A human arteriovenous differentiation roadmap reveals vein developmental mechanisms and vascular effects of viruses

Lay Teng Ang[1,2,19,21,*], Sherry Li Zheng[1,3,18,19], Kevin J. Liu[1,2,19], Anastasiia Masaltseva[1,3,19], June Winters[4,20], Isabel von Creytz[5,20], Sawan K. Jha[4,20], Qingqing Yin[1,3,20], Crystal Qian[1,3], Xiaochen Xiong[1,3,18], Amir Dailamy[1,3], Ellie Xi[1,2], Juan C. Alcocer[4], Daniel W. Sorensen[4], Richard She[6], Karina Smolyar[7], Dorota Szumska[8], Svanhild Nornes[8], Renata M. Martin[1,9,18], Benjamin J. Lesch[1,9,18], Nicole K. Restrepo[10], Wenfei Sun[11], Jonathan S. Weissman[7], Heiko Lickert[12,13,14], Matthew P. Porteus[1,9], Mark A. Skylar-Scott[15], Christian Mosimann[16], Saulius Sumanas[10], Sarah De Val[8,17], Joseph B. Prescott[5,*], Kristy Red-Horse[1,4,*], Kyle M. Loh[1,3,*]

[1]Institute for Stem Cell Biology & Regenerative Medicine, Stanford University, Stanford, CA 94305, USA
[2]Department of Urology, Stanford University, Stanford, CA 94305, USA
[3]Department of Developmental Biology, Stanford University, Stanford, CA 94305, USA
[4]Department of Biology, Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305, USA
[5]Center for Biological Threats and Special Pathogens, Robert Koch Institute, Berlin 13353, Germany
[6]School of Biological Sciences, Nanyang Technological University, Singapore, 637551, Singapore
[7]Whitehead Institute for Biomedical Research, Department of Biology, Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
[8]Institute of Developmental and Regenerative Medicine, Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford OX3 7TY, United Kingdom
[9]Department of Pediatrics, Stanford University, Stanford, CA 94305, USA
[10]Department of Pathology and Cell Biology, University of South Florida, Tampa, Florida, USA
[11]Division of Endocrinology, Department of Medicine, Stanford University, Stanford, CA 94305, USA
[12]Institute of Diabetes and Regeneration Research, Helmholtz Center Munich, Neuherberg, Germany
[13]German Center for Diabetes Research, Neuherberg, Germany
[14]School of Medicine, Technical University of Munich, Munich, Germany
[15]Department of Bioengineering, Basic Science and Engineering Initiative, Children's Heart Center, Chan Zuckerberg Biohub, Stanford University, Stanford, CA 94305, USA
[16]Department of Pediatrics, Section of Developmental Biology, University of Colorado School of Medicine, Anschutz Medical Campus, Aurora, CO, USA
[17]Ludwig Institute for Cancer Research Ltd, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7DQ, United Kingdom
[18]Current address: Stately Bio, Palo Alto, 94306 (S.L.Z.), Department of Biology, Stanford University, Stanford, CA 94305, USA (X.X.), Replay Bio, San Diego, CA 92121 (R.M.M.), Department of Surgery, UCSF, San Francisco, CA 94143 (B.J.L.)
[19]Equal contribution
[20]Equal contribution
[21]Lead contact
*Correspondence: layteng@stanford.edu (L.T.A.), PrescottJ@rki.de (J.B.P.), kredhors@stanford.edu (K.R.H.)., kyleloh@stanford.edu (K.M.L.)

SUMMARY

We map human artery and vein endothelial cell (EC) differentiation from pluripotent stem cells, and employ this roadmap to discover new mechanisms of vascular development (vein differentiation) and disease (viral infection). We discovered vein development unfolds in two steps driven by opposing signals: VEGF differentiates mesoderm into "pre-vein" ECs, but surprisingly, VEGF/ERK inhibition subsequently specifies vein ECs. Pre-vein ECs co-expressed certain arterial (*SOX17*) and venous (*APLNR*) markers, harbored poised chromatin at future venous genes, but completed venous differentiation only upon VEGF inhibition. Intersectional lineage tracing revealed that early *Sox17*+ *Aplnr*+ ECs also formed veins *in vivo*. Next, we compared how Ebola, Andes, and Nipah viruses infect artery and vein ECs under biosafety-level-4 containment. Each virus distinctly affected ECs. Interestingly, artery and vein ECs also responded divergently to the same virus, thus revealing that developmentally-specified cell identity impacts viral infection. Collectively, this arteriovenous differentiation roadmap illuminates vascular development and disease.

INTRODUCTION

Blood vessels provide essential nourishment to all tissues[1-6] and are also afflicted by deadly diseases, including viral diseases of the vasculature[7-9]. Endothelial cells (ECs) comprise the inner lining of blood vessels, and there are multiple types of ECs, including artery, vein, capillary, and lymphatic ECs[1-6]. Thus, considerable attention has been devoted to discovering how different EC types develop *in vivo* and to generating them *in vitro*. ECs are also infected by viruses that cause some of the deadliest vascular diseases known to humanity, including Ebola virus (~44% fatality rate), Andes virus (~24% fatality rate), and Nipah virus (~57% fatality rate)[7-9]. These viruses disproportionately affect underprivileged populations in African countries (Ebola virus), Argentina and Chile (Andes virus), and Bangladesh and India (Nipah virus)[7-9]. Here we chart a roadmap of human arteriovenous differentiation and demonstrate its applications to explore the development and viral disease of the vasculature.

Artery and vein ECs execute specialized functions and express different genes[1-6]. For instance, artery ECs express NOTCH ligands that specify the smooth muscle cells that ensconce arteries[10]. By contrast, vein ECs can construct valves that prevent backward blood flow[11]. The prevailing model suggests that artery ECs express SOXF family transcription factors (SOX7, SOX17, and SOX18) that impart arterial identity[12-16]. Conversely, vein ECs express the transcription factor NR2F2 and transmembrane proteins APLNR/APJ and FLRT2[17-23].

Extensive work across multiple model organisms has illuminated key precepts of vascular development, but mysteries continue to surround the precise developmental origins of artery and vein ECs[3-6,24-27]. One might assume that arteries and veins should arise contemporaneously, as both are required to form a functional circulatory system. Nevertheless, during the development of the first intraembryonic blood vessels, arteries emerge earlier than veins, as shown by anatomy and molecular markers in mouse and zebrafish embryos[21,28-33]. Why veins should emerge after arteries remains a longstanding curiosity. Additionally, the earliest veins transiently "mis-express" certain arterial markers[28]. This piques questions about the earliest steps of venous development and the exact identity of the progenitor cells that build veins.

The extracellular signals that specify venous identity also warrant further investigation[5,24]. VEGF is a master regulator of EC identity[4,34,35], but has paradoxical roles in venous development. On one hand, VEGF is *required* for vein EC formation, as ECs are absent in *Vegfr2*$^{-/-}$ mouse embryos[36]. On the other hand, VEGF also *represses* vein formation, as it instructs arterial identity in zebrafish embryos[37-39]. How can VEGF both promote and inhibit vein EC specification during vascular development?

Many questions likewise surround the viral infection of human vasculature. Ebola virus (family *Filoviridae*), Andes virus (family *Hantaviridae*), and Nipah virus (family *Paramyxoviridae*) share certain common properties: they cause deadly vascular diseases, and can only be studied in high-containment laboratories[40]. However, a fundamental question in comparative virology is whether different viruses exert similar effects on the same cell-type, or whether each virus elicits unique effects. Extensive studies have thus far explored the *individual* effects of either Ebola virus[41,42], Andes virus[43-46], or Nipah virus[47-53] on cultured ECs. To provide a foundation for comparative virology, it is important to systematically compare the effects of different viruses on the same human cell-type, in the same experimental system. The ability to mass-produce human artery and vein ECs would enable this, and many other, applications. To this end, we previously developed methods to differentiate human pluripotent stem cells (hPSCs) into artery and vein ECs, and studied how Nipah virus infects ECs[49]. However, do viruses from different families have common or disparate effects on the human vasculature?

Here we chart a roadmap of stepwise gene expression, chromatin, and cell-surface marker changes during human artery and vein EC differentiation from hPSCs, and we demonstrate the utility of this resource to explore vascular development and disease (**Fig. 1A**). First, this roadmap unveiled two separable steps of vein development. VEGF *activation* differentiates lateral mesoderm into a newly-discovered "pre-vein" EC intermediate, which co-expresses "arterial" marker SOX17 and "venous" marker APLNR. Future venous genes are poised at the chromatin level within pre-vein ECs. Subsequently, VEGF *inhibition* installs fully-fledged venous identity. Intersectional genetic tracing confirmed that *Sox17*+ *Aplnr*+ ECs form vein ECs *in vivo*. The discovery of pre-vein ECs reconciles several longstanding mysteries in vascular development, including why early veins transiently mis-express certain "arterial" markers[28], why veins tend to emerge after arteries *in vivo*[21,28-32], and the paradoxical roles of VEGF in vein development[36-39]. Second, to track arteriovenous identity *in vitro*, we furnish new fluorescent reporter hPSC lines and combinations of cell-surface markers. Third, utilizing hPSC-derived artery and vein ECs, we systematically compared the effects of Ebola, Andes, and Nipah viruses under biosafety-level-4 (BSL4) containment to investigate whether developmentally-specified cell identity (e.g., arteriovenous identity) impacts viral infection. This revealed striking differences in the cellular effects of these viruses (inter-virus differences), and in how artery and vein ECs responded to the same virus (intra-virus differences), thus enriching comparative virology.

RESULTS

A roadmap of human arteriovenous differentiation reveals unexpected "pre-vein" ECs

We charted a roadmap spanning hPSC differentiation into primitive streak (day 1) and lateral mesoderm (day 2), and subsequent bifurcation into artery ECs (day 3) and vein ECs (day 4)[49]. Every 24 hours, we mapped changes in gene expression (through single-cell [scRNAseq] and

bulk-population RNA-sequencing), chromatin accessibility (using OmniATACseq[54]), and cell-surface markers (through high-throughput screening of 332 cell-surface markers) (**Fig. 1A-B**, **Fig. S1A-B**, **Table S1-S2**). We defined marker genes for each of these cell-types (**Fig. S1C**) and created an interactive web browser to explore scRNAseq data of each differentiation stage (https://anglab.shinyapps.io/artery-vein-scrna-seq/), thereby furnishing resources for future studies. Alongside ECs, a small number of mesenchymal cells also emerged (**Fig. S1D-E**), which we did not pursue further. Crucially, our roadmap revealed a hitherto-unknown intermediate progenitor, which we term "pre-vein" ECs and describe further below.

During lateral mesoderm bifurcation into artery and vein ECs, scRNAseq analysis uncovered "pre-vein" EC intermediates that preceded vein ECs *in vitro*. Day-2 lateral mesoderm expressed the angioblast transcription factor *ETV2*[55] (**Fig. 1C-D**, **Fig. S2A-B**). Upon 24 hours of further venous differentiation, *ETV2* was sharply silenced, whereupon EC surface markers (e.g., *CD144/VE-CADHERIN*) became expressed (**Fig. 1C-D**, **Fig. S2B-C**). Intriguingly, these incipient day-3 "pre-vein" CD144+ ECs co-expressed the "arterial" marker SOX17[12-14,16] and "venous" marker APLNR[17,18] (**Fig. 1C-D**). By contrast, cells subject to arterial differentiation solely expressed SOX17, but not APLNR, as expected[12-23] (**Fig. 1D-E**). We confirmed that single pre-vein ECs co-expressed SOX17 and APLNR, as shown by combined *in situ* hybridization and immunostaining (**Fig. 1E**) and scRNAseq (**Fig. S2D**). However, pre-vein ECs did not express additional markers of arterial (*CXCR4* or *DLL4*), venous (*NR2F2* or *FLRT2*), or angioblast (*ETV2*) identity (**Fig. 1D**). Having discovered *SOX17*+ *APLNR*+ "pre-vein" EC intermediates, we subsequently explored the extracellular signals that drive entry into, and exit from, this intermediate state.

Two steps of vein EC differentiation *in vitro*: VEGF activation, followed by inhibition

First, we discovered that day-2 lateral mesoderm differentiation into day-3 pre-vein ECs is driven by VEGF (which specifies EC identity), alongside TGFβ and NOTCH inhibitors (to repress arterial specification)[49] (**Fig. 2A**). During this first 24-hour interval, VEGF/ERK activation was crucial to generate ECs; in its absence, EC formation *in vitro* was largely abrogated (**Fig. 2Bi**). This is consistent with how VEGF is crucial for EC formation *in vivo*[36-38].

While VEGF activation specified day-3 pre-vein ECs, VEGF/ERK *inhibition* was necessary for further progression into day-4 vein ECs (**Fig. 2Bii**). The requirement for VEGF/ERK inhibition to confer venous identity was surprising, as VEGF is typically construed as the master regulator of EC biology[4,34,35]. Nevertheless, we found that continued VEGF activation at this stage was arterializing, and repressed venous markers (**Fig. 2Bii**), as observed in zebrafish embryos[37,38]. VEGF/ERK inhibition, together with other vein-specifying signals (WNT activation, TGFβ inhibition, and NOTCH inhibition)[49], was crucial to activate venous markers *NR2F2* and *FLRT2*, and downregulate pre-vein marker *SOX17* (**Fig. 2C**, **Fig. S2C**).

We thus propose two sequential steps of vein differentiation driven by temporally-dynamic VEGF *activation* to specify pre-vein ECs, followed by VEGF/ERK *inhibition* to induce vein ECs (**Fig. 2A**). This temporally-dynamic signaling switch reconciles why VEGF/ERK signaling both promotes[36] and inhibits[37-39] vein formation *in vivo*: the same signal can do both, at different times. Extant methods to differentiate hPSCs into vein ECs typically entail prolonged VEGF activation[56-61], which may inadvertently deter venous differentiation. In sum, (1) discovery of

pre-vein EC intermediates and (2) temporally-dynamic VEGF/ERK activation followed by inhibition was critical to efficiently generate vein ECs in our *in vitro* system.

Pre-vein ECs harbor poised chromatin at future venous genes

How do pre-vein ECs acquire the developmental competence to differentiate into vein ECs? In progenitor cells, promoter elements of future differentiation genes can become bivalently decorated with H3K4me3 (activation-associated) and H3K27me3 (repression-associated) histone marks in anticipation of future developmental decisions[62]. In pre-vein ECs, the *NR2F2* promoter was bivalently marked by both H3K4me3 and H3K27me3, but was not transcribed, consistent with a state of chromatin poising (**Fig. 2D**). Upon 24 hours of venous differentiation and VEGF/ERK inhibition, the *NR2F2* promoter exclusively acquired active histone mark H3K4me3 and became transcribed in vein ECs (**Fig. 2D**). By contrast, in artery ECs, the *NR2F2* promoter was repressed by H3K27me3, and lacked H3K4me3 (**Fig. 2D**). The promoter elements of other venous genes were similarly bivalent (H3K4me3+ H3K27me3+) in pre-vein ECs, and resolved to acquire a fully active state (H3K4me3+ H3K27me3-) in vein ECs (**Fig. S2E**). In conclusion, pre-vein ECs represent an early intermediate in vein development wherein venous genes are poised for future activation at the chromatin level.

Sox17+ Aplnr+ pre-vein ECs exist in early embryos and contribute to veins

Having discovered *SOX17*+ *APLNR*+ ECs *in vitro*, we investigated their existence and fate *in vivo*. Sox17 marks arteries *in vivo*[12-16], but others suggest that it is more broadly expressed in the vasculature[63,64]. We found that *Sox17+ Aplnr+* ECs exist in E9.5 mouse embryos (**Fig. 3A**, **Fig. S3A**, **Table S3**) and Carnegie Stage 12 (CS12) human embryos (**Fig. 3B**, **Fig. S3B**), through analysis of published scRNAseq data[65,66]. Whole mount immunostaining revealed that Sox17 marked most, if not all, Erg+ ECs in the E8.5 mouse embryo, including those in the dorsal aorta and the vitelline vein (**Fig. 3C,D**, **Fig. S3C**). In E9.5 mouse embryos, Sox17 was expressed by both the dorsal aorta and cardinal vein (**Fig. 3E**). Indeed, permanent labeling of *Sox17*+ cells and all of their progeny with *Sox17-Cre* resulted in both artery and vein ECs being lineage traced (**Fig. S3D,E**), consistent with past observations[67,68]. These lineage tracing and immunostaining results demonstrate that vein ECs transiently express Sox17, although it is turned off later in vein development by E10.5[15,16]. Despite the prominent role of Sox17 in artery development[12-16], we conclude that it is expressed by most ECs in early vascular development, and these cells construct most of the vasculature, including arteries and veins.

Analogous to pre-vein ECs *in vitro*, we found that the common cardinal vein—the earliest intraembryonic vein[28]—co-expressed Sox17 and *Aplnr* in the E9.5-E10 mouse embryo. This was independently confirmed by two different means to detect *Aplnr* expression: *in situ* hybridization (**Fig. 3F**) and Aplnr-CreER activity (**Fig. S3F,G**). While the early cardinal vein anatomically appears as a vein, from a molecular perspective it transiently expresses Sox17 and *Aplnr*, mirroring pre-vein ECs *in vitro*, prior to downregulating Sox17[15,16]. Suggestive of evolutionary conservation, we found that *sox7* (a zebrafish *sox17* homolog[69]) was co-expressed alongside *aplnrb* in the zebrafish secondary vascular field (**Fig. 3G**), which represents future venous progenitors[70].

Finally, we employed intersectional lineage tracing to test whether *Sox17*+ *Aplnr*+ ECs contribute to vein ECs. *Sox17-Cre*[67] and *Aplnr-DreER*[71] driver mice were crossed with an

intersectional reporter mouse[72]. In this approach, both Cre- and Dre-mediated recombination in a single cell—reflecting expression of both Sox17 and Aplnr—ignites *GFP* expression[72]. *Sox17-Cre*; *Aplnr-DreER* labeled vein ECs in the common cardinal vein, supporting the hypothesis that pre-vein ECs form vein ECs *in vivo* (**Fig. 3H**). However, not every vein ECs was labeled, likely reflecting incomplete recombination efficiency inherent to intersectional lineage tracing, or alternative sources of vein ECs. Taken together, *Sox17*+ *Aplnr*+ ECs contribute to veins in mouse embryos, and cells with molecular features of pre-vein ECs may be evolutionarily conserved in human and zebrafish embryos.

SOXF transcription factors are required for both artery and vein differentiation

SOX17 is transiently expressed during vein differentiation *in vivo* and *in vitro*, but does it have any function at this early step of vein differentiation? The SOXF transcription family comprises SOX7, SOX17, and SOX18, which overlap in expression and function[5,69,73] (**Fig. S4A**). To overcome this genetic redundancy, we employed an enhanced CRISPR interference (CRISPRi) system[74] to simultaneously knockdown *SOX7*, *SOX17,* and *SOX18* (**Fig. 4A-C**, **Fig. S4B**). As expected[12-16], triple *SOXF* knockdown abrogated artery differentiation, with near-complete loss of arterial markers *CXCR4* and *DLL4* (**Fig. 4D-E**, **Fig. S4C**, **Table S4**).

Triple *SOXF* knockdown also strongly impaired vein differentiation: pan-endothelial marker *VE-CADHERIN/CDH5* and venous markers *APLNR*[17,18] and *NT5E/CD73*[56-58] were reduced (**Fig. 4D-E**, **Fig. S4C**, **Table S5**). Consistent with how SOXF is necessary for *VE-CADHERIN* and *APLNR* expression during venous differentiation, CUT&RUN showed that SOX17 directly bound the *VE-CADHERIN* and *APLNR* loci in pre-vein ECs, prior to venous differentiation (**Fig. 4F**, **Fig. S4D**).

Beyond individual markers, triple *SOXF* knockdown significantly impaired the venous transcriptional signature more broadly, as computed from known venous marker panels[21] (**Fig. 4G**). Gene sets associated with blood vessel development, migration, and junction formation were downregulated upon triple *SOXF* knockdown (**Fig. S4E**). Individually knocking down *SOX7*, *SOX17*, and *SOX18* led to modest effects (**Fig. S4F-G**), emphasizing the importance of combinatorial knockdowns.

Altogether, we confirm the conspicuous requirement for SOXF transcription factors in artery EC differentiation[12-16], but we additionally discover that SOXF factors are crucial for human vein EC differentiation, building on their roles in mouse and zebrafish vein development[75,76]. While SOXF is required for both artery and vein ECs, it turns on different genes in each. hPSC differentiation thus enables mechanistic studies of human arteriovenous differentiation and combinatorial genetic manipulations that would otherwise prove challenging *in vivo*.

Cell-surface marker combinations that define human arteriovenous identity

High-throughput antibody screening of 332 cell-surface markers revealed expanded cell-surface marker combinations that defined artery ECs (CD144+ CXCR4+ DLL4+) and vein ECs (CD144+ CD73+ CD317+), building on past work[56-58] (**Fig. 5A,B**, **Fig. S5A,B**, **Table S6**). With respect to these cell-surface marker combinations, pre-vein ECs displayed ambiguous arteriovenous identity, and were classified as neither arterial nor venous (**Fig. 5B**). We rigorously tested the ability of these cell-surface marker combinations to purify hPSC-derived

artery and vein ECs: scRNAseq of fluorescence activated cell sorting (FACS)-purified artery and vein ECs revealed that they were largely homogeneous, with the exception of cell-cycle status (**Fig. 5C**, **Fig. S5C-E**). As an independent means to visualize arteriovenous identity, we employed Cas9/AAV6[77] to engineer *SOX17-2A-mPlum*; *NR2F2-2A-GFP* double knock-in reporter hPSCs, which revealed mutually-exclusive generation of SOX17+ artery ECs and NR2F2+ vein ECs from the respective differentiation protocols (**Fig. 5D**).

Transcriptional and chromatin hallmarks of arteriovenous identity, including axon pathfinding molecules and enhancer elements

Leveraging these cell-surface markers to stringently purify hPSC-derived artery and vein ECs by FACS, we discovered striking differences in the gene expression and chromatin landscapes of these two EC types. Known arterial markers (*GJA4/CX37*, *UNC5B*, *DLL4, MECOM, HEY1, EFNB2, CXCR4*, and *IGFBP3*) and venous markers (*NR2F2*, *NRP2*, *APLNR*, and *FLRT2*) *in vivo*[20,21] were disparately expressed by hPSC-derived artery and vein ECs *in vitro*, attesting to their distinct arteriovenous identities (**Fig. S5F**).

We discovered engrained chromatin differences between hPSC-derived artery and vein ECs by using OmniATACseq[54] to map chromatin accessibility (**Fig. 5E**, **Table S7**) and CUT&RUN[78] to map H3K4me1, H3K4me3, and H3K27me3. At the genome-wide level, motifs for distinct transcription factor families were enriched in the accessible chromatin of hPSC-derived artery ECs (ETV and SOX) vs. vein ECs (NR2F and MAF), congruent with differential expression of these transcription factors in artery vs. vein ECs (**Fig. 5E**, **Fig. S5G**). The *DLL4* 12 kB upstream enhancer—an evolutionarily-conserved hallmark of arterial identity[79]—was accessible and bound by SOX17 in hPSC-derived artery ECs, but was not accessible in vein ECs (**Fig. 5F,G**). Functionally, this *Dll4* upstream enhancer sufficed to direct arterial expression of a *LacZ* reporter in mouse embryos[79] (**Fig. 5H**, **Fig. S5H**).

Our chromatin data revealed a novel, arterial-specific enhancer 13 kB upstream of *SEMA3G* gene, which encodes an arterially-expressed guidance cue required for vascular pathfinding[80,81] (**Fig. 5I**). Transgenic reporter assays delineated that this novel *Sema3g* enhancer was active in arteries, but not veins, within mouse[82,83] and zebrafish embryos, attesting to its arterial specificity (**Fig. 5J**, **Fig. S5I**). Regulatory elements flanking other arterial genes (*HEY1*, *CXCR4*)[84] and venous genes (*APLNR*) were respectively accessible in hPSC-derived artery vs. vein ECs (**Fig. 5K**, **Fig. S5J-K**). Taken together, hPSC-derived artery and vein ECs are distinguished by a wealth of arteriovenous regulatory elements at the chromatin level. Our genomic resource provides a means to identify additional such elements, as exemplified by the arterial *SEMA3G* enhancer.

Benchmarking EC differentiation methods and arteriovenous identity acquisition

To aid the community in selecting EC differentiation methods for each application, we systematically compared 8 differentiation methods by collating published scRNAseq datasets and generating new ones[49,60,61,85-87] (**Fig. 6A,B**, **Fig. S6A,B**). What is the diversity and arteriovenous character of cells generated by each hPSC differentiation protocol?

All 8 differentiation methods generated ECs (*VE-CADHERIN*+) and mesenchymal cells (*MEST*+) in varying proportions. Two differentiation protocols[49] yielded the highest EC

percentages (96.5% pure artery ECs and 84.0% pure vein ECs, respectively), consistent with the inhibition of signals that induce alternative cell-types at each differentiation step. In these two protocols, the remaining non-ECs comprised contaminating mesenchymal cells (**Fig. 6A,B**, **Fig. S6C**). Certain differentiation protocols[85,86] yielded a wider spectrum of non-endothelial cell-types, including blood (*PU.1/SPI1*+), kidney (*LHX1*+), heart (*NKX2.5*+), and liver (*APOA*+) (**Fig. 6A,B**, **Fig. S6C**). Emergence of mesenchymal cells alongside ECs in all differentiation protocols suggests that the bifurcation between EC vs. mesenchymal identities from lateral mesoderm remains incompletely understood[24]. Suppressing mesenchymal formation thus remains a future goal.

Next, we quantified the arteriovenous character of ECs emerging from each differentiation protocol *in vitro*, using *in vivo*-defined arteriovenous marker signatures[21]. Do all differentiation protocols culminate in a common EC "end state", or do different EC types emerge arise from each protocol? Multiple protocols generated *CXCR4*+ *VE-CADHERIN*+ artery ECs—perhaps reflecting the pervasive *in vitro* use of VEGF (**Fig. S6B**), which is arterializing *in vivo*[37,38]—although arterial marker intensity differed between protocols (**Fig. 6B,C**). One protocol[49] imparted the strongest arterial identity (**Fig. 6C**; $P<0.0001$), potentially attributable to the explicit inhibition of vein-inducing signal PI3K[39,88] to sharpen arterial identity. Only two differentiation protocols predominantly generated *NR2F2*+ *VE-CADHERIN*+ vein ECs[49,60], consistent with how it has proven more challenging to instate venous identity *in vitro*[24]. The protocol that conferred the strongest venous identity (**Fig. 6D**; $P<0.0001$) entailed VEGF-induced generation of pre-vein ECs, followed by VEGF/ERK inhibition to specify vein ECs[49]. Other methods relying on sustained VEGF activation less effectively induced venous character (**Fig. 6D**). This comparison of multiple differentiation protocols suggests that VEGF acts as a double-edged sword in venous differentiation, by inducing ECs but subsequently repressing venous identity[24,36-38,49]. Temporally-dynamic VEGF activation, followed by inhibition, may thus prove decisive in imparting venous identity *in vitro*.

Collectively, this provides a framework to select EC differentiation protocols based on the desired degree of cellular diversity (including mesenchymal cells that avail vascular network formation[89,90]) and arteriovenous character.

Finally, we developed scalable methods to differentiate hPSCs into 95-240 million artery or vein ECs in a single batch, while retaining >90% differentiation purity (**Fig. S7A-F**). We also developed media to maintain hPSC-derived vein ECs, entailing PKA activation and inhibition of artery-specifying signals TGFβ and NOTCH (**Fig. S7G-L**). The ability to generate hPSC-derived artery and vein ECs *en masse* and to maintain them for at least 6 days, while preserving arteriovenous marker expression, enables a range of practical applications, as detailed below.

Comparing the effects of Ebola, Nipah, and Andes viruses on hPSC-derived artery and vein ECs

Outstanding questions in comparative virology include: how do (1) different viruses affect the same cell-type, and (2) different cell-types respond to the same virus? To address both questions, we compared the effects of Ebola, Nipah, and Andes viruses on hPSC-derived artery and vein ECs. While Ebola[42,91-93], Nipah[94-96], and Andes[97-99] viruses all infect ECs *in vivo*

and cause severe vascular diseases, they belong to distinct virus families (filoviruses, paramyxoviruses, and hantaviruses, respectively) and cause different disease manifestations.

Under BSL4 containment, hPSC-derived artery and vein ECs were infected with high titers of Ebola, Nipah or, Andes viruses, or as a control, the non-pathogenic Sendai virus[100], followed by RNA-seq at 6, 12, 24, and 48 hours post-infection (**Fig. 7A**). Ebola and Andes viruses each replicated with similar kinetics in both artery and vein ECs (**Fig. 7B**). Nipah virus replicated more rapidly in artery ECs as previously reported[49], prior to destroying artery ECs by 24-48 hours post-infection (**Fig. 7B**). What are the transcriptional effects of these viruses?

Classically, virus-infected cells secrete both interferons (which activates antiviral genes that restrain viral replication) and inflammatory cytokines (which induce fever and other effects), thereby eliciting a balanced immune response[101]. Indeed, within 6 hours of infection, Sendai virus expectedly induced the expression of interferon, antiviral interferon-stimulated genes (ISGs)[102], and inflammatory cytokines and chemokines in both artery and vein ECs (**Fig. 7C-G**, **Fig. S8A**). Despite substantial intracellular Andes and Nipah virus replication at 6-12 hours post-infection (**Fig. S8B**), ECs failed to activate interferon or inflammatory responses at these early timepoints[49], likely reflecting how certain Andes[103-105] and Nipah[106-109] proteins suppress innate immunity. However, both the interferon and inflammatory programs were eventually activated upon 24-48 hours post-infection with Andes or Nipah viruses (**Fig. 7C-G**, **Fig. S8A**). This mirrors other pathogenic viruses, including SARS-CoV-2, which induce interferon and inflammatory programs in delayed fashion[110]. This delay in inducing interferon expression may provide a foothold for initial viral amplification in ECs, prior to engagement of innate immunity.

Ebola induces imbalanced innate immune responses in hPSC-derived ECs, typified by inflammatory cytokine production in the absence of interferon production

Ebola virus mRNAs occupied >50% of the artery EC transcriptome within 24-48 hours post-infection, but remarkably failed to induce type I interferon (*IFNβ*), type III interferon (*IFNλ1*) or antiviral ISGs to an appreciable extent (**Fig. 7C,F**, **Fig. S8A**). The quantitative degree of interferon suppression was profound. Upon Ebola infection, there was no statistically-significant increase in IFNβ secretion (**Fig. 7F**) nor upregulation of antiviral ISG *MX1* (**Fig. S8A**) relative to uninfected ECs, whereas *MX1* was elevated 937-fold in response to Sendai infection. Correspondingly, antigen presentation ISGs (*HLA-B*, *PSMB8*, *PSMB9*, and *TAP1*[102]) were not markedly induced by Ebola, although they were upregulated in response to other viruses (**Fig. 7C**). By suppressing antigen presentation gene expression, Ebola may impair links between innate and adaptive immunity in ECs, consistent with defective adaptive immune responses in patients with fatal Ebola disease[111-113].

Despite the profound lack of interferon induction, Ebola significantly induced inflammatory cytokines, including *IL6* and *CXCL8/IL8*, by 48 hours post-infection (**Fig. 7F**, **Fig. S8A**). Indeed, IL6 and CXCL8 are elevated 100- to 1000-fold in Ebola patients, and can reach extremely high bloodborne levels (>1 ng/mL) *in vivo*[111,114]. Taken together, Ebola seemingly triggers "immunological misfiring"[115] in human ECs by elevating inflammatory cytokines (e.g., IL6 and CXCL8), in the virtual absence of interferon and antiviral ISGs. This imbalanced innate immune response may underlie two clinical features of fatal Ebola disease: unremitting "cytokine

storm"[111,114], paired with extraordinary bloodborne viral loads[116], the latter perhaps attributable to failed antiviral gene induction.

hPSC-derived artery and vein ECs respond differently to the same virus

Finally, the same virus elicited different transcriptional responses in hPSC-derived artery vs. vein ECs (**Fig. S8C**). *TNFAIP6* was consistently elevated by Ebola, Andes, and Sendai viruses in artery ECs, relative to vein ECs (**Fig. 7H**). *TNFAIP6* encodes a stress-induced secreted protein that generally suppresses proteolysis of the extracellular matrix[117]. Conversely, *TRIM63* was upregulated by Ebola, Andes, and Sendai viruses in vein ECs, by comparison to artery ECs (**Fig. 7H**). *TRIM63* encodes an E3 ubiquitin ligase that regulates cell death[118]. Artery and vein ECs may thus have distinct and hardwired responses to viral infection, illustrating that developmental cell state can impact viral infection. This exemplifies how hPSC-derived artery and vein ECs enable comparative virology studies under BSL4 containment.

DISCUSSION

Mapping a developmental hierarchy—for instance, for ECs—relies on thorough description of intermediate states arising at each differentiation step. Here we chart a roadmap for human artery and vein EC differentiation *in vitro*, encompassing stepwise changes in gene expression, chromatin state, and cell-surface markers. We demonstrate the applications of this roadmap to illuminate vascular development and disease.

Pre-vein ECs: a hitherto-missing stage in vein development?

Our *in vitro* differentiation roadmap builds on influential scRNAseq surveys of embryonic ECs *in vivo*[20-22,65,66,119,120] and *in vitro*-generated ECs[49,60,61,85,87,90,121,122]. However, we specifically focus on how artery and vein ECs incipiently bifurcate from lateral mesoderm. This bifurcation is technically and ethically impossible to study in 3-week-old human embryos[123], but constitutes a pivotal time in life when the fundaments of the future circulatory system are constructed. Our *in vitro* modeling of this key developmental stage led to the discovery of pre-vein ECs, which constitute hitherto-missing stage in venous development.

Longstanding mysteries have surrounded vein development[24]. Why do veins emerge after arteries *in vivo*[21,28-32]? Why do early veins transiently "mis-express" certain artery markers[28]? *In vitro*, day-2 lateral mesoderm can rapidly acquire arterial identity (by day 3), but along the other developmental path, lateral mesoderm first transitions "pre-vein" ECs (day 3) before acquiring vein fate (day 4). Thus, artery ECs arise *prior* to vein ECs, both *in vitro* and *in vivo*[28]; these two EC types do not arise synchronously in early development. Pre-vein ECs co-express "arterial" marker *Sox17* and "venous" marker *Aplnr*, and can form vein ECs *in vitro* (during hPSC differentiation) and *in vivo* (in the mouse embryo, as shown by genetic lineage tracing). Cells with molecular features of pre-vein ECs also exist in human and zebrafish embryos. Collectively, the concept of pre-vein ECs unifies several observations concerning endothelial development, including why vein ECs emerge later and temporarily "mis-express" arterial markers[28].

Temporally dynamic VEGF activation, followed by inhibition, drives vein differentiation

We present evidence for a two-step model for vein development *in vitro*: VEGF/ERK *activation* is initially required to specify pre-vein ECs, but 24 hours later, VEGF/ERK *inhibition* is

subsequently required for vein EC specification. This reconciles the paradox of why VEGF both promotes[36] and represses[37-39] vein ECs *in vivo*: the same signal executes both roles, but at different times. The idea that VEGF/ERK inhibition is required for vein specification[37-39,49] may be surprising, as VEGF is often construed as the master regulator of EC identity[4,34,35]. However, timing is paramount: while VEGF induces ECs, we suggest that once a cell has acquired EC identity, it now re-interprets VEGF inhibition to signify that it should adopt vein fate. Indeed, in other developmental venues, the same signal can specify different identities at distinct times[124]. In any case, VEGF is ubiquitous in EC differentiation and culture media, and our findings suggest that it may inadvertently suppress venous identity in certain contexts.

Why do major intraembryonic arteries precede veins *in vivo*? The answer may lie in how VEGF induces both endothelial and arterial identity[36-39]. Perhaps cells can contemporaneously acquire endothelial and arterial identity[28] under the command of VEGF, but vein development must be parsed into two distinct steps wherein endothelial identity is first instated, followed by VEGF/ERK inhibition to specify veins.

In other developmental contexts, the rich detail afforded by scRNAseq suggests that differentiation paths are continua[125,126], piquing the philosophical question of where one cell state (e.g., pre-vein) ends and the other (e.g., vein) begins. Nevertheless, pre-vein ECs represent a clear intermediate step of venous development, as they differ from vein ECs in multiple ways. First, they are functionally induced by opposing signals: VEGF/ERK activation vs. inhibition, respectively. Second, pre-vein ECs do not express later-stage venous markers such as *NR2F2*, but harbor bivalent chromatin at future venous genes, implying developmental poising. Third, pre-vein ECs lack venous cell-surface markers CD73 and CD317.

SOXF transcription factors are required for both artery and vein EC specification

While the early cardinal vein anatomically appears as a vein, we find that it molecularly co-expresses Sox17 and *Aplnr*. The early cardinal vein may thus transiently resemble pre-vein ECs, prior to downregulating Sox17 at later stages[15,16]. Indeed, most if not all ECs temporarily express Sox17 in the mouse embryo, as shown by immunostaining and *Sox17-Cre* lineage tracing. Consistent with how the early vasculature broadly expresses Sox17 *in vivo*, simultaneous CRISPRi knockdown of all *SOXF* family members (*SOX7*, *SOX17*, and *SOX18*) compromised both human artery and vein EC differentiation *in vitro*.

Our findings support longstanding observations that SOXF transcription factors are required for arterial specification[12-16], but additionally suggest important roles in venous identity. Further evidence supports our conclusions. First, both arteries and veins are morphologically lost in the *Sox7*-/- *Sox17*-/- *Sox18*-/- mouse retinal vasculature[76]. Second, double *sox7* and *sox18* knockdown abrogates vein development in zebrafish[75]. Third, SOXF motifs reside in not just arterial enhancers, but also in pan-endothelial enhancers, implying that SOXF has broader roles in endothelial identity beyond arterialization[84]. Taken together, SOXF transcription factors are expressed in pre-vein ECs, and are crucial for human vein EC differentiation.

While SOXF is necessary for the specification of artery and vein ECs, it activates different genes in each. Our CRISPRi studies showed that SOXF upregulates arterial genes in artery ECs, whereas it drives expression of certain venous (e.g., *APLNR*) and pan-endothelial genes

(e.g., *VE-CADHERIN/CDH5*) in vein ECs. Indeed, transcription factors execute cell-type-specific roles[127]. Could differing SOXF levels contribute to these outcomes? Artery ECs express higher levels of SOX17 than pre-vein ECs (**Fig. 1D**). Perhaps higher vs. lower SOXF levels respectively bias ECs to arterial vs. venous identities, but wholesale SOXF depletion—as we achieved by CRISPRi—compromises both. SOXF transcription factors may also interact with differing cofactors in either artery vs. vein ECs to direct distinct transcriptional programs.

A foundation for comparative virology

We leveraged the ability to mass-produce human artery and vein ECs to address a central question in comparative virology: while multiple viruses (e.g., Ebola[42,91-93], Nipah[94-96], and Andes[97-99] viruses) infect the human vasculature, do they induce unique or common effects? Prior work examined each virus separately, in disparate EC model systems. These studies discovered that Andes induces interferon secretion[43,46], whereas conflicting conclusions were drawn regarding Ebola[41,42] and Nipah[48-50] viruses. We directly compared these viruses in the same experimental system and found they elicited starkly different effects on ECs. Nipah and Andes viruses induced both interferon and inflammatory cytokine expression in ECs, albeit in delayed fashion. Many pathogenic viruses repress innate immunity[49,101,110], but perhaps after extensive Nipah or Andes virus replication, this repression is outpaced by viral replication.

