell type, including EC subtype, number of detected proteins (proteins), number of independent samples per dataset (sample) and whether the proteome includes EC, immune cell (IC) and/or mesenchymal cell (MC) markers (cell type signature defined by key markers presented in Supplementary Table 3)

in these datasets (Table 3). By contrast, all other components of the VEGF hyperpermeability pathway examined were recorded across the 7 samples from the largest, 8210-, 8198- and 6876-protein HUVEC datasets, whereas most of them were detected across the 4 samples from the 6678- and 4591-protein HUVEC datasets and half of them in the 7 independent samples from the smallest, 4207-protein HUVEC dataset (Table 3). Thereafter, we extended analysis to whole cell proteomes from ECs of vascular permeability-relevant organs, including the heart [83], lung [79, 84, 85], retina [86], skin [87], brain and kidney [79]. These datasets recorded most components of the VEGF hyperpermeability pathway, including SDC2 in 1/1 skin and 3/3 heart EC samples, but all organ EC proteomes lacked TSAd (Table 3). Finally, we examined whether VEGF stimulation increased

TSAd detection, taking advantage of a published proteomics dataset from VEGF-treated versus control human lung ECs [85]. However, TSAd was also not detected in any VEGF-stimulated samples, similar to the controls (Supplemental Table 4). Although the above HUVEC and organ EC proteomes all lacked TSAd, 2/2 T cell samples [88] contained abundant reads for TSAd (Table 3), as expected [27].

We also examined the recorded list of 7625 proteins derived from a dataset of 7 independent HUVEC samples enriched for cytoplasmic and transmembrane proteins at the expense of secreted and nuclear proteins [46], as this approach potentially confers higher detection sensitivity for signal transducers. Similar to the HUVEC whole cell proteome analysis, all permeability pathway-relevant proteins examined were recorded in 4 out of 7 enriched

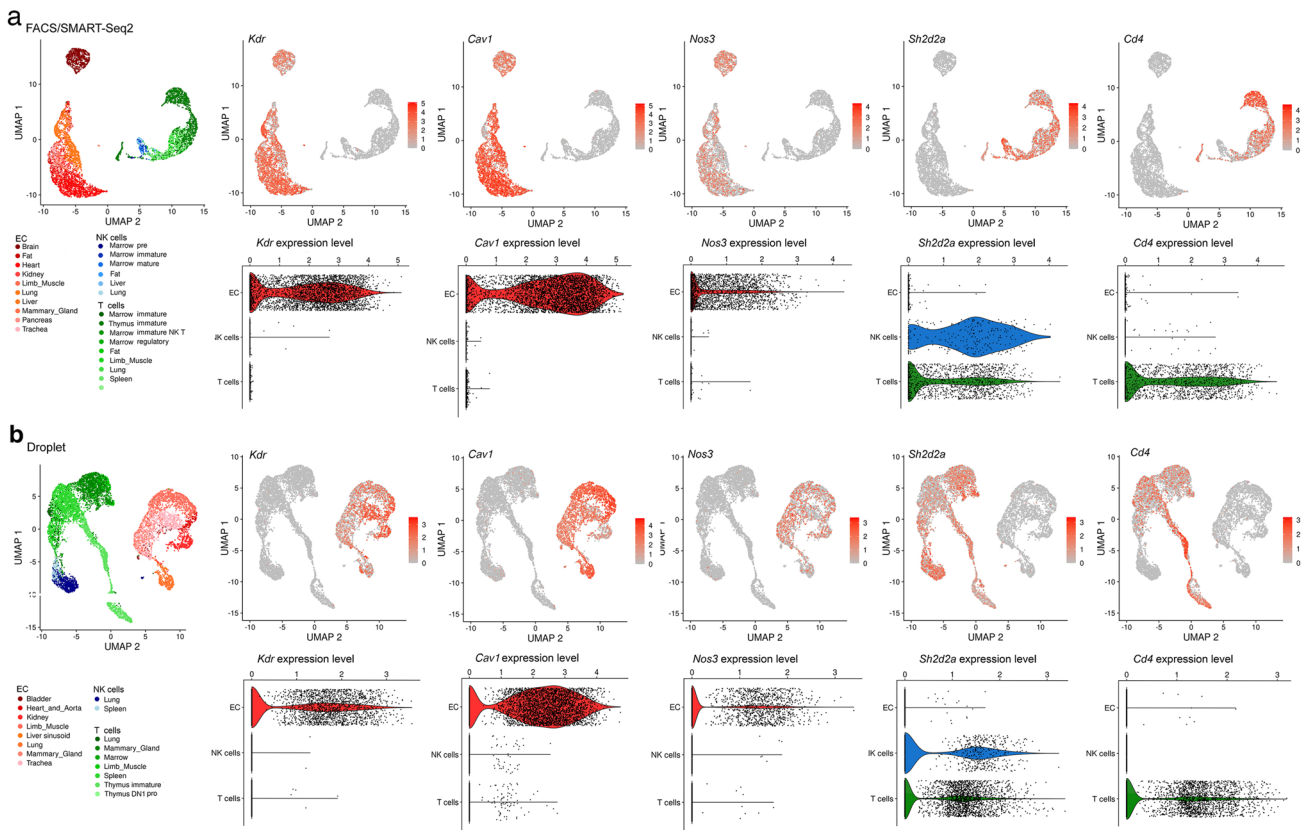


Fig. 2 *Sh2d2a* is rarely detected in EC sc-RNAseq data from *Tabula Muris*. *Tabula Muris* sc-RNAseq data for the indicated genes, shown for FACS-captured (a) and droplet-captured (b) cells from pooled EC and immune clusters. A colour coded UMAP projection of cell type

and tissue of origin is shown alongside UMAP projections illustrating the expression levels of the indicated genes per cell in each cluster. The corresponding violin plots illustrate the expression levels of each cell cluster. Each data point represents the value for one cell

HUVEC samples and most of them were recorded in the other 3 enriched HUVEC samples, except TSA_d and SDC2 (Supplemental Table 5). Specifically, SDC2 was recorded in only 1 and TSA_d in only 2 of the 7 HUVEC enriched samples; notably, the 2 HUVEC samples that recorded TSA_d also recorded immune cell markers (Supplemental Table 3). To more objectively assess the cell type signature of the two HUVEC samples that reported TSA_d in the enriched dataset, we first identified proteins exclusive to these two samples and then used them for overrepresentation analysis against the cell type signature gene sets from the Human Molecular Signatures Database, using all the proteins detected across the 7 HUVEC samples of the enriched dataset as the background universe (see Methods). This analysis demonstrated a significant enrichment of a foetal lymphoid cell signature (see methods; Supplemental Table 6). In the same study, 1/3 normal human dermal fibroblast (NHDF) and 3/3 human peripheral blood mononuclear cell (PBMC) cytoplasmic protein-enriched samples also recorded TSA_d (Supplemental Table 5). Unexpectedly, these non-EC proteomes also contained EC markers (Supplemental Table 3). It is not known why

12/13 of cytoplasmic fraction-enriched samples in this study contained markers typical of the other cell types examined.

In summary, TSA_d was not detected in 27/27 organ EC proteomes examined, and detected in just 2/21 HUVEC samples, both derived from a single dataset in which TSA_d expression was linked to an unexpected immune signature. Proteomic data therefore agree with epigenomic data to show that TSA_d is robustly expressed in T cells and PBMCs but is not detected in organ ECs, and is typically also absent from HUVEC proteomes.

Discussion

Here, we have mined publicly available sc-RNAseq, bulk RNAseq, epigenomic and proteomic data to investigate whether components of several proposed signalling axes for VEGF-induced vascular permeability induction are expressed in ECs of organs clinically affected by VEGF-induced oedema or in EC subtypes used for discovery science to elucidate vascular permeability signalling pathways.