Ebola virus was an outlier: we discovered that it *upregulated inflammatory cytokines* (e.g., IL6), while simultaneously *repressing interferon production* by infected ECs. Virus infection typically triggers two arms of innate immunity—interferon and inflammatory cytokine production[101]—yet, Ebola apparently decouples these two responses in opposite directions. Remarkably, Ebola-infected ECs failed to upregulate interferon or antiviral genes despite extraordinary viral loads (~1 in 2 mRNAs encoding Ebola virus). Substantial inflammatory cytokines, juxtaposed with defective interferon production and unchecked viral replication, may fuel an imbalanced immune response to Ebola infection. This is consistent with clinical observations. First, IL6 is massively induced in fatally-infected Ebola patients[111,114], even exceeding levels found in CAR-T induced "cytokine storm"[128]. Second, Ebola prolifically replicates in patients, reaching 2 billion viral genomes per mL of blood[116]. This may reflect how Ebola powerfully suppresses interferon production and antiviral gene expression *in vitro*.

This study represents one of the first systematic comparisons of how multiple Risk Group 3 and 4 viruses affect the same human cell-type, to the best of our knowledge. Direct comparison of these viruses in a single experimental system enabled discovery of the common and unique effects of each virus. This comparative virology approach also revealed that artery and vein ECs respond differently to the same virus, highlighting both inter-virus and intra-virus differences. Taken together, this—and the other examples in this manuscript—exemplifies how a human arteriovenous differentiation roadmap can help illuminate vascular development and disease.

Limitations of the study

First, do *all* vein ECs emerge via pre-vein intermediates, or can alternative differentiation paths also culminate in venous identity? Our *Sox17-Cre; Aplnr-DreER* system to label pre-vein ECs cannot decisively answer this question owing to incomplete recombination efficiency, but unbiased CRISPR barcoding[129,130] might provide future insights. Second, why is VEGF/ERK

inhibition crucial for venous identity? ERK2 transcriptionally pauses RNA polymerase II on future developmental genes[131]; perhaps ERK inhibition unleashes the transcription of paused venous genes. Third, we studied how Ebola, Nipah, and Andes affect different developmentally-specified types of human ECs (artery vs. vein ECs), but how do these viruses affect additional cell-types, including immune cells? For instance, Ebola-infected dendritic cells seemingly produce neither inflammatory cytokines nor interferon[132], and thus differ from Ebola-infected ECs studied here. Fourth, it will be fascinating to discover the mechanisms through which Ebola decouples two innate immunity pathways in ECs, eliciting inflammatory cytokine production while powerfully suppressing interferon expression. NFκB and IRF3/7 pathways respectively induce inflammatory cytokines and interferon[101]. The potential thus exists for each of these to be independently regulated by Ebola virus. Ebola proteins VP35 and VP24 inhibit interferon production[133-135], but how does Ebola activate inflammatory cytokine production in ECs? The ability to produce human artery and vein ECs in large numbers and to infect them under BSL4 containment may open the door to answering these, and other, questions.

ACKNOWLEDGMENTS

We thank Valerie Park, Liying Ou, Aaron McCarty, Catherine Carswell-Crumpton, the Stanford Institute for Stem Cell Biology & Regenerative Medicine, and the Stanford Stem Cell FACS Core Facility for infrastructure support. We thank Sophie Astrof for sharing *Aplnr-DreER* mice, and Ryan Corces, Arjun Rajan, Zhainib Amir-Ugokwe, Siva Vijayakumar, Len Pennacchio, Axel Visel, Stacey Lee, Avishek Ganguly, Yan Yang, Yimiao Qu, Alicia Wong, and Julie Chen for assistance. Genomics analyses were supported by the Stanford Diabetes Genomics Analysis Core Facility (NIH P30DK116074). This work was supported by NIH Director's Early Independence Award DP5OD024558 (K.M.L.), NIH R01HL128503 (K.R.-H.), NIH R01HL153005 (S.S.), NIH R01HL168097 (C.M.), Breakthrough T1D Northern California Center of Excellence (K.M.L.), Additional Ventures Expansion Award (K.M.L., L.T.A.), Siebel Stem Cell Institute (L.T.A.), Ludwig Cancer Research (K.M.L., S.D.V.), Stanford Maternal and Child Health Research Institute Chambers Family Foundation Innovation Grant (K.M.L., L.T.A., K.R.-H.), Advanced Research Projects Agency for Health (ARPA-H) under award number AY1AX000002 (M.A.S.-S., K.M.L.), Helmholtz Society (H.L.), German Research Foundation (H.L.), European Research Council (H.L.), British Heart Foundation FS/SBSRF/22/31037 and RE/18/3/34214 (S.D.V.), Robert Koch Institute intramural funds (J.B.P.), Amaranth Foundation (K.M.L., J.B.P.), and Anonymous and Gilbert families (K.M.L.). S.L.Z. was supported by NSF and Stanford Graduate Fellowships. R.M.M. was supported by NIH T32GM119995. C.M. is the The Helen and Arthur E. Johnson Chair For The Cardiac Research Director. L.T.A. is an Additional Ventures Catalyst to Independence Fellow and a Bladder Cancer Advocacy Network Career Development Awardee. K.R.-H. is a Howard Hughes Medical Institute Investigator. K.M.L. is a Human Frontier Science Program Young Investigator (RGY0069/2019), Packard Foundation Fellow, Pew Scholar, Baxter Foundation Faculty Scholar, and The Anthony DiGenova Endowed Faculty Scholar. Content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH, ARPA-H, or other funding organizations.

AUTHOR CONTRIBUTIONS

L.T.A., S.L.Z., K.J.L., A.M., Q.Y., C.Q., X.X., and K.M.L. differentiated and characterized hPSCs. S.L.Z., J.W., S.K.J., J.A., D.W.S., and K.R.-H. analyzed mouse embryos. I.v.C. and J.B.P. performed BSL4 virology. L.T.A., S.L.Z., K.J.L., A.M., X.X., A.D., E.X., W.S., C.M.,

S.D.V., and K.M.L. performed genomics analyses. R.S., K.S., R.M.M., B.J.L., J.S.W., H.L., M.A.S.-S., and M.H.P. developed and provided reagents and advice. N.K.R. and S.S. stained zebrafish embryos. D.S., S.N., and S.D.V. analyzed enhancer transgenic embryos. L.T.A., J.B.P., K.R.-H., and K.M.L. supervised the study, with input from S.L.Z., K.J.L., and A.M.

DECLARATION OF INTERESTS

Stanford University has filed patent applications related to endothelial differentiation. S.L.Z. and R.M.M. are presently employed in the biotechnology industry, but contributed to this work while affiliated with Stanford University.

REFERENCES

1. Rafii, S., Butler, J.M., and Ding, B.-S. (2016). Angiocrine functions of organ-specific endothelial cells. Nature *529*, 316-325. 10.1038/nature17040.
2. Augustin, H.G., and Koh, G.Y. (2017). Organotypic vasculature: From descriptive heterogeneity to functional pathophysiology. Science *357*. 10.1126/science.aal2379.
3. Potente, M., and Mäkinen, T. (2017). Vascular heterogeneity and specialization in development and disease. Nature Reviews Molecular Cell Biology *18*, 477-494. 10.1038/nrm.2017.36.
4. Coultas, L., Chawengsaksophak, K., and Rossant, J. (2005). Endothelial cells and VEGF in vascular development. Nature *438*, 937-945. 10.1038/nature04479.
5. Fish, J.E., and Wythe, J.D. (2015). The molecular regulation of arteriovenous specification and maintenance. Developmental Dynamics *244*, 391-409. 10.1002/dvdy.24252.
6. De Val, S., and Black, B.L. (2009). Transcriptional control of endothelial cell development. Dev Cell *16*, 180-195. 10.1016/j.devcel.2009.01.014.
7. Alonso, D.O., Iglesias, A., Coelho, R., Periolo, N., Bruno, A., Córdoba, M.T., Filomarino, N., Quipildor, M., Biondo, E., Fortunato, E., et al. (2019). Epidemiological description, case-fatality rate, and trends of Hantavirus Pulmonary Syndrome: 9 years of surveillance in Argentina. Journal of medical virology *91*, 1173-1181. 10.1002/jmv.25446.
8. Jacob, S.T., Crozier, I., Fischer, W.A., Hewlett, A., Kraft, C.S., de La Vega, M.-A., Soka, M.J., Wahl, V., Griffiths, A., Bollinger, L., and Kuhn, J.H. (2020). Ebola virus disease. Nature Reviews Disease Primers, 1-31. 10.1038/s41572-020-0147-3.
9. Ang, B.S.P., Lim, T.C.C., and Wang, L. (2018). Nipah Virus Infection. Journal of Clinical Microbiology *56*, e01875-01817-01810. 10.1128/JCM.01875-17.
10. High, F.A., Lu, M.M., Pear, W.S., Loomes, K.M., Kaestner, K.H., and Epstein, J.A. (2008). Endothelial expression of the Notch ligand Jagged1 is required for vascular smooth muscle development. Proc Natl Acad Sci U S A *105*, 1955-1959. 10.1073/pnas.0709663105.
11. Bazigou, E., and Makinen, T. (2013). Flow control in our vessels: vascular valves make sure there is no way back. Cell Mol Life Sci *70*, 1055-1066. 10.1007/s00018-012-1110-6.
12. Corada, M., Orsenigo, F., Morini, M.F., Pitulescu, M.E., Bhat, G., Nyqvist, D., Breviario, F., Conti, V., Briot, A., Iruela-Arispe, M.L., et al. (2013). Sox17 is indispensable for acquisition and maintenance of arterial identity. Nature Communications *4*, 2609. 10.1038/ncomms3609.

13. Sakamoto, Y., Hara, K., Kanai-Azuma, M., Matsui, T., Miura, Y., Tsunekawa, N., Kurohmaru, M., Saijoh, Y., Koopman, P., and Kanai, Y. (2007). Redundant roles of Sox17 and Sox18 in early cardiovascular development of mouse embryos. Biochem Biophys Res Commun *360*, 539-544.
14. Matsui, T., Kanai-Azuma, M., Hara, K., Matoba, S., Hiramatsu, R., Kawakami, H., Kurohmaru, M., Koopman, P., and Kanai, Y. (2006). Redundant roles of Sox17 and Sox18 in postnatal angiogenesis in mice. J Cell Sci *119*, 3513-3526.
15. Hosking, B., Francois, M., Wilhelm, D., Orsenigo, F., Caprini, A., Svingen, T., Tutt, D., Davidson, T., Browne, C., Dejana, E., and Koopman, P. (2009). Sox7 and Sox17 are strain-specific modifiers of the lymphangiogenic defects caused by Sox18 dysfunction in mice. Development *136*, 2385-2391. 10.1242/dev.034827.
16. Clarke, R.L., Yzaguirre, A.D., Yashiro-Ohtani, Y., Bondue, A., Blanpain, C., Pear, W.S., Speck, N.A., and Keller, G. (2013). The expression of Sox17 identifies and regulates haemogenic endothelium. Nature Cell Biology *15*, 502-510.
17. Chen, H.I., Sharma, B., Akerberg, B.N., Numi, H.J., Kivela, R., Saharinen, P., Aghajanian, H., McKay, A.S., Bogard, P.E., Chang, A.H., et al. (2014). The sinus venosus contributes to coronary vasculature through VEGFC-stimulated angiogenesis. Development *141*, 4500-4512. 10.1242/dev.113639.
18. Saint-Geniez, M., Argence, C.B., Knibiehler, B., and Audigier, Y. (2003). The msr/apj gene encoding the apelin receptor is an early and specific marker of the venous phenotype in the retinal vasculature. Gene Expr Patterns *3*, 467-472. 10.1016/s1567-133x(03)00062-0.
19. You, L.-R., Lin, F.-J., Lee, C.T., DeMayo, F.J., Tsai, M.-J., and Tsai, S.Y. (2005). Suppression of Notch signalling by the COUP-TFII transcription factor regulates vein identity. Nature *435*, 98-104.
20. Su, T., Stanley, G., Sinha, R., Amato, G.D.x., Das, S., Rhee, S., Chang, A.H., Poduri, A., Raftrey, B., Dinh, T.T., et al. (2018). Single-cell analysis of early progenitor cells that build coronary arteries. Nature, 1-29. 10.1038/s41586-018-0288-7.
21. Hou, S., Li, Z., Dong, J., Gao, Y., Chang, Z., Ding, X., Li, S., Li, Y., Zeng, Y., Xin, Q., et al. (2022). Heterogeneity in endothelial cells and widespread venous arterialization during early vascular development in mammals. Cell Res *32*, 333-348. 10.1038/s41422-022-00615-z.
22. Zeng, Y., He, J., Bai, Z., Li, Z., Gong, Y., Liu, C., Ni, Y., Du, J., Ma, C., Bian, L., et al. (2019). Tracing the first hematopoietic stem cell generation in human embryo by single-cell RNA sequencing. Cell Research, 1-14. 10.1038/s41422-019-0228-6.
23. Kalucka, J., de Rooij, L.P.M.H., Goveia, J., Rohlenova, K., Dumas, S.J., Meta, E., Conchinha, N.V., Taverna, F., Teuwen, L.-A., Veys, K., et al. (2020). Single-Cell Transcriptome Atlas of Murine Endothelial Cells. Cell, 1-37. 10.1016/j.cell.2020.01.015.
24. Loh, K.M., and Ang, L.T. (2023). Building human artery and vein endothelial cells from pluripotent stem cells, and enduring mysteries surrounding arteriovenous development. Semin Cell Dev Biol. 10.1016/j.semcdb.2023.06.004.
25. Fang, J., and Hirschi, K. (2019). Molecular regulation of arteriovenous endothelial cell specification. F1000Research *8*, 1208. 10.12688/f1000research.16701.1.
26. Lin, F.-J., Tsai, M.-J., and Tsai, S.Y. (2007). Artery and vein formation: a tug of war between different forces. EMBO Rep *8*, 920-924.

27. Trimm, E., and Red-Horse, K. (2023). Vascular endothelial cell development and diversity. Nat Rev Cardiol *20*, 197-210. 10.1038/s41569-022-00770-1.
28. Chong, D.C., Koo, Y., Xu, K., Fu, S., and Cleaver, O. (2011). Stepwise arteriovenous fate acquisition during mammalian vasculogenesis. Developmental Dynamics *240*, 2153-2165. 10.1002/dvdy.22706.
29. Drake, C.J., and Fleming, P.A. (2000). Vasculogenesis in the day 6.5 to 9.5 mouse embryo. Blood *95*, 1671-1679.
30. Kim, Y.H., Hu, H., Guevara-Gallardo, S., Lam, M.T., Fong, S.Y., and Wang, R.A. (2008). Artery and vein size is balanced by Notch and ephrin B2/EphB4 during angiogenesis. Development *135*, 3755-3764. 10.1242/dev.022475.
31. Herbert, S.P., Huisken, J., Kim, T.N., Feldman, M.E., Houseman, B.T., Wang, R.A., Shokat, K.M., and Stainier, D.Y.R. (2009). Arterial-venous segregation by selective cell sprouting: an alternative mode of blood vessel formation. Science (New York, NY) *326*, 294-298.
32. Lindskog, H., Kim, Y.H., Jelin, E.B., Kong, Y., Guevara-Gallardo, S., Kim, T.N., and Wang, R.A. (2014). Molecular identification of venous progenitors in the dorsal aorta reveals an aortic origin for the cardinal vein in mammals. Development *141*, 1120-1128. 10.1242/dev.101808.
33. Coffin, J.D., Harrison, J., Schwartz, S., and Heimark, R. (1991). Angioblast differentiation and morphogenesis of the vascular endothelium in the mouse embryo. Dev Biol *148*, 51-62. 10.1016/0012-1606(91)90316-u.
34. Carmeliet, P., Ferreira, V., Breier, G., Pollefeyt, S., Kieckens, L., Gertsenstein, M., Fahrig, M., Vandenhoeck, A., Harpal, K., Eberhardt, C., et al. (1996). Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. Nature *380*, 435-439. 10.1038/380435a0.
35. Ferrara, N., Carver-Moore, K., Chen, H., Dowd, M., Lu, L., O'Shea, K.S., Powell-Braxton, L., Hillan, K.J., and Moore, M.W. (1996). Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene. Nature *380*, 439-442. 10.1038/380439a0.
36. Shalaby, F., Rossant, J., Yamaguchi, T.P., Gertsenstein, M., Wu, X.F., Breitman, M.L., and Schuh, A.C. (1995). Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. Nature *376*, 62-66.
37. Casie Chetty, S., Rost, M.S., Enriquez, J.R., Schumacher, J.A., Baltrunaite, K., Rossi, A., Stainier, D.Y.R., and Sumanas, S. (2017). Vegf signaling promotes vascular endothelial differentiation by modulating etv2 expression. Developmental Biology *424*, 147-161. 10.1016/j.ydbio.2017.03.005.
38. Lawson, N.D., Vogel, A.M., and Weinstein, B.M. (2002). sonic hedgehog and vascular endothelial growth factor act upstream of the Notch pathway during arterial endothelial differentiation. Developmental Cell *3*, 127-136.
39. Hong, C.C., Peterson, Q.P., Hong, J.-Y., and Peterson, R.T. (2006). Artery/vein specification is governed by opposing phosphatidylinositol-3 kinase and MAP kinase/ERK signaling. Curr Biol *16*, 1366-1372.
40. CDC, and NIH. (2020). Biosafety in Microbiological and Biomedical Laboratories—6th Edition (U.S. Department of Health and Human Services).

41. Harcourt, B.H., Sanchez, A., and Offermann, M.K. (1999). Ebola virus selectively inhibits responses to interferons, but not to interleukin-1beta, in endothelial cells. Journal of Virology *73*, 3491-3496. 10.1128/JVI.73.4.3491-3496.1999.
42. Geisbert, T.W., Young, H.A., Jahrling, P.B., Davis, K.J., Larsen, T., Kagan, E., and Hensley, L.E. (2003). Pathogenesis of Ebola hemorrhagic fever in primate models: evidence that hemorrhage is not a direct effect of virus-induced cytolysis of endothelial cells. The American Journal of Pathology *163*, 2371-2382. 10.1016/S0002-9440(10)63592-4.
43. Shrivastava-Ranjan, P., Rollin, P.E., and Spiropoulou, C.F. (2010). Andes Virus Disrupts the Endothelial Cell Barrier by Induction of Vascular Endothelial Growth Factor and Downregulation of VE-Cadherin. Journal of Virology *84*, 11227-11234. 10.1128/JVI.01405-10.
44. Gavrilovskaya, I.N., Gorbunova, E.E., Mackow, N.A., and Mackow, E.R. (2008). Hantaviruses direct endothelial cell permeability by sensitizing cells to the vascular permeability factor VEGF, while angiopoietin 1 and sphingosine 1-phosphate inhibit hantavirus-directed permeability. J Virol *82*, 5797-5806. 10.1128/JVI.02397-07.
45. Spiropoulou, C.F., Albarino, C.G., Ksiazek, T.G., and Rollin, P.E. (2007). Andes and Prospect Hill hantaviruses differ in early induction of interferon although both can downregulate interferon signaling. J Virol *81*, 2769-2776. 10.1128/JVI.02402-06.
46. Williams, E.P., Nandi, A., Nam, V., Allen, L.J.S., Trindade, A.A., Kosiewicz, M.M., and Jonsson, C.B. (2023). Modeling the Immune Response for Pathogenic and Nonpathogenic Orthohantavirus Infections in Human Lung Microvasculature Endothelial Cells. Viruses *15*. 10.3390/v15091806.
47. Mathieu, C., Guillaume, V., Sabine, A., Ong, K.C., Wong, K.T., Legras-Lachuer, C., and Horvat, B. (2012). Lethal Nipah virus infection induces rapid overexpression of CXCL10. PLoS ONE *7*, e32157. 10.1371/journal.pone.0032157.
48. Schountz, T., Campbell, C., Wagner, K., Rovnak, J., Martellaro, C., DeBuysscher, B., Feldmann, H., and Prescott, J. (2019). Differential Innate Immune Responses Elicited by Nipah Virus and Cedar Virus Correlate with Disparate In Vivo Pathogenesis in Hamsters. Viruses *11*, 291-211. 10.3390/v11030291.
49. Ang, L.T., Nguyen, A.T., Liu, K.J., Chen, A., Xiong, X., Curtis, M., Martin, R.M., Raftry, B.C., Ng, C.Y., Vogel, U., et al. (2022). Generating human artery and vein cells from pluripotent stem cells highlights the arterial tropism of Nipah and Hendra viruses. Cell *185*, 2523-2541 e2530. 10.1016/j.cell.2022.05.024.
50. Lo, M.K., Miller, D., Aljofan, M., Mungall, B.A., Rollin, P.E., Bellini, W.J., and Rota, P.A. (2010). Characterization of the antiviral and inflammatory responses against Nipah virus in endothelial cells and neurons. Virology *404*, 78-88. 10.1016/j.virol.2010.05.005.
51. Aljofan, M., Saubern, S., Meyer, A.G., Marsh, G., Meers, J., and Mungall, B.A. (2009). Characteristics of Nipah virus and Hendra virus replication in different cell lines and their suitability for antiviral screening. Virus research *142*, 92-99. 10.1016/j.virusres.2009.01.014.
52. Erbar, S., Diederich, S., and Maisner, A. (2008). Selective receptor expression restricts Nipah virus infection of endothelial cells. Virology journal *5*, 142-146. 10.1186/1743-422X-5-142.

53. DeBuysscher, B.L., Scott, D.P., Rosenke, R., Wahl, V., Feldmann, H., and Prescott, J. (2021). Nipah Virus Efficiently Replicates in Human Smooth Muscle Cells without Cytopathic Effect. Cells *10*. 10.3390/cells10061319.
54. Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nature Methods *528*, 142. 10.1038/nmeth.4396.
55. Koyano-Nakagawa, N., and Garry, D.J. (2017). Etv2 as an essential regulator of mesodermal lineage development. Cardiovascular Research *113*, 1294-1306. 10.1093/cvr/cvx133.
56. Ditadi, A., Sturgeon, C.M., Tober, J., Awong, G., Kennedy, M., Yzaguirre, A.D., Azzola, L., Ng, E.S., Stanley, E.G., French, D.L., et al. (2015). Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. Nature Cell Biology *17*, 580-591. 10.1038/ncb3161.
57. Gage, B.K., Liu, J.C., Innes, B.T., MacParland, S.A., McGilvray, I.D., Bader, G.D., and Keller, G.M. (2020). Generation of Functional Liver Sinusoidal Endothelial Cells from Human Pluripotent Stem-Cell-Derived Venous Angioblasts. Cell Stem Cell. 10.1016/j.stem.2020.06.007.
58. Park, M.A., Kumar, A., Jung, H.S., Uenishi, G., Moskvin, O.V., Thomson, J.A., and Slukvin, I.I. (2018). Activation of the Arterial Program Drives Development of Definitive Hemogenic Endothelium with Lymphoid Potential. Cell Reports *23*, 2467-2481. 10.1016/j.celrep.2018.04.092.
59. Rosa, S., Praça, C., Pitrez, P.R., Gouveia, P.J., Aranguren, X.L., Ricotti, L., and Ferreira, L.S. (2019). Functional characterization of iPSC-derived arterial- and venous-like endothelial cells. Scientific Reports *9*, 3826-3815. 10.1038/s41598-019-40417-9.
60. Pan, Z., Yao, Q., Kong, W., Ma, X., Tian, L., Zhao, Y., Zhu, S., Chen, S., Sun, M., Liu, J., et al. (2024). Generation of iPSC-derived human venous endothelial cells for the modeling of vascular malformations and drug discovery. Cell Stem Cell. 10.1016/j.stem.2024.10.015.
61. Nikolova, M.T., He, Z., Seimiya, M., Jonsson, G., Cao, W., Okuda, R., Wimmer, R.A., Okamoto, R., Penninger, J.M., Camp, J.G., and Treutlein, B. (2025). Fate and state transitions during human blood vessel organoid development. Cell. 10.1016/j.cell.2025.03.037.
62. Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.
63. González-Hernández, S., Gomez, M.J., Sanchez-Cabo, F., Mendez-Ferrer, S., Munoz-Canoves, P., and Isern, J. (2020). Sox17 Controls Emergence and Remodeling of Nestin-Expressing Coronary Vessels. Circ Res *127*, e252-e270. 10.1161/CIRCRESAHA.120.317121.
64. Sharma, B., Ho, L., Ford, G.H., Chen, H.I., Goldstone, A.B., Woo, Y.J., Quertermous, T., Reversade, B., and Red-Horse, K. (2017). Alternative Progenitor Cells Compensate to Rebuild the Coronary Vasculature in Elabela- and Apj-Deficient Hearts. Dev Cell *42*, 655-666 e653. 10.1016/j.devcel.2017.08.008.
65. Chen, J., Zhang, X., DeLaughter, D.M., Trembley, M.A., Saifee, S., Xiao, F., Chen, J., Zhou, P., Seidman, C.E., Seidman, J.G., and Pu, W.T. (2024). Molecular and Spatial

Signatures of Mouse Embryonic Endothelial Cells at Single-Cell Resolution. Circ Res *134*, 529-546. 10.1161/CIRCRESAHA.123.323956.

66. Calvanese, V., Capellera-Garcia, S., Ma, F., Fares, I., Liebscher, S., Ng, E.S., Ekstrand, S., Aguade-Gorgorio, J., Vavilina, A., Lefaudeux, D., et al. (2022). Mapping human haematopoietic stem cells from haemogenic endothelium to birth. Nature *604*, 534-540. 10.1038/s41586-022-04571-x.
67. Engert, S., Liao, W.P., Burtscher, I., and Lickert, H. (2009). Sox17-2A-iCre: a knock-in mouse line expressing Cre recombinase in endoderm and vascular endothelial cells. Genesis *47*, 603-610.
68. Choi, E., Kraus, M.R.C., Lemaire, L.A., Yoshimoto, M., Vemula, S., Potter, L.A., Manduchi, E., Stoeckert, C.J., Grapin-Botton, A., and Magnuson, M.A. (2012). Dual lineage-specific expression of Sox17 during mouse embryogenesis. Stem cells (Dayton, Ohio) *30*, 2297-2308.
69. Lilly, A.J., Lacaud, G., and Kouskoff, V. (2017). SOXF transcription factors in cardiovascular development. Semin Cell Dev Biol *63*, 50-57. 10.1016/j.semcdb.2016.07.021.
70. Metikala, S., Warkala, M., Casie Chetty, S., Chestnut, B., Rufin Florat, D., Plender, E., Nester, O., Koenig, A.L., Astrof, S., and Sumanas, S. (2022). Integration of vascular progenitors into functional blood vessels represents a distinct mechanism of vascular growth. Dev Cell *57*, 767-782 e766. 10.1016/j.devcel.2022.02.015.
71. Ramirez, A., Vyzas, C.A., Zhao, H., Eng, K., Degenhardt, K., and Astrof, S. (2024). Buffering Mechanism in Aortic Arch Artery Formation and Congenital Heart Disease. Circ Res *134*, e112-e132. 10.1161/CIRCRESAHA.123.322767.
72. Plummer, N.W., Evsyukova, I.Y., Robertson, S.D., de Marchena, J., Tucker, C.J., and Jensen, P. (2015). Expanding the power of recombinase-based labeling to uncover cellular diversity. Development *142*, 4385-4393. 10.1242/dev.129981.
73. Payne, S., Neal, A., and De Val, S. (2024). Transcription factors regulating vasculogenesis and angiogenesis. Dev Dyn *253*, 28-58. 10.1002/dvdy.575.
74. Alerasool, N., Segal, D., Lee, H., and Taipale, M. (2020). An efficient KRAB domain for CRISPRi applications in human cells. Nat Methods *17*, 1093-1096. 10.1038/s41592-020-0966-x.
75. Swift, M.R., Pham, V.N., Castranova, D., Bell, K., Poole, R.J., and Weinstein, B.M. (2014). SoxF factors and Notch regulate nr2f2 gene expression during venous differentiation in zebrafish. Dev Biol *390*, 116-125. 10.1016/j.ydbio.2014.03.018.
76. Zhou, Y., Williams, J., Smallwood, P.M., and Nathans, J. (2015). Sox7, Sox17, and Sox18 Cooperatively Regulate Vascular Development in the Mouse Retina. PLoS One *10*, e0143650. 10.1371/journal.pone.0143650.
77. Martin, R.M., Ikeda, K., Cromer, M.K., Uchida, N., Nishimura, T., Romano, R., Tong, A.J., Lemgart, V.T., Camarena, J., Pavel-Dinu, M., et al. (2019). Highly Efficient and Marker-free Genome Editing of Human Pluripotent Stem Cells by CRISPR-Cas9 RNP and AAV6 Donor-Mediated Homologous Recombination. Cell Stem Cell *24*, 821-828.e825. 10.1016/j.stem.2019.04.001.
78. Skene, P.J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife *6*. 10.7554/eLife.21856.
79. Sacilotto, N., Monteiro, R., Fritzsche, M., Becker, P.W., Sanchez-Del-Campo, L., Liu, K., Pinheiro, P., Ratnayaka, I., Davies, B., Goding, C.R., et al. (2013). Analysis of Dll4

regulation reveals a combinatorial role for Sox and Notch in arterial development. Proceedings of the National Academy of Sciences of the United States of America *110*, 11893-11898. 10.1073/pnas.1300805110.

80. Liu, X., Uemura, A., Fukushima, Y., Yoshida, Y., and Hirashima, M. (2016). Semaphorin 3G Provides a Repulsive Guidance Cue to Lymphatic Endothelial Cells via Neuropilin-2/PlexinD1. Cell Rep *17*, 2299-2311. 10.1016/j.celrep.2016.11.008.
81. Kutschera, S., Weber, H., Weick, A., De Smet, F., Genove, G., Takemoto, M., Prahst, C., Riedel, M., Mikelis, C., Baulande, S., et al. (2011). Differential endothelial transcriptomics identifies semaphorin 3G as a vascular class 3 semaphorin. Arterioscler Thromb Vasc Biol *31*, 151-159. 10.1161/ATVBAHA.110.215871.
82. Kosicki, M., Baltoumas, F.A., Kelman, G., Boverhof, J., Ong, Y., Cook, L.E., Dickel, D.E., Pavlopoulos, G.A., Pennacchio, L.A., and Visel, A. (2025). VISTA Enhancer browser: an updated database of tissue-specific developmental enhancers. Nucleic Acids Res *53*, D324-D330. 10.1093/nar/gkae940.
83. Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser--a database of tissue-specific human enhancers. Nucleic Acids Res *35*, D88-92. 10.1093/nar/gkl822.
84. Nornes, S., Bruche, S., Adak, N., McCracken, I.R., and De Val, S. (2025). Evaluating the transcriptional regulators of arterial gene expression via a catalogue of characterized arterial enhancers. Elife *14*. 10.7554/eLife.102440.
85. Paik, D.T., Tian, L., Lee, J., Sayed, N., Chen, I.Y., Rhee, S., Rhee, J.-W., Kim, Y., Wirka, R.C., Buikema, J.W., et al. (2018). Large-Scale Single-Cell RNA-Seq Reveals Molecular Signatures of Heterogeneous Populations of Human Induced Pluripotent Stem Cell-Derived Endothelial Cells. Circulation Research. 10.1161/CIRCRESAHA.118.312913.
86. Zhang, J., Chu, L.-F., Hou, Z., Schwartz, M.P., Hacker, T., Vickerman, V., Swanson, S., Leng, N., Nguyen, B.K., Elwell, A., et al. (2017). Functional characterization of human pluripotent stem cell-derived arterial endothelial cells. Proceedings of the National Academy of Sciences of the United States of America, 201702295. 10.1073/pnas.1702295114.
87. McCracken, I.R., Taylor, R.S., Kok, F.O., de la Cuesta, F., Dobie, R., Henderson, B.E.P., Mountford, J.C., Caudrillier, A., Henderson, N.C., Ponting, C.P., and Baker, A.H. (2019). Transcriptional dynamics of pluripotent stem cell-derived endothelial cell differentiation revealed by single-cell RNA sequencing. European Heart Journal *385*, 9963117. 10.1093/eurheartj/ehz351.
88. Sabata, H., Roca-Coll, A., Dengra, J.A., Gouveia, L., Martinez-Larrinaga, A., Collado-Remacha, A., Castillo, S.D., Castillo, E., Tisch, N., De Andrés-Laguillo, M., et al. (2025). Context-dependent response of endothelial cells to PIK3CA mutation. bioRxiv. 10.1101/2025.02.25.640041.
89. Wimmer, R.A., Leopoldi, A., Aichinger, M., Wick, N., Hantusch, B., Novatchkova, M., Taubenschmid, J., mmerle, M.H.x.E., Esk, C., Bagley, J.A., et al. (2019). Human blood vessel organoids as a model of diabetic vasculopathy. Nature, 1-34. 10.1038/s41586-018-0858-8.
90. Gong, L., Zhang, Y., Zhu, Y., Lee, U., Luo, A.C., Li, X., Wang, X., Chen, D., Pu, W.T., Lin, R.Z., et al. (2025). Rapid generation of functional vascular organoids via simultaneous transcription factor activation of endothelial and mural lineages. Cell Stem Cell. 10.1016/j.stem.2025.05.014.

91. Baskerville, A., Fisher-Hoch, S.P., Neild, G.H., and Dowsett, A.B. (1985). Ultrastructural pathology of experimental Ebola haemorrhagic fever virus infection. J Pathol *147*, 199-209. 10.1002/path.1711470308.
92. Jaax, N.K., Davis, K.J., Geisbert, T.J., Vogel, P., Jaax, G.P., Topper, M., and Jahrling, P.B. (1996). Lethal experimental infection of rhesus monkeys with Ebola-Zaire (Mayinga) virus by the oral and conjunctival route of exposure. Arch Pathol Lab Med *120*, 140-155.
93. Davis, K.J., Anderson, A.O., Geisbert, T.W., Steele, K.E., Geisbert, J.B., Vogel, P., Connolly, B.M., Huggins, J.W., Jahrling, P.B., and Jaax, N.K. (1997). Pathology of experimental Ebola virus infection in African green monkeys. Involvement of fibroblastic reticular cells. Arch Pathol Lab Med *121*, 805-819.
94. Wong, K.T., Shieh, W.-J., Kumar, S., Norain, K., Abdullah, W., Guarner, J., Goldsmith, C.S., Chua, K.B., Lam, S.K., Tan, C.T., et al. (2002). Nipah virus infection: pathology and pathogenesis of an emerging paramyxoviral zoonosis. The American Journal of Pathology *161*, 2153-2167. 10.1016/S0002-9440(10)64493-8.
95. Geisbert, T.W., Daddario-DiCaprio, K.M., Hickey, A.C., Smith, M.A., Chan, Y.-P., Wang, L.-F., Mattapallil, J.J., Geisbert, J.B., Bossart, K.N., and Broder, C.C. (2010). Development of an acute and highly pathogenic nonhuman primate model of Nipah virus infection. PLoS ONE *5*, e10690. 10.1371/journal.pone.0010690.
96. Baseler, L., de Wit, E., Scott, D.P., Munster, V.J., and Feldmann, H. (2015). Syrian hamsters (Mesocricetus auratus) oronasally inoculated with a Nipah virus isolate from Bangladesh or Malaysia develop similar respiratory tract lesions. Veterinary pathology *52*, 38-45. 10.1177/0300985814556189.
97. Safronetz, D., Zivcec, M., Lacasse, R., Feldmann, F., Rosenke, R., Long, D., Haddock, E., Brining, D., Gardner, D., Feldmann, H., and Ebihara, H. (2011). Pathogenesis and host response in Syrian hamsters following intranasal infection with Andes virus. PLoS Pathog *7*, e1002426. 10.1371/journal.ppat.1002426.
98. Wahl-Jensen, V., Chapman, J., Asher, L., Fisher, R., Zimmerman, M., Larsen, T., and Hooper, J.W. (2007). Temporal analysis of Andes virus and Sin Nombre virus infections of Syrian hamsters. J Virol *81*, 7449-7462. 10.1128/JVI.00238-07.
99. Hooper, J.W., Larsen, T., Custer, D.M., and Schmaljohn, C.S. (2001). A lethal disease model for hantavirus pulmonary syndrome. Virology *289*, 6-14. 10.1006/viro.2001.1133.
100. López, C.B., Yount, J.S., Hermesh, T., and Moran, T.M. (2006). Sendai virus infection induces efficient adaptive immunity independently of type I interferons. Journal of Virology *80*, 4538-4545. 10.1128/JVI.80.9.4538-4545.2006.
101. Honda, K., Takaoka, A., and Taniguchi, T. (2006). Type I interferon [corrected] gene induction by the interferon regulatory factor family of transcription factors. Immunity *25*, 349-360. 10.1016/j.immuni.2006.08.009.
102. Shaw, A.E., Hughes, J., Gu, Q., Behdenna, A., Singer, J.B., Dennis, T., Orton, R.J., Varela, M., Gifford, R.J., Wilson, S.J., and Palmarini, M. (2017). Fundamental properties of the mammalian innate immune system revealed by multispecies comparison of type I interferon responses. PLoS Biology *15*, e2004086. 10.1371/journal.pbio.2004086.
103. Cimica, V., Dalrymple, N.A., Roth, E., Nasonov, A., and Mackow, E.R. (2014). An innate immunity-regulating virulence determinant is uniquely encoded by the Andes virus nucleocapsid protein. mBio *5*. 10.1128/mBio.01088-13.

104. Matthys, V.S., Cimica, V., Dalrymple, N.A., Glennon, N.B., Bianco, C., and Mackow, E.R. (2014). Hantavirus GnT elements mediate TRAF3 binding and inhibit RIG-I/TBK1-directed beta interferon transcription by blocking IRF3 phosphorylation. J Virol *88*, 2246-2259. 10.1128/JVI.02647-13.
105. Levine, J.R., Prescott, J., Brown, K.S., Best, S.M., Ebihara, H., and Feldmann, H. (2010). Antagonism of type I interferon responses by new world hantaviruses. Journal of Virology *84*, 11790-11801. 10.1128/JVI.00916-10.
106. Shaw, M.L., Garcia-Sastre, A., Palese, P., and Basler, C.F. (2004). Nipah virus V and W proteins have a common STAT1-binding domain yet inhibit STAT1 activation from the cytoplasmic and nuclear compartments, respectively. J Virol *78*, 5633-5641. 10.1128/JVI.78.11.5633-5641.2004.
107. Bharaj, P., Wang, Y.E., Dawes, B.E., Yun, T.E., Park, A., Yen, B., Basler, C.F., Freiberg, A.N., Lee, B., and Rajsbaum, R. (2016). The Matrix Protein of Nipah Virus Targets the E3-Ubiquitin Ligase TRIM6 to Inhibit the IKKepsilon Kinase-Mediated Type-I IFN Antiviral Response. PLoS Pathog *12*, e1005880. 10.1371/journal.ppat.1005880.
108. Park, M.S., Shaw, M.L., Munoz-Jordan, J., Cros, J.F., Nakaya, T., Bouvier, N., Palese, P., Garcia-Sastre, A., and Basler, C.F. (2003). Newcastle disease virus (NDV)-based assay demonstrates interferon-antagonist activity for the NDV V protein and the Nipah virus V, W, and C proteins. J Virol *77*, 1501-1511. 10.1128/jvi.77.2.1501-1511.2003.
109. Becker, N., and Maisner, A. (2023). Nipah Virus Impairs Autocrine IFN Signaling by Sequestering STAT1 and STAT2 into Inclusion Bodies. Viruses *15*. 10.3390/v15020554.
110. Huang, J., Hume, A.J., Abo, K.M., Werder, R.B., Villacorta-Martin, C., Alysandratos, K.-D., Beermann, M.L., Simone-Roach, C., Lindstrom-Vautrin, J., Olejnik, J., et al. (2020). SARS-CoV-2 Infection of Pluripotent Stem Cell-Derived Human Lung Alveolar Type 2 Cells Elicits a Rapid Epithelial-Intrinsic Inflammatory Response. Cell Stem Cell. 10.1016/j.stem.2020.09.013.
111. Wauquier, N., Becquart, P., Padilla, C., Baize, S., and Leroy, E.M. (2010). Human fatal zaire ebola virus infection is associated with an aberrant innate immunity and with massive lymphocyte apoptosis. PLoS Negl Trop Dis *4*. 10.1371/journal.pntd.0000837.
112. Baize, S., Leroy, E.M., Georges-Courbot, M.C., Capron, M., Lansoud-Soukate, J., Debre, P., Fisher-Hoch, S.P., McCormick, J.B., and Georges, A.J. (1999). Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients. Nat Med *5*, 423-426. 10.1038/7422.
113. Ruibal, P., Oestereich, L., Ludtke, A., Becker-Ziaja, B., Wozniak, D.M., Kerber, R., Korva, M., Cabeza-Cabrerizo, M., Bore, J.A., Koundouno, F.R., et al. (2016). Unique human immune signature of Ebola virus disease in Guinea. Nature *533*, 100-104. 10.1038/nature17949.
114. Reynard, S., Journeaux, A., Gloaguen, E., Schaeffer, J., Varet, H., Pietrosemoli, N., Mateo, M., Baillet, N., Laouenan, C., Raoul, H., et al. (2019). Immune parameters and outcomes during Ebola virus disease. JCI Insight *4*. 10.1172/jci.insight.125106.
115. Lucas, C., Wong, P., Klein, J., Castro, T.B.R., Silva, J., Sundaram, M., Ellingson, M.K., Mao, T., Oh, J.E., Israelow, B., et al. (2020). Longitudinal analyses reveal immunological misfiring in severe COVID-19. Nature *584*, 463-469. 10.1038/s41586-020-2588-y.
116. Uyeki, T.M., Mehta, A.K., Davey, R.T., Jr., Liddell, A.M., Wolf, T., Vetter, P., Schmiedel, S., Grunewald, T., Jacobs, M., Arribas, J.R., et al. (2016). Clinical Management of Ebola

Virus Disease in the United States and Europe. N Engl J Med *374*, 636-646. 10.1056/NEJMoa1504874.
117. Wisniewski, H.G., and Vilcek, J. (1997). TSG-6: an IL-1/TNF-inducible protein with anti-inflammatory activity. Cytokine Growth Factor Rev *8*, 143-156. 10.1016/s1359-6101(97)00008-7.
118. Li, Y., Ozaki, T., Kikuchi, H., Yamamoto, H., Ohira, M., and Nakagawara, A. (2008). A novel HECT-type E3 ubiquitin protein ligase NEDL1 enhances the p53-mediated apoptotic cell death in its catalytic activity-independent manner. Oncogene *27*, 3700-3709. 10.1038/sj.onc.1211032.
119. Cao, J., O'Day, D.R., Pliner, H.A., Kingsley, P.D., Deng, M., Daza, R.M., Zager, M.A., Aldinger, K.A., Blecher-Gonen, R., Zhang, F., et al. (2020). A human cell atlas of fetal gene expression. Science *370*. 10.1126/science.aba7721.
120. Guo, F.H., Guan, Y.N., Guo, J.J., Zhang, L.J., Qiu, J.J., Ji, Y., Chen, A.F., and Jing, Q. (2022). Single-Cell Transcriptome Analysis Reveals Embryonic Endothelial Heterogeneity at Spatiotemporal Level and Multifunctions of MicroRNA-126 in Mice. Arterioscler Thromb Vasc Biol *42*, 326-342. 10.1161/ATVBAHA.121.317093.
121. Miao, Y., Pek, N.M., Tan, C., Jiang, C., Yu, Z., Iwasawa, K., Shi, M., Kechele, D.O., Sundaram, N., Pastrana-Gomez, V., et al. (2025). Co-development of mesoderm and endoderm enables organotypic vascularization in lung and gut organoids. Cell. 10.1016/j.cell.2025.05.041.
122. Chen, D., Fan, X., Sun, N., Wang, K., Gong, L., Melero-Martin, J.M., and Pu, W.T. (2025). Pioneer factor ETV2 safeguards endothelial cell specification by recruiting the repressor REST to restrict alternative lineage commitment. Nat Cardiovasc Res *4*, 689-709. 10.1038/s44161-025-00660-y.
123. O'Rahilly, R., and Müller, F. (1987). Developmental stages in human embryos (Carnegie Institute of Washington).
124. Fowler, J.L., Ang, L.T., and Loh, K.M. (2019). A critical look: Challenges in differentiating human pluripotent stem cells into desired cell types and organoids. Wiley Interdisciplinary Reviews: Developmental Biology *113*, 891-823. 10.1002/wdev.368.
125. Wagner, D.E., and Klein, A.M. (2020). Lineage tracing meets single-cell omics: opportunities and challenges. Nature Reviews Genetics, 1-18. 10.1038/s41576-020-0223-2.
126. Domcke, S., and Shendure, J. (2023). A reference cell tree will serve science better than a reference cell atlas. Cell *186*, 1103-1114. 10.1016/j.cell.2023.02.016.
127. Kim, S., and Wysocka, J. (2023). Deciphering the multi-scale, quantitative cis-regulatory code. Mol Cell *83*, 373-392. 10.1016/j.molcel.2022.12.032.
128. Pabst, T., Joncourt, R., Shumilov, E., Heini, A., Wiedemann, G., Legros, M., Seipel, K., Schild, C., Jalowiec, K., Mansouri Taleghani, B., et al. (2020). Analysis of IL-6 serum levels and CAR T cell-specific digital PCR in the context of cytokine release syndrome. Exp Hematol *88*, 7-14 e13. 10.1016/j.exphem.2020.07.003.
129. Chan, M.M., Smith, Z.D., Grosswendt, S., Kretzmer, H., Norman, T.M., Adamson, B., Jost, M., Quinn, J.J., Yang, D., Jones, M.G., et al. (2019). Molecular recording of mammalian embryogenesis. Nature, 1. 10.1038/s41586-019-1184-5.
130. Bowling, S., Sritharan, D., Osorio, F.G., Nguyen, M., Cheung, P., Rodriguez-Fraticelli, A., Patel, S., Yuan, W.C., Fujiwara, Y., Li, B.E., et al. (2020). An Engineered CRISPR-

Cas9 Mouse Line for Simultaneous Readout of Lineage Histories and Gene Expression Profiles in Single Cells. Cell *181*, 1410-1422 e1427. 10.1016/j.cell.2020.04.048.

131. Tee, W.W., Shen, S.S., Oksuz, O., Narendra, V., and Reinberg, D. (2014). Erk1/2 activity promotes chromatin features and RNAPII phosphorylation at developmental promoters in mouse ESCs. Cell *156*, 678-690. 10.1016/j.cell.2014.01.009.
132. Bosio, C.M., Aman, M.J., Grogan, C., Hogan, R., Ruthel, G., Negley, D., Mohamadzadeh, M., Bavari, S., and Schmaljohn, A. (2003). Ebola and Marburg viruses replicate in monocyte-derived dendritic cells without inducing the production of cytokines and full maturation. J Infect Dis *188*, 1630-1638. 10.1086/379199.
133. Basler, C.F., Wang, X., Muhlberger, E., Volchkov, V., Paragas, J., Klenk, H.D., Garcia-Sastre, A., and Palese, P. (2000). The Ebola virus VP35 protein functions as a type I IFN antagonist. Proc Natl Acad Sci U S A *97*, 12289-12294. 10.1073/pnas.220398297.
134. Basler, C.F., Mikulasova, A., Martinez-Sobrido, L., Paragas, J., Muhlberger, E., Bray, M., Klenk, H.D., Palese, P., and Garcia-Sastre, A. (2003). The Ebola virus VP35 protein inhibits activation of interferon regulatory factor 3. J Virol *77*, 7945-7956. 10.1128/jvi.77.14.7945-7956.2003.
135. Guito, J.C., Albarino, C.G., Chakrabarti, A.K., and Towner, J.S. (2017). Novel activities by ebolavirus and marburgvirus interferon antagonists revealed using a standardized in vitro reporter system. Virology *501*, 147-165. 10.1016/j.virol.2016.11.015.
136. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902.e1821. 10.1016/j.cell.2019.05.031.
137. Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nature Neuroscience *13*, 133-140. 10.1038/nn.2467.
138. Muzumdar, M.D., Tasic, B., Miyamichi, K., Li, L., and Luo, L. (2007). A global double-fluorescent Cre reporter mouse. Genesis *45*, 593-605. 10.1002/dvg.20335.
139. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550. 10.1073/pnas.0506580102.
140. Avižinienė, A., Kucinskaite-Kodze, I., Petraityte-Burneikiene, R., Zvirbliene, A., Mertens, M.L., Schmidt, S., Schlegel, M., Lattwein, E., Koellner, B., and Ulrich, R.G. (2023). Characterization of a Panel of Cross-Reactive Hantavirus Nucleocapsid Protein-Specific Monoclonal Antibodies. Viruses *15*. 10.3390/v15020532.
141. Johnson, K.M., Lange, J.V., Webb, P.A., and Murphy, F.A. (1977). Isolation and partial characterisation of a new virus causing acute haemorrhagic fever in Zaire. Lancet *1*, 569-571. 10.1016/s0140-6736(77)92000-1.
142. Toro, J., Vega, J.D., Khan, A.S., Mills, J.N., Padula, P., Terry, W., Yadon, Z., Valderrama, R., Ellis, B.A., Pavletic, C., et al. (1998). An outbreak of hantavirus pulmonary syndrome, Chile, 1997. Emerg Infect Dis *4*, 687-694. 10.3201/eid0404.980425.

143. Harcourt, B.H., Tamin, A., Ksiazek, T.G., Rollin, P.E., Anderson, L.J., Bellini, W.J., and Rota, P.A. (2000). Molecular characterization of Nipah virus, a newly emergent paramyxovirus. Virology *271*, 334-349. 10.1006/viro.2000.0340.
144. Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.
145. Kreitzer, F.R., Salomonis, N., Sheehan, A., Huang, M., Park, J.S., Spindler, M.J., Lizarraga, P., Weiss, W.A., So, P.L., and Conklin, B.R. (2013). A robust method to derive functional neural crest cells from human pluripotent stem cells. Am J Stem Cells *2*, 119-131.
146. Guna, A., Page, K.R., Replogle, J.M., Esantsi, T.K., Wang, M.L., Weissman, J.S., and Voorhees, R.M. (2023). A dual sgRNA library design to probe genetic modifiers using genome-wide CRISPRi screens. BMC Genomics *24*, 651. 10.1186/s12864-023-09754-y.
147. Goldstein, A.S., Drake, J.M., Burnes, D.L., Finley, D.S., Zhang, H., Reiter, R.E., Huang, J., and Witte, O.N. (2011). Purification and direct transformation of epithelial progenitor cells from primary human prostate. Nature Protocols *6*, 656-667. 10.1038/nprot.2011.317.
148. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676-682. 10.1038/nmeth.2019.
149. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573-3587 e3529. 10.1016/j.cell.2021.04.048.
150. Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nature Communications *8*, 14049. 10.1038/ncomms14049.
151. Rainer, J. (2025). EnsDb.Hsapiens.v86. https://bioconductor.org/packages/release/data/annotation/html/EnsDb.Hsapiens.v86.html.
152. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*, 550. 10.1186/s13059-014-0550-8.
153. Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284-287. 10.1089/omi.2011.0118.
154. Ouyang, J.F., Kamaraj, U.S., Cao, E.Y., and Rackham, O.J.L. (2021). ShinyCell: Simple and sharable visualisation of single-cell gene expression data. Bioinformatics. 10.1093/bioinformatics/btab209.
155. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
156. Krueger, F. (2012). Trim Galore. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
157. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nature Biotechnology *34*, 525-527. 10.1038/nbt.3519.

158. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140. 10.1093/bioinformatics/btp616.
159. Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol *15*, R29. 10.1186/gb-2014-15-2-r29.
160. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res *43*, e47. 10.1093/nar/gkv007.
161. Wickham, H., François, R., Henry, L., and Müller, K. (2022). dplyr: A Grammar of Data Manipulation. https://dplyr.tidyverse.org.
162. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag New York).
163. Wickham, H., Averick, M., Bryan, J., and Chang, W. (2019). Welcome to the Tidyverse. The Journal of Open Source Software. 10.21105/joss.01686.
164. Wingett, S.W., and Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality control. F1000Res *7*, 1338. 10.12688/f1000research.15931.2.
165. Gaspar, J.M. (2018). NGmerge: merging paired-end reads via novel empirically-derived models of sequencing errors. BMC Bioinformatics *19*, 536. 10.1186/s12859-018-2579-2.
166. Faust, G.G., and Hall, I.M. (2014). SAMBLASTER: fast duplicate marking and structural variant read extraction. Bioinformatics *30*, 2503-2505. 10.1093/bioinformatics/btu314.
167. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079. 10.1093/bioinformatics/btp352.
168. Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160-165. 10.1093/nar/gkw257.
169. Soneson, C., Love, M.I., and Robinson, M.D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res *4*, 1521. 10.12688/f1000research.7563.2.
170. Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389-393. 10.1038/nature10730.
171. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.
172. Kolde, R. (2025). pheatmap. https://github.com/raivokolde/pheatmap.
173. Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R., Propson, N.E., et al. (2011). Chemically defined conditions for human iPSC derivation and culture. Nature Methods *8*, 424-429.
174. Loh, K.M., Zheng, S.L., Liu, K.J., Yin, Q., Amir-Ugokwe, Z.A., Jha, S.K., Qi, Y., Wazny, V.K., Nguyen, A.T., Chen, A., et al. (2024). Protocol for efficient generation of human

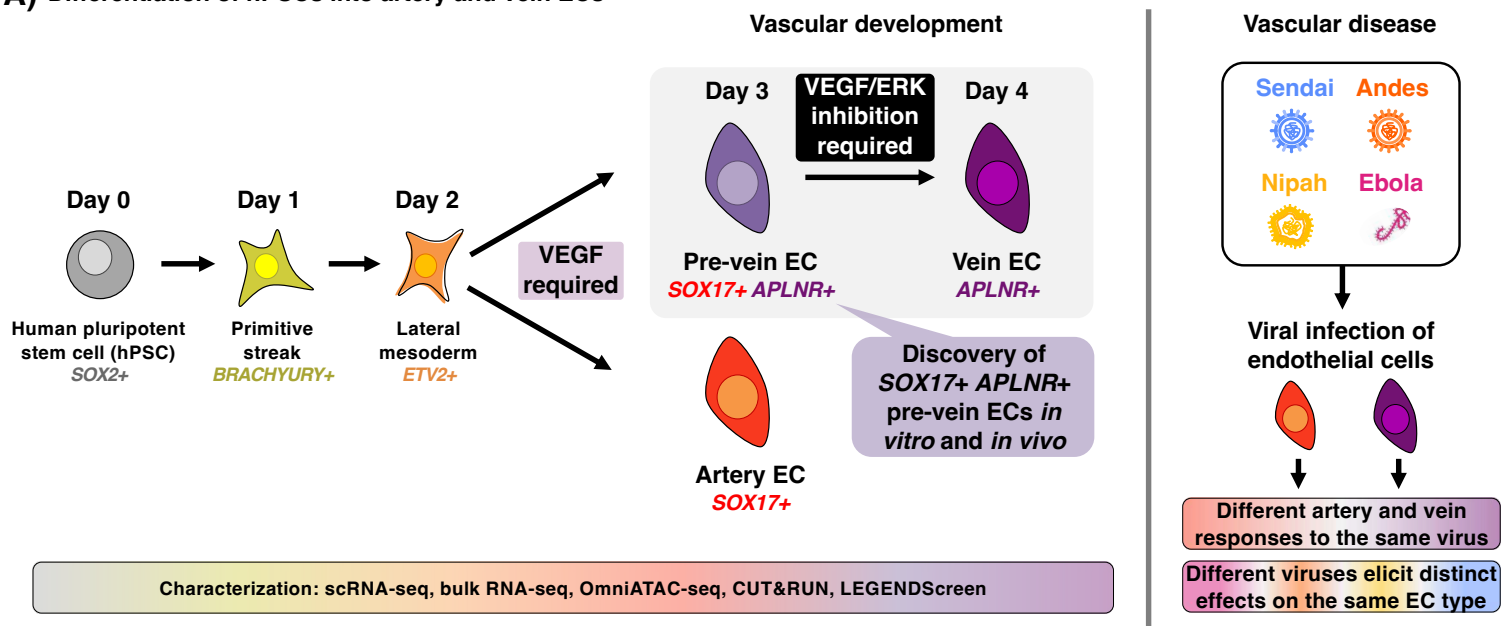
artery and vein endothelial cells from pluripotent stem cells. STAR Protoc *6*, 103494. 10.1016/j.xpro.2024.103494.
175. Biedenkopf, N., Bukreyev, A., Chandran, K., Di Paola, N., Formenty, P.B.H., Griffiths, A., Hume, A.J., Muhlberger, E., Netesov, S.V., Palacios, G., et al. (2023). Renaming of genera Ebolavirus and Marburgvirus to Orthoebolavirus and Orthomarburgvirus, respectively, and introduction of binomial species names within family Filoviridae. Arch Virol *168*, 220. 10.1007/s00705-023-05834-2.
176. Alonso, D.O., Perez-Sautu, U., Bellomo, C.M., Prieto, K., Iglesias, A., Coelho, R., Periolo, N., Domenech, I., Talmon, G., Hansen, R., et al. (2020). Person-to-Person Transmission of Andes Virus in Hantavirus Pulmonary Syndrome, Argentina, 2014. Emerg Infect Dis *26*, 756-759. 10.3201/eid2604.190799.
177. Cantell, K., and Hirvonen, S. (1981). Preparation and assay of Sendai virus. Methods in enzymology *78*, 299-301. 10.1016/0076-6879(81)78131-x.
178. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26. 10.1038/nbt.1754.
179. Fowler, J.L., Zheng, S.L., Nguyen, A., Chen, A., Xiong, X., Chai, T., Chen, J.Y., Karigane, D., Banuelos, A.M., Niizuma, K., et al. (2024). Lineage-tracing hematopoietic stem cell origins in vivo to efficiently make human HLF+ HOXA+ hematopoietic progenitors from pluripotent stem cells. Dev Cell. 10.1016/j.devcel.2024.03.003.
180. Choi, H.M.T., Schwarzkopf, M., Fornace, M.E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A., and Pierce, N.A. (2018). Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. Development *145*, dev165753-165122. 10.1242/dev.165753.
181. Loh, K.M., Ang, L.T., Zhang, J., Kumar, V., Ang, J., Auyeong, J.Q., Lee, K.L., Choo, S.H., Lim, C Y.Y., Nichane, M., et al. (2014). Efficient Endoderm Induction from Human Pluripotent Stem Cells by Logically Directing Signals Controlling Lineage Bifurcations. Cell Stem Cell *14*, 237-252.
182. Loh, K.M., Chen, A., Koh, P.W., Deng, T.Z., Sinha, R., Tsai, J.M., Barkal, A.A., Shen, K.Y., Jain, R., Morganti, R.M., et al. (2016). Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types. Cell *166*, 451-467. 10.1016/j.cell.2016.06.011.
183. Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nature Biotechnology *37*, 38-44. 10.1038/nbt.4314.
184. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods *9*, 357-359. 10.1038/nmeth.1923.
185. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.
186. Gerstein, M.B., Kundaje, A.B., Hariharan, M., Landt, S.G., Yan, K.-K., Cheng, C., Mu, X.J., Khurana, E., Rozowsky, J., Alexander, R., et al. (2012). Architecture of the human regulatory network derived from ENCODE data. Nature *488*, 91-100. 10.1038/nature11245.
187. Horlbeck, M.A., Gilbert, L.A., Villalta, J.E., Adamson, B., Pak, R.A., Chen, Y., Fields, A.P., Park, C.Y., Corn, J.E., Kampmann, M., and Weissman, J.S. (2016). Compact and

highly active next-generation libraries for CRISPR-mediated gene repression and activation. eLife *5*, 914. 10.7554/eLife.19760.

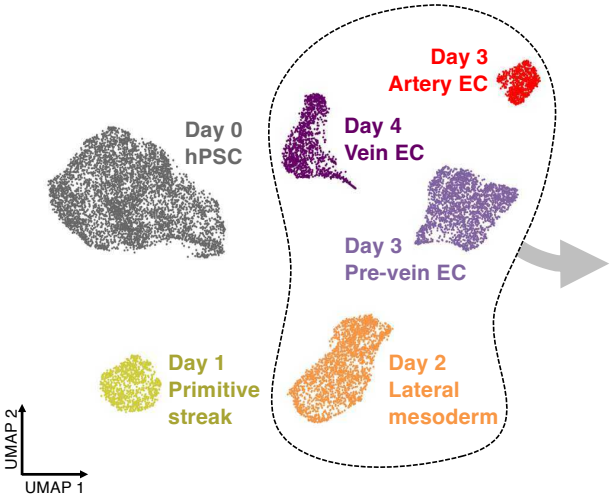
188. Yoshida, A., Kawabata, R., Honda, T., Sakai, K., Ami, Y., Sakaguchi, T., and Irie, T. (2018). A Single Amino Acid Substitution within the Paramyxovirus Sendai Virus Nucleoprotein Is a Critical Determinant for Production of Interferon-Beta-Inducing Copyback-Type Defective Interfering Genomes. Journal of Virology *92*. 10.1128/JVI.02094-17.

Figure 1

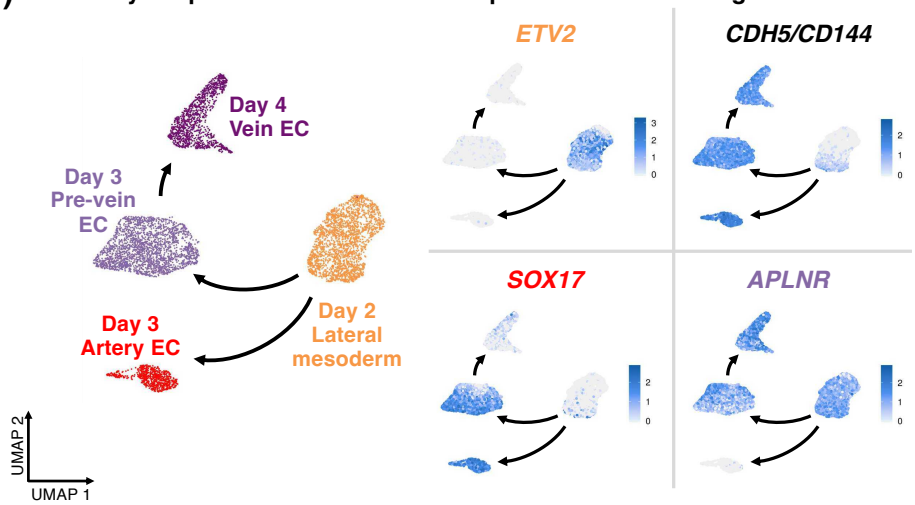
A) Differentiation of hPSCs into artery and vein ECs



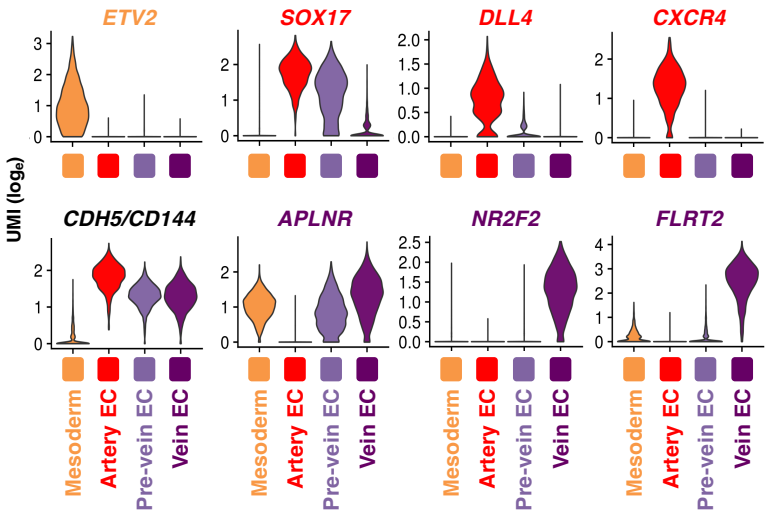
B) scRNA-seq profiling of hPSC differentiation



C) Discovery of "pre-vein" EC intermediate prior to vein EC emergence *in vitro*



D) Co-expression of *SOX17* and *APLNR* in pre-vein ECs (scRNA-seq)



E) Pre-vein ECs co-express SOX17 and *APLNR*

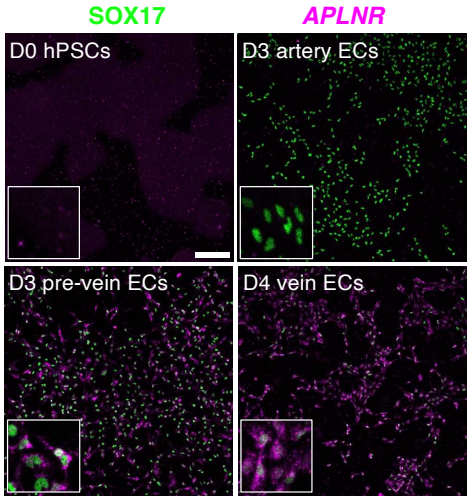
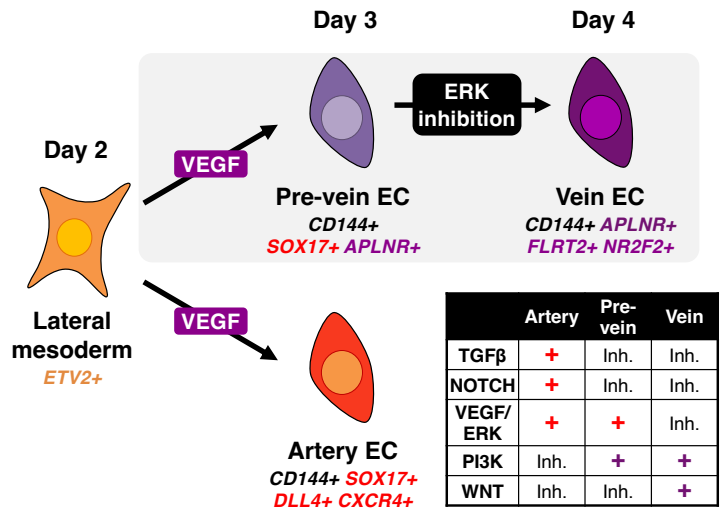


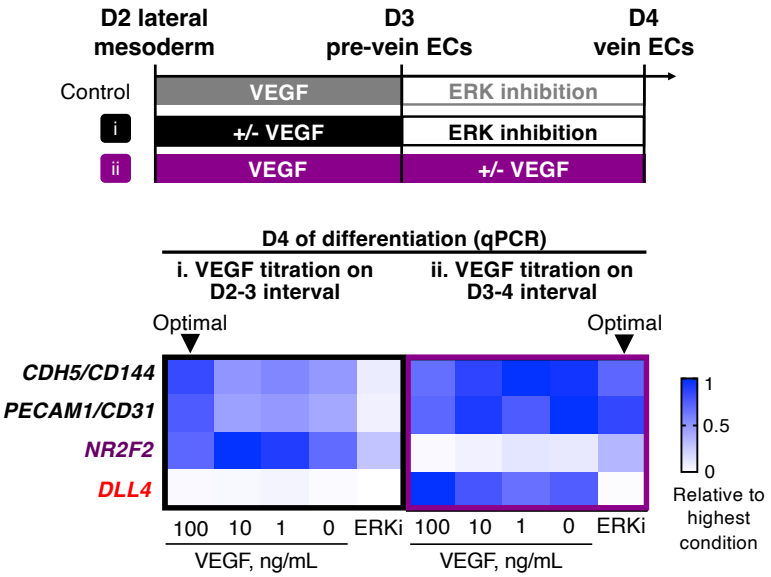
Figure 2

A) Two vein differentiation steps: VEGF activation, then inhibition

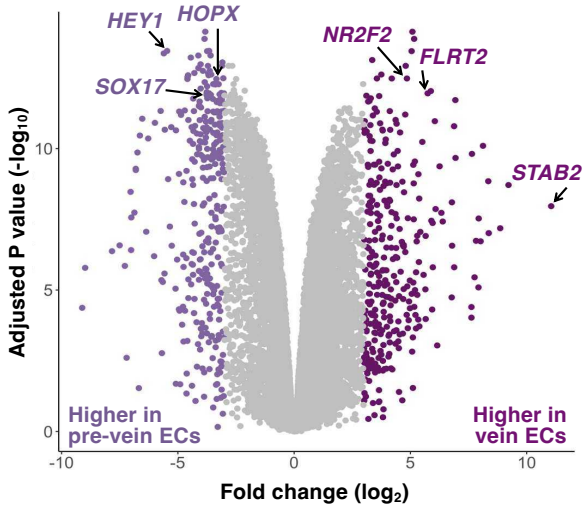


	Artery	Pre-vein	Vein
TGFβ	+	Inh.	Inh.
NOTCH	+	Inh.	Inh.
VEGF/ ERK	+	+	Inh.
PI3K	Inh.	+	+
WNT	Inh.	Inh.	+

B) VEGF activation, followed by inhibition, specifies vein ECs



C) Vein genes activated during pre-vein to vein transition



D) *NR2F2* promoter is bivalently marked in pre-vein ECs

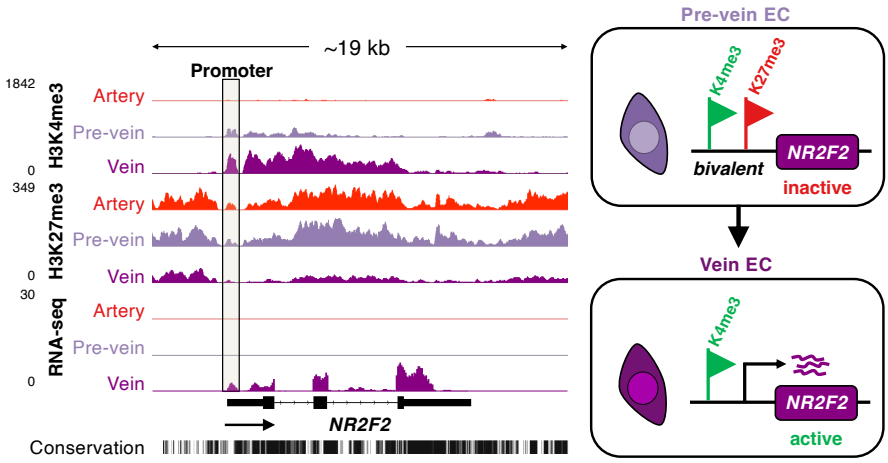
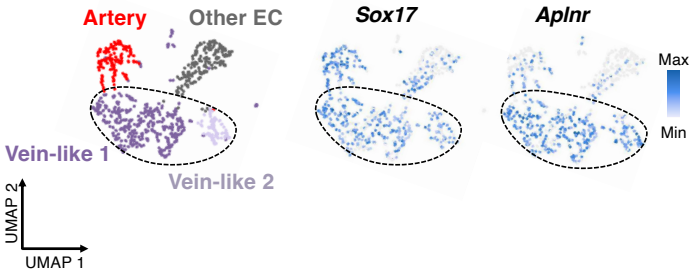
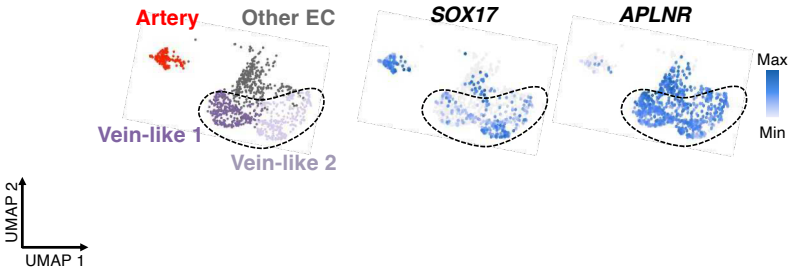


Figure 3

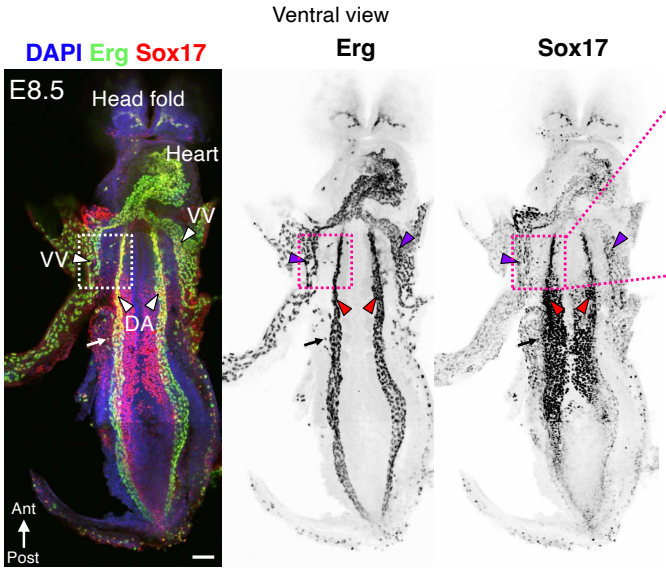
A) E9.5 mouse vasculature (Chen et al.)



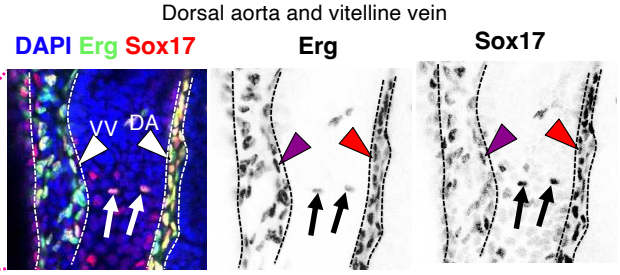
B) CS12 human vasculature (Calvanese et al.)



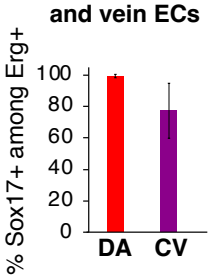
C) Broad Sox17 expression in E8.5 mouse embryo ECs



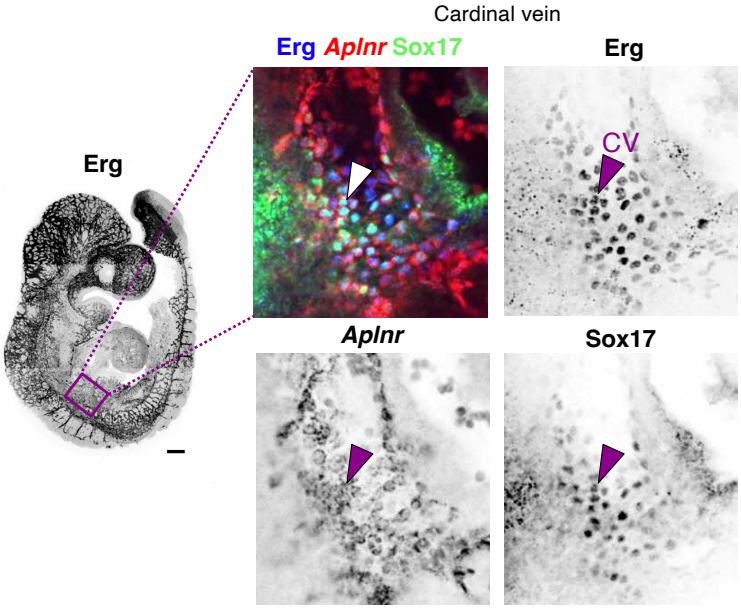
D) Early artery and vein ECs at E8.5 express Sox17



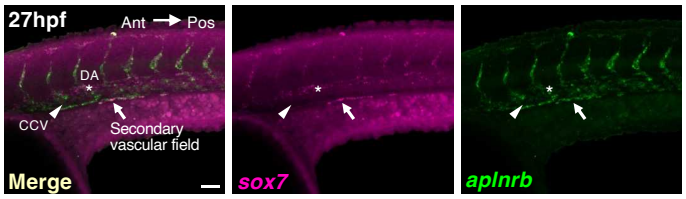
E) Sox17 marks E9.5 artery and vein ECs



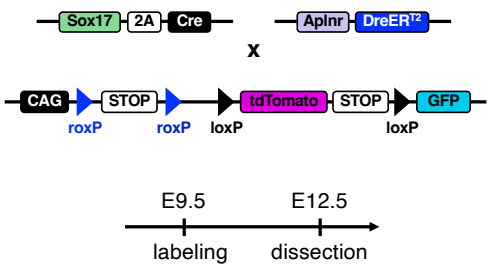
F) E9.5 cardinal vein ECs co-express Sox17 and *Aplnr*

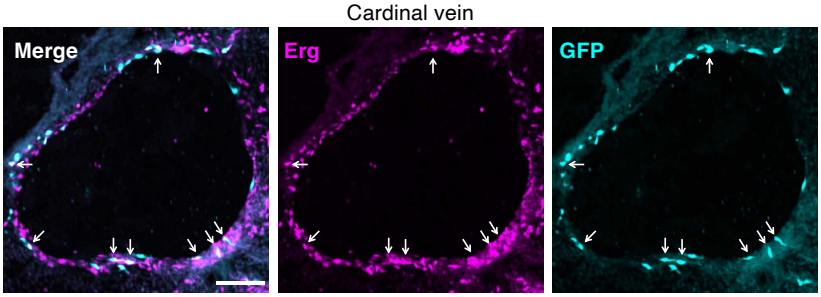


G) *In situ* hybridization (27hpf zebrafish embryo)



H) *Sox17-Cre;Aplnr-DreERT2* intersectional lineage tracing (E12.5 mouse embryo)





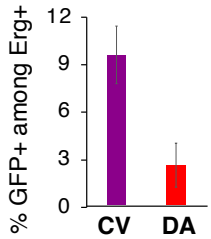
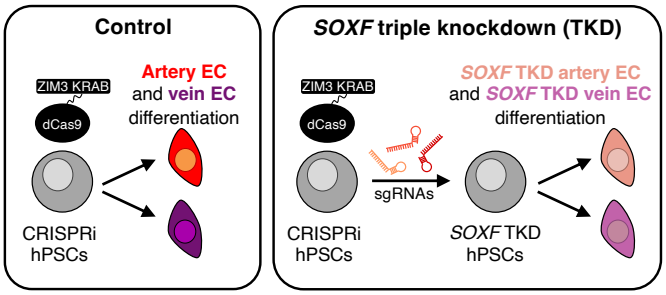
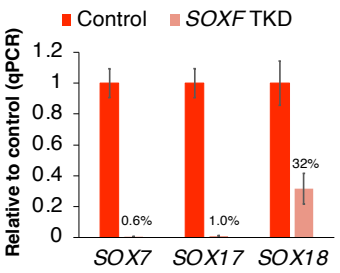