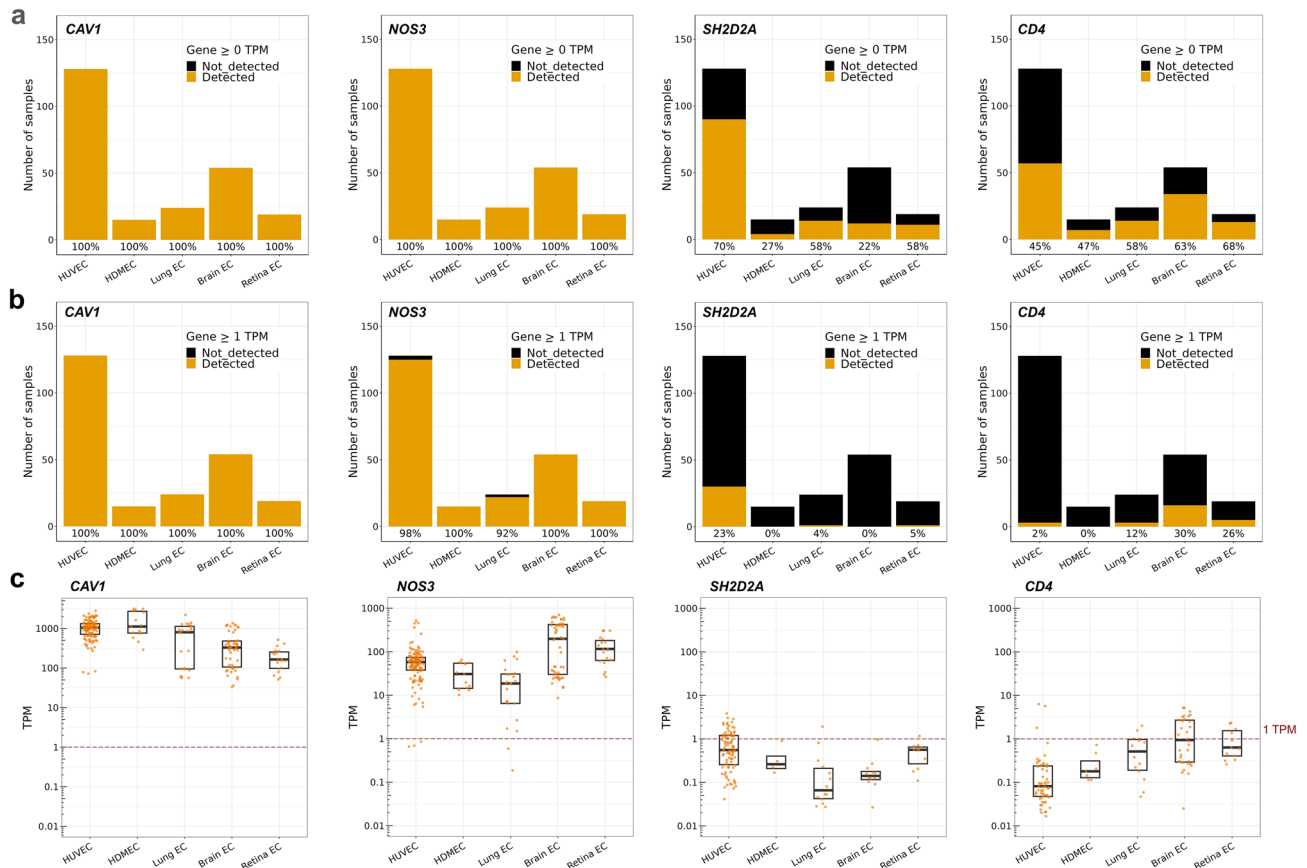


Fig. 3 Bulk RNAseq reveals absent or low *SH2D2A* expression in human and mouse ECs. Expression of *CAV1*, *NOS3*, *SH2D2A* and *CD4* in endothelial bulk RNAseq datasets (HUVEC n=128, HDMEC n=15, mouse lung EC=24, mouse brain EC=54, mouse retina EC n=19). **a**, **b** Stacked bar charts depict the total number of bulk RNAseq datasets analysed and the frequency at which transcripts for the indicated genes were (a) detected (>0 TPM) versus not

detected ($=0$ TPM), or (b) detected (≥ 1 TPM) versus not detected (<1 TPM) above the 1 TPM threshold; data are resolved by EC subtype. **c** Transcript levels for the indicated genes with expression >0 TPM in each dataset for the indicated EC subtypes, including boxplots to illustrate the median and interquartile range; each data point represents one dataset. The dashed red line indicates the 1 TPM threshold. For n, see corresponding source data file

We found that all signal transducers examined were expressed in the ECs of the organs examined, with organ selectivity for some components such as *SDC2*, but very low or complete lack of TSA detection. As these findings are incompatible with a role for TSA in vascular permeability signalling, we discuss below the evidence and potential limitations of our investigation for each method we employed.

We found that sc-RNAseq datasets of ECs from such human organs consistently contained transcripts for all VEGF permeability pathway-relevant signal transducers examined, with exception of transcripts for TSA (Fig. 1, Tables 1 and 2). Moreover, TSA transcripts were also lacking from all EC transcriptomes contained in the EC atlas and barely present in all other EC transcriptomes included in *Tabula Muris* (Fig. 2, Tables 1 and 2). We also used the Bulk-ECexplorer as a complementary resource to determine whether relevant signal transducers were expressed in ECs, and to predict whether detectable gene expression was

likely at functional levels [35]. Transcripts for *VEGFR2* and *CDH5* are defining features of Bulk-ECexplorer datasets, and transcripts for *YES1* and *SRC* are detected as actively expressed genes in these datasets [35]. Except for *SH2D2A* encoding TSA, all other VEGF permeability pathway-relevant signal transducers examined were consistently and robustly detected, and, accordingly, predicted to be actively expressed and therefore functional (Figs. 3, 4, 5, Supplemental Table 1). By contrast, *SH2D2A* detection in half of the EC datasets of the Bulk-ECexplorer datasets was low and mostly in HUVECs (Figs. 3a, c, S1B), agreeing with prior work in HUVECs, which reported that *SH2D2A* detection in reverse transcribed mRNA required a large number (45) of polymerase chain reaction (PCR) cycles [22]. Moreover, as reported for porcine aortic endothelial (PAE) cells [22], *SH2D2A* transcript levels did not increase in HUVECs after VEGF stimulation (Supplemental Table S2).