Figure 4

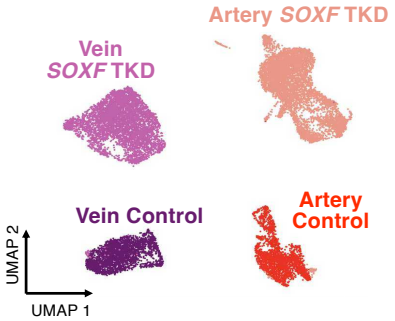
A) Experimental strategy



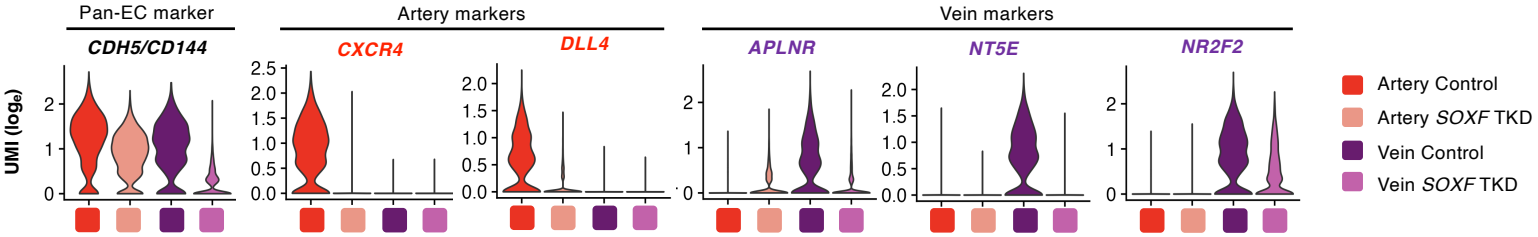
B) *SOXF* TKD in artery ECs



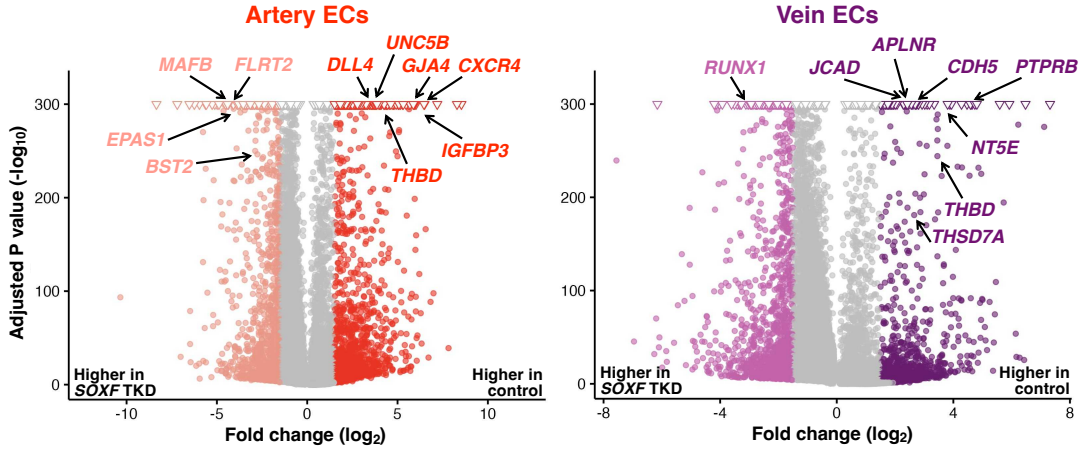
C) scRNA-seq of control and *SOXF* TKD artery and vein ECs



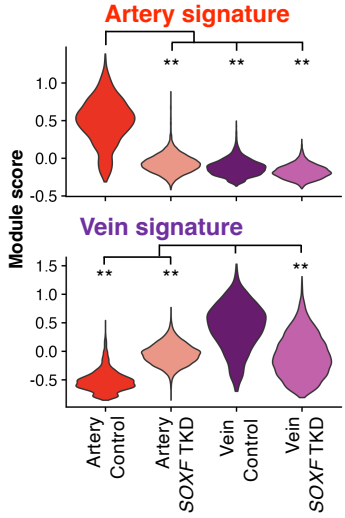
D) *SOXF* transcription factors are required for both artery and vein EC differentiation



E) Gene expression changes upon *SOXF* loss



G) Artery and vein signatures



F) SOX17 binds *APLNR* and *CDH5* in pre-vein ECs

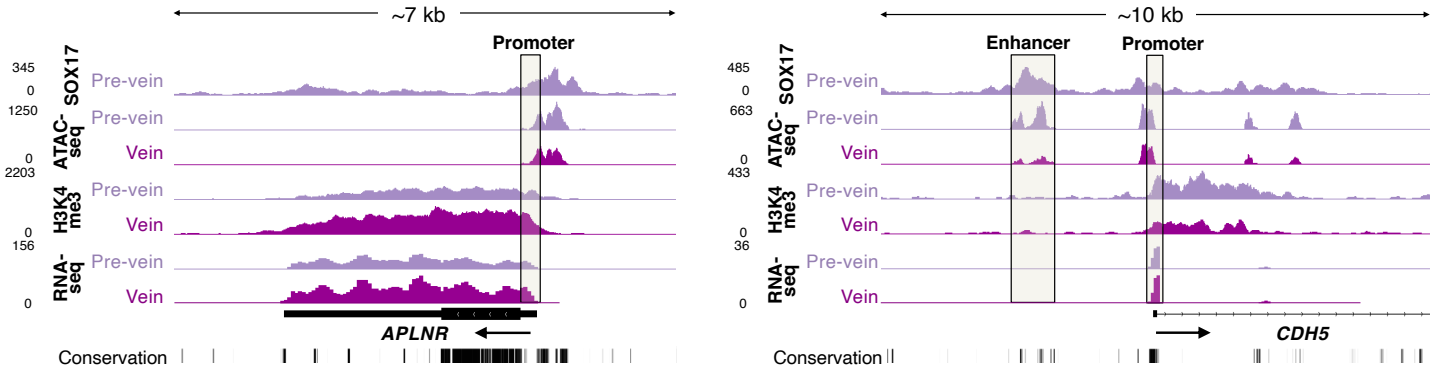
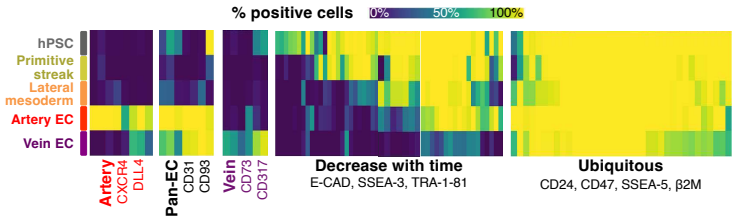
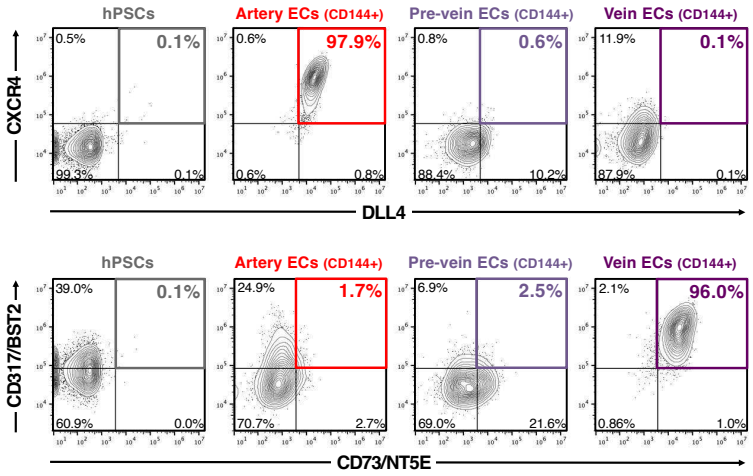


Figure 5

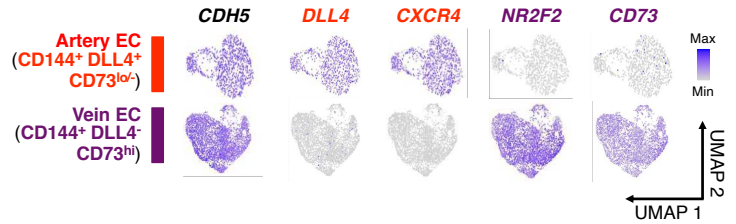
A) LEGENDScreen flow cytometry of cell-surface markers



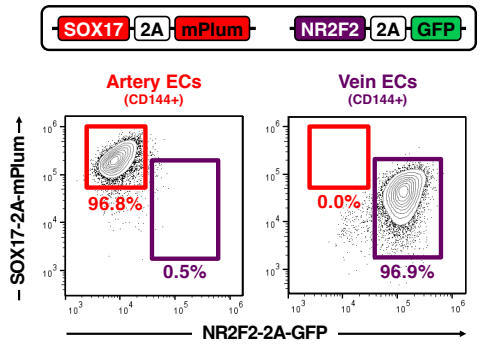
B) Surface markers to track arteriovenous identity



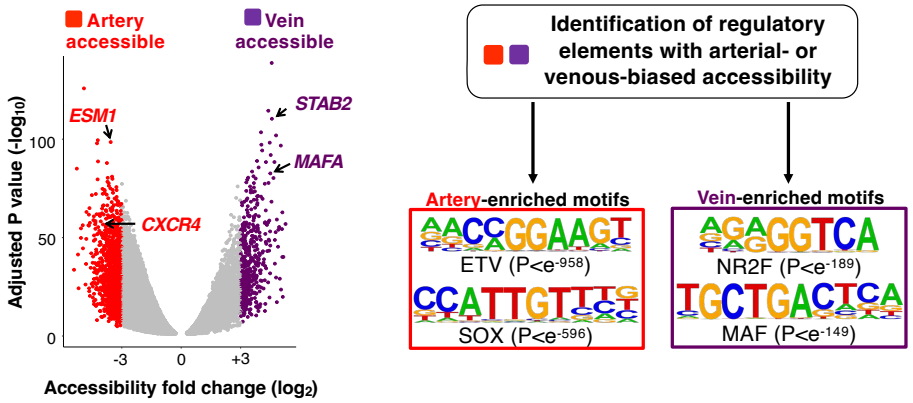
C) scRNA-seq of FACS-purified artery and vein ECs



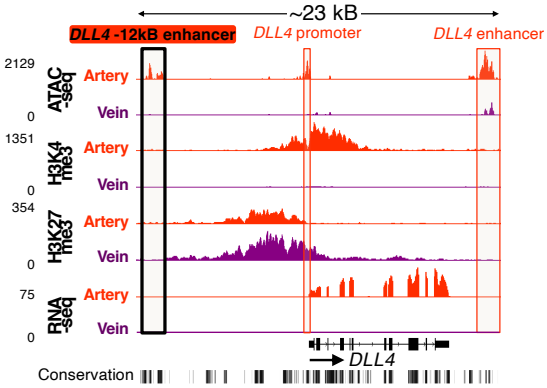
D) Reporter hPSCs track arteriovenous fate



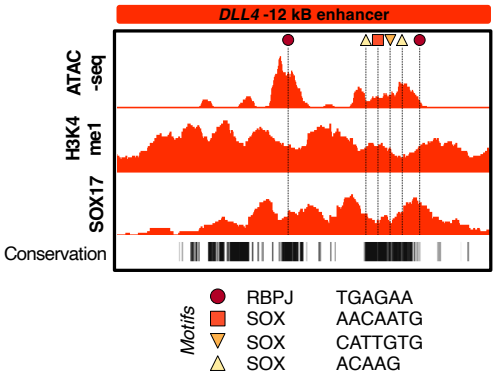
E) OmniATAC-seq of artery and vein ECs



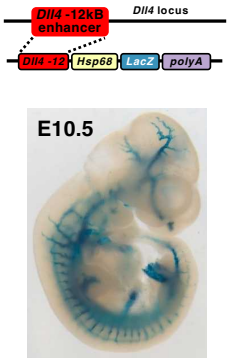
F) Chromatin state of artery marker gene



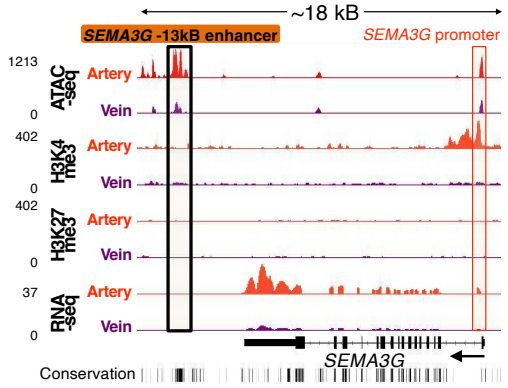
G) *DLL4* enhancer is bound by SOX17



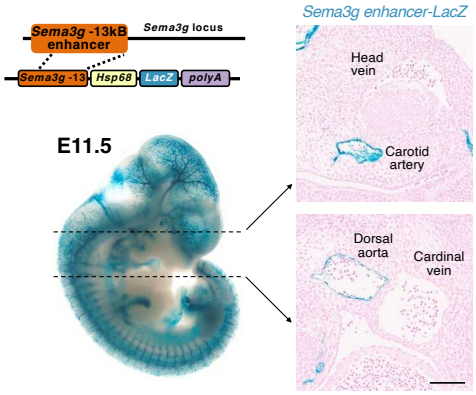
H) *Dll4* enhancer activity *in vivo*



I) Arterial *SEMA3G* enhancer discovery



J) *Sema3g* enhancer activity *in vivo*



K) Chromatin state of vein marker gene

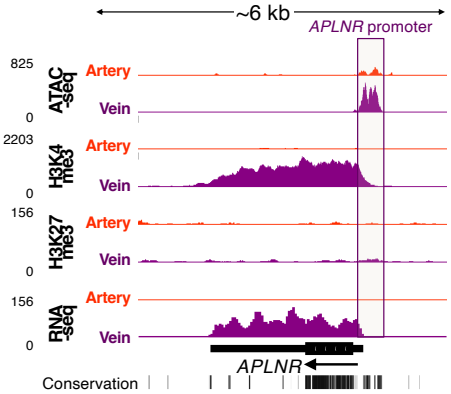
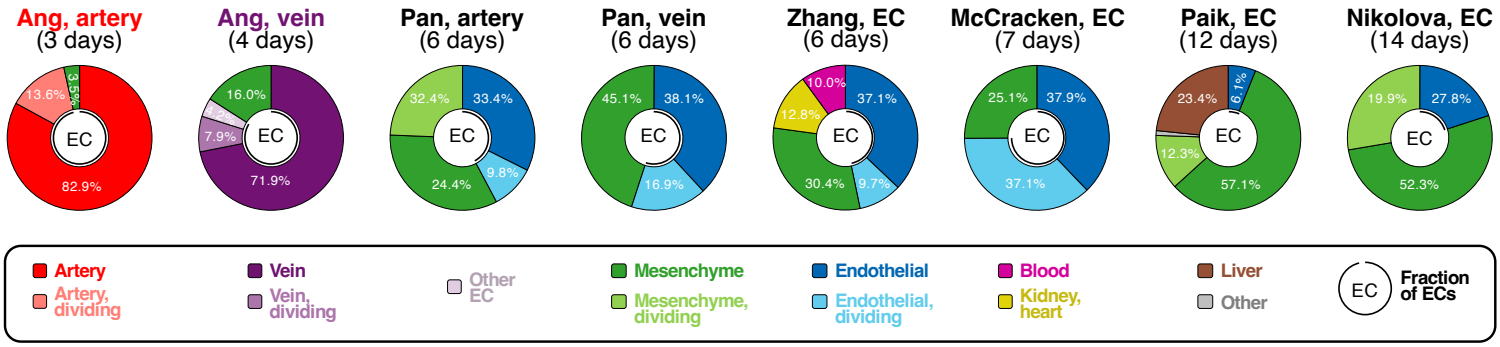
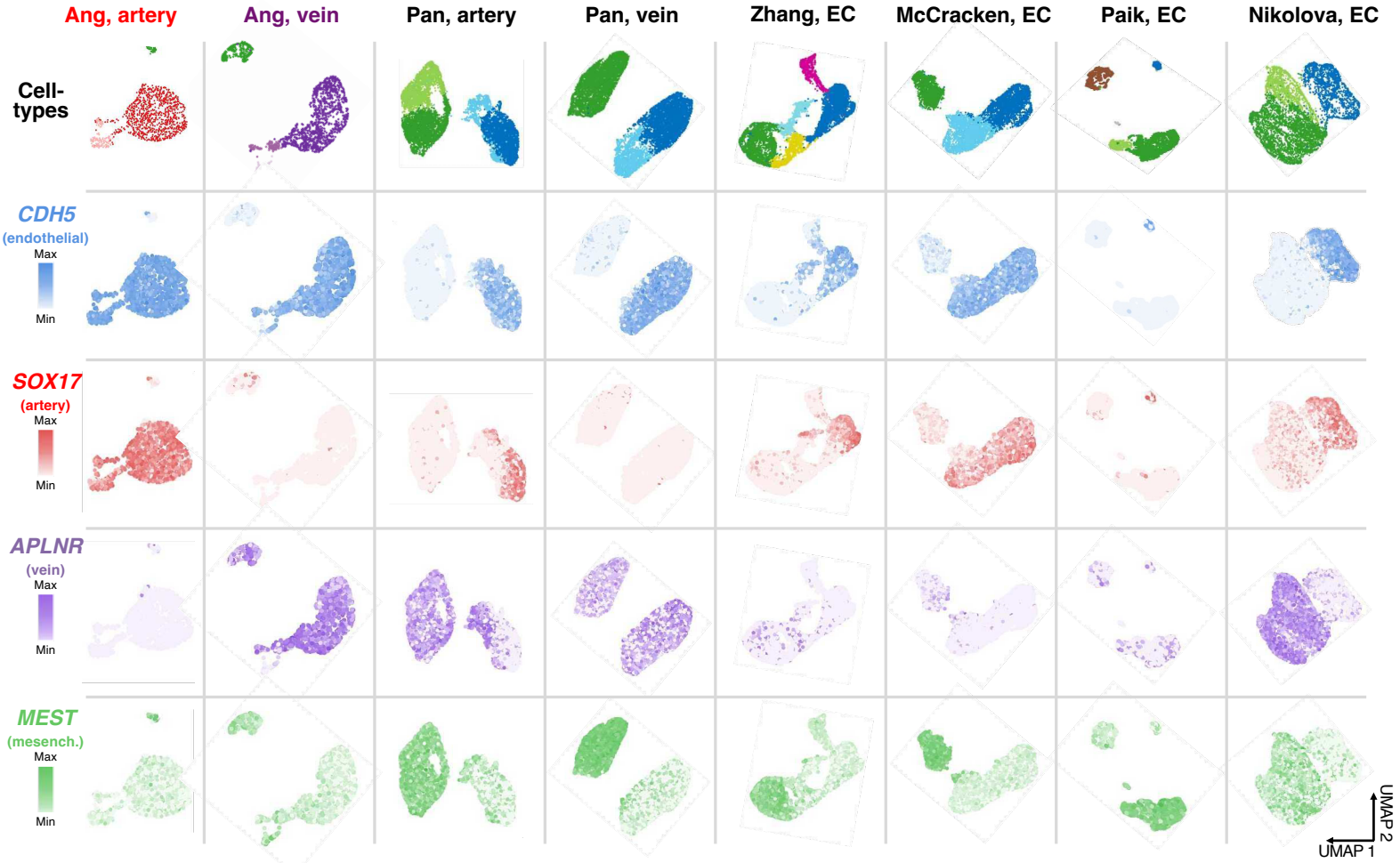


Figure 6

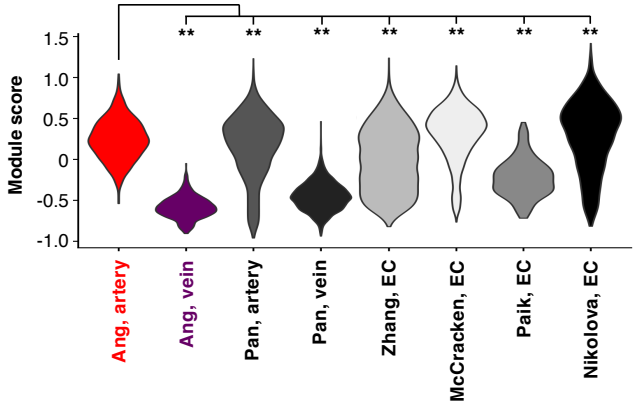
A) Comparison of EC differentiation methods (scRNA-seq)



B) Comparison of EC differentiation methods (scRNA-seq)



C) Artery signature (scRNA-seq)



D) Vein signature (scRNA-seq)

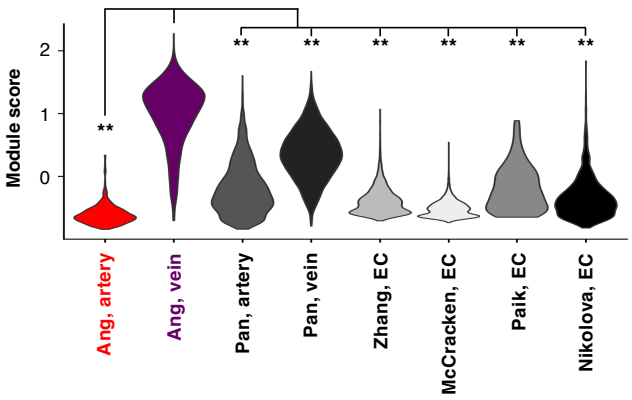
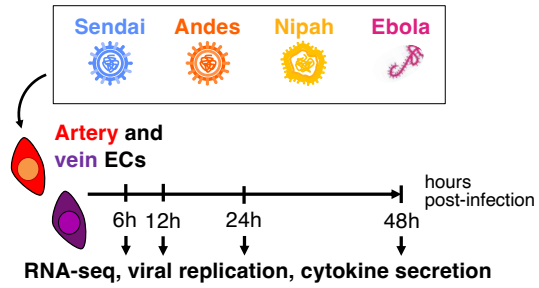
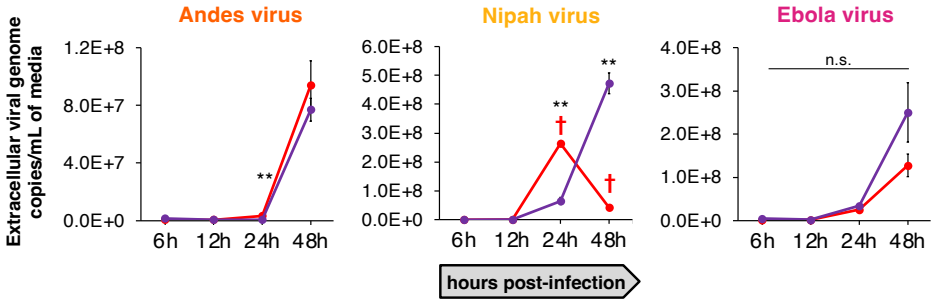


Figure 7

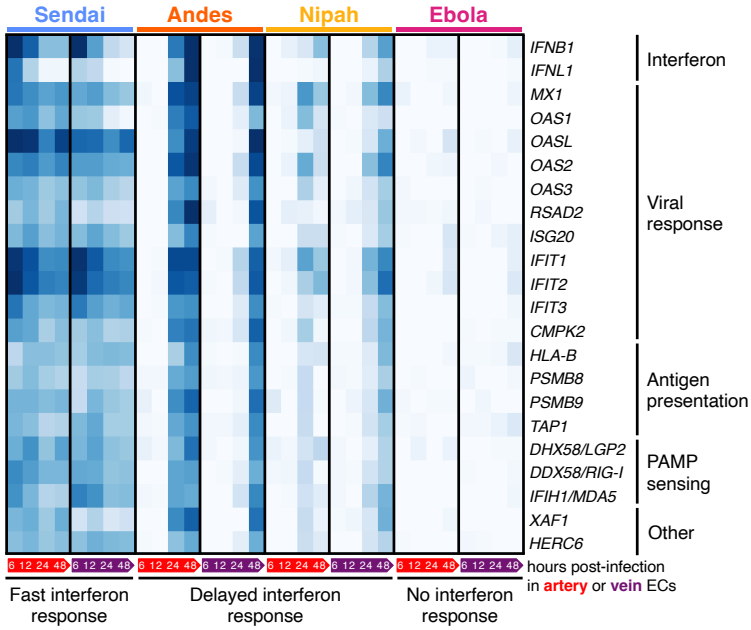
A) Viral infection under BSL4 containment



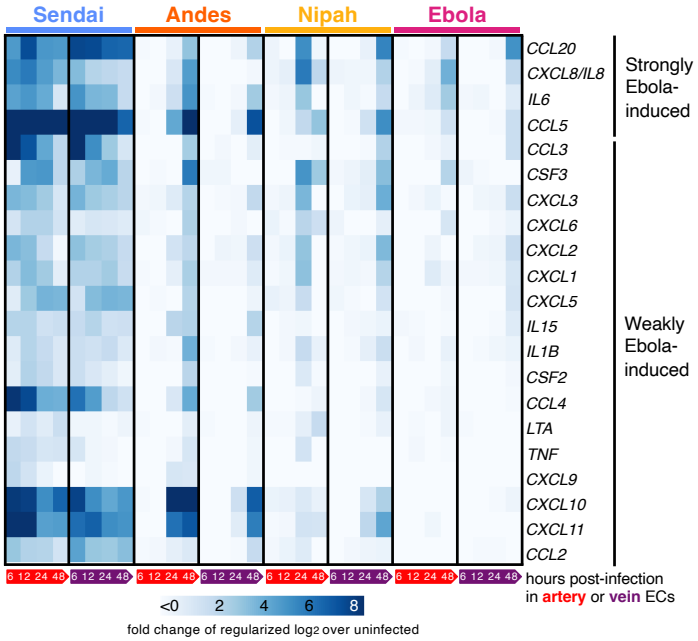
B) Viral replication



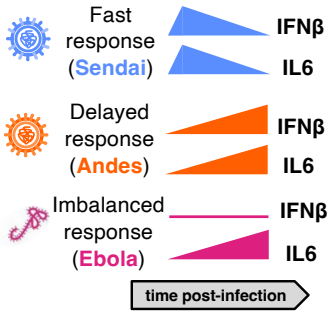
C) Interferon and interferon-stimulated genes



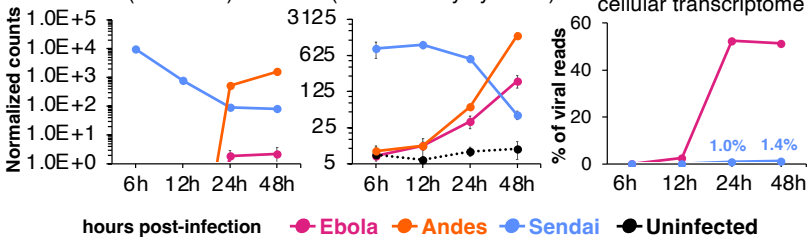
D) Inflammatory cytokines



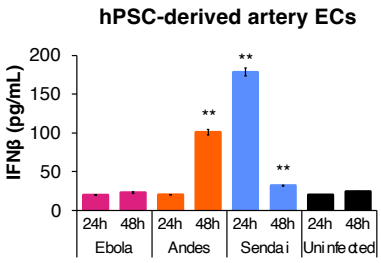
E) Virus-specific effects on ECs



F) Interferon pathway and inflammatory cytokine gene expression



G) Interferon secretion



H) Artery- or vein-specific responses to viral infection

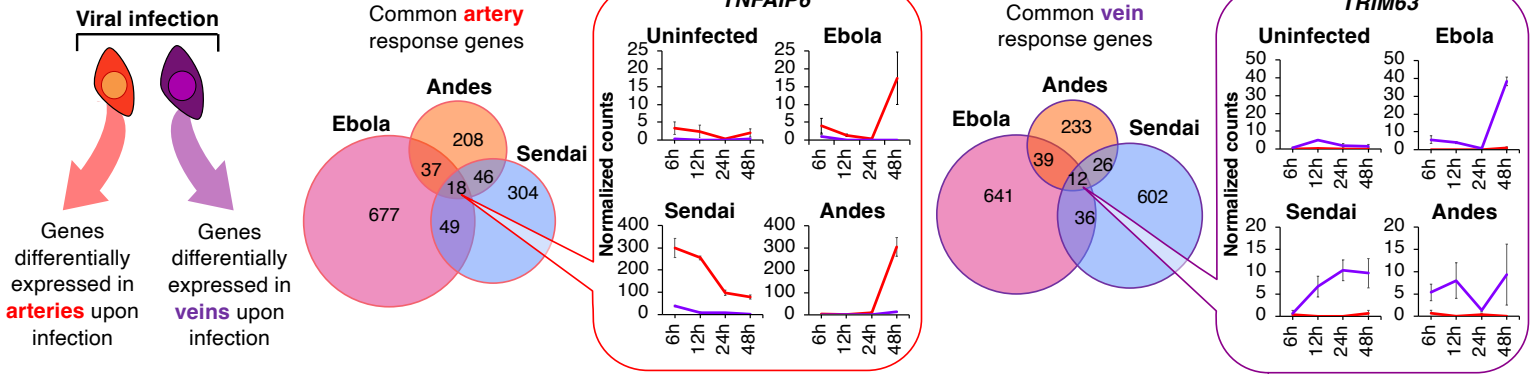


FIGURE LEGENDS

Figure 1: A roadmap for human arteriovenous differentiation reveals pre-vein endothelial cells

A) Summary of the present study. H1 hPSCs were sequentially differentiated[49] into day-1 primitive streak, day-2 lateral mesoderm, day-3 artery ECs, day-3 pre-vein ECs, and day-4 vein ECs. Cells were characterized by scRNA-seq, bulk RNA-seq, OmniATAC-seq, CUT&RUN, and LEGENDScreen.
B) scRNA-seq of day-0 H1 hPSCs, day-1 primitive streak, day-2 lateral mesoderm, CD144+ FACS-purified day-3 artery ECs, CD144+ FACS-purified day-3 pre-vein ECs, and CD144+ FACS-purified day-4 vein ECs.
C) scRNA-seq of day-2 lateral mesoderm, CD144+ FACS-purified day-3 artery ECs, CD144+ FACS-purified day-3 pre-vein ECs, and CD144+ FACS-purified day-4 vein ECs.
D) scRNA-seq of day-2 lateral mesoderm, CD144+ FACS-purified day-3 artery ECs, CD144+ FACS-purified day-3 pre-vein ECs, and CD144+ FACS-purified day-4 vein ECs. Gene expression is depicted in $\log_e$ unique molecular identifier (UMI) counts.
E) Combined immunostaining for SOX17 protein and HCR3 *in situ* hybridization for *APLNR* mRNA in the indicated H1 hPSC-derived cell-types. Scale: 200 μm.

Related to Figure S1

Figure 2: Two separable steps of vein differentiation driven by temporally-dynamic VEGF/ERK activation, followed by inhibition

A) Summary of the present study.
B) As a positive control, H1 hPSCs were sequentially differentiated into day-2 lateral mesoderm, day-3 pre-vein ECs, and day-4 vein ECs (positive control)[49]. Alternatively, during the 24-hour interval of either pre-vein EC induction (day 3, experiment i) or vein EC induction (day 4, experiment ii), either VEGF (0-100 ng/mL) or ERK inhibitor (ERKi; PD0325901, 500 nM) was added, with all other signals kept constant. qPCR was conducted on day 4 of hPSC differentiation, and expression is normalized to the sample with the highest expression.
C) Bulk-population RNA-seq of FACS-purified CD144+ day-3 pre-vein ECs and day 4-vein ECs generated from H1 hPSCs. Differentially expressed genes are colored purple.
D) CUT&RUN profiling of H3K4me3 and H3K27me3 and bulk RNA-seq H1 hPSC-derived day-3 artery ECs, day-3 pre-vein ECs, and day-4 vein ECs. Conservation: Phastcons evolutionary conservation of genome sequence across 46 vertebrate species.

Related to Figure S2

Figure 3: *Sox17*+ *Aplnr*+ endothelial cells exist in the early embryo, and contribute to vein endothelial cells

A) scRNAseq of E9.5 mouse embryo ECs. scRNAseq data were obtained from Chen et al., 2024[65].
B) scRNAseq of Carnegie Stage 12 (CS12) human embryo ECs. scRNAseq data were obtained from Calvanese et al., 2022[66].

C) Whole-mount Erg and Sox17 immunostaining of E8.5 mouse embryo. Ant: anterior. Post: posterior. DA: dorsal aorta (arrowhead). VV: vitelline vein (arrowhead). Arrow: single migrating Erg+ Sox17+ ECs. Scale: 100 μm.
D) Whole-mount Erg and Sox17 immunostaining of E8.5 mouse embryo.
E) Percentage of Erg+ ECs in the E9.5 dorsal aorta (DA) or cardinal vein (CV) that express Sox17. Error bars: SD.
F) Left: whole-mount Erg immunostaining of E9.5 mouse embryo. Right: E9.5 mouse embryo sectioned and immunostained for Sox17 and Erg proteins, alongside *in situ* staining for *Aplnr* mRNA. Arrow: cardinal vein. Scale: 200 μm.
G) *sox7* and *aplnrb* staining of 27-hour post fertilization zebrafish embryos by HCR3 *in situ* hybridization. Lateral view of the mid-trunk region. Scale: 50 μm.
H) 4OHT was administered *in utero* to E9.5 *Sox17-Cre*; *Aplnr-DreER*; *RLTG* mouse embryos[67,71,72], which were then isolated at E12.5, cryosectioned, and immunostained for Erg and GFP proteins. Arrows: GFP+ vein ECs. Scale: 100 μm. Error bars: SD.

Related to Figure S3

Figure 4: SOXF transcription factors are required for human vein EC specification *in vitro*

A) H1 CRISPRi-expressing hPSCs were transduced with sgRNAs targeting *SOX7*, *SOX17*, and *SOX18* (*SOXF* triple knockdown [TKD]), and then subsequently differentiated into artery and vein ECs.
B) qPCR of day-3 artery ECs generated from H1 control vs. *SOXF* TKD hPSCs. Gene expression normalized to control artery ECs. Percentages indicate remaining gene expression in *SOXF* TKD artery ECs, relative to controls. Statistics: unpaired t-test. Error bars: standard error of the mean (SEM).
C) scRNA-seq of day-3 artery ECs and day-4 vein ECs generated from H1 control vs. *SOXF* TKD hPSCs. Mesenchymal cells were computationally excluded.
D) scRNA-seq of day-3 artery ECs and day-4 vein ECs generated from H1 control vs. *SOXF* TKD hPSCs.
E) scRNA-seq of day-3 artery ECs and day-4 vein ECs generated from H1 control vs. *SOXF* TKD hPSCs. Differentially expressed genes between control vs. *SOXF* TKD ECs are colored.
F) OmniATAC-seq, CUT&RUN, and bulk RNA-seq of H1 hPSC-derived day-3 pre-vein ECs and day-4 vein ECs. Conservation: Phastcons evolutionary conservation of genome sequence across 46 vertebrate species.
G) scRNA-seq of day-3 artery ECs and day-4 vein ECs generated from H1 control vs. *SOXF* TKD hPSCs. Transcriptional module scores[136] computed from known panels of arterial and venous marker genes *in vivo*[21] are shown. Statistics: Wilcoxon rank sum test. **P<0.01.

Related to Figure S4.

Figure 5: Transcriptional, chromatin, and surface-marker hallmarks of arteriovenous identity

A) LEGENDScreen high-throughput flow cytometry of 332 cell-surface markers in day-0 hPSCs, day-1 primitive streak, day-2 lateral mesoderm, day-3 artery ECs, and day-4 vein ECs. Only strongly positive markers are shown. Day-3 artery and day-4 vein

populations were pre-gated on the CD144+ EC subset before depicting marker expression.

B) Flow cytometry of H1 CRISPRi-expressing hPSCs, day-3 artery ECs, day-3 pre-vein ECs, and day-4 vein ECs. Day 3-4 populations were pre-gated on the CD144+ EC subset before depicting marker expression.
C) scRNAseq of FACS-purified CD144+ DLL4+ CD73lo/- day-3 artery ECs and CD144+ DLL4- CD73hi day-4 vein ECs generated from H1 hPSCs.
D) Flow cytometry of H1 *SOX17-2A-mPlum; NR2F2-2A-GFP* hPSCs differentiated into day-3 artery ECs or day-5 vein ECs. Day 3-5 populations were pre-gated on the CD144+ EC subset before depicting marker expression.
E) OmniATAC-seq of H1 hPSC-derived day-3 artery ECs and day-4 vein ECs. Left: genetic loci preferentially accessible in either artery or vein ECs are colored. Right: transcription factor motifs enriched in artery- or vein-accessible chromatin elements.
F) OmniATAC-seq, CUT&RUN, and bulk RNA-seq of H1 hPSC-derived day-3 artery ECs and day-4 vein ECs. Conservation: Phastcons evolutionary conservation of genome sequence across 46 vertebrate species.
G) OmniATAC-seq and CUT&RUN analysis of the *DLL4* -12kb enhancer[79] in H1 hPSC-derived day-3 artery ECs, with transcription factor motifs labeled.
H) LacZ staining of E10.5 mouse embryo bearing a *Dll4* -12 kB enhancer transgene driving *LacZ* expression[79].
I) OmniATAC-seq, CUT&RUN, and bulk RNA-seq of H1 hPSC-derived day-3 artery ECs and day-4 vein ECs. Conservation: Phastcons evolutionary conservation of genome sequence across 46 vertebrate species.
J) LacZ staining of E11.5 mouse embryo bearing a *Sema3g* -13 kB enhancer transgene driving *LacZ* expression (VISTA Enhancer Browser ID hs2179)[82,83]. Scale: 100 μm.
K) OmniATAC-seq, CUT&RUN, and bulk RNA-seq of H1 hPSC-derived day-3 artery ECs and day-4 vein ECs. Conservation: Phastcons evolutionary conservation of genome sequence across 46 vertebrate species.

Related to Figure S5.

Figure 6: Comparison of methods to differentiate hPSCs into ECs

A) scRNA-seq of hPSCs differentiated into ECs using various protocols, with the number of days of differentiation indicated. scRNA-seq datasets were subclustered to identify population heterogeneity. Clusters were annotated by marker expression. The proportion of ECs is indicated. scRNA-seq data were obtained from this study, Ang et al., 2022[49], Pan et al., 2024[60], McCracken et al., 2019[87], Paik et al., 2018[85], and Nikolova et al., 2025[61].
B) scRNA-seq of differentiated hPSC populations described in **Fig. 6A**, depicting expression of pan-endothelial, arterial, venous, and mesenchymal marker genes. Top row: clusters are annotated and colored as described in **Fig. 6A**.
C) scRNA-seq of differentiated hPSC populations described in **Fig. 6A**, and ECs were computationally isolated. An expression module score[136] of arterial markers defined *in vivo* by Hou et al., 2022[21] is shown. Statistics: Wilcoxon rank sum test. **P<0.01.
D) scRNA-seq of differentiated hPSC populations described in **Fig. 6A**, and ECs were computationally isolated. An expression module score[136] of venous markers defined *in vivo* by Hou et al., 2022[21] is shown. Statistics: Wilcoxon rank sum test. **P<0.01.

Related to Figures S6 and S7.