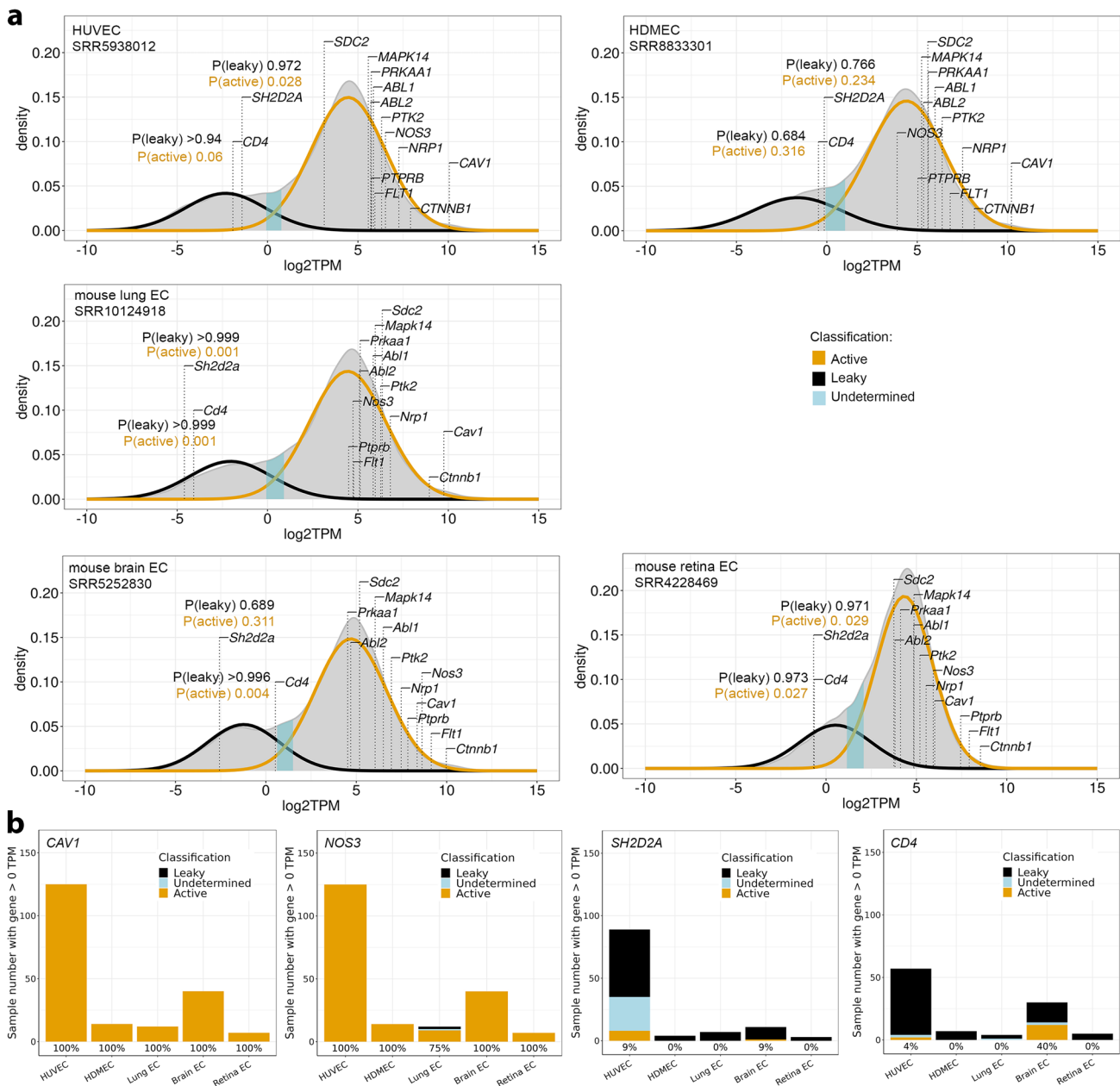


Fig. 4 GMM-based classification links *SH2D2A* detection to leaky transcription. **a** Illustrative kernel density estimates of \log_2 -transformed TPM values for protein-coding genes in the indicated bulk RNA seq datasets from the Bulk-ECexplorer. The expectation maximisation algorithm was used to estimate the parameters of a low Gaussian distribution (predicted leaky expression, black line) and a high Gaussian distribution (predicted active expression, gold line). Specific \log_2 TPM values are shown for known vascular permeability signal transducers as well as *SH2D2A* and *CD4*. The $P(\text{active})$ and $P(\text{leaky})$ values are shown for *SH2D2A* and *CD4* and for all indicated

genes listed in Supplemental Table 1. **b** Stacked bar charts show the number of Bulk-ECexplorer datasets for each EC subtype in which the indicated genes were classified with the GMM method as actively expressed (gold), leakily expressed (black) or in which the gene was not classified (undetermined, blue). Only datasets with a bimodal distribution of gene expression values analysed. Number of bimodally distributed datasets with detectable transcript for the examined gene: *CAV1* and *NOS3* $n=198$; *SH2D2A* $n=114$, *CD4* $n=103$. The percentage of datasets in which each gene was classified as actively expressed is reported for each EC subtype below each bar

Low or absent *SH2D2A* detection in both EC single cell and bulk RNAseq data is incompatible with TSAd’s proposed central role in VEGF/VEGFR2-mediated vascular permeability signalling. For this reason, we considered

technical parameters that could have, conceivably, impacted our analysis. Unsuitable sc-RNAseq methods could not account for a lack of *SH2D2A* transcript detection, because detection was robust in immune cells (Fig. 2). For bulk

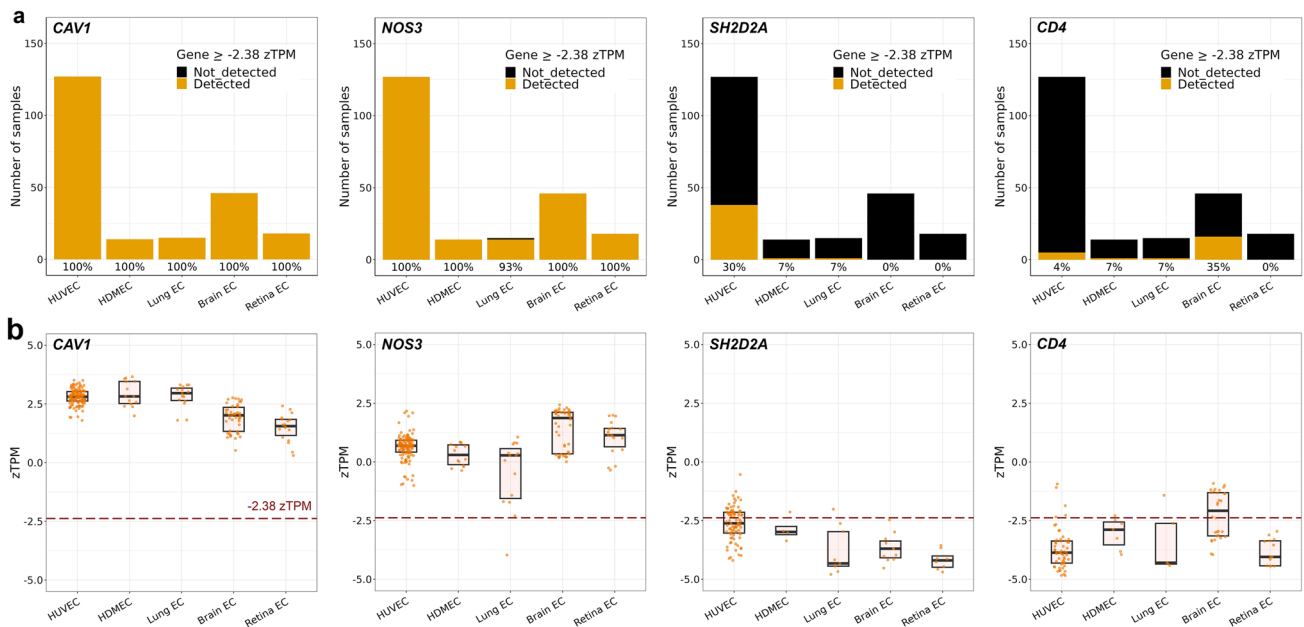


Fig. 5 zTPM standardisation links *SH2D2A* detection in ECs to leaky transcription. Number of datasets analysed across EC subtypes for each of the four genes $n=220$. The active expression threshold of -2.38 zTPM was previously determined for HUVEC. **a** Stacked bar charts show the number of datasets for each EC subtype in which the indicated genes were detected above (gold) or below (black) the -2.38 zTPM threshold. The percentage of datasets in which the indicated gene was expressed above the -2.38 zTPM threshold is reported below each bar for the corresponding EC subtype. HUVEC $n=127$,

HDMEC $n=14$, mouse lung EC $n=15$, mouse brain EC $n=46$, mouse retina EC $n=18$. **b** Range of zTPM values for the indicated genes in each of the indicated EC subtypes, including boxplots to illustrate the interquartile range. Each data point represents one dataset; values are shown together with boxplots to illustrate the interquartile range (for TPM values and n , see corresponding source data file). The dashed red line indicates the -2.38 zTPM threshold. Only datasets with gene detection >0 TPM and without a unimodal distribution were included

RNAseq analysis, we considered whether the HISAT2-Stringtie pipeline used to quantify transcript abundance in the Bulk-ECexplorer may have impacted *SH2D2A* transcript detection. However, many of the studies from which we sourced EC bulk RNAseq data used alternative pipelines in their own analyses, and inspection of their published summary data similarly indicated absent or very low *SH2D2A* expression (although these studies did not comment on *SH2D2A* detection levels per se). For example, TopHat-Cufflinks was used to process bulk RNAseq data for HDMECs and HUVECs [89], TopHat-HTSeq for retina ECs [90], TopHat-Bowtie for brain ECs [91] and TopHat2-Seqsolve for lung ECs [92]. Further, we considered that insufficient read depth can sometimes explain the absence of a transcript in bulk RNAseq data [61], and, in agreement, datasets with higher read depth were more likely to return reads for *SH2D2A* (Supplemental Fig. S2). However, transcripts detectable only at high read depth are inherently rare and may arise from leaky transcription, as previously observed in many different cell types [62, 69–72], including in ECs [35]. Therefore, we used the GMM and zTPM threshold tools included with the Bulk-ECexplorer to calculate the probabilities of *SH2D2A* being a leaky transcript in ECs. Both

methods predicted that *SH2D2A* transcript presence in EC datasets likely reflects leaky transcription (Figs. 4 and 5).

Drawing on the concept of leaky transcription to explain low *SH2D2A* levels in ECs will be controversial, because it implies that TSA_d is unlikely to be a central, functional player in vascular permeability signalling as held in current vascular permeability models. For this reason, we have re-considered the strength of evidence that supports a two-component transcriptome that is represented by a bimodal distribution of transcript abundance in bulk RNAseq data and reflects the leaky versus active transcriptome. Supporting this concept, several studies in different cell types reported that LE transcripts encode proteins not expected to be functional in that cell type; for instance, olfactory-related LE transcripts were identified in cervical carcinoma-derived HeLa cells [70], whereas T helper cell (Th) 1-specific LE transcripts were identified in Th2 cells [73]. In analogy, the T cell marker *CD4* was detected at very low levels in approximately half of the 240 EC RNAseq datasets included in the Bulk-ECexplorer (Fig. 3), which was classified as leaky expression (Figs. 4, 5), akin to prior observations for the erythroid marker *KLF1* in ECs [35]. Further, the expression profiles of *SH2D2A*, *CD4* and *KLF1* were similar to several other genes not expected to be functional within ECs,

including other immune genes and osteoblast or sex cell-specific genes [35].

Analysis of epigenomic data across different cell types has previously provided additional and decisive support for the concept of a two-component transcriptome model arising from leaky and active gene expression. For example, a comparison of DNase-seq and RNAseq data for Th2 cells revealed that the activating histone modification H3K9/14ac is enriched near genes classified by the GMM approach as HE/active, but this modification was largely absent from genes classified as LE/leaky [73]. In another example, a study analysing ENCODE RNAseq and ChIP-seq data reported that the promoters of GMM-classified HE/active genes are consistently associated with active chromatin markers in several different cell types, whereas the promoters of GMM-classified LE/leaky genes are frequently associated with repressive chromatin markers in the same cell types [62]. Here, we have also examined ENCODE epigenetic data obtained by DNase-seq and found that the human *SH2D2A* promoter was positioned within open chromatin in 15 leukocyte datasets but within closed chromatin in 15 EC datasets (Fig. 7), including in HUVECs and HDMECs that have previously been reported to contain TSAAd protein. This finding is most consistent with *SH2D2A* being an immune-related gene that is either not expressed in ECs or is expressed as a leaky gene in ECs.

Agreeing with the idea that TSAAd is a non-expressed or leaky gene in ECs, TSAAd was not recorded in 53/55 EC proteome samples examined and the 2 remaining samples that recorded TSAAd were HUVECs also enriched in other immune cell markers (Table 3, Supplemental Tables 3 and 6). By contrast, 2/2 T cell and 3/3 PBMC proteome samples examined detected TSAAd (Table 3, Supplemental Table 5), thus arguing against the possibility that poor detectability of TSAAd peptides accounted for TSAAd's absence in the recorded EC proteomes. Across the EC proteomes examined, our findings link TSAAd to immune marker expression rather than a pure EC phenotype. Conceivably, HUVECs may have an inherent propensity for immune cell marker expression under specific environmental conditions. Alternatively, we need to consider the possibility of immune cell contamination, because these samples were part of a study that directly compared endothelial, PBMC and fibroblast proteomes, and several samples from these cell types contained markers of the other cell types included in the study (Supplemental Table 3).

Aside from the finding that HUVEC samples from one laboratory had an immune signature that included TSAAd, two other laboratories have reported that TSAAd is detectable by immunoblotting in HUVECs and HDMECs [22, 24, 93]. Immunoblotting differs from mass spectrometry by its reliance on highly specific antibodies for protein detection and is likely more sensitive for detecting low protein

levels than mass spectrometry-based methods. Therefore, we cannot exclude that TSAAd protein could become detectable in EC proteomic datasets generated with technologies that have increased sensitivity for detecting and reporting proteins translated from rare transcripts. Nevertheless, transcripts or protein detected at very low levels may not be of biological significance [62, 70–73]. Furthermore, it is surprising that a protein whose transcripts are either not detectable or detected at levels below 1 TPM in HDMECs (Fig. 3) should be detectable by Western blotting in these cells. Therefore, we have considered that leaky transcripts may not be uniformly distributed across a cell population, but instead can be present in a very small proportion of cells within the overall cell population, as reported for the *Tbx21* gene in Th2 cells [73]. If a very small proportion of ECs were to express TSAAd at high enough levels to lead to protein production detectable with very sensitive methods, a hypothetical scenario arises in which rare TSAAd-expressing ECs might selectively respond to VEGF with vascular permeability signalling. For example, it has been proposed that vascular leakage occurs at distinct focal points of a blood vessel [94]. However, the sc-RNAseq dataset from human skin, the only human organ reported to contain vessel-associated TSAAd protein detectable by immunostaining, lacked an obvious population of *SH2D2A*-positive ECs (Fig. 1). As our analysis of human dermis and trachea as well as mouse organ sc-RNAseq data did not identify a subset of *SH2D2A*-positive ECs in any of these organs, it is difficult to conceive how focal activity of an EC subset that is too rare to be identified by sc-RNAseq could then cause measurable vascular leakage across an organ.

We have also considered whether published immunostaining data for TSAAd support the possibility of an EC subset with functional TSAAd protein levels. However, we found that no prior study was compatible with the existence of TSAAd hotspots in vascular endothelium. Firstly, one study [28] found TSAAd lacking from blood vessels in all human organs examined except in the dermis, where continuous staining of putative vessels was reported. Secondly, another immunostaining study suggested continuous staining of kidney endothelium [22]. However, it should also be considered that neither immunostaining study had specifically validated the EC identity of the stained cell type (as opposed to an endothelium-associated cell type) nor have prior studies excluded cross-reactivity of the antibody with an EC epitope that does not belong to TSAAd, even though antibodies were clearly shown to be specific for TSAAd in immune cells. In agreement with antibody cross-reactivity to a non-TSAAd epitope in ECs, the supplemental data of a prior study showed positive reactivity in the lung lysates of constitutive TSAAd-null mice by immunoblotting [95]. Additional evidence from more