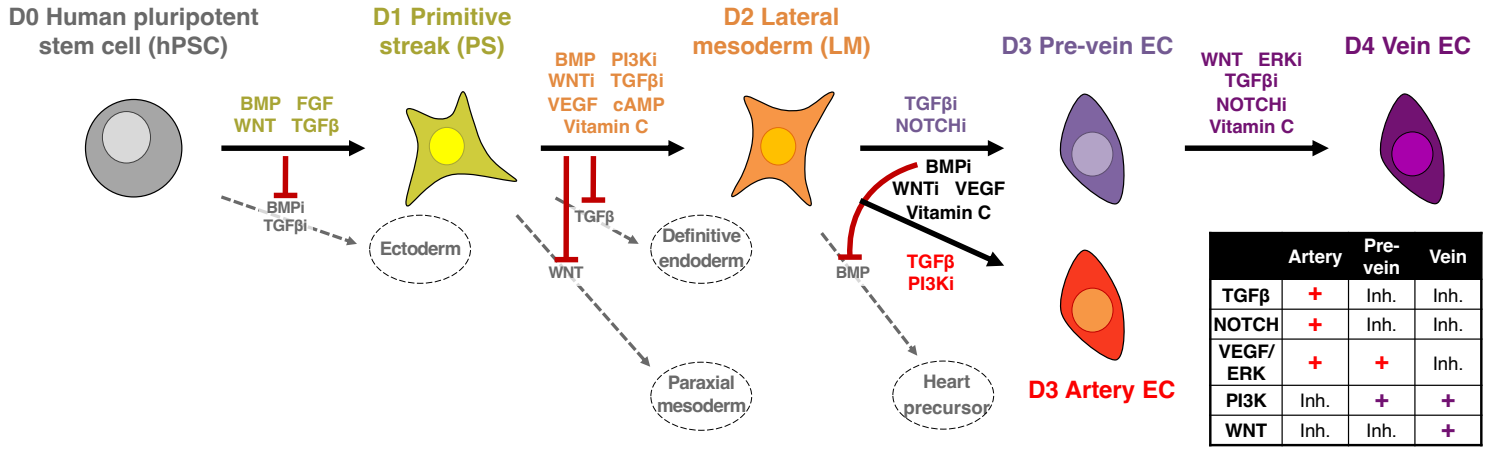
Figure 7: Ebola, Nipah, and Andes viruses each exert different effects on hPSC-derived ECs

A) Experimental summary. h: hour.
B) Replication of Andes, Nipah, and Ebola viruses in hPSC-derived artery and vein ECs, as assayed by qPCR for viral genomes in the culture media. Statistics: unpaired t-test. Error bars: SEM. **P<0.01. n.s.: not significant. †: extensive cell death.
C) Bulk RNA-seq of interferon and interferon-stimulated gene expression in hPSC-derived artery and vein ECs 6, 12, 24, and 48 hours after infection with Sendai, Andes, Nipah, and Ebola viruses. Fold change relative to uninfected cells is depicted.
D) Bulk RNA-seq of inflammatory cytokine gene expression in hPSC-derived artery and vein ECs 6, 12, 24, and 48 hours after infection with Sendai, Andes, Nipah, and Ebola viruses. Fold change relative to uninfected cells is depicted.
E) Summary of the present study.
F) Bulk-RNA-seq of hPSC-derived artery ECs 6, 12, 24, and 48 hours after infection with Ebola, Andes, and Sendai viruses, or left uninfected (mock control). Error bars: SEM.
G) IFNβ protein secretion by hPSC-derived artery ECs after 24 or 48 hours of infection by Ebola, Andes, and Sendai viruses, or left uninfected (mock control), as measured by ELISA. Statistics: unpaired t-test. **P<0.01.
H) Bulk-RNA-seq of hPSC-derived artery ECs 6, 12, 24, and 48 hours after infection with Ebola, Andes, and Sendai viruses, or left uninfected (mock control). Error bars: SEM.

Related to Figures S7 and S8.

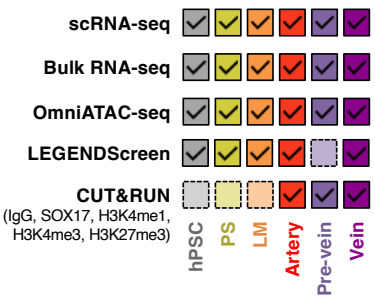
Figure S1

A) Differentiation of hPSCs into artery and vein ECs

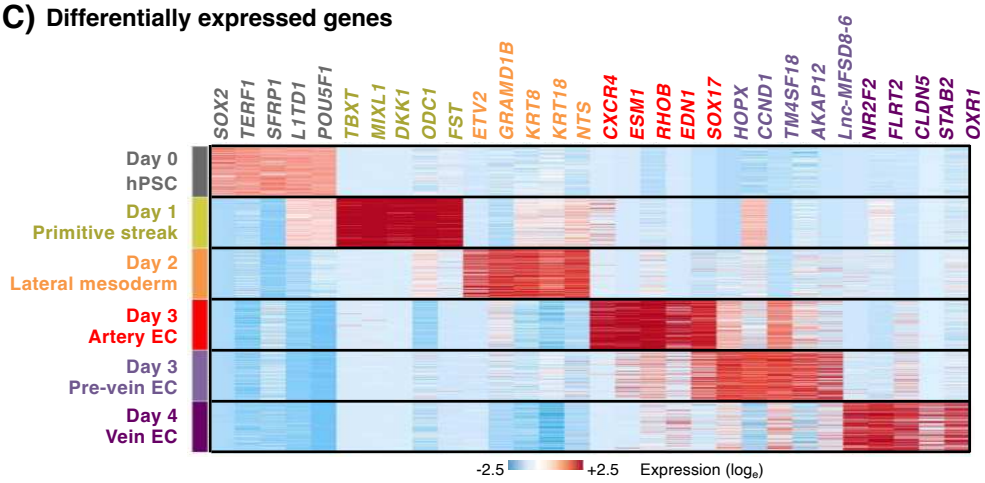


	Artery	Pre-vein	Vein
TGFβ	+	Inh.	Inh.
NOTCH	+	Inh.	Inh.
VEGF/ ERK	+	+	Inh.
PI3K	Inh.	+	+
WNT	Inh.	Inh.	+

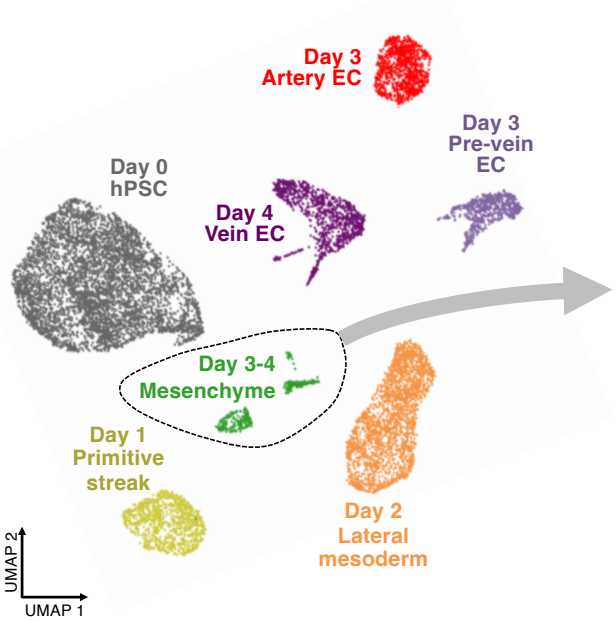
B) Characterization of differentiated cell-types



C) Differentially expressed genes



D) scRNA-seq of unsorted cell-types



E) Diverse mesenchymal cells arising on days 3-4 of differentiation

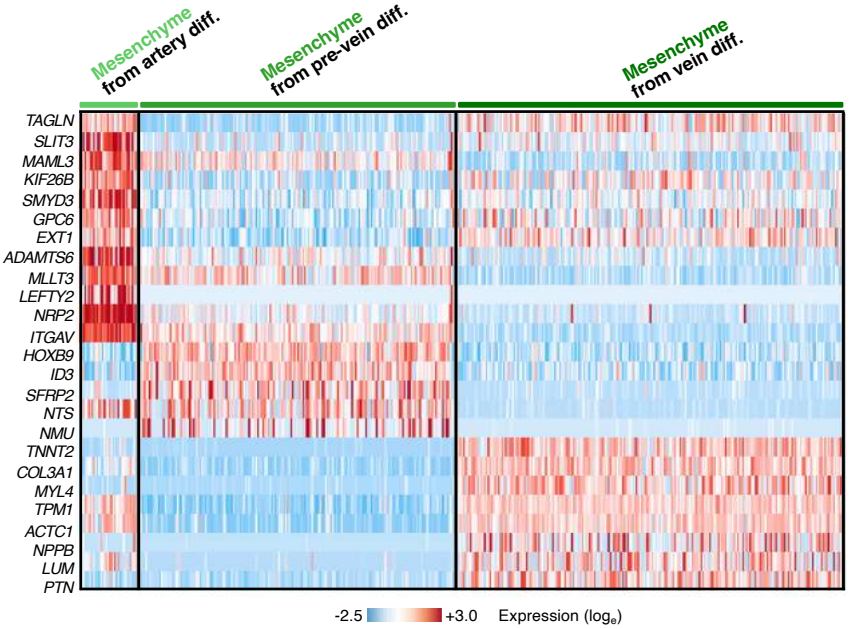
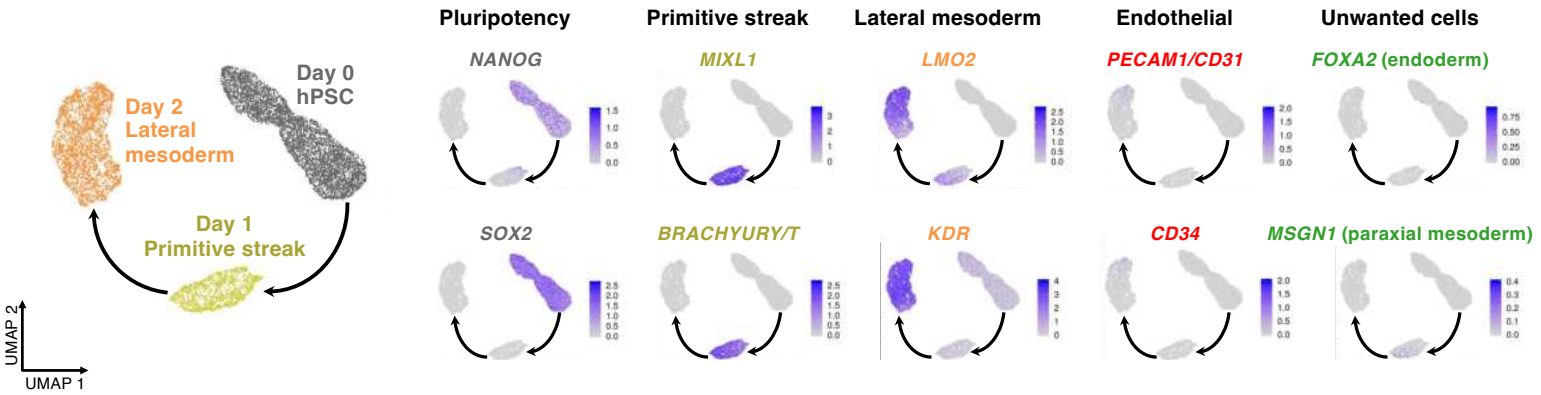
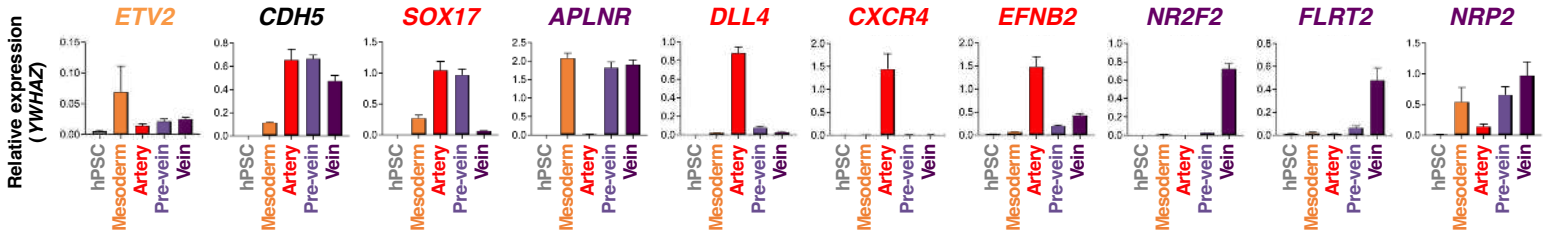


Figure S2

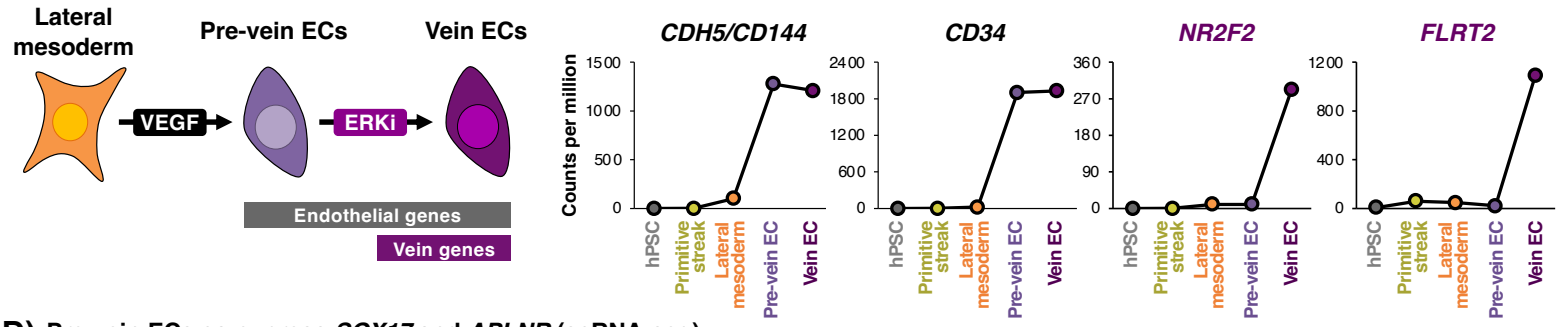
A) Stepwise differentiation of primitive streak and lateral mesoderm



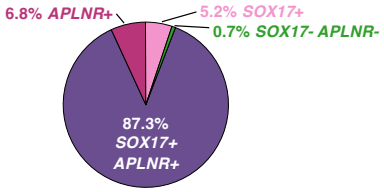
B) qPCR validation of scRNA-seq



C) Bulk RNA-seq of vein differentiation



D) Pre-vein ECs co-express *SOX17* and *APLNR* (scRNA-seq)



E) *NT5E* and *MAF* promoters are bivalently marked in pre-vein ECs

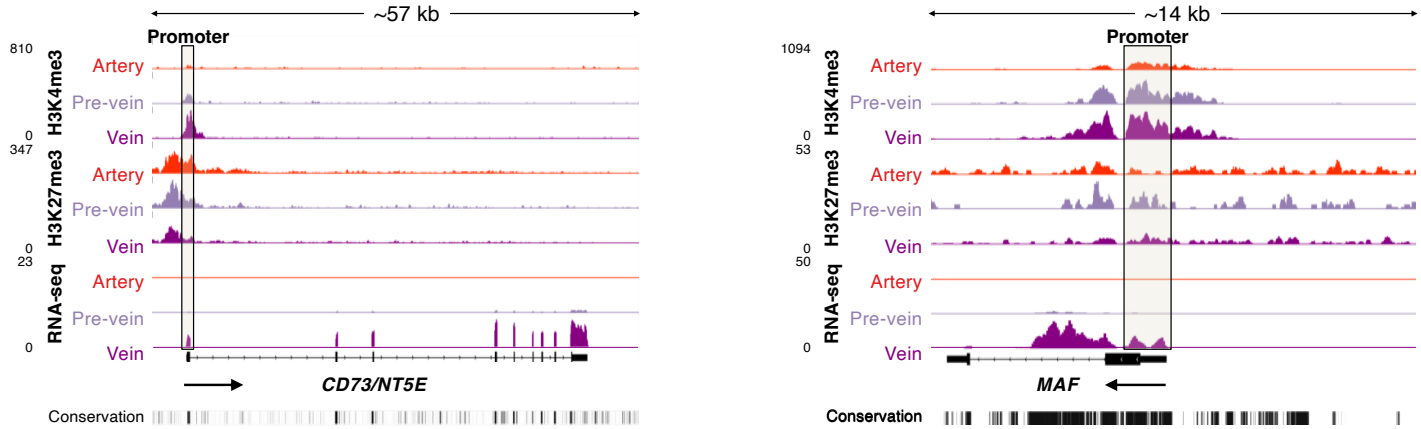
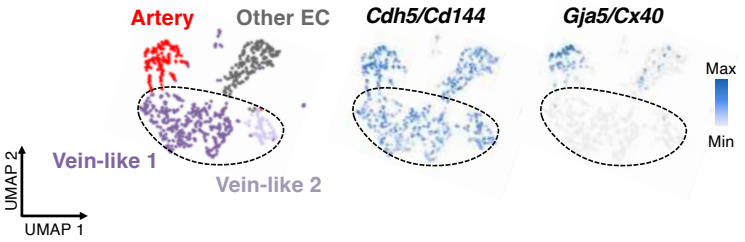
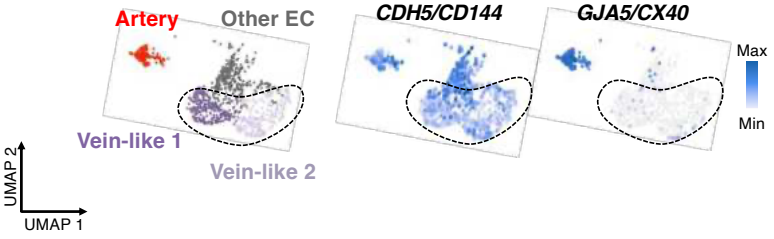


Figure S3

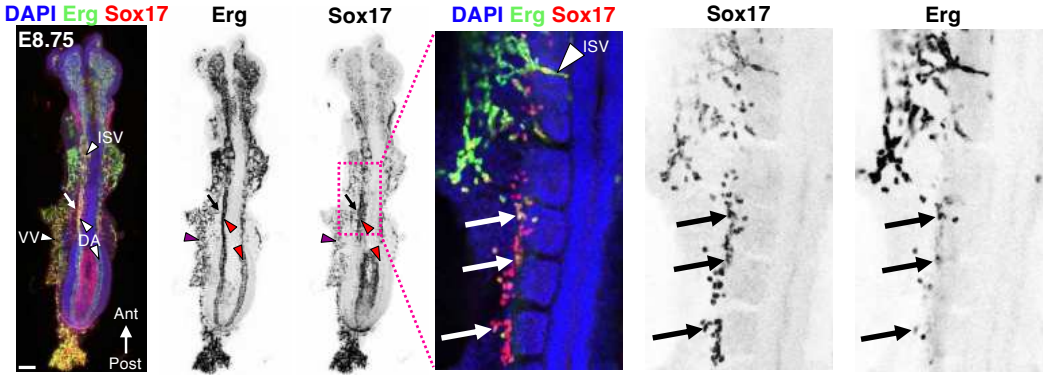
A) E9.5 mouse vasculature (Chen et al.)



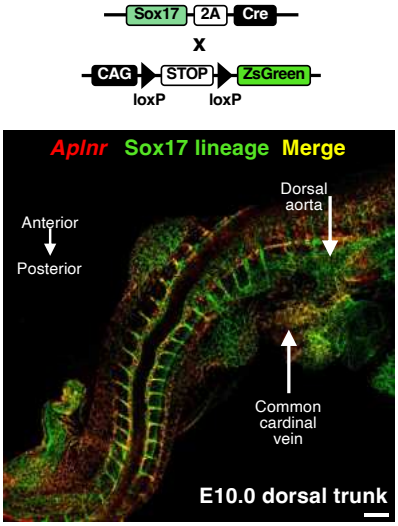
B) CS12 human vasculature (Calvanese et al.)



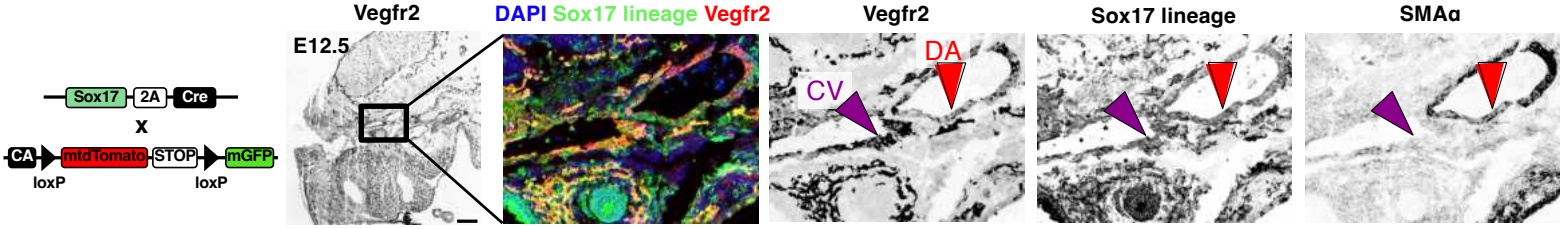
C) Widespread Sox17 expression in E8.75 mouse embryo ECs



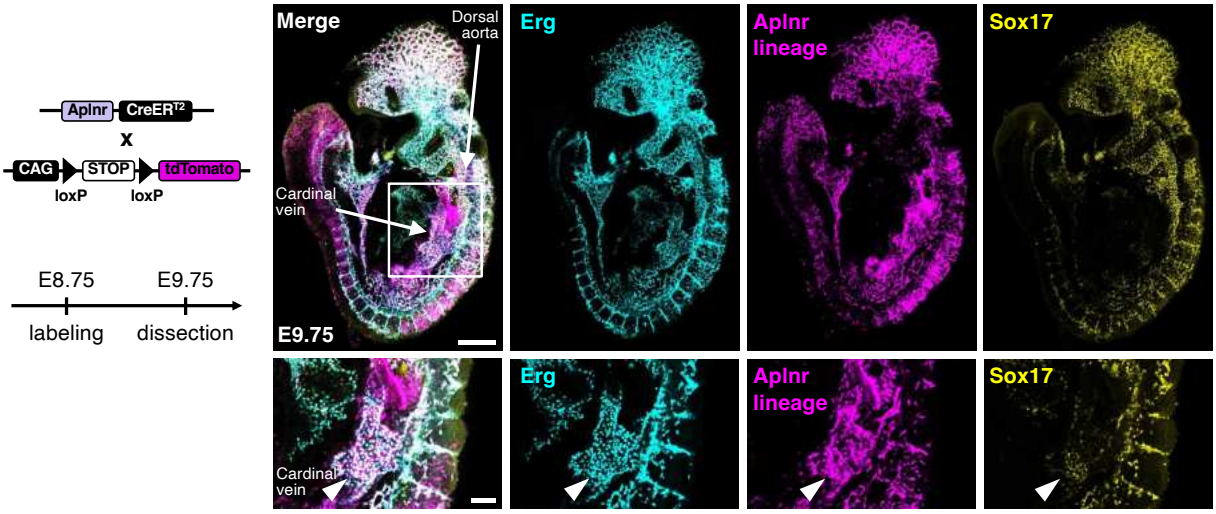
D) *Sox17-Cre* labels artery and vein ECs



E) *Sox17-Cre* lineage tracing of E12.5 mouse embryo



F) *Aplnr-CreER* tracing (E9.75 mouse embryo)



G) Co-expression of Sox17 and *Aplnr-CreER* in E9.5 mouse embryos

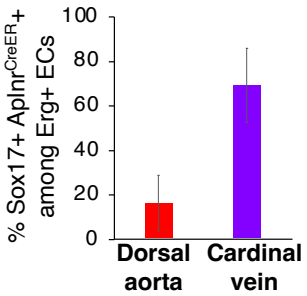
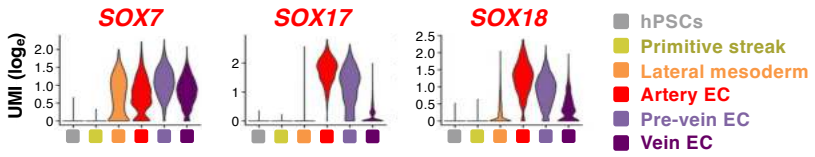
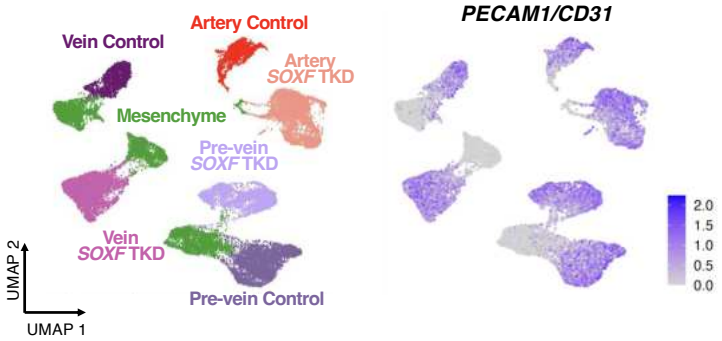


Figure S4

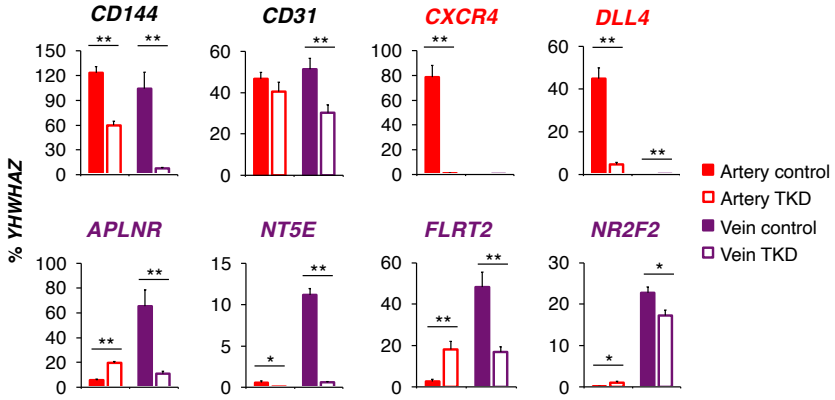
A) *SOXF* expression during EC differentiation (scRNA-seq)



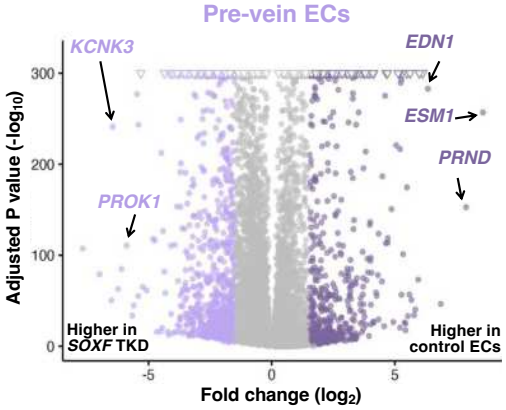
B) Control and *SOXF* TKD differentiated cells (scRNA-seq)



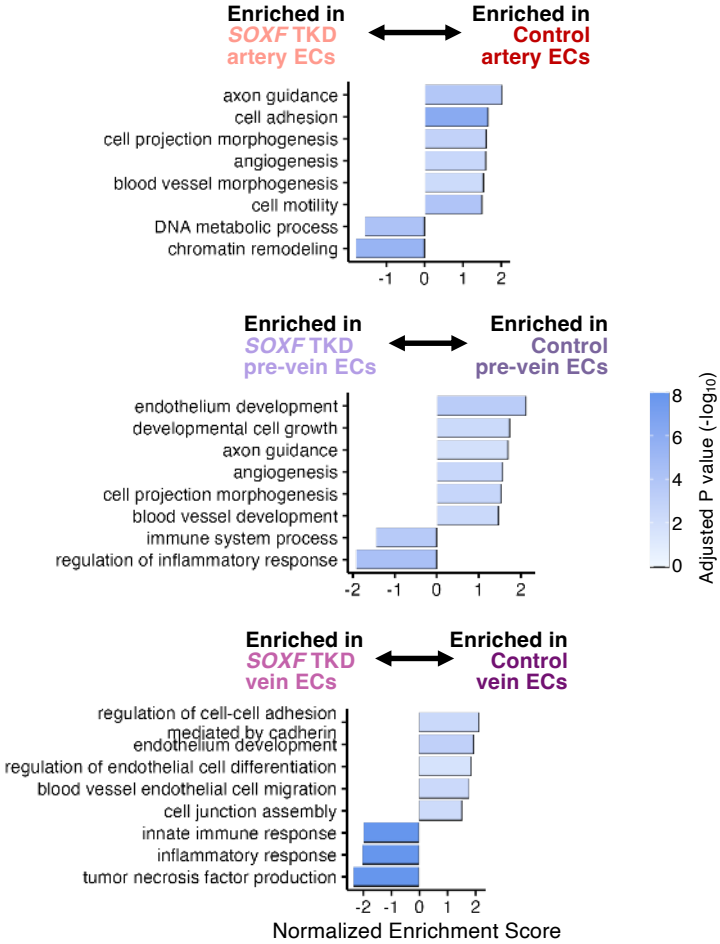
C) Gene expression changes in *SOXF* TKD artery and vein ECs (qPCR)



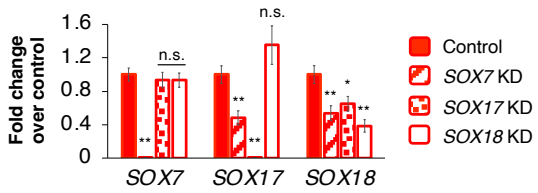
D) Gene expression changes upon *SOXF* loss in pre-vein ECs



E) Gene ontology analysis of *SOXF*-deficient ECs (scRNA-seq)



F) Interdependency of *SOXF* factors in artery ECs (qPCR)



G) Single knockdown of *SOX7, SOX17,* and *SOX18* (qPCR)

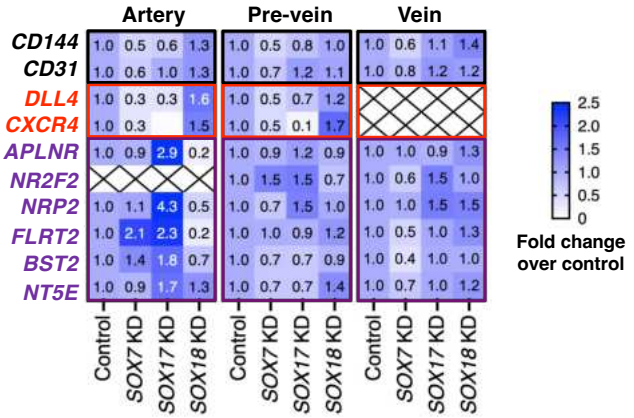
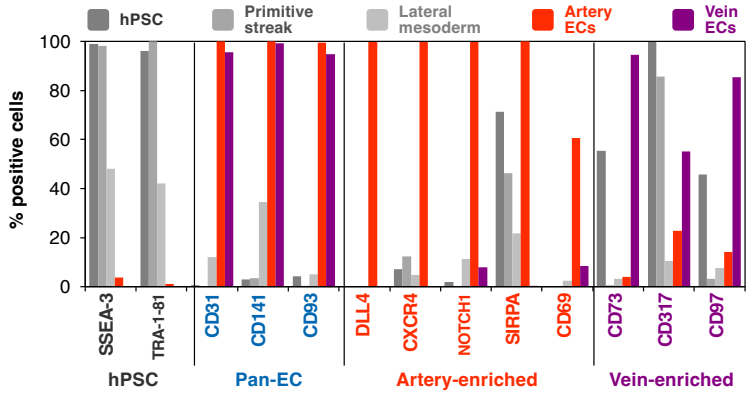
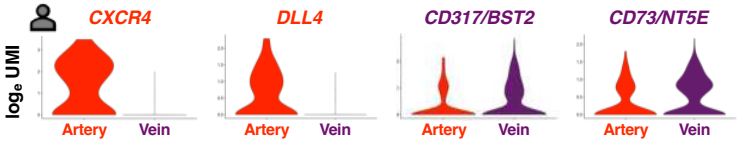


Figure S5

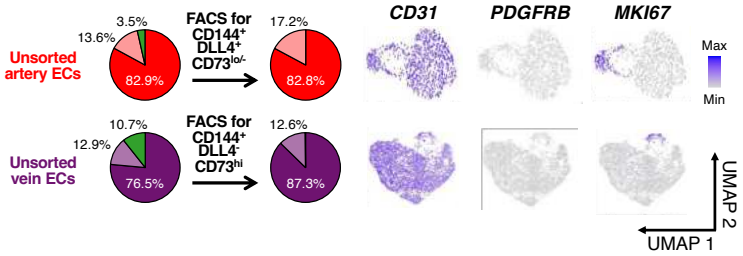
A) High-throughput screening of cell-surface markers



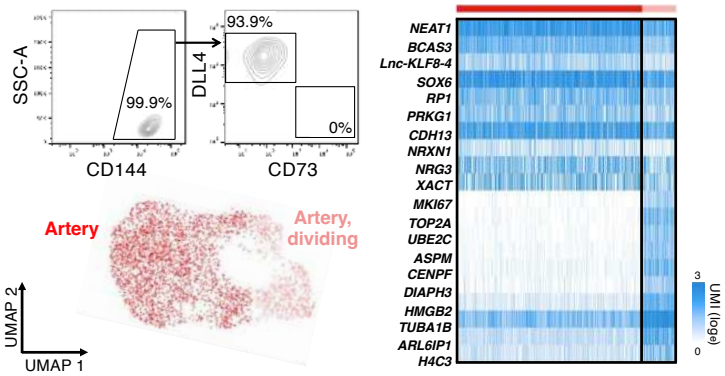
B) CS17 human fetal ECs (scRNA-seq; Calvanese et al.)



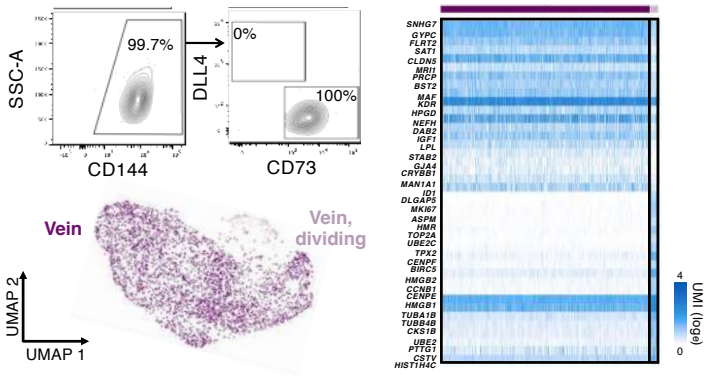
C) scRNA-seq of FACS-purified CD144+ artery and vein ECs



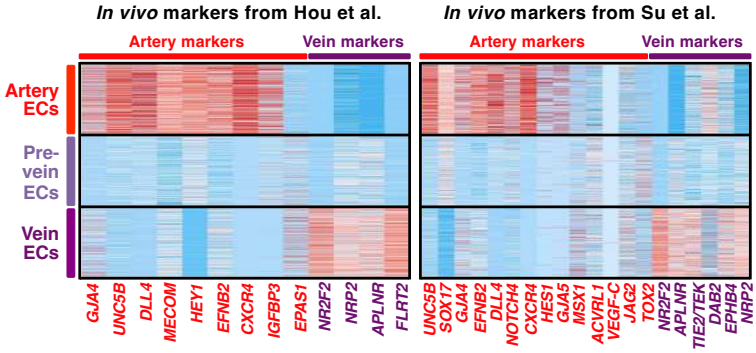
D) scRNA-seq of FACS-purified CD144+ artery ECs



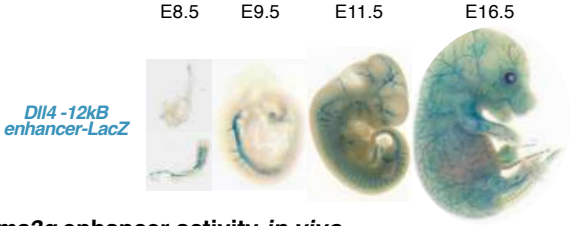
E) scRNA-seq of FACS-purified CD144+ vein ECs



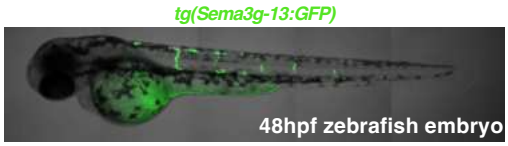
F) Arteriovenous markers in hPSC-derived ECs



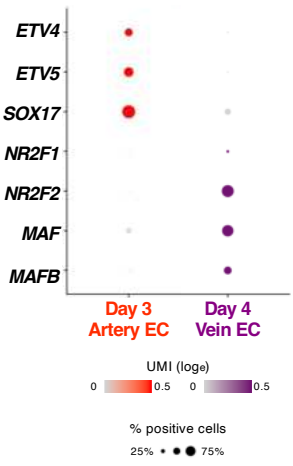
H) *Dll4* enhancer activity *in vivo*



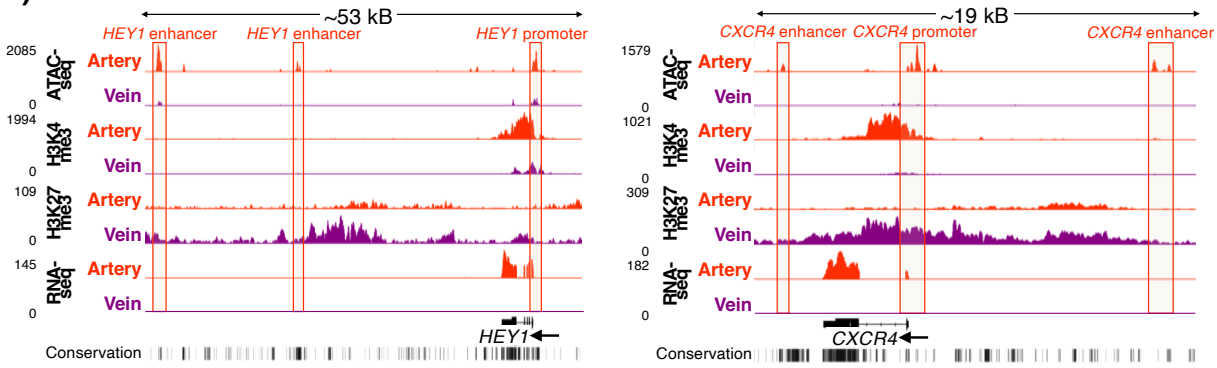
I) *Sema3g* enhancer activity *in vivo*



G) Transcription factor expression



J) Chromatin state of artery genes



K) *CXCR4* +135 kB enhancer

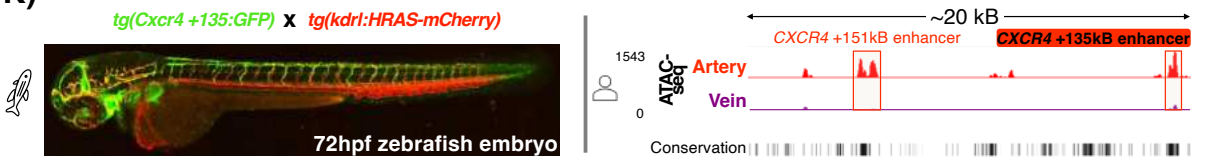
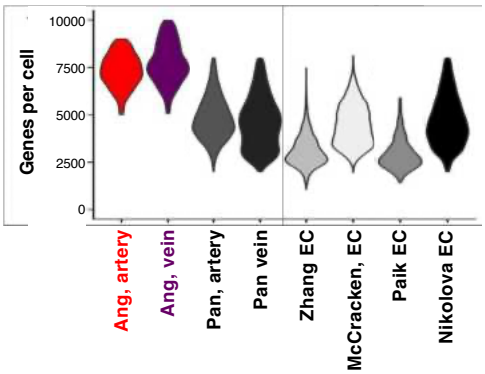
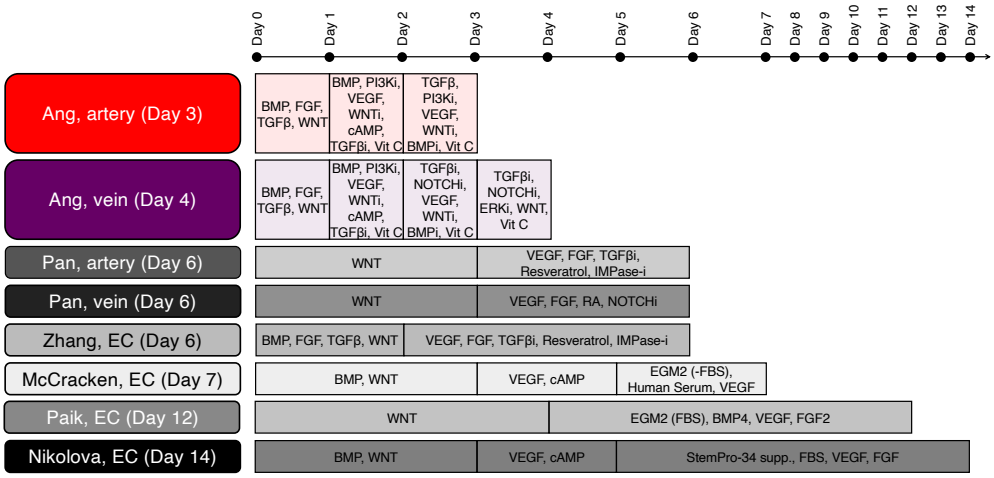


Figure S6

A) Quality control of scRNA-seq datasets



B) Comparison of EC differentiation methods



C) Comparison of EC differentiation methods (scRNA-seq)

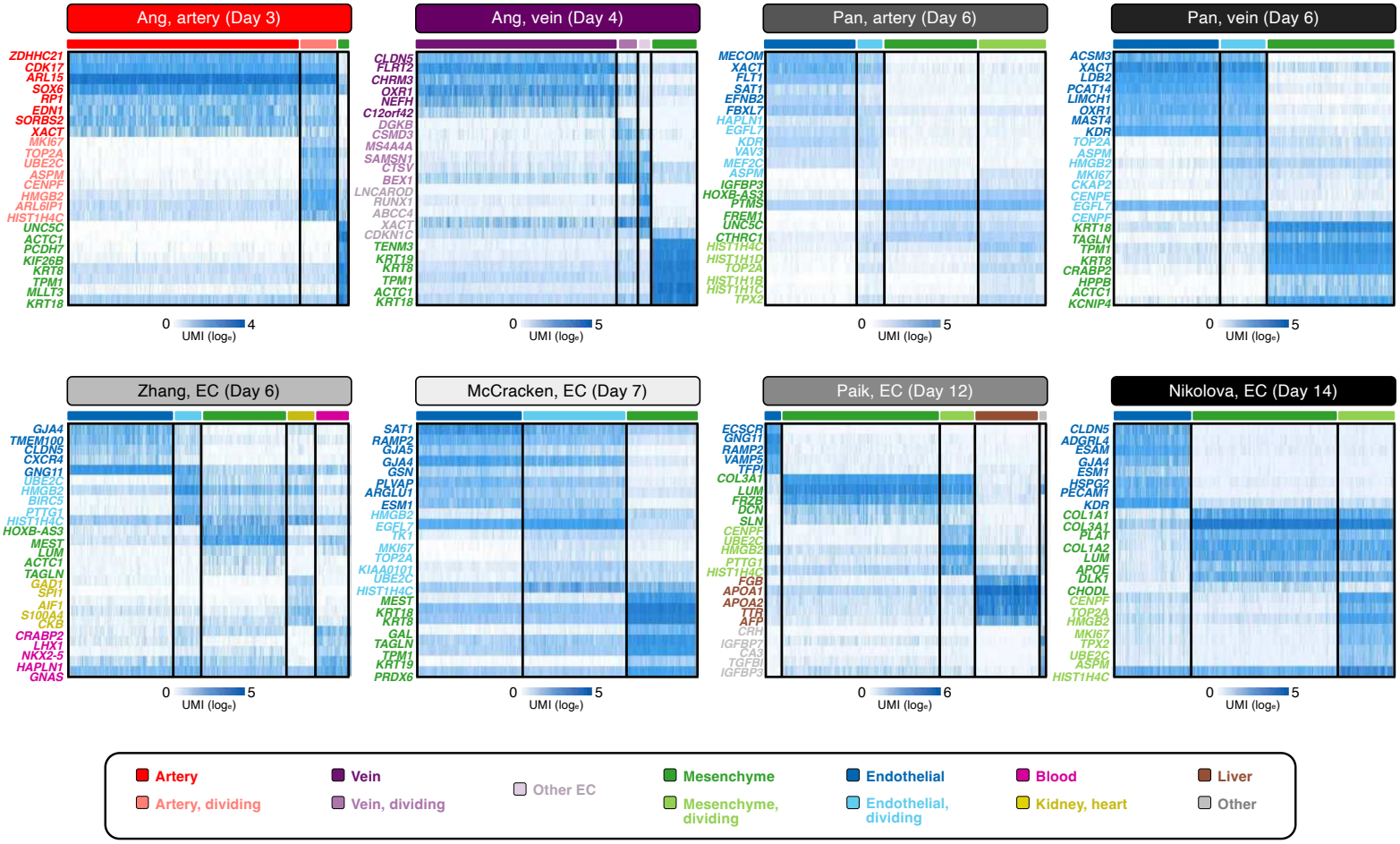
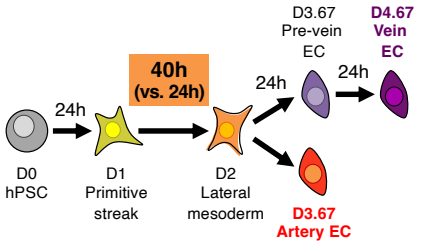
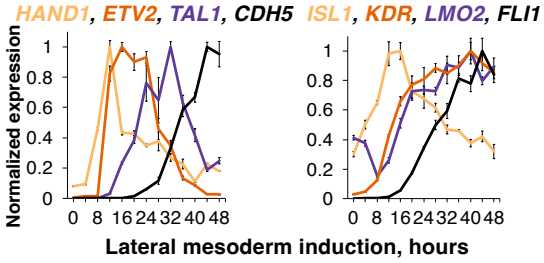


Figure S7

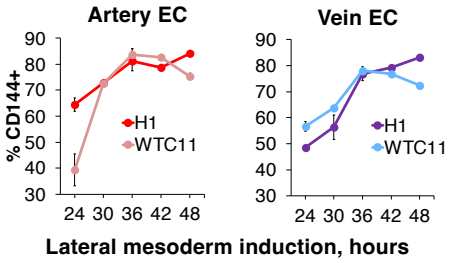
A) Lateral mesoderm extension



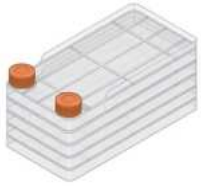
B) Lateral mesoderm genes (qPCR)



C) Artery and vein EC purity (flow cytometry)



D) Large-scale EC differentiation



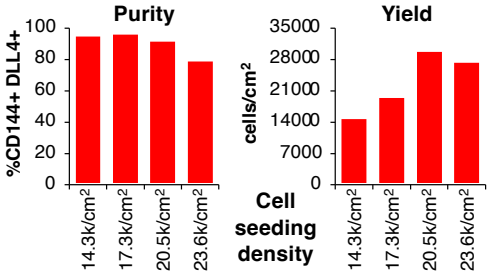
Chambers: 5 (or more) chambers per stack

Area: 636cm^2 per chamber

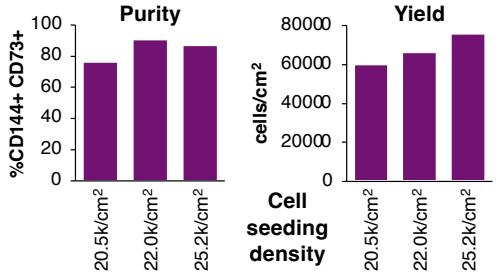
Seed: 9-16 million hPSCs per chamber

Yield: up to ~200 million ECs total

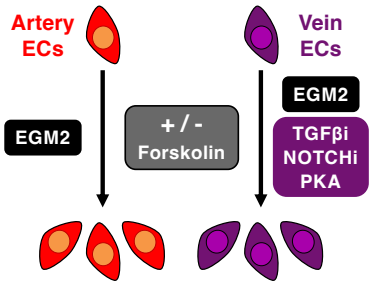
E) Large-scale artery EC differentiation



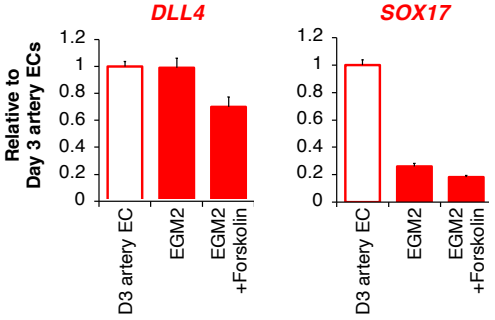
F) Large-scale vein EC differentiation



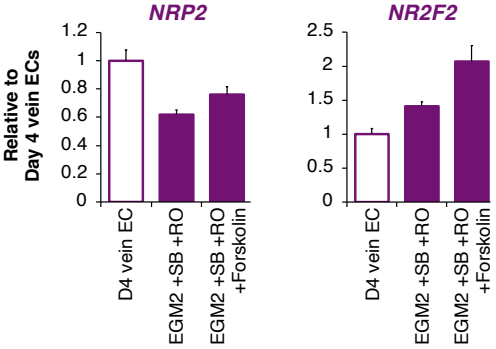
G) Artery and vein expansion



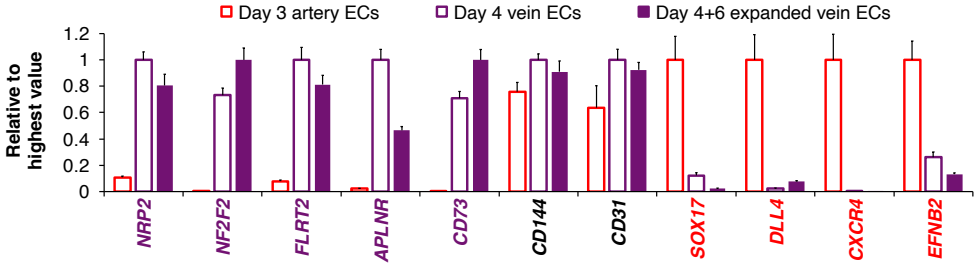
H) Artery expansion (qPCR)



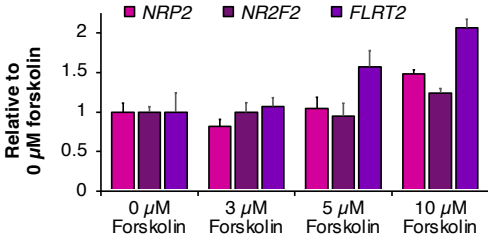
I) Vein expansion (qPCR)



J) 6 days of vein EC expansion (qPCR)



K) PKA dosage titration (qPCR)



L) 6 days of artery or vein EC expansion (qPCR)

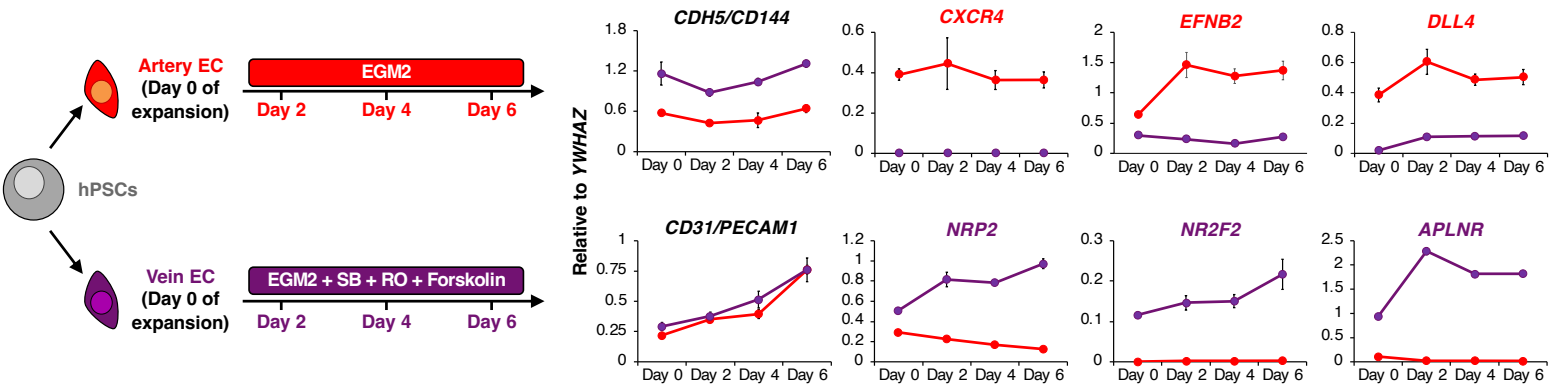
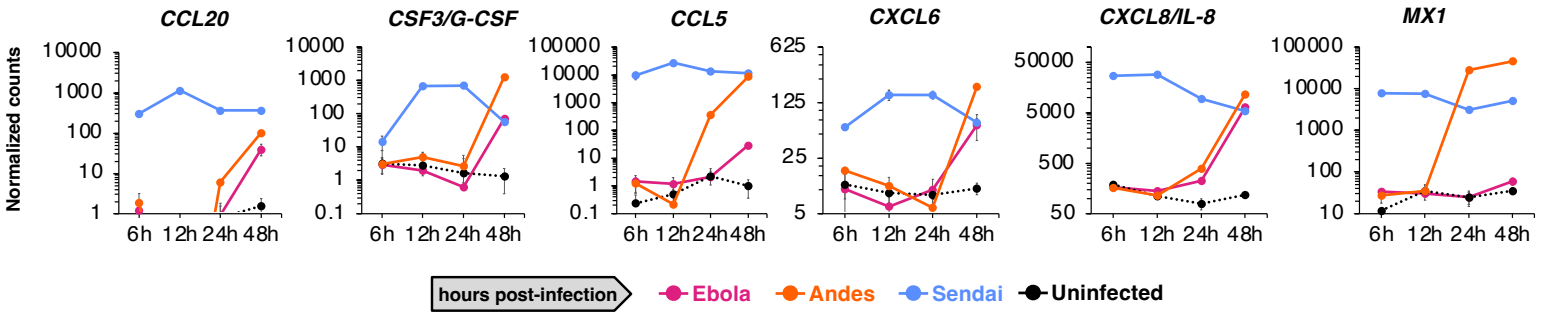
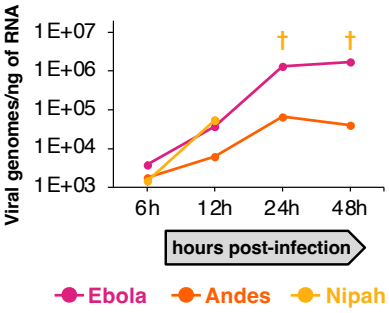


Figure S8

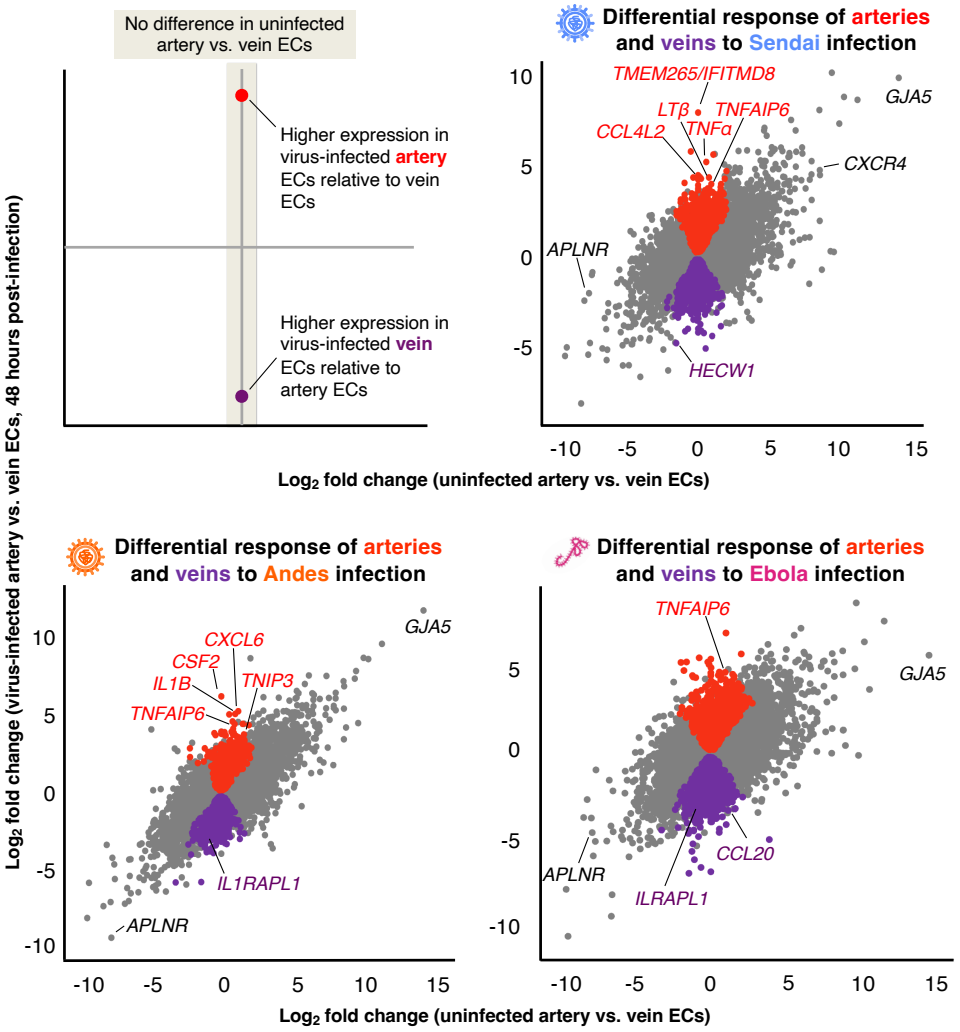
A) Gene expression of virally infected hPSC-derived artery ECs



B) Intracellular viral replication in hPSC-derived artery ECs



C) Artery EC vs. vein EC responses to viral infection (bulk RNA-seq)



SUPPLEMENTAL FIGURE LEGENDS

Figure S1: Cellular diversity during early steps of hPSC differentiation, related to Figure 1

A) Summary of hPSC differentiation approach[49]. Inh. and i: inhibition.
B) Summary of gene expression, chromatin state, and cell-surface marker profiling conducted as part of this study. Dotted lines indicate that profiling was not conducted on a given cell-type.
C) Differentially expressed genes that distinguish each hPSC-derived cell-type, as detected by scRNAseq. scRNA-seq was performed on day-0 H1 hPSCs, day-1 primitive streak, day-2 lateral mesoderm, CD144+ FACS-purified day-3 artery ECs, CD144+ FACS-purified day-3 pre-vein ECs, and CD144+ FACS-purified day-4 vein ECs.
D) scRNAseq of day-0 H1 hPSCs that were differentiated into day-1 primitive streak, day-2 lateral mesoderm, day-3 artery ECs, day-3 pre-vein ECs, and day-4 vein ECs. None of cell populations shown here were FACS purified and consequently, mesenchymal cells were detected alongside artery, pre-vein, and vein ECs on days 3-4 of differentiation. The same day-0 hPSC, day-1 primitive streak, and day-2 primitive streak scRNAseq datasets were also used in **Fig. 1B**.
E) scRNAseq of H1 hPSC-derived mesenchymal cells that arose alongside day-3 artery ECs, day-3 pre-vein ECs, and day-4 vein ECs. Differentially expressed genes that distinguish these three different types of mesenchyme are shown. Diff: differentiation.

Figure S2: Characterization of pre-vein endothelial cells *in vitro*, related to Figure 2

A) Gene expression on days 0, 1, and 2 of H1 hPSC differentiation, as detected by scRNAseq.
B) qPCR of the indicated H1 hPSC-derived cell-types. qPCR data normalized to reference gene *YWHAZ* (i.e., *YWHAZ* levels = 1.0). Error bars: standard deviation (SD).
C) Bulk RNA-seq of the indicated H1 hPSC-derived cell-types. FACS was used to purify CD144+ pre-vein and vein ECs for RNA-seq. i: inhibition.
D) scRNAseq of H1 hPSC-derived day-3 pre-vein EC populations, showing the percentage of cells that expressed *APLNR* and/or *SOX17* mRNAs.
E) CUT&RUN profiling of H3K4me3 and H3K27me3 and bulk RNA-seq H1 hPSC-derived day-3 artery ECs, day-3 pre-vein ECs, and day-4 vein ECs. Conservation: Phastcons evolutionary conservation of genome sequence across 46 vertebrate species.

Figure S3: Existence and fate of *Sox17*+ *Aplnr*+ endothelial cells in the mouse embryo, related to Figure 3

A) scRNAseq of E9.5 mouse embryo ECs. scRNAseq data were obtained from Chen et al., 2024[65].
B) scRNAseq of Carnegie Stage 12 (CS12) human embryo ECs. scRNAseq data were obtained from Calvanese et al., 2022[66].
C) Whole-mount Erg and Sox17 immunostaining of E8.75 mouse embryo. DA: dorsal aorta (arrowhead). VV: vitelline vein (arrowhead). ISV: intersomitic vessel (arrowhead). Arrows: solitary Sox17+ Erg+ ECs. Ant: anterior. Post: posterior. Scale: 100 μm.
D) Whole-mount *Aplnr* mRNA staining of E10 *Sox17*^{Cre}*; R26*^{zsGreen} mouse embryos[67,137] by HCR3 *in situ* hybridization. Scale: 100 μm.

E) E12.5 *Sox17*^{Cre}*; R26*^{mTmG} mouse embryos[67,138] were sectioned and stained for Vegfr2, SMAα, and GFP proteins. Scale: 200 μm.
F) 4-hydroxytamoxifen (4OHT) was administered *in utero* to E8.75 *Aplnr-CreER*; *R26-tdTomato* mouse embryos[17,137], which were then isolated at E9.75 and immunostained for Sox17 and Erg proteins. Bottom row arrows: cardinal vein. Scale: 200 μm (top row), 50 μm (bottom row).
G) 4-hydroxytamoxifen (4OHT) was administered *in utero* to E8.75 *Aplnr-CreER*; *R26-tdTomato* mouse embryos[17,137], which were then isolated at E9.75 and immunostained for Sox17 and Erg proteins. Quantification of Erg+ ECs in the dorsal aorta and cardinal vein that co-expressed tdTomato (indicative of *Aplnr-CreER* activity) and Sox17. Error bars: SD. **P<0.01.

Figure S4: Triple CRISPRi knockdown of *SOXF* transcription factors impairs both human artery and vein EC differentiation *in vitro*, related to Figure 4.

A) scRNA-seq of day-2 lateral mesoderm, CD144+ FACS-purified day-3 artery ECs, CD144+ FACS-purified day-3 pre-vein ECs, and CD144+ FACS-purified day-4 vein ECs.
B) *Left*: scRNA-seq of day-3 artery ECs, day-3 pre-vein ECs and day-4 vein ECs generated from either H1 control or *SOXF* TKD CRISPRi hPSCs. *Right*: *PECAM1*+ ECs and *PECAM1*- mesenchymal cells were detected. The same scRNAseq datasets are shown here as in **Fig. 4C**, except that mesenchymal cells were computationally excluded in **Fig. 4C**.
C) qPCR of day-3 artery ECs and day-4 vein ECs generated from H1 control vs. *SOXF* TKD CRISPRi hPSCs. Gene expression normalized to reference gene *YWHAZ* (i.e., *YWHAZ* = 100%). Statistics: unpaired t-test. Error bars: SEM. *P<0.05. **P<0.01.
D) Differentially expressed genes between day-3 pre-vein ECs generated from H1 control vs. *SOXF* TKD CRISPRi hPSCs are colored.
E) Gene Set Enrichment analysis (GSEA)[139] of scRNA-seq data from H1 CRISPRi control vs. *SOXF* TKD hPSCs differentiated into day-3 artery ECs, day-3 pre-vein ECs, and day-4 vein ECs. Color represents the $\log_{10}$-transformed adjusted P-value.
F) qPCR of day-3 artery ECs derived from control and single *SOX7*-, *SOX17*-, or *SOX18*-knockdown H1 CRISPRi hPSC lines. Gene expression normalized to levels in control hPSC-derived artery ECs, which was set as “1.0”. Statistics: unpaired t-test. Error bars: SEM. n.s.: not significant. *P<0.05. **P<0.01.
G) qPCR of day-3 artery ECs derived from control and single *SOX7*-, *SOX17*-, or *SOX18*-knockdown H1 CRISPRi hPSC lines. Gene expression normalized to levels in control hPSC-derived artery ECs, which was set as “1.0”. “X” indicates that qPCR data were not shown, because gene expression was under 2% of *YWHAZ* in control samples. Statistics: unpaired t-test. **P<0.01.

Figure S5: Transcriptional, chromatin, and surface-marker differences between artery and vein ECs, related to Figure 5

A) LEGENDScreen high-throughput flow cytometry of cell-surface markers in day-0 hPSCs, day-1 primitive streak, day-2 lateral mesoderm, day-3 artery ECs, and day-4 vein ECs. Day-3 artery and day-4 vein populations were pre-gated on the CD144+ EC subset before depicting marker expression.

B) scRNAseq of human Carnegie Stage 17 (CS17) fetal ECs. scRNAseq data were obtained from Calvanese et al., 2022[66].
C) scRNAseq of FACS-purified CD144+ DLL4+ CD73lo/- day-3 artery ECs and CD144+ DLL4- CD73hi day-4 vein ECs generated from H1 hPSCs.
D) scRNAseq of FACS-purified CD144+ DLL4+ CD73lo/- day-3 artery ECs generated from H1 hPSCs. Left: FACS isolation strategy. Right: subclustering was performed to assess any potential population heterogeneity, and differentially expressed genes that distinguished cell subsets are shown.
E) scRNAseq of FACS-purified CD144+ DLL4- CD73hi day-4 vein ECs generated from H1 hPSCs. Left: FACS isolation strategy. Right: subclustering was performed to assess any potential population heterogeneity, and differentially expressed genes that distinguished cell subsets are shown.
F) scRNAseq of FACS-purified CD144+ DLL4+ CD73lo/- day-3 artery ECs, CD144+ day-3 pre-vein ECs, and CD144+ DLL4- CD73hi day-4 vein ECs generated from H1 hPSCs. In each of these *in vitro* cell-types, expression of *in vivo* arterial and venous markers reported by Hou et al., 2022[21] and Su et al., 2018[20] is shown.
G) Expression of *ETV*, *SOX*, *NR2F*, and *MAF* family transcription factors at the mRNA level, as shown by scRNAseq of FACS-purified CD144+ DLL4+ CD73lo/- day-3 artery ECs and CD144+ DLL4- CD73hi day-4 vein ECs generated from H1 hPSCs.
H) LacZ staining of E8.5-E16.5 mouse embryos bearing a *Dll4* -12 kB enhancer transgene driving *LacZ* expression[79].
I) Image of a 2 day-post-fertilization (dpf) zebrafish bearing a *Sema3g* -13 kB enhancer driving *GFP* expression.
J) OmniATAC-seq, CUT&RUN, and bulk RNA-seq of H1 hPSC-derived day-3 artery ECs and day-4 vein ECs. Conservation: Phastcons evolutionary conservation of genome sequence across 46 vertebrate species.
K) Left: image of a 3 day-post-fertilization (dpf) zebrafish bearing a *Cxcr4* +135 kB enhancer driving *GFP* expression, together with a *kdrl:HRAS-mCherry* transgene to label ECs[84]. Right: ATAC-seq of H1 hPSC-derived day-3 artery ECs and day-4 vein ECs, highlighting the *CXCR4* +135 kB enhancer whose ortholog was tested in the zebrafish transgenic assay[84]. Conservation: Phastcons evolutionary conservation of genome sequence across 46 vertebrate species.

Figure S6: Comparison of methods to differentiate hPSCs into ECs, related to Figure 6

A) scRNA-seq of hPSCs differentiated into ECs using various protocols, depicting the number of genes detected per single cell. This represents the quality control assessment for scRNA-seq datasets described in **Fig. 6A**.
B) Summary of protocols to differentiate hPSCs into ECs[49,60,61,85-87] used to generate the scRNA-seq datasets described in **Fig. 6A**.
C) scRNA-seq of differentiated hPSC populations described in **Fig. 6A**. Subclustering was performed to assess population heterogeneity, and differentially expressed genes that distinguished cell subsets are shown.

Figure S7: Scaling up hPSC differentiation into artery and vein ECs, and maintenance of arteriovenous identity, related to Figures 6 and 7

A) Schematic of differentiation protocol, which includes prolongation of lateral mesoderm induction from 24 hours (V1 protocol) to 40 hours (V2 protocol).
B) qPCR of lateral mesoderm and vascular marker expression in H1 hPSC-derived day-1 primitive streak cells challenged with lateral mesoderm induction media for up to 48 hours. 0 hours corresponds to day-1 primitive streak cells, prior to addition of lateral mesoderm induction media. Cells were lysed every 4 hours and qPCR was performed. For each gene, expression is normalized to the sample with the highest value, which was set to “1.0”. Error bars: SEM.
C) Effect of lateral mesoderm induction duration on the subsequent production of CD144+ artery and vein ECs from the H1 and WTC11 hPSC lines, as measured by flow cytometry. Duration of lateral mesoderm induction was varied from 24 hours (V1 protocol) to 48 hours, with all other parameters of the differentiation protocol unchanged. Error bars: SEM.
D) Schematic of CellSTACK system for large-scale differentiation of hPSCs into ECs.
E) Yield and percentage of CD144+ DLL4+ day-3.67 artery ECs generated in the large-scale CellSTACK differentiation system, depending on initial hPSC seeding density. Yield and differentiation purity were respectively measured by flow cytometry and cell counting.
F) Yield and percentage of CD73+ DLL4+ day-4.67 vein ECs generated in the large-scale CellSTACK differentiation system, depending on initial hPSC seeding density. Yield and differentiation purity were respectively measured by flow cytometry and cell counting.
G) Strategy for expanding hPSC-derived artery and vein ECs. Artery ECs are cultured in EGM2 medium[49], and vein ECs are cultured in EGM2 medium + TGFβ inhibitor + NOTCH inhibitor + PKA activator. i: inhibitor.
H) qPCR of H1 hPSC-derived artery ECs that were thawed in EGM2 medium[49] and cultured for 6 days. ROCK inhibitor (Thiazovivin, 2 μM) was added for the first day of thawing to improve cell survival, and was subsequently removed in later days. Forskolin (3 μM) was added as indicated. For each gene, expression is normalized to the sample with the highest value, which was set to “1.0”. Error bars: SEM.
I) qPCR of H1 hPSC-derived vein ECs that were thawed in EGM2 medium + SB505124 (2 μM, TGFβ inhibitor) + RO4929097 (1 μM, NOTCH inhibitor)[49] and cultured for 6 days. ROCK inhibitor (Thiazovivin, 2 μM) was added for the first day of thawing to improve cell survival, and was subsequently removed in later days. Forskolin (3 μM) was added as indicated. For each gene, expression is normalized to the sample with the highest value, which was set to “1.0”. Error bars: SEM.
J) qPCR of H1 hPSC-derived vein ECs that were thawed in EGM2 medium + SB505124 (2 μM, TGFβ inhibitor) + RO4929097 (1 μM, NOTCH inhibitor) + Forskolin (10 μM, adenylate cyclase agonist) and cultured for 6 days. ROCK inhibitor (Thiazovivin, 2 μM) was added for the first day of thawing to improve cell survival, and was subsequently removed in later days. As a control, H1 hPSC-derived artery ECs prior to expansion were also analyzed. qPCR data are normalized to the sample with the highest expression, which was set to “1.0”. Error bars: SEM.
K) qPCR of H1 hPSC-derived vein ECs that were thawed in EGM2 medium + SB505124 (2 μM, TGFβ inhibitor) + RO4929097 (1 μM, NOTCH inhibitor), in the presence or absence of Forskolin (3-10 μM, adenylate cyclase agonist) and cultured for 6 days.

qPCR data are normalized to the sample lacking forskolin, which was set to “1.0”. Error bars: SEM.

L) qPCR of pan-endothelial, arterial, and venous marker expression over the course of 6-day expansion of SUN004.2 *CAG-mScarlet* hPSC-derived artery and vein ECs. Artery ECs were expanded in EGM2 medium, whereas vein ECs were expanded in EGM2 medium + SB505124 (2 μM, TGFβ inhibitor) + RO4929097 (1 μM, NOTCH inhibitor) + Forskolin (10 μM, adenylate cyclase agonist). qPCR data normalized to reference gene *YWHAZ* (i.e., *YWHAZ* levels = 1.0). Error bars: SEM.

Figure S8: Responses of artery and vein ECs to infection by Risk Group 4 viruses, related to Figure 7

A) Bulk RNA-seq of hPSC-derived artery ECs 6, 12, 24, and 48 hours after infection with Ebola, Andes, and Sendai viruses, or left uninfected (mock control). Error bars: SEM.

B) Replication of Andes, Nipah, and Ebola viruses in hPSC-derived artery and vein ECs, as assayed by qPCR for intracellular viral genomes. †: extensive cell death, and therefore RNA was not quantified at the indicated timepoints.

C) Bulk RNA-seq of artery and vein ECs 48 hours after infection with Sendai, Andes, Nipah, and Ebola viruses. Genes that are differentially expressed between infected artery and vein ECs, but are not differentially expressed between uninfected artery and vein ECs are correspondingly colored. Red: genes more highly induced by viral infection in hPSC-derived artery relative to vein ECs, and which do not show substantially different expression in uninfected artery vs. vein ECs. Purple: genes more highly induced by viral infection in hPSC-derived vein relative to artery ECs, and which do not show substantially different expression in uninfected artery vs. vein ECs.

INVENTORY OF SUPPLEMENTAL INFORMATION

Table S1: **Single-cell RNA-seq of hPSC-derived cell-types**
Differentially expressed genes that distinguish day 0 hPSCs ("H1"), day 1 mid primitive streak ("d1ps"), day 2 lateral mesoderm ("d2dlm"), day 3 artery ECs ("d3aus"), day 3 pre-vein ECs ("d3pvus"), day 4 vein ECs ("d4vus"), and day 3-4 mesenchyme ("mes1"). Gene expression was measured by single-cell RNA-seq.

Table S2: **Bulk RNA-seq of hPSC-derived cell-types**
Count matrix of gene expression of day-0 hPSCs ("H1"), day-1 mid primitive streak ("D1_PS"), day-2 lateral mesoderm ("D2_DLM"), day 3 artery ECs, CD144+ FACS-purified ("D3_Artery"), day 3 pre-vein ECs, CD144+ FACS-purified ("D3_Prevein"), and day 4 vein ECs, CD144+ FACS-purified ("D4_Vein"). Gene expression was measured by bulk-population RNA-seq, and is quantified in $\log_2$ counts per million (CPM) units.

Table S3: **Single-cell RNA-seq of E9.5 mouse embryo ECs *in vivo***
Differentially expressed genes that distinguish 4 different types of ECs identified in the E9.5 mouse embryo. Raw single-cell RNA-sequencing data were obtained from a previous study[65]. We provisionally annotated the following clusters based on expression of characteristic marker genes: 0 (vein-like 1 ECs), 1 (other ECs), 2 (artery ECs), and 3 (vein-like 2 ECs).

Table S4: **Single-cell RNA-seq of gene expression changes upon triple *SOXF* knockdown during artery EC differentiation**
Genes that are differentially expressed between artery ECs generated from control vs. *SOX7/SOX17/SOX18*-CRISPRi hPSCs. Positive fold change indicates higher expression in control artery ECs, and $\log_2$ fold change is shown. Gene expression was measured by single-cell RNA-seq.

Table S5: **Single-cell RNA-seq of gene expression changes upon triple *SOXF* knockdown during vein EC differentiation**
Genes that are differentially expressed between vein ECs generated from control vs. *SOX7/SOX17/SOX18*-CRISPRi hPSCs. Positive fold change indicates higher expression in control vein ECs, and $\log_2$ fold change is shown. Gene expression was measured by single-cell RNA-seq.

Table S6: **Cell-surface marker screening of hPSC-derived cell-types**
For each hPSC-derived cell-type, the number indicates the percentage of cells that expressed a given cell-surface marker by high-throughput flow cytometry. Day 3 artery EC and day 4 vein EC populations were pre-gated on the CD144+ endothelial subset before quantifying the expression of other cell-surface markers. This was done to avoid being confounded by CD144- non-endothelial cells in the culture.

Table S7: **Differentially accessible genomic elements in hPSC-derived artery vs. vein ECs, as assayed by OmniATAC-seq**
Differentially accessible genomic elements in hPSC-derived artery vs. vein ECs, as assayed by OmniATAC-seq. Fold change indicates accessibility in vein ECs relative to artery ECs (i.e.,

a positive value indicates that a genomic element is more accessible in vein ECs, whereas a negative value indicates greater chromatin accessibility in artery ECs). The gene closest to each genomic element is also indicated.

Table S8: Bulk RNA-seq of hPSC-derived artery and vein ECs infected by Risk Group 4 viruses

Tab 1: count matrix of gene expression of hPSC-derived artery and vein ECs infected with either Sendai, Andes, Nipah, or Ebola viruses at 6, 12, 24, and 48 hours post-infection. Uninfected cells are designated as "Mock". Each condition was assessed in triplicate. Tab 2: sample annotations. Gene expression was measured by bulk-population RNA-seq, and count units are shown.