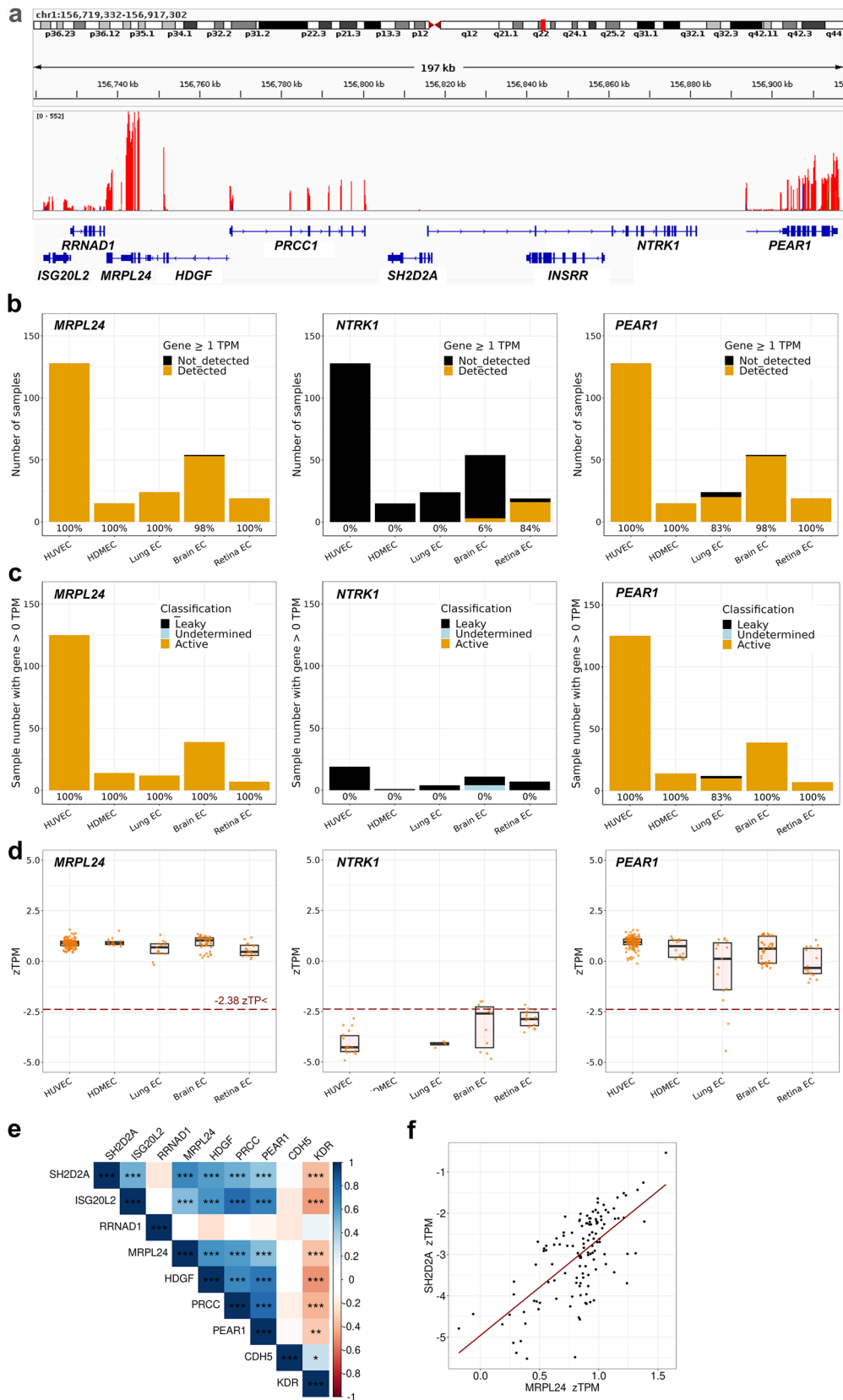


Fig. 6 *SH2D2A* expression correlates with neighbouring gene expression in ECs. **a** Read coverage for *SH2D2A* and its neighbouring genes (± 100 kilo base pairs) on chromosome 1 in the HUVEC sample SRR1640083 that has 0.95 TPM for *SH2D2A*; human genome build (hg) 38; Y axis axis: read coverage range 0–575. **b** Stacked bar charts show the number of Bulk-ECexplorer datasets (n=240) in which the indicated genes were detected above or below the 1 TPM threshold. The percentage of datasets in which the indicated gene was expressed above the threshold is reported for each EC subtype below each bar. **c** Stacked bar charts show the number of Bulk-ECexplorer datasets in which the indicated genes were classified with the GMM method as actively expressed (gold), leakily expressed (black) or in which the gene was not classified (undetermined, blue). Only datasets with a bimodal distribution of gene expression values and detectable transcript were examined; *MRPL24* n=197, *NTRK1* n=42, *PEAR1* n=195. The percentage of datasets in which each gene was classified as actively expressed is reported for each EC subtype below each bar. **d** Range of zTPM values for the indicated genes in Bulk-ECexplorer datasets, including boxplots to illustrate the interquartile range. Each data point represents one dataset; values are shown together with boxplots to illustrate the interquartile range (for TPM values and n, see corresponding source data file). The dashed red line indicates the -2.38 zTPM threshold in HUVEC. Each data point represents one sample; *MRPL24* n=219, *NTRK1* n=52, *PEAR1* n=219 samples across all EC subtypes for each gene (only datasets with gene detection >0 TPM and without a unimodal distribution were included). **e** Correlation of zTPM values for *SH2D2A* to neighbouring HE genes as well as *KDR* and *CDH5* across all EC types; Pearson’s correlation (r) with Holm adjusted p-values, * < 0.05, ** < 0.01, *** < 0.001; n = 126 (except *RRNAD1*, n = 125); unimodal samples were excluded. **f** Linear regression of *SH2D2A* zTPM values onto *MRPL24* zTPM values. Each data point represents the zTPM value of one sample (n = 126; unimodal samples were excluded and only samples with detectable *SH2D2A* and *MRPL24* expression were included; coefficient of determination $R^2=0.45$; slope co-efficient $\beta_1 = 2.33$ and p-value for regression $< 2.2^{-16}$

informative immunostaining protocols, immunoblotting with conclusively validated antibodies for ECs and/or the analysis of cell-type specific TSAd knockout mice would therefore be required to uphold an EC cell autonomous role for TSAd in vascular permeability. Alternatively, it should be considered that TSAd expression in non-ECs, such as vessel-associated immune cells, contributes to EC barrier regulation to explain impaired VEGF-induced vascular permeability induction in TSAd knockout mice, or that, in such genetically-modified mice, another genomic element was inadvertently disrupted during gene targeting to affect vascular barrier function, such as, hypothetically, a regulatory element for one of the HE genes in the *SH2D2A* neighbourhood. Notably, *SH2D2A* neighbouring genes have not yet been studied for potential functions in vascular permeability, although its closest genomic neighbour, *HDGF*, has been described as an angiogenesis regulator [96, 97].

In summary, EC expression of TSAd transcripts is either not detectable, or detectable only at very low levels by RNAseq assays. When combined with EC epigenomic and proteomic data, these findings support the conclusion that TSAd is either not expressed, or a leaky and non-functional gene in ECs. By extension, our findings also argue against

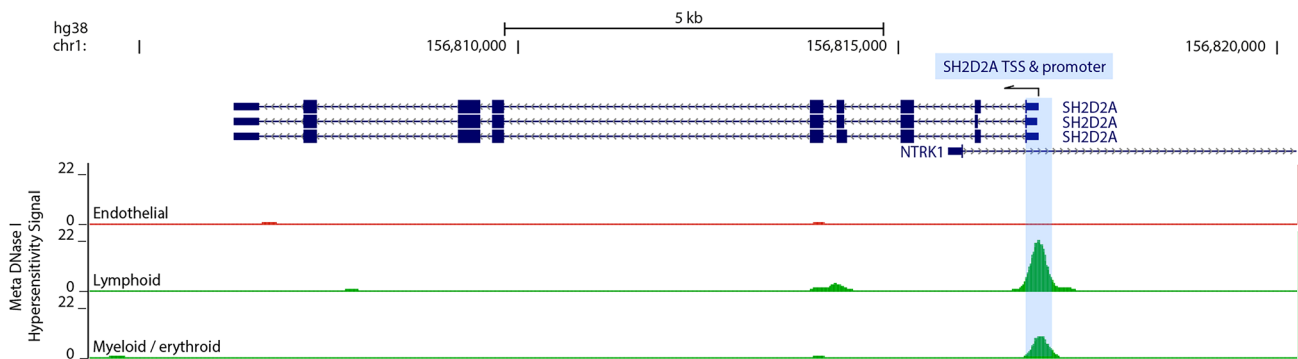


Fig. 7 The *SH2D2A* promoter resides in closed chromatin in ECs. The ENCODE meta-DNase I HS signal for the human *SH2D2A* locus is shown for endothelial, lymphoid and myeloid cells (n = 15 samples for each of the three meta-profile). The intron-exon structure of 3

known *SH2D2A* transcripts and the first exon and intron of the neighbouring *NTRK1* transcript are shown. The *SH2D2A* transcription start site (TSS) and promoter region are highlighted in blue.

a central role for TSA α in the endothelial vascular permeability signalling cascade, thereby excluding it as useful target for therapeutic approaches that seek to inhibit VEGF-induced oedema. By contrast, all other genes implicated in VEGF-induced permeability signalling were robustly expressed in ECs at both transcript and protein level, with some organ specificity. These findings should be considered in studies seeking to identify therapeutic tools for modulating vascular barriers. Further, our study may provide the impetus for future work to examine vascular permeability-relevant gene expression in additional contexts, such as tumour biology, lung and eye disease, where vascular permeability mediators other than VEGF contribute to vascular barrier breakdown.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10456-025-09971-x>.

Acknowledgements We thank Martina Rudnicki and Tiago V. Pereira for advice on statistical analysis and Blake Moore for technical advice on Linux virtual machines.

Author's contributions J.T.B., A.F., G.D.P. and C.R. conceived and designed the study, analysed the data and co-wrote the manuscript. J.T.B., A.F., L.R., R.F.P.C. and G.D.P. performed bioinformatic analyses. All authors read and approved the submitted manuscript.

Funding This study was supported by grants from the British Heart Foundation [PG/19/37/34399], Medical Research Council [MR/N011511/1] and Wellcome [205099/Z/16/Z] to CR, the Fondazione Cariplo (2018–2028) and the Fondazione Associazione Italiana per la Ricerca sul Cancro (AIRC) (22905) to AF and a PhD fellowship from the BHF to GDP [FS/4yPhD/F/21/34157].

Data Availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Dolgin E (2017) The most popular genes in the human genome. *Nature* 551(7681):427–431. <https://doi.org/10.1038/441586-017-07291-9>
- Carmeliet P, Ferreira V, Breier G, Pollefeyt S, Kieckens L, Gertsenstein M et al (1996) Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. *Nature* 380(6573):435–439
- Apte RS, Chen DS, Ferrara N (2019) VEGF in signaling and disease: beyond discovery and development. *Cell* 176(6):1248–1264. <https://doi.org/10.1016/j.cell.2019.01.021>
- Weis SM, Cheresh DA (2005) Pathophysiological consequences of VEGF-induced vascular permeability. *Nature* 437(7058):497–504. <https://doi.org/10.1038/nature03987>
- Campochiaro PA, Akhlaq A (2021) Sustained suppression of VEGF for treatment of retinal/choroidal vascular diseases. *Prog Retin Eye Res* 83:100921. <https://doi.org/10.1016/j.preteyeres.2020.100921>
- Senger DR, Galli SJ, Dvorak AM, Perruzzi CA, Harvey VS, Dvorak HF (1983) Tumor cells secrete a vascular permeability factor that promotes accumulation of ascites fluid. *Science* 219(4587):983–985
- Campochiaro PA (2015) Molecular pathogenesis of retinal and choroidal vascular diseases. *Prog Retin Eye Res* 49:67–81. <https://doi.org/10.1016/j.preteyeres.2015.06.002>
- Komarova YA, Kruse K, Mehta D, Malik AB (2017) Protein interactions at endothelial junctions and signaling mechanisms regulating endothelial permeability. *Circ Res* 120(1):179–206. <https://doi.org/10.1161/CIRCRESAHA.116.306534>
- Wettschureck N, Strilic B, Offermanns S (2019) Passing the vascular barrier: endothelial signaling processes controlling extravasation. *Physiol Rev* 99(3):1467–1525. <https://doi.org/10.1152/physrev.00037.2018>
- Claesson-Welsh L, Dejana E, McDonald DM (2021) Permeability of the endothelial barrier: identifying and reconciling controversies. *Trends Mol Med* 27(4):314–331. <https://doi.org/10.1016/j.molmed.2020.11.006>
- Domingues A, Fantin A (2021) Neuropilin 1 regulation of vascular permeability signaling. *Biomolecules* 11(5). <https://doi.org/10.3390/biom11050666>
- Acevedo LM, Barillas S, Weis SM, Gothert JR, Cheresh DA (2008) Semaphorin 3A suppresses VEGF-mediated angiogenesis yet acts as a vascular permeability factor. *Blood* 111(5):2674–2680. <https://doi.org/10.1182/blood-2007-08-110205>
- Fantin A, Lampropoulou A, Senatore V, Brash JT, Prahst C, Lange CA et al (2017) VEGF165-induced vascular permeability requires NRP1 for ABL-mediated SRC family kinase activation. *J Exp Med* 214(4):1049–1064. <https://doi.org/10.1084/jem.20160311>
- Corti F, Ristori E, Rivera-Molina F, Toomre D, Zhang J, Mihailovic J et al (2022) Syndecan-2 selectively regulates VEGF-induced vascular permeability. *Nat Cardiovasc Res* 1(5):518–528. <https://doi.org/10.1038/s44161-022-00064-2>
- Lambeng N, Wallez Y, Rampon C, Cand F, Christe G, Gulino-Debrac D et al (2005) Vascular endothelial-cadherin tyrosine phosphorylation in angiogenic and quiescent adult tissues. *Circ Res* 96(3):384–391. <https://doi.org/10.1161/01.RES.0000156652.99586.9f>
- Orsenigo F, Giampietro C, Ferrari A, Corada M, Galaup A, Sigismund S et al (2012) Phosphorylation of VE-cadherin is modulated by haemodynamic forces and contributes to the regulation of vascular permeability in vivo. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. *Nat Commun* 3:1208. <https://doi.org/10.1038/ncomms2199>
- Weis S, Shintani S, Weber A, Kirchmair R, Wood M, Cravens A et al (2004) Src blockade stabilizes a Flk/cadherin complex, reducing edema and tissue injury following myocardial infarction. [Research Support, Non-U.S. Gov't Research Support, U.S.

- Gov't, P.H.S.]. *J Clin Invest* 113(6):885–894. <https://doi.org/10.1172/JCI20702>.
18. Wallez Y, Cand F, Cruzalegui F, Wernstedt C, Souchelnytskyi S, Vilgrain I et al (2007) Src kinase phosphorylates vascular endothelial-cadherin in response to vascular endothelial growth factor: identification of tyrosine 685 as the unique target site. *Oncogene* 26(7):1067–1077. <https://doi.org/10.1038/sj.onc.1209855>
 19. Chislock EM, Pendergast AM (2013) Abl family kinases regulate endothelial barrier function in vitro and in mice. *PLoS ONE* 8(12):e85231. <https://doi.org/10.1371/journal.pone.0085231>
 20. Eliceiri BP, Paul R, Schwartzberg PL, Hood JD, Leng J, Cheresh DA (1999) Selective requirement for Src kinases during VEGF-induced angiogenesis and vascular permeability. *Mol Cell* 4(6):915–924
 21. Jin Y, Ding Y, Richards M, Kaakinen M, Giese W, Baumann E et al (2022) Tyrosine-protein kinase Yes controls endothelial junctional plasticity and barrier integrity by regulating VE-cadherin phosphorylation and endocytosis. *Nat Cardiovasc Res* 1(12):1156–1173. <https://doi.org/10.1038/s44161-022-00172-z>
 22. Matsumoto T, Bohman S, Dixelius J, Berge T, Dimberg A, Magnusson P et al (2005) VEGF receptor-2 Y951 signaling and a role for the adapter molecule TSAd in tumor angiogenesis. *EMBO J* 24(13):2342–2353. <https://doi.org/10.1038/sj.emboj.7600709>
 23. Wu LW, Mayo LD, Dunbar JD, Kessler KM, Ozes ON, Warren RS et al (2000) VRAP is an adaptor protein that binds KDR, a receptor for vascular endothelial cell growth factor. *J Biol Chem* 275(9):6059–6062
 24. Sun Z, Li X, Massena S, Kutschera S, Padhan N, Gualandi L et al (2012) VEGFR2 induces c-Src signaling and vascular permeability in vivo via the adaptor protein TSAd. [Research Support, Non-U.S. Gov't]. *J Exp Med* 209(7):1363–1377. <https://doi.org/10.1084/jem.20111343>
 25. Li X, Padhan N, Sjoström EO, Roche FP, Testini C, Honkura N et al (2016) VEGFR2 pY949 signalling regulates adherens junction integrity and metastatic spread. *Nat Commun* 7:11017. <https://doi.org/10.1038/ncomms11017>
 26. Smith RO, Ninchoji T, Gordon E, Andre H, Dejana E, Vestweber D et al (2020) Vascular permeability in retinopathy is regulated by VEGFR2 Y949 signaling to VE-cadherin. *Elife*, 9. <https://doi.org/10.7554/eLife.54056>.
 27. Spurkland A, Brinchmann JE, Markussen G, Pedeutour F, Munthe E, Lea T et al (1998) Molecular cloning of a T cell-specific adapter protein (TSAd) containing an Src homology (SH) 2 domain and putative SH3 and phosphotyrosine binding sites. *J Biol Chem* 273(8):4539–4546. <https://doi.org/10.1074/jbc.273.8.4539>
 28. Pandya AD, Leergaard TB, Dissen E, Haraldsen G, Spurkland A (2014) Expression of the T cell-specific adapter protein in human tissues. *Scand J Immunol* 80(3):169–179. <https://doi.org/10.1111/sji.12199>
 29. Simons M, Toomre D (2022) YES to junctions, No to Src. *Nat Cardiovasc Res* 1(12):1116–1118. <https://doi.org/10.1038/s44161-022-00185-8>
 30. Uemura A, Fruttiger M, D'Amore PA, De Falco S, Jousen AM, Sennlaub F et al (2021) VEGFR1 signaling in retinal angiogenesis and microinflammation. *Prog Retin Eye Res* 84:100954. <https://doi.org/10.1016/j.preteyeres.2021.100954>
 31. Takahashi H, Hattori S, Iwamatsu A, Takizawa H, Shibuya M (2004) A novel snake venom vascular endothelial growth factor (VEGF) predominantly induces vascular permeability through preferential signaling via VEGF receptor-1. *J Biol Chem* 279(44):46304–46314. <https://doi.org/10.1074/jbc.M403687200>
 32. Dragoni S, Caridi B, Karatsai E, Burgoyne T, Sarker MH, Turowski P (2021) AMP-activated protein kinase is a key regulator of acute neurovascular permeability. *J Cell Sci* 134(7). <https://doi.org/10.1242/jcs.253179>.
 33. Chen XL, Nam JO, Jean C, Lawson C, Walsh CT, Goka E et al (2012) VEGF-induced vascular permeability is mediated by FAK. *Dev Cell* 22(1):146–157. <https://doi.org/10.1016/j.devcel.2011.11.002>
 34. Vestweber D (2021) Vascular endothelial protein tyrosine phosphatase regulates endothelial function. *Physiology (Bethesda)* 36(2):84–93. <https://doi.org/10.1152/physiol.00026.2020>
 35. Brash JT, Diez-Pinel G, Colletto C, Castellan RFP, Fantin A, Ruhrberg C (2024) The BulkECexplorer compiles endothelial bulk transcriptomes to predict functional versus leaky transcription. *Nat Cardiovasc Res*. <https://doi.org/10.1038/s44161-024-00436-w>
 36. Li Q, Zhu Z, Wang L, Lin Y, Fang H, Lei J et al (2021) Single-cell transcriptome profiling reveals vascular endothelial cell heterogeneity in human skin. *Theranostics* 11(13):6461–6476. <https://doi.org/10.7150/thno.54917>
 37. Han X, Zhou Z, Fei L, Sun H, Wang R, Chen Y et al (2020) Construction of a human cell landscape at single-cell level. *Nature* 581(7808):303–309. <https://doi.org/10.1038/s41586-020-2157-4>
 38. Tacconi C, Plein A, Colletto C, Villa E, Denti L, Barone C et al (2022) KIT is dispensable for physiological organ vascularisation in the embryo. *Angiogenesis* 25(3):343–353. <https://doi.org/10.1007/s10456-022-09837-6>
 39. Kalucka J, de Rooij L, Goveia J, Rohlenova K, Dumas SJ, Meta E et al (2020) Single-cell transcriptome atlas of murine endothelial cells. *Cell* 180(4):764–779.e720. <https://doi.org/10.1016/j.cell.2020.01.015>
 40. The Tabula Muris Consortium, Overall coordination, Logistical coordination, Organ collection and processing, Library preparation and sequencing, Computational data analysis, Cell type annotation, Writing group, Supplemental text writing group & Principal investigators (2018) Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* 562(7727):367–372. <https://doi.org/10.1038/s41586-018-0590-4>
 41. RStudio_Team_(2020) (2020). RStudio: integrated development for R. RStudio, PBC, Boston, MA.
 42. Satija R, Farrell JA, Gennert D, Schier AF, Regev A (2015) Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* 33(5):495–502. <https://doi.org/10.1038/nbt.3192>
 43. Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C (2017) Nextflow enables reproducible computational workflows. *Nat Biotechnol* 35(4):316–319. <https://doi.org/10.1038/nbt.3820>
 44. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550. <https://doi.org/10.1186/s13059-014-0550-8>
 45. Meuleman W, Muratov A, Rynes E, Halow J, Lee K, Bates D et al (2020) Index and biological spectrum of human DNase I hypersensitive sites. *Nature* 584(7820):244–251. <https://doi.org/10.1038/s41586-020-2559-3>
 46. Slany A, Bileck A, Kreutz D, Mayer RL, Muqaku B, Gerner C (2016) Contribution of human fibroblasts and endothelial cells to the hallmarks of inflammation as determined by proteome profiling. *Mol Cell Proteomics* 15(6):1982–1997. <https://doi.org/10.1074/mcp.M116.058099>
 47. Bileck A, Kreutz D, Muqaku B, Slany A, Gerner C (2014) Comprehensive assessment of proteins regulated by dexamethasone reveals novel effects in primary human peripheral blood mononuclear cells. *J Proteome Res* 13(12):5989–6000. <https://doi.org/10.1021/pr5008625>
 48. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression

- profiles. *Proc Natl Acad Sci U S A* 102(43):15545–15550. <https://doi.org/10.1073/pnas.0506580102>
49. Murohara T, Horowitz JR, Silver M, Tsurumi Y, Chen D, Sullivan A et al (1998) Vascular endothelial growth factor/vascular permeability factor enhances vascular permeability via nitric oxide and prostacyclin. *Circulation* 97(1):99–107
 50. Six I, Kureishi Y, Luo Z, Walsh K (2002) Akt signaling mediates VEGF/VPF vascular permeability in vivo. *FEBS Lett* 532(1–2):67–69
 51. Brash JT, Ruhrberg C, Fantin A (2018) Evaluating vascular hyperpermeability-inducing agents in the skin with the miles assay. *J Vis Exp*, 136. <https://doi.org/10.3791/57524>.
 52. Moussa P, Abrahamsen G, Fodil N, Gopalakrishnan RP, Mancini M, Dissen E et al (2015) Reduced MCMV Deltam157 viral clearance in the absence of TSA. *Sci Rep* 5:9219. <https://doi.org/10.1038/srep09219>
 53. Berge T, Gronningsaeter IH, Lorvik KB, Abrahamsen G, Granum S, Sundvold-Gjerstad V et al (2012) SH2D2A modulates T cell mediated protection to a B cell derived tumor in transgenic mice. *PLoS ONE* 7(10):e48239. <https://doi.org/10.1371/journal.pone.0048239>
 54. Thomas ML, Green JR (1983) Molecular nature of the W3/25 and MRC OX-8 marker antigens for rat T lymphocytes: comparisons with mouse and human antigens. *Eur J Immunol* 13(10):855–858. <https://doi.org/10.1002/eji.1830131014>
 55. Wilson GL, Fox CH, Fauci AS, Kehrl JH (1991) cDNA cloning of the B cell membrane protein CD22: a mediator of B-B cell interactions. *J Exp Med* 173(1):137–146. <https://doi.org/10.1084/jem.173.1.137>
 56. Baluk P, McDonald DM (2022) Imaging blood vessels and lymphatics in mouse trachea wholemounts. *Methods Mol Biol* 2441:115–134. https://doi.org/10.1007/978-1-0716-2059-5_10
 57. Schubert W, Frank PG, Woodman SE, Hyogo H, Cohen DE, Chow CW et al (2002) Microvascular hyperpermeability in caveolin-1 (-/-) knock-out mice. Treatment with a specific nitric-oxide synthase inhibitor, L-NAME, restores normal microvascular permeability in Cav-1 null mice. *J Biol Chem* 277(42):40091–40098. <https://doi.org/10.1074/jbc.M205948200>
 58. Marinov GK, Williams BA, McCue K, Schroth GP, Gertz J, Myers RM et al (2014) From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing. *Genome Res* 24(3):496–510. <https://doi.org/10.1101/gr.161034.113>
 59. Esser S, Lampugnani MG, Corada M, Dejana E, Risau W (1998) Vascular endothelial growth factor induces VE-cadherin tyrosine phosphorylation in endothelial cells. *J Cell Sci* 111(Pt 13):1853–1865
 60. Simard JM, Kent TA, Chen M, Tarasov KV, Gerzanich V (2007) Brain oedema in focal ischaemia: molecular pathophysiology and theoretical implications. *Lancet Neurol* 6(3):258–268. [https://doi.org/10.1016/S1474-4422\(07\)70055-8](https://doi.org/10.1016/S1474-4422(07)70055-8)
 61. Toung JM, Morley M, Li M, Cheung VG (2011) RNA-sequence analysis of human B-cells. *Genome Res* 21(6):991–998. <https://doi.org/10.1101/gr.116335.110>
 62. Hart T, Komori HK, LaMere S, Podshivalova K, Salomon DR (2013) Finding the active genes in deep RNA-seq gene expression studies. *BMC Genomics* 14:778. <https://doi.org/10.1186/1471-2164-14-778>
 63. Rapp J, Jung M, Klar RFU, Wolf J, Arnold J, Gorka O et al (2023) STAT3 signaling induced by the IL-6 family of cytokines modulates angiogenesis. *J Cell Sci* 136(1). <https://doi.org/10.1242/jcs.260182>.
 64. Wang S, Chen J, Garcia SP, Liang X, Zhang F, Yan P et al (2019) A dynamic and integrated epigenetic program at distal regions orchestrates transcriptional responses to VEGFA. *Genome Res* 29(2):193–207. <https://doi.org/10.1101/gr.239053.118>
 65. Rapp J, Ness J, Wolf J, Hospach A, Liang P, Hug MJ et al (2024) 2D and 3D in vitro angiogenesis assays highlight different aspects of angiogenesis. *Biochim Biophys Acta Mol Basis Dis* 1870(3):167028. <https://doi.org/10.1016/j.bbadis.2024.167028>
 66. Ebisuya M, Yamamoto T, Nakajima M, Nishida E (2008) Ripples from neighbouring transcription. *Nat Cell Biol* 10(9):1106–1113. <https://doi.org/10.1038/ncb1771>
 67. Hebenstreit D, Deonarine A, Babu MM, Teichmann SA (2012) Duel of the fates: the role of transcriptional circuits and noise in CD4+ cells. *Curr Opin Cell Biol* 24(3):350–358. <https://doi.org/10.1016/j.ccb.2012.03.007>
 68. Gray JM, Harmin DA, Boswell SA, Cloonan N, Mullen TE, Ling JJ et al (2014) SnapShot-Seq: a method for extracting genome-wide, in vivo mRNA dynamics from a single total RNA sample. *PLoS ONE* 9(2):e89673. <https://doi.org/10.1371/journal.pone.0089673>
 69. Miika K, Marinic M, Singh M, Muter J, Brosens JJ, Lynch VJ (2021) Evolutionary transcriptomics implicates new genes and pathways in human pregnancy and adverse pregnancy outcomes. *Elife*, 10. <https://doi.org/10.7554/eLife.69584>.
 70. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J et al (2011) Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol* 7:548. <https://doi.org/10.1038/msb.2011.81>
 71. Sharma A, Shiras A (2015) Cancer stem cell-vascular endothelial cell interactions in glioblastoma. *Biochem Biophys Res Commun*. <https://doi.org/10.1016/j.bbrc.2015.12.022>
 72. Walley JW, Sartor RC, Shen Z, Schmitz RJ, Wu KJ, Urlich MA et al (2016) Integration of omic networks in a developmental atlas of maize. *Science* 353(6301):814–818. <https://doi.org/10.1126/science.aag1125>
 73. Hebenstreit D, Fang M, Gu M, Charoensawan V, van Oudenaarden A, Teichmann SA (2011) RNA sequencing reveals two major classes of gene expression levels in metazoan cells. *Mol Syst Biol* 7:497. <https://doi.org/10.1038/msb.2011.28>
 74. Hebenstreit D, Teichmann SA (2011) Analysis and simulation of gene expression profiles in pure and mixed cell populations. *Phys Biol* 8(3):035013. <https://doi.org/10.1088/1478-3975/8/3/035013>
 75. ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489(7414):57–74. <https://doi.org/10.1038/nature11247>
 76. Madugundu AK, Na CH, Nirujogi RS, Renuse S, Kim KP, Burns KH et al (2019) Integrated transcriptomic and proteomic analysis of primary human umbilical vein endothelial cells. *Proteomics* 19(15):e1800315. <https://doi.org/10.1002/pmic.201800315>
 77. Patella F, Neilson LJ, Athineos D, Erami Z, Anderson KI, Blyth K et al (2016) In-depth proteomics identifies a role for autophagy in controlling reactive oxygen species mediated endothelial permeability. *J Proteome Res* 15(7):2187–2197. <https://doi.org/10.1021/acs.jproteome.6b00166>
 78. Afshar Y, Ma F, Quach A, Jeong A, Sunshine HL, Freitas V et al (2023) Transcriptional drifts associated with environmental changes in endothelial cells. *Elife*, 12. <https://doi.org/10.7554/eLife.81370>.
 79. Kostyunina DS, Pakhomov NV, Jouida A, Dillon E, Baugh JA, McLoughlin P (2024) Transcriptomics and proteomics revealed sex differences in human pulmonary microvascular endothelial cells. *Physiol Genomics* 56(2):194–220. <https://doi.org/10.1152/physiolgenomics.00051.2023>
 80. Sacnun JM, Hoogenboom R, Eibensteiner F, Sobieszek JJ, Unterwurzacher M, Wagner A et al (2022) Proteome-wide differential effects of peritoneal dialysis fluid properties in an in vitro human endothelial cell model. *Int J Mol Sci* 23(14). <https://doi.org/10.3390/ijms23148010>.
 81. Zanivan S, Maione F, Hein MY, Hernandez-Fernaund JR, Ostasiewicz P, Giraudo E et al (2013) SILAC-based proteomics of