STAR METHODS

KEY RESOURCES TABLE

Reagent or Resource	Source	Identifier
Antibodies		
CD144 FITC antibody (used 1:50 for *in vitro* studies)	BD Biosciences	560411
CD144 Alexa Fluor 647 antibody (used 1:50 for *in vitro* studies)	BD Biosciences	561567
DLL4 APC antibody (used 1:5 for *in vitro* studies)	BioLegend	346508
CXCR4 PE-Cy7 antibody (used 1:50 for *in vitro* studies)	BD Biosciences	560669
CD73 APC antibody (used 1:10 for *in vitro* studies)	BD Biosciences	560847
CD73 PE antibody (used 1:10 for *in vitro* studies)	Biolegend	344004
BST2/CD317 PE-Cy7 antibody (used 1:20 for *in vitro* studies)	BioLegend	348416
SOX17 antibody (used 1:1000 for *in vitro* studies and 1:200 for *in vivo* studies)	R&D Systems	AF-1924
ERG antibody (used 1:500 for *in vivo* studies)	Abcam	ab92513
VEGFR2 antibody (used 1:125 for *in vivo* studies)	R&D Systems	AF644
SMAα antibody Cy3 (used 1:500 for *in vivo* studies)	Sigma	C6198
GFP antibody (used 1:300 for *in vivo* studies)	Abcam	ab13970

GFP antibody (used 1:1000 for *in vivo* studies)	Aves Labs	NC9510598
Ebola virus NP antibody (used 1:2500 for *in vitro* studies)	Robert Koch Institute	Clone 173/303/109
TULV1 N antibody, cross-reactive against Andes virus (used 1:1000 for *in vitro* studies)	Rainer Ulrich's laboratory	Avižinienė et al.[140]
Goat anti-mouse IgG secondary antibody Alexa Fluor 488 (used 1:500-1:1000 for *in vitro* studies)	Jackson ImmunoResearch	115-545-003
Donkey anti-chicken IgY secondary antibody Alexa Fluor 488 (used 1:250-1:400 for *in vivo* studies)	Thermo Fisher Scientific	A78948
Donkey anti-goat IgG secondary antibody Alexa Fluor 594 (used 1:500 for *in vivo* studies)	Thermo Fisher Scientific	A32758
Donkey anti-rabbit IgG secondary antibody Alexa Fluor 647 (used 1:500 for *in vivo* studies)	Thermo Fisher Scientific	A32795
Bacterial and viral strains		
Ebola virus, Yambuku variant, Mayinga isolate (Ebola virus/Human/COD/1976/Yambuku-Mayinga)	NIH, National Institute of Allergy and Infectious Diseases, Rocky Mountain Laboratories	Johnson et al.[141]
Andes virus, South variant, Chile-9717869 isolate (Andes virus/Oligoryzomys longicaudatus/CHL/1997/Chile-9717869)	NIH, National Institute of Allergy and Infectious Diseases, Rocky Mountain Laboratories	Toro et al.[142] and Hooper et al.[99]
Nipah virus (Malaysia variant): Nipah virus/human/MY/1999/CDC	Centers for Disease Control and Prevention, Viral Special Pathogens Branch	Harcourt et al.[143]
Sendai virus (Cantell strain)	ATCC	VR-907
Mix & Go! *E. coli* Competent Cells	Zymo Research	T3020
Chemicals, peptides and recombinant proteins		
mTesR Plus medium	STEMCELL Technologies	100-0276

Essential 8 medium	Thermo Fisher Scientific	A1517001
StemFlex medium	Thermo Fisher Scientific	A3349401
Penicillin/streptomycin	Thermo Fisher Scientific	15140163
Geltrex LDEV-free, hESC-qualified, reduced growth factor basement membrane matrix	Thermo Fisher Scientific	A1413302
Versene (EDTA-based dissociation buffer)	Thermo Fisher Scientific	15040066
Accutase (dissociation buffer)	Thermo Fisher Scientific	00-4555-56
TrypLE Express (dissociation buffer)	Thermo Fisher Scientific	12604013
EDTA (for flow cytometry buffer)	Thermo Fisher Scientific	15575020
DMEM/F12 + GlutaMAX	Thermo Fisher Scientific	10565042
IMDM + GlutaMAX	Thermo Fisher Scientific	31980-097
F12 + GlutaMAX	Thermo Fisher Scientific	31765-092
Polyvinyl alcohol	Sigma	P8136-250G
Chemically defined lipid concentrate	Thermo Fisher Scientific	11905-031
1-thioglycerol	Sigma	M6145-100ML
Recombinant human insulin	Sigma	11376497001
Human transferrin	Sigma	10652202001
Recombinant human Activin A	R&D Systems	338-AC-050
Recombinant human BMP4	R&D Systems	314-BP-050
Recombinant human FGF2	R&D Systems	233-FB-01M
Recombinant human VEGF	R&D Systems	293-VE-0500
CHIR99201	Tocris	4423
GDC-0941	Cellagen Technology	C4321-25
XAV939	Tocris	3748
SB505124	Tocris	3263
Forskolin	Tocris	1099
Ascorbic acid-2-phosphate (AA2P, Vitamin C analog)	Sigma	49752-10G
RO4929097	Cellagen Technology	C7649-10
DMH1	Tocris	4126
PD0325901	Tocris	4192
Thiazovivin	Tocris	3845
Dimethyl sulfoxide (DMSO)	Sigma	D2650

UltraPure DNase/RNase-free distilled water (H_2O)	Thermo Fisher Scientific	10977023
Bovine albumin fraction V (7.5% solution), to reconstitute recombinant growth factors	Thermo Fisher Scientific	15260037
Endothelial Cell Growth Medium 2 (EGM2)	Lonza	CC-3162
Triton X-100	Sigma	X100-500ML
PBS, without Ca^{2+} or Mg^{2+}, for non-virological experiments	Thermo Fisher Scientific	10010049
DPBS, without Ca^{2+} or Mg^{2+}, for virological experiments	Thermo Fisher Scientific	14190144
DAPI, for *in vitro* studies	Thermo Fisher Scientific	D1306
DAPI, for *in vivo* studies	Sigma	D9542-1MG
DMEM, for lentiviral packaging	Thermo Fisher Scientific	10566024
Fetal bovine serum, for lentiviral packaging	Atlanta Biologicals (later sold by R&D Systems)	S11550H
FuGENE HD Transfection Reagent, for lentiviral packaging	Promega	E2311
Poly-L-lysine, for lentiviral packaging	R&D Systems	3438-200-01
Opti-MEM, for lentiviral packaging	Thermo Fisher Scientific	31985070
Wizard Plus SV Minipreps DNA Purification System	Promega	A1460
Quick Ligation Kit	New England Biolabs	M2200
Hygromycin B	Thermo Fisher Scientific	10687010
Puromycin	Thermo Fisher Scientific	A1113803
Bovine serum albumin (BSA)	Sigma	A2153
Fluoromount-G mounting medium, for mouse embryology	SouthernBiotech	0100-01
32% paraformaldehyde (PFA), for mouse embryo immunostaining	Electron Microscopy Sciences	15714
32% paraformaldehyde (PFA), for mouse embryo *in situ* hybridization	Fisher Scientific	50-980-495
Methanol, for mouse embryo *in situ* hybridization	Fisher Scientific	A412-1

Tween 20, for mouse embryo *in situ* hybridization	Sigma	P9416-100ML
Glycerol, for mouse embryology	Thermo Fisher Scientific	327255000
Sucrose, for mouse embryology	Sigma	S0389
Normal donkey serum, freeze-dried, for mouse embryology	Jackson ImmunoResearch	102644-006
Tissue-Plus optimal cutting temperature (OCT) embedding medium, for mouse embryology	Fisher Scientific	23-730-571
(*Z*)-4-hydroxytamoxifen (4OHT), for *Aplnr-CreER* lineage tracing	Cayman Chemical	14854
(*Z*)-4-hydroxytamoxifen (4OHT), for *Sox17-Cre;Aplnr-DreER* lineage tracing	Sigma	H7904
Ethanol, for lineage tracing	Fisher Scientific	BP2818500
Corn oil, for lineage tracing	Sigma	C8267
DMEM, for viral titration experiments	Robert Koch Institute	Produced in-house
Fetal bovine serum, for viral titration experiments	PAN Biotech	P30-3306
Stable glutamine, for viral titration experiments	PAN Biotech	P04-82100
Carboxymethylcellulose sodium salt, for viral titration experiments	Sigma	C4888-500G
Crystal violet, for viral titration experiments	Roth	T123.1
Methanol, for viral titration experiments	Roth	CP43.1
Tween-20, for viral titration experiments	Sigma	P9416
10% Formalin, for viral titration experiments	Sigma	F5554
Critical commercial assays		
Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3	10x Genomics	PN-1000075

Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1	10x Genomics	PN-1000121
LEGENDScreen PE-Conjugated Human Antibody Plates	BioLegend	700007
Human IFNβ Quantikine QuicKit ELISA	R&D Systems	QK410
QIAamp Viral RNA Kit, for virological experiments	Qiagen	52904
AgPath-ID One-Step RT-PCR Kit, for virological experiments	Thermo Fisher Scientific	4387391
RNeasy Mini Kit, for virological experiments	Qiagen	74106
RNeasy Micro Kit, for virological experiments	Qiagen	74004
RNeasy Plus Mini Kit	Qiagen	10010049
High-Capacity cDNA Reverse Transcription Kit	Applied Biosystems	4368814
SensiFAST SYBR Lo-ROX Kit	Thomas Scientific	BIO-94050
Quick-RNA Microprep Kit	Zymo Research	R1051
Tagment DNA Enzyme 1 (TDE1)	Illumina	15027865
Tagment DNA Buffer	Illumina	15027866
MinElute PCR Purification Kit, for Omni-ATAC-seq	Qiagen	28204
AMPure XP Beads, for Omni-ATAC-seq	Beckman Coulter	A63880
CellBIND polystyrene CellSTACK chamber	Corning	3311
Deposited data		
Single-cell RNA-sequencing dataset of hPSCs differentiated into day 1 mid primitive streak, day 2 lateral mesoderm, day 3 artery ECs (unsorted cell population), day 3 ($CD144^{+}$ $DLL4^{+}$ $CD73^{lo/-}$ FACS-purified), day 3 pre-vein ECs (unsorted cell population), day 3 pre-vein ECs ($CD144^{+}$	Kyle Loh's and Lay Teng Ang's laboratories (Stanford University)	This study

FACS-purified), and day 4 vein ECs (CD144^{+} DLL4^{-} CD73^{hi} FACS-purified)		
Single-cell RNA-sequencing dataset of hPSCs differentiated into day 4 vein ECs (unsorted cell population) using the Ang et al. protocol	Kyle Loh's and Lay Teng Ang's laboratories (Stanford University)	NCBI PRJNA837932 (Ang et al.[49])
Single-cell RNA-seq of CRISPRi control vs. *SOX7/SOX17/SOX18*-knockdown hPSCs differentiated into day 3 artery ECs, day 3 pre-vein ECs, and day 4 vein ECs	Kyle Loh's and Lay Teng Ang's laboratories (Stanford University)	This study
Single-cell RNA-sequencing dataset of hPSCs differentiated into day 6 artery ECs using the Pan et al. protocol	Wei Kong's, Xi Wang's, and Kai Wang's laboratories (Peking University)	NCBI PRJNA1114402 (Pan et al.[60])
Single-cell RNA-sequencing dataset of hPSCs differentiated into day 6 vein ECs using the Pan et al. protocol	Wei Kong's, Xi Wang's, and Kai Wang's laboratories (Peking University)	NCBI PRJNA1114402 (Pan et al.[60])
Single-cell RNA-sequencing dataset of hPSCs differentiated into day 6 ECs using the Zhang et al. protocol	Andrew Baker's laboratory (University of Edinburgh)	NCBI GSE131736 (McCracken et al.[87])
Single-cell RNA-sequencing dataset of hPSCs differentiated into day 7 ECs using the McCracken et al. protocol	Andrew Baker's laboratory (University of Edinburgh)	NCBI GSE131736 (McCracken et al.[87])
Single-cell RNA-sequencing dataset of hPSCs differentiated into day 12 ECs using the Paik et al. protocol	Joseph Wu's laboratory (Stanford University)	NCBI GSE116555 (Paik et al.[85])
Single-cell RNA-sequencing dataset of hPSCs differentiated into day 14 ECs	Josef Penninger's, J. Gray Camp's, and Barbara Treutlein's laboratories (ETH Zurich, IMBA, and Roche)	Array Express E-MTAB-14807 (Nikolova et al.[61])

using the Nikolova et al. protocol		
Single-cell RNA-sequencing dataset of E9.5 mouse embryonic vasculature	William Pu's laboratory (Boston Children's Hospital)	NCBI GSE216970 (Chen et al.[65])
Single-cell RNA-sequencing dataset of Carnegie Stage 12 human fetal vasculature	Vincenzo Calvanese's and Hanna Mikkola's laboratories (University of California, Los Angeles)	NCBI GSE162950 (Calvanese et al.[66])
Bulk-population RNA-sequencing dataset of hPSCs differentiated into day 1 mid primitive streak, day 2 lateral mesoderm, and day 3 pre-vein ECs (CD144^{+} FACS-purified)	Kyle Loh's and Lay Teng Ang's laboratories (Stanford University)	This study
Bulk-population RNA-sequencing dataset of hPSCs differentiated into day 3 artery ECs (CD144^{+} DLL4^{+} CD73$^{lo/-}$ FACS-purified) and day 4 vein ECs (CD144^{+} DLL4^{-} CD73^{hi} FACS-purified)	Kyle Loh's and Lay Teng Ang's laboratories (Stanford University)	NCBI PRJNA837932 (Ang et al.[49])
OmniATAC-seq dataset of hPSCs differentiated into day 3 artery ECs (CD144^{+} DLL4^{+} CD73$^{lo/-}$ FACS-purified), day 3 pre-vein ECs (CD144^{+} FACS-purified), and day 4 vein ECs (CD144^{+} DLL4^{-} CD73^{hi} FACS-purified)	Kyle Loh's and Lay Teng Ang's laboratories (Stanford University)	This study
CUT&RUN dataset of hPSCs differentiated into day 3 artery ECs (CD144^{+} DLL4^{+} CD73$^{lo/-}$ FACS-purified), day 3 pre-vein ECs (CD144^{+} FACS-purified), and day 4 vein ECs (CD144^{+} DLL4^{-} CD73^{hi} FACS-purified)	Kyle Loh's and Lay Teng Ang's laboratories (Stanford University)	This study
Experimental models: cell lines		
H1 hESCs	WiCell	WiCell, WA01 (Thomson et al.[144])

H1 *NR2F2-2A-GFP; SOX17-2A-mPlum* knock-in reporter hESCs	This study	N/A
H1 CRISPRi-expressing hESCs	Richard She's and Jonathan Weissman's laboratories	N/A
WTC11 hiPSCs	Coriell Institute for Medical Research	Coriell Institute for Medical Research, GM25256, Kreitzer et al.[145]
SUN004.2 *CAG-mScarlet* hiPSCs	Hiro Nakauchi's laboratory	Ang et al.[49]
HEK 293T/17 cells	ATCC	CRL-11268
Vero C1008 (Vero 76, clone E6, Vero E6)	ECACC	85020206
Experimental models: organisms/strains		
Wild-type CD-1 mice	Charles River	022
Sox17-2A-iCre mouse (otherwise known as *Sox17-Cre*)	Heiko Lickert's laboratory	Engert et al.[67]
Aplnr-CreER mouse (otherwise known as *Apj-CreER*)	Kristy Red-Horse's laboratory	Chen et al.[17]
Aplnr-DreER mouse (otherwise known as *Apj-DreER*)	Sophie Astrof's laboratory	Ramirez et al.[71]
ROSA26-LSL-zsGreen mouse (otherwise known as Ai6)	The Jackson Laboratory (originally developed by the Allen Brain Institute)	JAX stock 007906 (Madisen et al.[137])
ROSA26-LSL-tdTomato mouse (otherwise known as Ai14)	The Jackson Laboratory (originally developed by the Allen Brain Institute)	JAX stock 007914 (Madisen et al.[137])
RC::RLTG mouse	The Jackson Laboratory (originally developed by Patricia Jensen's laboratory)	JAX stock 026931 (Plummer et al.[72])
mTmG mouse	The Jackson Laboratory (originally developed by Liqun Luo's laboratory)	JAX stock 007676 (Muzumdar et al.[138])
Sema3g -13kB enhancer-LacZ mouse	Len Pennacchio's and Axel Visel's laboratories	Kosicki et al.[82]
Dll4 -12 kB enhancer-LacZ mouse	Sarah De Val's laboratory	Sacilotto et al.[79]
Wild-type AB strain zebrafish	Zebrafish International Resource Center	ZL1

Cxcr4 +135 enhancer-GFP; *kdrl-HRAS-mCherry* zebrafish	Sarah De Val's laboratory	Nornes et al.[84]
Oligonucleotides		
Primers	See **Method details** section	This study
Pool of HCR3 *in situ* hybridization probes for human *APLNR* mRNA	Molecular Instruments	Custom order for human *APLNR* mRNA (NM_005161.6)
Pool of HCR3 *in situ* hybridization probes for zebrafish *sox7* mRNA	Molecular Instruments	Custom order
Pool of HCR3 *in situ* hybridization probes for zebrafish *aplnrb* mRNA	Molecular Instruments	Custom order
Recombinant DNA		
mU6-sgRNA 1; hU6-sgRNA 2; EF1A-PuroR-2A-GFP backbone plasmid, fragment 1	Addgene (originally developed by Jonathan Weissman's laboratory)	Addgene, 187243 (Guna et al.[146])
mU6-sgRNA 1; hU6-sgRNA 2; EF1A-PuroR-2A-GFP backbone plasmid, fragment 2	Addgene (originally developed by Jonathan Weissman's laboratory)	Addgene, 187239 (Guna et al.[146])
mU6-sgRNA; EF1A-HygroR plasmid	Karina Smolyar and Jonathan Weissman	This study
pVSV-G plasmid	Owen Witte's laboratory	Goldstein et al.[147]
pMDL plasmid	Owen Witte's laboratory	Goldstein et al.[147]
pREV plasmid	Owen Witte's laboratory	Goldstein et al.[147]
Software and algorithms		
ImageJ/Fiji	[148]	https://imagej.net/software/fiji/
Seurat v3	[136]	https://satijalab.org/seurat/
Seurat v4	[149]	https://satijalab.org/seurat/
RStudio	RStudio Team	https://www.rstudio.com/
Cell Ranger	10x Genomics (ref. [150])	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger
EnsDb.Hsapiens.v86	[151]	https://bioconductor.org/packages/release/data/annotation/html/EnsDb.Hsapiens.v86.html
DESeq2	[152]	https://github.com/mikelove/DESeq2
clusterProfiler	[153]	https://guangchuangyu.github.io/software/clusterProfiler/

ShinyCell	154	https://github.com/SGDDNB/ShinyCell
FastQC	155	http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Trim Galore	156	http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
Kallisto	157	https://pachterlab.github.io/kallisto/
EdgeR	158	http://bioconductor.org/packages/release/bioc/html/edgeR.html
Voom	159	https://rdrr.io/bioc/limma/man/voom.html
Limma	160	https://bioconductor.org/packages/release/bioc/html/limma.html
Dplyr	161	https://dplyr.tidyverse.org/
ggplot2	162	https://ggplot2.tidyverse.org/
Tidyverse	163	https://www.tidyverse.org/
FastQ Screen	164	https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/
NGmerge	165	https://github.com/harvardinformatics/NGmerge
SAMBLASTER	166	https://github.com/GregoryFaust/samblaster
SAMtools	167	https://github.com/samtools/samtools
deepTools	168	https://github.com/deeptools/deepTools
tximport	169	https://github.com/thelovelab/tximport
Diffbind	170	https://bioconductor.org/packages/release/bioc/html/DiffBind.html
HOMER	171	http://homer.ucsd.edu/homer/
pheatmap	172	https://github.com/raivokolde/pheatmap

RESOURCE AVAILABILITY

Lead contact

Requests for further information should be directed to and will be fulfilled by the Lead Contact, Lay Teng Ang (layteng@stanford.edu).

Materials availability

H1 *SOX17-2A-mPlum*; *NR2F2-2A-GFP* double reporter hESCs were generated as part of this study, and will be made freely available upon request and the completion of applicable Material Transfer Agreements.

Ebola virus, Andes virus, and Nipah virus reference stocks were obtained from the CDC Viral Special Pathogens Branch and NIH/NIAID Rocky Mountain Laboratories. In accordance with U.S. federal law, requests for these viruses from these repositories must be approved by the U.S. Federal Select Agent Program, in compliance with biological security procedures that regulate the secure transfer of these viruses to authorized recipients within biosafety level 4 (BSL4) containment facilities.

Data and code availability

Bulk-population RNA-seq, single-cell RNA-seq, bulk-population ATAC-seq, and bulk-population CUT&RUN datasets generated in this study will be available upon acceptance. An interactive portal to access the single-cell RNA-seq dataset of hPSC differentiation into artery and vein ECs is available at https://anglab.shinyapps.io/artery-vein-scrna-seq/. Computational scripts used for genomics analyses conducted in this study are available at GitHub: https://github.com/LayTengAngLab/Prevein_endothelial.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture

All cells in this study were cultured in standard incubator conditions (20% O_2, 5% CO_2 and 37 °C).

Human pluripotent stem cell lines

Wild-type H1 hESCs[144], H1 *NR2F2-2A-GFP; SOX17-2A-mPlum* knock-in reporter hESCs, CRISPRi-expressing H1 hESCs, wild-type WTC11 hiPSCs[145], and SUN004.2 *CAG-mScarlet* hiPSCs[49] were used in this study.

All hPSC lines in this study, except for WTC11 and H1 CRISPRi hPSCs, were cultured in mTeSR Plus medium (STEMCELL Technologies). WTC11 hPSCs were cultured in Essential 8 medium[173] (Thermo Fisher Scientific). H1 CRISPRi hPSCs were cultured in StemFlex medium (Thermo Fisher Scientific).

Methods to culture undifferentiated hPSCs have been described previously[174]. In brief: undifferentiated hPSCs were propagated in mTeSR Plus medium (STEMCELL Technologies) + 1% penicillin/streptomycin in monolayer cultures, on Geltrex basement membrane matrix-coated plates (described below). To maintain undifferentiated hPSCs, mTeSR Plus was changed either every day or every other day as per the manufacturer's recommendations. In order to maintain cultures of undifferentiated hPSCs, when they became partially confluent,

undifferentiated hPSCs were passaged by treating them for 7 minutes with Versene (Thermo Fisher; an EDTA-based dissociation buffer) at room temperature. Subsequently, Versene was removed, mTeSR Plus was added, and then hPSCs were manually scraped off the plate to generate clumps. hPSC clumps were then seeded onto new plates that had been precoated with Geltrex basement membrane matrix (described below) in mTeSR Plus medium + 1% penicillin/streptomycin. To reiterate, during Versene-based maintenance passaging of undifferentiated hPSCs as clumps, ROCK inhibitor was not added.

Mouse models

Wild-type CD-1 mice (Charles River, catalog number 022), *Sox17-Cre* mice (provided by Heiko Lickert's laboratory[67]), *Aplnr-CreER* mice (provided by Kristy Red-Horse's laboratory[17]), *Aplnr-DreER* mice (provided by Sophie Astrof's laboratory[71]), *ROSA26-LSL-zsGreen* mice (Ai6; JAX stock 007906, and developed by the Allen Brain Institute[137]), *ROSA26-LSL-tdTomato* mice (Ai14; JAX stock 007914, and developed by the Allen Brain Institute[137]), *RC::RLTG* mice (JAX stock 026931, and developed by Patricia Jensen's laboratory[72]), *Sema3g -13kB enhancer-LacZ* mice (developed by Len Pennacchio's and Axel Visel's laboratories[82]), and *Dll4 -12 kB enhancer-LacZ* mice (developed by Sarah De Val's laboratory[79]), were used in this study.

Zebrafish models

Wild-type AB zebrafish and *Cxcr4 +135 enhancer-GFP*; *kdrl-HRAS-mCherry* zebrafish (developed by Sarah De Val's laboratory[84]) were used in this study.

Ebola virus, Yambuku variant, Mayinga isolate

Ebola virus, Yambuku variant, Mayinga isolate (Ebola virus/Human/COD/1976/Yambuku-Mayinga; genomic sequence reported in NCBI accession number AF086833.2) was originally isolated by the CDC from a fatally-infected human in the Democratic Republic of the Congo in 1976, and was passaged on Vero E6 cells[141]. Ebola virus is a member of the species *Orthoebolavirus zairense*; genus *Orthoebolavirus*; family *Filoviridae*; order *Mononegavirales*[175]. Ebola virus, Yambuku variant, Mayinga isolate was obtained from NIH Rocky Mountain Laboratories reference stocks.

Andes virus, South variant, Chile-9717869 isolate

Andes virus, South variant, Chile-9717869 isolate (Andes virus/Oligoryzomys longicaudatus/CHL/1997/Chile-9717869; genome sequence reported in NCBI accession numbers AF291704 [L genome segment], AF291703 [M genome segment], and AF291702 [S genome segment]) was originally isolated by the CDC from an infected *O. longicaudatus* rodent in Chile in 1997, and was passaged on Vero E6 cells[99,142]. Isolate Chile-9717869 has been genomically assigned to the South variant[176]. Andes virus is a member of the species *Orthohantavirus andesense*; genus *Orthohantavirus*; family *Hantaviridae*; order *Bunyavirales*. Andes virus, South variant, Chile-9717869 isolate was obtained from NIH Rocky Mountain Laboratories reference stocks.

Nipah virus, Malaysia variant, CDC isolate

Nipah virus, Malaysia variant (Nipah virus/human/MY/1999/CDC; genomic sequence reported in NCBI accession number NC_002728) was originally isolated by the CDC from the cerebrum of an infected human in Malaysia in 1999, and passaged on Vero E6 cells [143]. Nipah virus is a

member of the species *Nipah henipavirus*; genus *Henipavirus*; family *Paramyxoviridae*; order *Mononegavirales*. Nipah virus, Malaysia variant was obtained from CDC Viral Special Pathogens Branch reference stocks.

Sendai virus, Cantell strain

The original provenance of Sendai virus, Cantell strain is unknown, but it was passaged over 100 times in chicken embryonated eggs at the Central Public Health Laboratory in Helsinki, Finland[177]. Sendai virus is a member of the species *Murine respirovirus*; genus *Respirovirus*; family *Paramyxoviridae*; order *Mononegavirales*. Sendai virus, Cantell strain was obtained from ATCC (VR-907).

METHOD DETAILS

Data processing and visualization

Standard plots were prepared with Microsoft Excel, Microsoft PowerPoint, or GraphPad Prism. Flow cytometry data were visualized with FlowJo. Microscope images were visualized with Fiji[148]. Genomics data were respectively wrangled and plotted using dplyr[161] and ggplot2[162] in the tidyverse[163], in the RStudio environment. Genomics tracks were visualized with the Integrated Genomics Viewer (IGV)[178]. Color palettes were chosen with the assistance of https://colorbrewer2.org/.

Mouse husbandry

Adult mice of the indicated genotypes were mated to generate timed pregnancies. Females were checked each morning for a vaginal plug; noon on the day a plug was observed was defined as embryonic day 0.5 (E0.5).

Lineage tracing

- ***Sox17-Cre;Aplnr-DreER* lineage tracing**: As described previously[179], 25 mg of (*Z*)-4-hydroxytamoxifen (4OHT; Sigma, H7904) was dissolved in 1250 µL ethanol (Fisher Scientific, BP2818500) by vortexing and heating at 60 °C. This yielded a 20 mg/mL stock of 4OHT, which was aliquoted and stored at -20 °C. Prior to dosing mice, 50 µL aliquots (containing 1 mg of 4OHT) were heated for 10 minutes at 65 °C, and then combined with pre-warmed corn oil (250 µL, Sigma, C8267). This mixture of 4OHT and corn oil was thoroughly vortexed. Pregnant females were intraperitoneally injected at the specified labeling timepoint with 1 mg of 4OHT per mouse.
- ***Aplnr-CreER* lineage tracing**: 25 mg of (*Z*)-4-hydroxytamoxifen (4OHT; Cayman Chemical, 14854) was dissolved at 10mg/mL in ethanol (Fisher Scientific, BP2818500) by vortexing, and then was aliquoted and stored at -80 °C. Pregnant female mice were intraperitoneally injected with 28 mg 4OHT per kg of mouse weight at the specified labeling timepoint.

***In situ* hybridization and immunostaining of whole-mount mouse embryos**

Fluorescent *in situ* hybridization (FISH) of whole-mount mouse embryos was performed using hybridization chain reaction v3.0 (HCR3)[180], and some instances, immunostaining was simultaneously performed. HCR3 was performed as per the Molecular Instruments protocol (https://www.molecularinstruments.com/hcr-v3-protocols), and is briefly summarized here.

Mouse embryos were dissected in ice-cold 4% paraformaldehyde (Fisher Scientific, 50-980-495, diluted in PBS). They were then subsequently fixed overnight in 4% paraformaldehyde, sequentially dehydrated in methanol solutions of increasing concentration (Fisher Scientific, A412-1), and then incubated in 100% methanol overnight.

Embryos were subsequently permeabilized in PBS + 0.1% Triton X-100 for 1 hour at room temperature, and then blocked in blocking buffer (PBS + 0.05% Triton X-100 + 5% BSA) overnight at 4 °C. Embryos were then stained with primary antibodies (diluted in blocking buffer) for 24-72 hours at 4 °C, washed three times with PBS, and stained with secondary antibodies (diluted in blocking buffer) for 24 hours at 4 °C. Subsequently, embryos were washed three times with PBS, and then HCR3 was performed as per the Molecular Instruments protocol. Hybridization mRNA probes, amplifiers, and buffers were obtained from Molecular Instruments. Embryos were incubated in DAPI + SSCT (sodium chloride, sodium citrate, and Tween buffer) prior to mounting. Images were captured on an Olympus FV3000 confocal microscope.

Immunostaining of whole-mount mouse embryos

Mouse embryos were dissected in ice-cold PBS, and then fixed in 4% paraformaldehyde (Electron Microscopy Sciences, 15714; diluted in PBS) for 1 hour at 4 °C. Subsequently, embryos were washed in PBS and stained with primary antibodies (diluted in PBS + 0.1% Triton X-100) overnight at 4 °C. Following three washes in PBS + 0.1% Triton X-100, embryos were incubated in secondary antibodies (diluted in PBS + 0.1% Triton X-100) overnight at 4 °C. Embryos were then washed in PBS + 0.1% Triton X-100, counterstained with 1 µg/ml DAPI (Sigma, D9542), and washed twice in PBS. Embryos were then equilibrated sequentially in 25%, 50%, and 75% glycerol in PBS for 30 minutes at room temperature without agitation. Embryos were then mounted in Fluoromount-G (SouthernBiotech, 0100-010) and imaged on a Zeiss LSM 980 confocal microscope.

Immunostaining of cryosectioned mouse embryos

Mouse embryos were dissected in ice-cold PBS, and then fixed in 4% paraformaldehyde (Electron Microscopy Sciences, 15714; diluted in PBS) for 1 hour at 4 °C. Subsequently, embryos were washed in PBS, and dehydrated in 30% sucrose at 4°C overnight. Embryos were embedded in OCT medium (Fisher Scientific, 23-730-571), and then frozen overnight at -80 °C. Embryos were sectioned to a thickness of 20 µm and incubated at room temperature for 30 minutes. Sections were subsequently permeabilized in PBS + 0.1% Triton X-100 for 15 minutes at room temperature and blocked in 5% donkey serum + PBS + 0.05% Triton X-100 at room temperature. Sections were then stained with primary antibodies (diluted in blocking buffer) at 4 °C overnight, washed with PBS + 0.05% Triton X-100. They were then stained with secondary antibodies (diluted in blocking buffer) at room temperature, followed by additional washes in PBS + 0.05% Triton X-100, counterstained with 1 µg/ml DAPI (Sigma, D9542) for 10 minutes at room temperature, and then washed with PBS. Slides were then mounted with Fluoromount-G (SouthernBiotech, 0100-010) and imaged on a Zeiss LSM 980 confocal microscope.

Quantification of mouse embryo images

Fiji[148] was used to generate maximum-intensity projections from z-stacks of whole-mount mouse embryos or sectioned mouse embryos.

- Sox17 immunostaining (**Fig. 3E**): Erg+ or Sox17+ cells were manually counted in the dorsal aorta (DA) and cardinal vein (CV). The percentage of Erg+ cells that expressed Sox17 is shown.
- *Sox17-Cre;Aplnr-DreER* lineage tracing (**Fig. 3H**): Erg+ or GFP+ cells were manually counted in the DA and CV. The percentage of Erg+ cells that expressed GFP is shown.
- *Aplnr-CreER* lineage tracing (**Fig. S3F**): Erg+, Sox17+, or tdTomato+ cells were manually counted in the DA and CV. The percentage of Erg+ cells that co-expressed both Sox17 and tdTomato is shown.

Functional testing of *Dll4* enhancer element in mouse embryos
All animal procedures were approved by a local ethical review committee at Oxford University and licensed by the UK Home Office.

A stable transgenic mouse line bearing a genomic integration of the *Dll4* -12 kb enhancer element driving *LacZ* reporter expression, known as *Tg(Dll4-12:lacZ)*, was generated as previously described[79]. Embryos were fixed in 2% paraformaldehyde + 0.2% glutaraldehyde + 1x PBS for 60 minutes. After fixation, embryos were rinsed in 0.1% sodium deoxycholate, 0.2% Nonidet P-40, 2 mM $MgCl_2$ and 1x PBS, then stained for 2-24 hours in 1 mg/ml 5-bromo-4-chloro-3-indolyo-β-D-galactoside solution (X-gal) containing 5 mM potassium ferrocyanide, 5 mM ferricyanide, 0.1% sodium deoxycholate, 0.2% Nonidet P-40, 2 mM $MgCl_2$, and 1x PBS. After staining, embryos were rinsed through a series of 1x PBS washes, then fixed overnight in 4% paraformaldehyde at 4°C. All embryos were imaged using a Leica M165C stereo microscope equipped with a ProGres CF Scan camera and CapturePro software (Jenoptik).

Functional testing of *Sema3g* enhancer element in mouse embryos
The activity of the *Sema3g* -13 kb enhancer element was tested in mouse embryos through a transgenic reporter assay by Len Pennacchio's and Axel Visel's laboratories, and was reported as part of the VISTA Enhancer Browser[82,83] (VISTA Enhancer Browser ID hs2179: https://enhancer.lbl.gov/vista/element?vistaId=hs2179&alleleId=0&backbone=hZR&stage=e11.5). In brief, a reporter construct containing the *Sema3g* -13 kb enhancer element and the *Hsp68* minimal promoter driving *LacZ* reporter expression was randomly integrated into the mouse genome; whole-mount staining of mouse embryos for *LacZ* reporter expression was then performed[82,83].

Fluorescent *in situ* hybridization of zebrafish embryos
FISH of whole-mount zebrafish embryos was performed using HCR3[180].

Wild-type AB strain embryos were fixed overnight with 4% paraformaldehyde (BT-Fix) in PBS at 4 °C, dehydrated through a sequential ethanol series, and stored at -20°C. Embryos were rehydrated and washed three times with PBT (PBS + 0.2% bovine serum albumin + 0.2% Tween 20). A prehybridization step was done using hybridization buffer (30% formamide + 5x SSC + 9 mM citric acid at pH 6.0 + 0.1% Tween 20 + 50 µg/ml heparin + 1x Denhardt's solution, + 10% dextran sulfate) for 30 mins at 37 °C.

Following this, each 2 pmol of each HCR3 probe was combined with 500 µL of hybridization buffer and zebrafish embryos were incubated overnight in this solution at 37°C, while being gyrated. Following incubation, embryos were washed four times with 30% formamide + 5x SSC

+ 9 mM citric acid at pH 6.0 + 0.1% Tween 20 + 50 μg/mL heparin at 37°C, for 15 minutes each. This was then followed with three washes with 5x SSCT (5x SSC + 0.1% Tween 20) at room temperature for 5 minutes each. Subsequently, 30 minutes of incubation in amplification buffer (5× SSC + 0.1% Tween 20 + 10% dextran sulfate) at room temperature was performed. Hairpin probes (30 pmol each), fluorescently labeled through snap cooling of 3 μM stock solution, were added to the embryos in amplification buffer and incubated overnight at room temperature in the dark. Subsequently, samples were washed five times with 5x SSCT. Embryos were mounted in 0.6% low-melting agarose and imaged using a Nikon A1R confocal microscope and a 20× objective. Images were processed using the denoiseAI function in the NIS Elements software (Nikon) to reduce noise. Maximum intensity images were obtained with Fiji[180].

Differentiation of hPSCs into artery and vein ECs

Two differentiation protocols were used throughout this study, which we refer to "Version 1" (V1, which was used for most experiments) and "Version 2" (V2, which was only used in **Fig. S7A-F**). The V1 differentiation protocol has been described previously[49,174]. All RNA-seq, ATAC-seq, CUT&RUN, and LEGENDScreen profiling in this study was conducted on cells generated by the V1 differentiation protocol.

The V2 differentiation protocol was developed as part of this study and is more efficient at generating both artery and vein ECs than the V1 protocol (**Fig. S7C**). The V2 differentiation protocol is identical to the V1 protocol, with the exception that the V1 protocol entails 24 hours of lateral mesoderm differentiation, whereas the V2 protocol entails 40 hours of lateral mesoderm differentiation (**Fig. S7A-C**).

All differentiation was conducted in defined, serum-free CDM2 basal media[181,182]. Detailed instructions to how to prepare CDM2 are available[174].

- **Seeding hPSCs for differentiation** (**Step 0**). In preparation for differentiation, hPSCs were dissociated into single cells using Accutase (Thermo Fisher), because sparse seeding of cells is important for differentiation. Accutase-dissociated hPSCs were plated into recipient wells in mTeSR medium supplemented with the ROCK inhibitor thiazovivin (1 μM, Tocris; to enhance hPSC survival after passaging) onto plates precoated with Geltrex basement membrane matrix. hPSCs were seeded at a density of 25,000-50,000 cells/cm^2 were seeded (i.e., 95,000-190,000 hPSCs/well of a 12-well plate)[49,174]. To clarify, long-term maintenance of undifferentiated hPSCs entailed passaging as clumps using Versene (an EDTA-based dissociation buffer; to maintain normal karyotype), but hPSCs were dissociated using Accutase to seed single cells for differentiation. 24 hours after seeding in mTeSR + 1 μM thiazovivin, during which the hPSCs re-formed small clumps, differentiation was initiated as described below.
- **Step 1: Mid primitive streak induction, 24 hours**. Day 0 hPSCs were briefly washed (DMEM/F12, Thermo Fisher) to remove all traces of mTeSR + thiazovivin. Then, they were differentiated towards mid primitive streak in CDM2 media supplemented with Activin A (30 ng/mL, R&D Systems), BMP4 (40 ng/mL, R&D Systems), CHIR99021 (6 μM, Tocris), FGF2 (20 ng/mL, Thermo Fisher), as previously described[49,174,182]. In both the V1 and V2 protocols, mid primitive streak induction was conducted for 24 hours.

- **Step 2: Lateral mesoderm induction, 24-40 hours**. Day 1 mid primitive streak cells were briefly washed (DMEM/F12) and then differentiated towards lateral mesoderm in CDM2 media supplemented with BMP4 (40 ng/mL), GDC-0941 (2.5 μM, Cellagen Technology), Forskolin (10 μM, Tocris), SB-505124 (2 μM, Tocris), VEGF (100 ng/mL, R&D Systems), XAV939 (1 μM, Tocris) and ascorbic acid-2-phosphate (AA2P; 200 μg/mL, Sigma), as previously described[49,174]. In the V1 protocol, lateral mesoderm induction was performed for 24 hours. In the V2 protocol, lateral mesoderm induction was performed for 40 hours. Subsequently, lateral mesoderm was subjected to either artery EC induction (Step 3A; below) or pre-vein EC induction (Step 3A; below).
- **Step 3A: Artery EC induction, 24 hours**. Day 2 lateral mesoderm cells (V1 protocol) or day 2.67 lateral mesoderm cells (V2 protocol) were briefly washed (DMEM/F12) and then differentiated towards artery ECs in CDM2 media supplemented with Activin A (15 ng/mL), DMH1 (250 nM, Tocris), GDC-0941 (2.5 μM), VEGF (100 ng/mL), XAV939 (1 μM) and AA2P (200 μg/mL), as previously described[49,174]. In both the V1 and V2 protocols, artery EC induction was performed for 24 hours.
- **Step 3B: Pre-vein EC induction, 24 hours**. Day 2 lateral mesoderm cells (V1 protocol) or day 2.67 lateral mesoderm cells (V2 protocol) were briefly washed (DMEM/F12) and then differentiated into pre-vein ECs in CDM2 media supplemented with SB505124 (2 μM), DMH1 (250 nM), RO4929097 (2 μM, Cellagen Technology), VEGF (100 ng/mL), XAV939 (1 μM) and AA2P (200 μg/mL), as previously described[49,174]. In both the V1 and V2 protocols, pre-vein EC induction was performed for 24 hours.
- **Step 4B: Vein EC induction, 24 hours**. Day 3 pre-vein ECs (V1 protocol) or day 3.67 pre-vein ECs (V2 protocol) were briefly washed (DMEM/F12) and then differentiated into vein ECs in CDM2 media supplemented with SB505124 (2 μM), RO4929097 (2 μM), PD0325901, (500 nM, Tocris), CHIR99021 (1 μM) and AA2P (200 mg/mL), as previously described[49,174]. In both the V1 and V2 protocols, vein EC induction was performed for 24 hours.

Detailed methods to reconstitute each differentiation-inducing small molecule and recombinant growth factor and to prepare stocks of each are available[174]. In brief, (1) all recombinant growth factors were reconstituted in PBS + 0.1% bovine albumin fraction V (both from Thermo Fisher Scientific), (2) all small molecules except for AA2P were reconstituted in DMSO (Sigma), and (3) AA2P was reconstituted in H_2O (Thermo Fisher Scientific), as described previously[174].

Large-scale differentiation of hPSCs into artery and vein ECs

Large-scale differentiation of hPSCs was performed using the V2 differentiation protocol described above. For artery EC differentiation, WTC11 hPSCs were seeded at a density of 20.5K cells/cm^2 in a Geltrex-coated 5-stack CellSTACK device (Corning), and yielded 94.5 million cells. For vein EC differentiation, WTC11 hPSCs were seeded at a density of 25.2K cells/cm^2, and 239.6 million cells were generated per 5-stack CellSTACK device.

Cryopreservation and maintenance of hPSC-derived artery and vein ECs

After hPSC differentiation into artery ECs or vein ECs as described above, they could be maintained for at least 6 additional days *in vitro* on Geltrex-coated cell culture plates. As described previously[49], hPSC-derived artery ECs were expanded in EGM2 (Endothelial Cell Growth Medium 2, Lonza CC-3162), which was refreshed every 24 hours. By contrast, hPSC-

derived vein ECs were expanded in EGM2 + SB505124 (2 μM) + RO4929097 (2 μM) + Forskolin (10 μM), which was refreshed every 24 hours.