- human primary endothelial cell morphogenesis unveils tumor angiogenic markers. *Mol Cell Proteomics* 12(12):3599–3611. <https://doi.org/10.1074/mcp.M113.031344>
82. Groten SA, Smit ER, van den Biggelaar M, Hoogendijk AJ (2024) The proteomic landscape of in vitro cultured endothelial cells across vascular beds. *Commun Biol* 7(1):989. <https://doi.org/10.1038/s42003-024-06649-w>
83. Cao Z, Minnier J, Liu L, Scott KLL, Reddy AP, Wilmarth PA et al (2022) Proteomic profiling of concurrently isolated primary microvascular endothelial cells, pericytes, and vascular smooth muscle cells from adult mouse heart. *Sci Rep* 12(1):8835. <https://doi.org/10.1038/s41598-022-12749-6>
84. Kostyunina DS, Rowan SC, Pakhomov NV, Dillon E, Rochfort KD, Cummins PM et al (2023) Shear stress markedly alters the proteomic response to hypoxia in human pulmonary endothelial cells. *Am J Respir Cell Mol Biol* 68(5):551–565. <https://doi.org/10.1165/rcmb.2022-03400C>
85. Alonso F, Dong Y, Li L, Jahjah T, Dupuy JW, Fremaux I et al (2023) Fibrillin-1 regulates endothelial sprouting during angiogenesis. *Proc Natl Acad Sci U S A* 120(23):e2221742120. <https://doi.org/10.1073/pnas.2221742120>
86. Smith JR, David LL, Appukuttan B, Wilmarth PA (2018) Angiogenic and immunologic proteins identified by deep proteomic profiling of human retinal and choroidal vascular endothelial cells: potential targets for new biologic drugs. *Am J Ophthalmol* 193:197–229. <https://doi.org/10.1016/j.ajo.2018.03.020>
87. Dyring-Andersen B, Lovendorf MB, Coscia F, Santos A, Moller LBP, Colaco AR et al (2020) Spatially and cell-type resolved quantitative proteomic atlas of healthy human skin. *Nat Commun* 11(1):5587. <https://doi.org/10.1038/s41467-020-19383-8>
88. Subbannayya Y, Haug M, Pinto SM, Mohanty V, Meas HZ, Flo TH et al (2020) The proteomic landscape of resting and activated CD4+ T cells reveal insights into cell differentiation and function. *Int J Mol Sci* 22(1). <https://doi.org/10.3390/ijms22010275>
89. Sandler VM, Lis R, Liu Y, Kedem A, James D, Elemento O et al (2014) Reprogramming human endothelial cells to haematopoietic cells requires vascular induction. *Nature* 511(7509):312–318. <https://doi.org/10.1038/nature13547>
90. Jeong HW, Hernandez-Rodriguez B, Kim J, Kim KP, Enriquez-Gasca R, Yoon J et al (2017) Transcriptional regulation of endothelial cell behavior during sprouting angiogenesis. *Nat Commun* 8(1):726. <https://doi.org/10.1038/s41467-017-00738-7>
91. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O’Keeffe S et al (2014) An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci* 34(36):11929–11947. <https://doi.org/10.1523/JNEUROSCI.1860-14.2014>
92. Ochoa-Callejero L, Pozo-Rodrigalvarez A, Martinez-Murillo R, Martinez A (2016) Lack of adrenomedullin in mouse endothelial cells results in defective angiogenesis, enhanced vascular permeability, less metastasis, and more brain damage. *Sci Rep* 6:33495. <https://doi.org/10.1038/srep33495>
93. Ruan GX, Kazlauskas A (2012) Axl is essential for VEGF-A-dependent activation of PI3K/Akt. *EMBO J* 31(7):1692–1703. <https://doi.org/10.1038/emboj.2012.21>
94. Honkura N, Richards M, Lavina B, Sainz-Jaspeado M, Betsholtz C, Claesson-Welsh L (2018) Intravital imaging-based analysis tools for vessel identification and assessment of concurrent dynamic vascular events. *Nat Commun* 9(1):2746. <https://doi.org/10.1038/s41467-018-04929-8>
95. Gordon EJ, Fukuhara D, Westrom S, Padhan N, Sjostrom EO, van Meeteren L et al (2016) The endothelial adaptor molecule TSA1 is required for VEGF-induced angiogenic sprouting through junctional c-Src activation. *Sci Signal* 9(437):ra72. <https://doi.org/10.1126/scisignal.aad9256>
96. Everett AD, Narron JV, Stoops T, Nakamura H, Tucker A (2004) Hepatoma-derived growth factor is a pulmonary endothelial cell-expressed angiogenic factor. *Am J Physiol Lung Cell Mol Physiol* 286(6):L1194–L1201. <https://doi.org/10.1152/ajplung.00427.2003>
97. Okuda Y, Nakamura H, Yoshida K, Enomoto H, Uyama H, Hirofumi T et al (2003) Hepatoma-derived growth factor induces tumorigenesis in vivo through both direct angiogenic activity and induction of vascular endothelial growth factor. *Cancer Sci* 94(12):1034–1041. <https://doi.org/10.1111/j.1349-7006.2003.tb01397.x>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.