Alternatively, as previously described[49], hPSC-derived day 3 artery ECs and day 4 vein ECs were dissociated, cryopreserved in freezing media (90% PBS + 10% DMSO), and stored in liquid nitrogen. hPSC-derived artery and vein ECs were then thawed in their respective media (EGM2 for artery ECs and EGM2 + SB505124 + RO4929097 + Forskolin for vein ECs) and cultured for up to 6 days as described above, with Thiazovivin (1 μM) added for the first 24 hours post-thawing to improve cell survival.

Flow cytometry

Cultured cells were dissociated by incubation in TrypLE Express (Thermo Fisher) for 5 minutes at 37 °C. Following dissociation, the cells were diluted with 5-10 times excess volume of FACS buffer (PBS + 1 mM EDTA [Thermo Fisher] + 2% v/v FBS [Atlanta Bio] + 1% v/v Penicillin/Streptomycin [Thermo Fisher]) and centrifuged at 500g for 5 minutes to pellet them. Each cell pellet was then resuspended in FACS buffer and incubated with fluorescently-conjugated primary antibodies for 15-30 minutes in the dark at 4°C. After staining, cells were washed twice with FACS buffer and resuspended in 100 μL FACS buffer containing DAPI (1 μg/mL) for live/dead discrimination. Flow cytometry was conducted on a Beckman Coulter CytoFlex analyzer in the Stanford Stem Cell Institute FACS Core Facility. For data analysis, cells were gated based on forward and side scatter area, followed by height and width parameters for doublet discrimination. Subsequently, live cells that were negative for DAPI were gated for marker analyses and calculations of population frequency.

In this study, we defined hPSC-derived artery and vein ECs using the following cell-surface marker combinations:

- **Artery ECs**: $CD144^+$ $CXCR4^+$ $DLL4^+$ (and, in some experiments, $CD144^+$ $DLL4^+$ $CD73^{lo/-}$). The following antibody combination was used to define arterial identity: CD144 FITC (BD Biosciences, 560411 [1:50 concentration]), DLL4 APC (Biolegend, 346508 [1:5 concentration]), and CXCR4 PE-Cy7 (BD Biosciences, 560669 [1:50 concentration]).
- **Vein ECs**: $CD144^+$ $CD317^+$ $CD73^+$ (and, in some experiments, $CD144^+$ $DLL4^-$ $CD73^{hi}$). The following antibody combination was used to define venous identity: CD144 FITC (BD Biosciences, 560411 [1:50 concentration]), CD73 APC (BD Biosciences, 560847 [1:10 concentration]), and CD317 PE-Cy7 (BioLegend, 348416 [1:20 concentration]).

The CD73 protein is encoded by the *NT5E* gene. The CD317 protein is encoded by the *BST2* gene.

Quantitative PCR

Methods for RNA extraction, reverse transcription, and qPCR have been described previously[49]. In brief, undifferentiated or differentiated hPSCs were first lysed in 350 μL of RLT Plus Buffer and RNA was extracted using the RNeasy Plus Mini Kit (Qiagen) according to the manufacturer's protocol. Second, 300 ng of total RNA was reverse transcribed into cDNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's protocol. Third, qPCR was performed in 384-well format using the SensiFAST SYBR Lo-ROX Kit (Thomas Scientific) as previously described[49,182], using gene-specific

forward and reverse primers on a QuantStudio 5 qPCR machine (Thermo Fisher). Expression of all genes was normalized to the levels of the reference gene *YWHAZ*.

This qPCR procedure was only used for biosafety level 2 (BSL2) cell cultures. A separate qPCR procedure was used for biosafety level 4 (BSL4) materials, and is detailed further below.

Human gene	Forward primer	Reverse primer
YWHAZ	GAGCTGGTTCAGAAGGCCAAAC	CCTTGCTCAGTTACAGACTTCATGCA
HAND1	GTGCGTCCTTTAATCCTCTTC	GTGAGAGCAAGCGGAAAAG
ISL1	AGATTATATCAGGTTGTACGGGATCA	ACACAGCGGAAACACTCGAT
VEGFR2/FLK1/KDR	TTTTTGCCCTTGTTCTGTCC	TCATTGTTCCCAGCATTTCA
ETV2	CCGACGGCGATACCTACTG	CGGTGGTTAGTTTTGGGGCAT
SCL/TAL1	CAAAGTTGTGCGGCGTATCTT	TCATTCTTGCTGAGCTTCTTGTC
LMO2	ATTGGGGACCGCTACTTC	GCCCAAAAAGCCTGAGATAGT
FLI1	ACCTCCCACACCGACCAAT	GGACTTTTGTTGAGGCCAGAA
CDH5/VE-CADHERIN	AACGAGCAGGGCGAGTTCACCTTC	TAGGTGACCAGCTGCTCGTGGATC
PECAM1/CD31	AACAGTGTTGACATGAAGAGCC	TGTAAAACAGCACGTCATCCTT
DLL4	GTCTCCACGCCGGTATTGG	CAGGTGAAATTGAAGGGCAGT
CXCR4	CACCGCATCTGGAGAACCA	GCCCATTTCCTCGGTGTAGTT
EFNB2	AAGGACTGGTACTATACCCACAG	TGTCTGCTTGGTCTTTATCAACC
SOX7	AGCCGGAGCAGACCTTCTT	GCCGGGGAGTAATAGGCAG
SOX17	CGCACGGAATTTGAACAGTA	GGATCAGGGACCTGTCACAC
SOX18	AAGCGTGGAAGGAGCTGAAC	CGCGGCCGGTACTTGTAGTT
NR2F2	GCCATAGTCCTGTTCACCTCA	AATCTCGTCGGCTGGTTG
APLNR	CTCTGGACCGTGTTTCGGAG	GGTACGTGTAGGTAGCCCACA
FLRT2	CGCTGCGACAGGAACTTTG	TGGAGGTAGAGTACGGTTACG
NRP2	GCTGGCTATATCACCTCTCCC	TCTCGATTTCAAAGTGAGGGTTG
NT5E/CD73	CCAGTACCAGGGCACTATCTG	TGGCTCGATCAGTCCTTCCA

Combined immunostaining and *in situ* hybridization of cultured cells

Combined immunostaining and *in situ* hybridization of cultured cells was performed as described by Molecular Instruments (https://files.molecularinstruments.com/MI-Protocol-2%C2%BAIF-RNAFISH-GenericSolution-Rev6.pdf). First, cultured monolayer cells were fixed, permeabilized and immunostained with primary and secondary antibodies. Next, probe hybridization, amplification, and wash steps were performed using the HCR3 protocol[180]. Imaging was conducted using an FV3000 confocal microscope (Olympus).

Bulk-population RNA-seq of uninfected cells

We performed bulk-population RNA-seq on the following cell populations that were generated from H1 hPSCs in the same biological experiment: *1*) day-0 hPSCs (entire cell population,

unsorted); *2*) day-1 mid primitive streak (entire cell population, unsorted); *3*) day-2 lateral mesoderm (entire cell population, unsorted); *4*) day-3 artery ECs (FACS-purified for CD144$^+$ DLL4$^+$ CD73$^{lo/-}$ cells; 10 μM forskolin was added during day-3 artery EC induction); *5*) day-3 pre-vein ECs (FACS-purified for CD144$^+$ cells); and *6*) day-4 vein ECs (FACS-purified for CD144$^+$ cells). Bulk-population RNA-seq datasets of day-3 artery ECs and day-4 vein ECs were previously reported by Ang et al., 2022 (NCBI PRJNA837932)[49].

Bulk-population RNA-seq was performed and computationally analyzed as described previously[49]. Cells were lysed in Zymo RNA lysis buffer, and RNA was purified using the Zymo Quick-RNA Microprep Kit (Zymo, R1051). RNA integrity was assessed by Agilent Bioanalyzer on-chip electrophoresis. High-quality RNA samples with RNA integrity number ≥ 8 underwent poly(A) enrichment, and libraries were prepared with indexed adaptors for multiplexing. RNA-seq libraries were then sequenced on the DNBSEQ-G400 sequencer by BGI Global Genomic Services to generate 150-bp paired-end reads. To limit batch effects, all libraries were pooled prior to sequencing and distributed across multiple planes.

FastQC[155] was used to perform quality control of raw RNA-seq reads. Adapters and low-quality bases were trimmed with Trim Galore[156]; a Phred quality threshold of ≥ 33 was used, and reads shorter than 20 nucleotides after trimming were discarded. Reads from each library—which was originally sequenced on multiple lanes—were then concatenated to yield one file per library. RNA-seq reads were pseudoaligned to human reference genome hg38, and gene-level RNA-seq counts were then quantified, using Kallisto[157]. After quantification of gene-level RNA-seq counts, two separate computational workflows in the RStudio environment were used:

- **Workflow 1** (**Generation of volcano plots**): Gene counts were filtered with edgeR[158]. Counts were then transformed to $\log_2$ counts per million (CPM) using voom[159]. Differentially expressed genes were determined by limma[160], and P values were adjusted for multiple hypothesis testing using the Benjamini-Hochberg method to control the false discovery rate.
- **Workflow 2** (**Generation of gene expression matrices**): Ensembl transcript IDs were mapped to Ensembl gene IDs using EnsDb.Hsapiens.v86[151]. Transcript-level abundance estimates were summarized to gene-level counts with tximport[169], and then imported into DESeq2[152] for differential expression analysis.

A separate procedure was used to perform bulk-population RNA-seq of cells that underwent viral infection under BSL4 containment, which is described below.

Single-cell RNA-sequencing of hPSC-derived cell populations

The 10x Genomics Chromium platform was used to perform single-cell RNA-seq profiling every 24 hours during the differentiation of H1 hPSCs into artery and vein ECs using the V1 differentiation protocol.

In one experiment, the Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3 was used to profile the following samples: *1*) day-0 hPSCs (entire cell population, unsorted); *2*) day-1 mid primitive streak (entire cell population, unsorted); *3*) day-2 lateral mesoderm (entire cell population, unsorted); *4*) day-3 pre-vein EC (entire cell population, unsorted); *5*) day-3 pre-vein EC (FACS-purified for CD144$^+$ cells); *6*) day-4 vein EC (entire cell population, unsorted); and *7*) day-4 vein EC (FACS-purified for CD144$^+$ DLL4$^-$ CD73^{hi} cells). In a second experiment, the Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3.1 was used to profile the following

samples: *8*) day-3 artery EC (entire cell population, unsorted); and *9*) day-3 artery EC (FACS-purified for CD144^{+} DLL4^{+} CD73$^{lo/-}$ cells). This scRNAseq dataset of unsorted day-4 vein ECs was previously reported by Ang et al., 2022 (NCBI PRJNA837932)[49]; all other scRNAseq datasets were generated as part of this study.

The rationale to perform scRNAseq of FACS-purified ECs was to rigorously test whether cell-surface marker combinations would enable the isolation of transcriptionally-homogeneous cell populations.

As per the manufacturer, no batch effects have been detected between the Chromium v3 and v3.1 chemistries (https://kb.10xgenomics.com/hc/en-us/articles/360047373071-Does-Cell-Ranger-distinguish-between-v3-and-v3-1-chemistry).

Sequencing libraries were prepared using Chromium Single Cell 3' GEM Gene Expression v3 or v3.1 kits as per the manufacturer's guidelines. Sequencing libraries were diluted in Buffer EB. Libraries were prepared with 10-nucleotide indices compatible with Illumina sequencers.

Computational analysis of single-cell RNA-sequencing data from hPSC-derived cell populations

scRNAseq libraries were sequenced across multiple lanes on an Illumina HiSeq 4000 sequencer by Novogene. The first and last 8 nucleotides of the i7 indices were unique and thus used for demultiplexing. FASTQ sequencing files were input into Cell Ranger[150], which was used to align reads to the hg38 reference genome (version GRCh38-2024-A), followed by filtering, barcode counting, and unique molecular identifier (UMI) counting.

Subsequent analyses were performed using Seurat v3[136] in the RStudio environment. Cell matrix files generated from Cell Ranger were imported into R using the Seurat function "Read10x_h5". For each individual scRNAseq dataset, we performed quality control by excluding dying/dead cells that *1*) exhibited low numbers of expressed genes, *2*) displayed anomalously low or high mitochondrial counts, or *3*) did not express reference genes *ACTB* or *YWHAZ*; additionally, we also *4*) computationally excluded likely doublets by removing cells that had significantly higher counts of expressed genes. High-quality single-cell transcriptomes that passed these quality control metrics were used for subsequent analyses.

Seurat objects from all scRNAseq datasets were merged using the "Merge" function in Seurat. scRNAseq data were then normalized using the "LogNormalized" function and scaled using a linear transformation in Seurat. Marker gene expression was depicted on UMAP plots[183] by coloring each single cell (i.e., a dot) according to the levels of marker gene expression. Dots were randomly ordered, without visually superimposing dots that were positive for a given marker gene, thereby avoiding visual stacking bias.

For scRNAseq analysis of control vs. *SOXF*-deficient CRISPRi hPSCs that were subject to EC differentiation, first we computationally identified ECs and excluded non-ECs from the scRNAseq dataset. This was performed via Louvain clustering, which clearly distinguished *PECAM1*+ ECs vs. *PECAM1*- non-EC clusters (**Fig. S4B**). EC clusters were used for subsequent analyses, using the aforementioned scRNAseq analysis workflows. Gene set

enrichment analysis of gene ontology terms was performed on genes that were differentially expressed between control vs. *SOXF*-CRISPRi ECs using clusterProfiler[153]. Selected gene ontology terms were shown.

Data wrangling and plotting were respectively performed using dplyr[161] and ggplot2[162] in the tidyverse[163]. An interactive web browser to explore scRNAseq data was constructed using ShinyCell[154].

Comparing different endothelial differentiation protocols using single-cell RNA-seq
We analyzed scRNAseq datasets of hPSCs that were subjected to 8 different EC differentiation protocols. The 8 differentiated cell populations analyzed were:

- **Ang et al., day 3 artery ECs**: hPSCs were differentiated into artery ECs in 3 days using the V1 differentiation protocol described by Ang et al., 2022[49]. This new scRNAseq dataset was generated as part of this study.
- **Ang et al., day 4 vein ECs**: hPSCs were differentiated into vein ECs in 4 days using the V1 differentiation protocol described by Ang et al., 2022[49]. This scRNAseq dataset was deposited by the Ang et al., 2022 study[49] and is publicly available from NCBI PRJNA837932.
- **Pan et al., day 6 artery ECs**: hPSCs were differentiated into artery ECs in 6 days using the Pan et al., 2024 protocol[60]. This scRNAseq dataset was deposited by the Pan et al., 2024 study[60], and is publicly available from NCBI PRJNA1114402.
- **Pan et al., day 6 vein ECs**: hPSCs were differentiated into vein ECs in 6 days using the Pan et al., 2024 protocol[60]. This scRNAseq dataset was deposited by the Pan et al., 2024 study [60], and is publicly available from NCBI PRJNA1114402.
- **Zhang et al., day 6 ECs**: hPSCs were differentiated into ECs in 6 days using the Zhang et al., 2017 differentiation protocol[86]. This scRNAseq dataset was deposited by the McCracken et al., 2019 study[87], and is publicly available from NCBI GSE131736.
- **McCracken et al., day 7 ECs**: hPSCs were differentiated into ECs in 7 days using the McCracken et al., 2019 protocol[87]. This scRNAseq dataset was deposited by the McCracken et al., 2019 study[87], and is publicly available from NCBI GSE131736. Data from 3 experimental replicates were deposited to the Gene Expression Omnibus, and our preliminary analysis revealed batch effects among these three experimental replicates. To reduce batch effects, we selected replicate 3 for analysis, as it was the replicate that contained the largest number of cells. Of note, the original McCracken et al. study refers to cells being harvested on day 8 of differentiation[87]. However, because the first day of differentiation in their procedure entails cell seeding in hPSC medium[87], to be consistent with the nomenclature used here to describe other differentiation protocols, here we refer to these cells being differentiated for 7 days.
- **Paik et al., day 12 ECs**: hPSCs were differentiated into ECs in 12 days using the Paik et al., 2018 protocol[85]. This scRNAseq dataset was deposited by the Paik et al., 2018 study[85], and is publicly available from NCBI GSE116555.
- **Nikolova et al., day 14 ECs**: hPSCs were differentiated into vascular organoids in 14 days using the Nikolova et al., 2025 protocol[61], which in turn was modified from the Wimmer et al., 2019 protocol[89] to generate vascular organoids. This scRNAseq dataset was deposited by the Nikolova et al., 2025 study[61], and is publicly available from ArrayExpress E-MTAB-14807.

scRNAseq datasets were analyzed using Seurat v3[136]. For all scRNAseq datasets, first we selected single-cell transcriptomes that passed well-established quality control metrics, by excluding dying/dead cells that *1*) exhibited low numbers of expressed genes, *2*) displayed anomalously low or high mitochondrial counts, or *3*) did not express reference genes *ACTB* or *YWHAZ*; additionally, we also *4*) computationally excluded likely doublets by removing cells that had significantly higher counts of expressed genes. All single-cell transcriptomes across all datasets that passed quality control are shown in **Fig. S6A**. For initial analyses, we assessed all single-cell transcriptomes that passed these quality control criteria; we did not pre-select a given subpopulation of cells based on marker gene expression that might bias further analyses.

Seurat objects from all scRNAseq datasets were then merged using the "Merge" function in Seurat v3[136]. To quantify the degree of cellular heterogeneity generated by each differentiation protocol, Louvain clustering was applied at the same resolution (0.1) across all datasets, thus decomposing each scRNAseq dataset into multiple constituent cell-types. To assign the identity of each "cell-type" within each dataset, we analyzed genes that were differentially expressed between each cell-type within a given dataset, which we annotated based on the expression of known marker genes:

- **Ang et al., day 3 artery ECs** (3 constituent clusters): artery EC (*EDN1*+ *MKI67*-), dividing artery EC (*EDN1*+ *MKI67+*), mesenchyme (*ACTC1*+ *TPM1*+)
- **Ang et al., day 4 vein ECs** (4 constituent clusters): vein EC *(FLRT2*+ *NEFH+)*, dividing vein EC (*FLRT2*+ *DGKB+),* other EC (*LNCAROD*+), mesenchyme (*ACTC1*+ *TPM1*+)
- **Pan et al., day 6 artery ECs** (4 constituent clusters): EC (*EGFL7*+ *ASPM*-), dividing EC (*EGFL7*+ *ASPM*+), mesenchyme (*IGFBP3*+ *TOP2A*-), dividing mesenchyme (*IGFBP3*+ *TOP2A*+)
- **Pan et al., day 6 vein ECs** (3 constituent clusters): EC (*KDR*+ *TOP2A*-), dividing EC (*KDR*+ *TOP2A*+), mesenchyme (*ACTC1*+)
- **Day 6 Zhang et al.** (5 constituent clusters): EC (*CLDN5*+ *UBE2C*-), dividing EC (*CLDN5*+ *UBE2C*+), mesenchyme (*MEST*+ *ACTC1*+), blood-like (*SPI1/PU.1*+), heart/kidney (*NKX2.5*+ *LHX1*+)
- **McCracken et al., day 7 ECs** (3 constituent clusters): EC (*PLVAP*+ *MKI67*-), dividing EC (*PLVAP*+ *MKI67*+), mesenchyme (*MEST*+ *TAGLN*+)
- **Paik et al., day 12 ECs** (5 constituent clusters): EC (*ECSCR*+), mesenchyme (*LUM*+ *CENPF*-), proliferating mesenchyme (*LUM*+ *CENPF*+), liver-like (*FGB+ APOA2*+ *TTR*+ *AFP*+), unknown (co-expression of epithelial marker *EPCAM* and mesenchymal marker *LUM*)
- **Nikolova et al., day 14 ECs** (3 constituent clusters): EC (*CLDN5*+), mesenchyme (*LUM*+ *TOP2A*-), dividing mesenchyme (*LUM*+ *TOP2A*+)

Assessing arteriovenous identity of cells generated from each hPSC differentiation protocol, using single-cell RNA-seq

We quantified the arteriovenous identity of hPSC-derived ECs that were profiled by aforementioned scRNAseq studies. Each dataset comprises a mixture of ECs and non-ECs at varying proportions. First, we computationally selected the EC cluster generated by each differentiation protocol, as per the above cluster annotations. (For differentiation protocols that generated "EC" and "dividing EC" clusters, all EC clusters were combined.) Then, in these EC populations obtained from distinct differentiation protocols, we analyzed the expression of arterial vs. venous marker gene modules, which are referred to as the "artery signature" or

"vein signature" in this study. The Hou et al., 2022 study[21] previously demonstrated that the expression of these arteriovenous gene modules is evolutionarily conserved across ECs obtained from both human and mouse embryos:

- **Arterial gene module**: *GJA4*, *UNC5B*, *DLL4*, *MECOM*, *HEY1*, *EFNB2*, *EPAS1*, *CXCR4*, *IGFBP3*
- **Venous gene module**: *NR2F2*, *NRP2*, *APLNR*, *FLRT2*

We implemented the AddModuleScore function of Seurat v3[136] to calculate the average expression of these arterial module genes and venous module genes in hPSC-derived ECs generated from each differentiation protocol.

Single-cell RNA-seq analysis of endothelial heterogeneity in human and mouse embryos
We downloaded scRNAseq datasets of ECs isolated from the E9.5 mouse embryo (generated by Chen et al., 2024[65]; NCBI GSE216970) or the Carnegie Stage 12 human embryo (generated by Calvanese et al., 2022[66]; NCBI GSE162950, sample GSM4968831). Computational analysis was performed as described above, with the exception that Seurat v4[149] was used.

OmniATAC-seq library construction
We performed OmniATAC-seq profiling of hPSC-derived day 3 artery ECs (FACS-purified for $CD144^+$ $DLL4^+$ $CD73^{lo/-}$ cells) and day 4 vein ECs (FACS-purified for $CD144^+$ $DLL4^-$ $CD73^{hi}$ cells). A slightly modified version (https://www.med.upenn.edu/kaestnerlab/assets/user-content/documents/ATAC-seq-Protocol-(Omni)-Kaestner-Lab.pdf) of the original OmniATAC-seq protocol[54] was employed here, and is briefly summarized below.

First, we prepared resuspension buffer (10 mM Tris-HCl, pH 7.5 + 10 mM NaCl + 3 mM $MgCl_2$ + nuclease-free H_2O), cold lysis buffer (resuspension buffer + 0.1% v/v NP-40 + 0.1% v/v Tween-20 + 0.01% v/v digitonin), and wash buffer (99.9% resuspension buffer + 0.1% v/v Tween-20).

50,000 cells from each hPSC-derived cell-type were pelleted and washed with 500 µL of cold PBS, before lysis in 100 µL cold lysis buffer for 3 minutes on ice. To the cell lysate, 1 mL of cold wash buffer was added. The mixture was centrifuged at 500g for 10 minutes at 4 °C. Then, the supernatant (cytoplasm) was discarded and the pellet (nuclei) was retained. Transposition reaction mix from the Nextera DNA library prep kit (1x Tagment DNA [TD] Buffer + 1x PBS + 0.1% v/v Tween-20 + 0.01% v/v Digitonin + Tn5 Transposase [Tagment DNA Enzyme 1] + nuclease-free H_2O) was added to the pellet to resuspend nuclei. The transposition reaction was incubated at 37 °C for 30 minutes on a thermal mixer, with shaking at 1,000 rpm.

DNA was purified using the Qiagen MinElute Reaction Cleanup Kit, and then PCR-amplified using Illumina i5 and i7 index primers on a thermal cycler. Then, the libraries were purified using AMPure XP beads. All OmniATAC-seq libraries were multiplexed such that they could be sequenced as a pool on a single lane. After quality control was performed on pooled libraries, deep sequencing was performed on an Illumina HiSeq sequencer (~350 million reads/lane) and an Illumina NovaSeq S4 sequencer (~2500 million reads/lane). According to general guidelines, a minimum of 50 million reads are needed to identify accessible chromatin elements and 200 million reads are needed to identify enriched transcription factor motifs

(https://www.illumina.com/techniques/popular-applications/epigenetics/atac-seq-chromatin-accessibility.html) by ATAC-seq. In this study, OmniATAC-seq libraries were sequenced at a depth of 292-1908 million raw reads per library.

OmniATAC-seq computational analysis

OmniATAC-seq data were computationally processed using the standardized ENCODE ATAC-seq analysis pipeline (https://www.encodeproject.org/atac-seq/). First, reads were aligned to human reference genome hg38 using Bowtie2[184]. Then, MACS2[185] was used to call peaks for each library. A unified peak list for each cell-type was generated by selecting only peaks that were reproducible between the two replicates. This was achieved through an irreproducible discovery rate (IDR) analysis at the threshold of 0.05 described by the ENCODE Consortium[186]. Finally, peaks that overlapped with a "black list" of artifactual regions in hg38 (https://sites.google.com/site/anshulkundaje/projects/blacklists) were removed.

Diffbind[170] was used to identify chromatin regions that exhibited >8-fold differential accessibility between artery and vein ECs. HOMER[171] was used to discover DNA motifs overrepresented in these artery- or vein-accessible elements. HOMER analyses were run on repeat-masked hg38 sequences extracted in 200-nucleotide windows centered on peak summits, using the "findMotifsGenome.pl" function.

CUT&RUN library construction

CUT&RUN profiling of H3K4me1, H3K4me3, H3K27me3, and SOX17 was conducted on H1 hPSCs differentiated into day 3 artery ECs, day 3 pre-vein ECs, and day 4 vein ECs. For most CUT&RUN experiments, CD144+ ECs were purified by FACS to exclude any contaminating mesenchymal cells, with the exception for SOX17 CUT&RUN, as SOX17 is not expressed in mesenchymal cells. As a negative control, an isotype IgG control antibody was also separately included.

CUT&RUN profiling and library construction was performed as previously described[78]. CUT&RUN libraries were sequenced on an Illumina NovaSeq X Plus by Novogene.

CUT&RUN computational analysis

FastQC[155] and FastQ Screen[164] were used to perform quality control of raw CUT&RUN reads. Paired-end reads were merged using NGmerge[165] and aligned to human reference genome hg38 with Bowtie2[64], using the --very-sensitive mode and fragment length parameters -I 10 -X 2000. Resulting SAM files were processed with SAMBLASTER[166] to remove PCR duplicates. SAMtools[167] was then used to convert files into the sorted BAM format. Bigwig signal tracks were generated using the "bamCoverage" function from deepTools[168], with 5-nucleotide bin size and values displayed in reads per kilobase million (RPKM).

High-throughput surface marker screen by flow cytometry

The expression of 332 cell-surface markers was assessed across undifferentiated hPSCs (day 0), primitive streak (day 1), lateral mesoderm (day 2), artery ECs (day 3) and vein ECs (day 4) through the use of high-throughput flow cytometry as described previously[182]. In brief, hPSCs or their differentiated mesoderm progeny were dissociated using TrypLE Express. They were then plated into individual wells of four 96-well LEGENDScreen PE-Conjugated Human

Antibody Plates (Biolegend, 700001). Each well containing a distinct antibody against a human cell-surface antigen, altogether totaling 332 unique cell-surface markers across four 96-well plates. High-throughput cell-surface marker staining was largely done as per the manufacturer's recommendations, and cells were stained with a viability dye (DAPI) prior to robotically-enabled plate-based analysis on an BD FACSCanto II (Stanford Stem Cell Institute FACS Core). Stained cells were not fixed prior to FACS analysis. LEGENDScreen data for undifferentiated H7 hPSCs (day 0) and H7-derived anterior primitive streak (day 1) were published previously[182]. LEGENDScreen data for H1-derived lateral mesoderm (day 2), H1-derived artery ECs (day 3) and H1-derived vein ECs (day 4) was generated in this study. Day 3 artery ECs and day 4 vein ECs were both co-stained with an anti-CD144 Alexa Fluor 647 antibody (BD Biosciences, 561567) to identify CD144+ ECs, and surface-marker expression was evaluated specifically in the CD144+ population.

Assembling CRISPRi constructs

sgRNAs targeting the human *SOX7*, *SOX17*, or *SOX18* genes were selected from genome-wide libraries of CRISPRi sgRNAs[187].

- *SOX7* CRISPRi sgRNA 1: TCGCCTCGCTTCGCCTGGCG
- *SOX7* CRISPRi sgRNA 2: GAAGCGAGGCGACCCGCGTG
- *SOX17* CRISPRi sgRNA: GCGACAGGCCAGAACACGGG
- *SOX18* CRISPRi sgRNA: GCGGATGGCGGTGGGGACGG

To prepare sgRNA inserts, we synthesized the following oligonucleotides in preparation for introduction into the single sgRNA plasmid (harboring BstXI and BmtI restriction sites):

- 5' ttg + top strand sgRNA + gtttaagagc 3'
- 5' ttagctcttaaac + bottom strand sgRNA + caacaag 3'

To prepare sgRNA inserts, we synthesized the following oligonucleotides in preparation for introduction into the dual sgRNA plasmid (harboring BstXI, BsmBI, and BlpI restriction sites):

- 5' ttg + position A top strand sgRNA + tctca 3'
- 5' ctcttgaga + position A bottom strand sgRNA + caacaag 3'
- 5' gaaaggag + position B top strand sgRNA + gtttaagagc 3'
- 5' ttagctcttaaac + position B top strand sgRNA + ctcc 3'

These oligonucleotides were annealed using 20 μL of each oligonucleotide at 100 μM concentration + 10 μL of 10x annealing buffer (100 μM Tris HCl (pH 7.5), 500 mM NaCl, 10 mM EDTA (pH 8.0)) + 50 μL water. Annealing was performed at 99 °C for 5 minutes and then brought to 25 °C for 5 minutes in a thermal cycler.

To individually knock down each of these genes, we cloned each individual sgRNA into a separate *mU6-sgRNA; EF1A-HygroR* plasmid using the Quick Ligation Kit (New England Biolabs, M2200), generally using the manufacturer's protocol but with modified reaction volumes: 0.33 μL digested plasmid at 25 ng/μL concentration, 1.67 μL 2x quick ligase buffer, 0.166 μL quick ligase, and 1.33 μL annealed sgRNA at 20 nM concentration per reaction.

In experiments where we sought to simultaneously knockdown all three genes, we assembled a dual sgRNA plasmid using a previously-described strategy[146]. A dual sgRNA plasmid was assembled from two backbone fragments (Addgene, 187243 and 187239, respectively). We assembled a *mU6-SOX7-sgRNA 1; hU6-SOX17-sgRNA; EF1A-PuroR-2A*-GFP plasmid using the Quick Ligation Kit (New England Biolabs, M2200). The manufacturer's protocol was generally followed, but modified reaction volumes were used: 0.33 μL of backbone plasmid 1 at 25-33 ng/μL concentration, 0.33 μL of backbone plasmid 2 at 25-33 ng/μL concentration, 1.67 μL 2x quick ligase buffer, 0.166 μl quick ligase, 0.5 μL of annealed sgRNA 1 at 200 nM concentration per reaction, and 0.5 μL of annealed sgRNA 2 at 200 nM concentration per reaction. As described below, to achieve triple *SOXF* knockdown, we transduced CRISPRi hPSCs with a *mU6-SOX18-sgRNA*; *EF1A-HygroR* plasmid, and subsequently transduced them with this dual *SOX7/SOX17* sgRNA construct.

All plasmids were transformed into Mix & Go! *E. coli* Competent Cells (Zymo Research) and purified using the Wizard Plus SV Minipreps DNA Purification System (Promega).

CRISPRi knockdown in hPSCs

CRISPRi plasmids were packaged into VSV-G pseudotyped lentiviruses in HEK293T/17 cells using a 3rd generation lentiviral packaging system as described previously[49]. 24 hours prior to transfection, HEK293T/17 cells were seeded at a density of 105,000 cells/cm^2 in 6-well plates coated with 0.01% poly-L-Lysine. Each well was transfected with 1.39 μg pMDL plasmid + 0.78 μg VSV-G plasmid + 0.53 μg pREV plasmid + 11.3 μL FuGENE HD transfection reagent + 2.1 μg sgRNA plasmid in Opti-MEM medium. 18 hours post-transfection, media was changed to DMEM with 10% FBS. Supernatant was collected 42 hours post-transfection, filtered with 0.45 μM polyethersulfone filter, and stored at -80 °C until used for transduction.

In parallel, H1 hPSCs were engineered to constitutively express CRISPR interference (CRISPRi) machinery, namely nuclease-dead Cas9 (dCas9) fused to the transcriptional repressor ZIM3 KRAB[74]. Validation of these CRISPRi-expressing H1 hPSCs will be described in a forthcoming manuscript.

To transduce them with lentiviruses carrying single sgRNAs, H1 CRISPRi hPSCs were dissociated with Accutase, and then re-seeded at 150,000 cells/well in 6-well plates in StemFlex (Thermo Fisher, A3349401) supplemented with 1% penicillin/streptomycin and 1 μM thiazovivin and lentivirus-containing supernatant (15 μL for *SOX17*-sgRNA, 10 μL for *SOX7*-sgRNA, and 50 μL for *SOX18*-sgRNA lentiviruses, respectively). After reaching 80% confluency after 2-3 days, cells were re-plated using Accutase and cultured in StemFlex with Hygromycin B (Thermo Fisher, 10687010) at 50-100 μg/ml concentration for at least 3 weeks. 1 μM thiazovivin was supplemented for the first day of selection.

To generate the triple *SOXF* knockdown line, first we transduced H1 CRISPRi hPSCs with the *mU6-SOX18-sgRNA*; *EF1A-HygroR* lentivirus and performed hygromycin selection for at least 3 weeks, as described above. These H1 CRISPRi *SOX18*-sgRNA hPSCs were then dissociated with Accutase, re-seeded at 200,000 cells/well in 12-well plates in StemFlex supplemented with 1% penicillin/streptomycin, 50 μg/ml Hygromycin B and 1 μM thiazovivin and transduced with 50 μl of lentivirus-containing supernatant. The following day, media was

replaced with StemFlex supplemented with 1% penicillin/streptomycin and 50 μg/ml Hygromycin. After 3 days, cells were re-plated using Accutase and selected with 1 μg/ml Puromycin (Thermo Fisher, A1113803) and 50 μg/mL Hygromycin B for at least 3 days. 1 μM thiazovivin was supplemented for the first day of antibiotic selection.

H1 CRISPRi hPSC lines were then further differentiated into artery and vein ECs using the V1 protocol described above. qPCR and scRNAseq was performed on control and *SOXF* knockdown artery ECs, pre-vein ECs, and vein ECs as described above.

Construction of *SOX17-2A-mPlum; NR2F2-2A-GFP* reporter hPSC line

The Cas9/AAV6 knock-in strategy[77] was used to generate a double *SOX17-2A-mPlum*; *NR2F2-2A-GFP* knock-in reporter hPSC line. We previously generated single *SOX17-2A-mPlum* or *NR2F2-2A-GFP* reporter hPSC lines[49]. In this study, starting from a *NR2F2-2A-GFP* reporter line (clone 10), we used our previously-described approach to knock-in a *2A-mPlum* cassette downstream of the endogenous *SOX17* gene such that the *SOX17* stop codon was removed[49].

Generating Andes, Ebola, and Nipah virus stocks

All experiments with Risk Group 3 and Risk Group 4 viruses were conducted under maximum containment conditions in the biosafety-level-4 (BSL4) laboratory of the Robert Koch Institute, in accordance with approved standard operating procedures. Work at the Robert Koch Institute was performed using only terminally-differentiated artery and vein ECs.

Andes virus (Andes virus/Oligoryzomys longicaudatus/CHL/1997/Chile-9717869)[99,142], Ebola virus (Ebola virus/Human/COD/1976/Yambuku-Mayinga)[141], and Nipah virus (Nipah virus/human/MY/1999/CDC)[143] were obtained from NIH and CDC reference stocks, as described above. These viruses were then subsequently propagated on Vero E6 cells (Vero (Vero 76, clone E6, Vero E6), from the European Collection of Authenticated Cell Cultures (ECACC)) in DMEM + 2% FBS to generate archive stocks, which were frozen at -80 °C. Archive stocks were thawed and passaged once more on Vero E6 to generate virus working stocks for experiments, which were stored at -80 °C.

In certain experiments, Sendai virus, Cantell strain[177]—a paramyxovirus not known to cause disease[100,188]—was used as a control. Sendai virus is classified in Risk Group 2, but was handled in the BSL4 laboratory for consistency. Sendai virus, Cantell strain was obtained from ATCC (VR-907).

Quantification of Andes and Ebola virus stocks

To determine the infectious units of Ebola and Andes virus stocks by focus assay, viral titration was performed using the fluorescent focus-forming unit (FFU) assay. In brief, Vero E6 cells were seeded into 96-well tissue culture plates (for Ebola virus titration) or 24-well plates (for Andes virus titration). Cells were seeded in DMEM containing 10% FBS, 2 mM glutamine and 1% penicillin/streptomycin, at a density such that after overnight incubation, the cells were 95-100% confluent. Separately, virus stocks were diluted in DMEM in serial 10-fold dilutions. Then, the Vero E6 culture medium was aspirated and 100 μL of diluted Ebola virus or 200 μL of diluted Andes virus was added to triplicate wells of Vero E6 cells. After 1 hour of incubation,

viral supernatants were removed and medium comprising DMEM + 2% FBS + 1% carboxymethylcellulose (CMC) was added. After 3 days for Ebola virus and 5 days for Andes virus, cells were stained with virus-specific antibodies. For this, the CMC-containing medium was removed, and cells were fixed in 10% formalin for 10-30 minutes, permeabilized in DPBS + 0.1% Triton X-100 at room temperature for 5 minutes, washed once in DPBS, and then incubated in blocking buffer containing 2% BSA for 1 hour at 37 °C.

Subsequently, 100 μL of primary antibody used to stain the cells for 1 hour at 37 °C. For Ebola virus, anti-Ebola NP antibody (clone 173/303/109, diluted 1:2500) was added. For Andes virus, anti-TULV1 N antibody (diluted 1:1000)—which is known to cross-react against Andes N protein[140]—was added. After primary antibody staining, cells were then washed and then incubated with 100 μL of secondary antibody (goat anti-mouse IgG Alexa 488, Jackson ImmunoResearch, 115-545-003, diluted 1:500-1:1000) for 30 minutes at 37 °C, washed once, and then foci of infected cells were enumerated. The titer/mL of the input viral supernatant was calculated using the Spearman-Karber method.

Quantification of Nipah virus stocks

To determine the infectious units of Nipah virus stocks by plaque assay, Vero E6 cells were seeded into 12-well tissue culture plates in DMEM containing 10% FBS, 2 mM glutamine and 1% penicillin/streptomycin, at a density that after overnight incubation, the cells were 95-100% confluent. Separately, the virus stock was diluted in DMEM in serial 2-fold dilutions. Then, the Vero E6 culture medium was aspirated, and 400 μL of diluted Ebola virus was added to triplicate wells of Vero E6 cells. After 1 hour of incubation, viral supernatant was removed and media comprising DMEM + 2% FBS + 1% carboxymethylcellulose (CMC) was added. After 3 days, plaques of infected cells were detected by crystal violet staining. In brief, the CMC overlay was removed and cells were fixed in 10% formalin for 10 minutes. Cells were then washed with PBS and stained using 0.3% crystal violet in 20% methanol/H_2O for 10 minutes. After a final wash step with PBS, the titer/mL of the input viral supernatant was calculated using the Spearman-Karber method.

Viral infection of hPSC-derived artery and vein ECs

As described above, cryopreserved H1 hPSC-derived artery and vein ECs were thawed in Geltrex-coated 48-well plates. They were cultured for 4 days in their respective growth media (EGM2 for artery ECs and EGM2 + SB505124 + RO4929097 + Forskolin for vein ECs) until they were confluent, with Thiazovivin (1 μM) added for the first 24 hours post-thaw to enhance survival. hPSC-derived artery and vein ECs were then inoculated with the following viruses, or left uninfected (negative control), for 1 hour:

- Andes virus: 2.2×10^4 FFU units/well
- Ebola virus: 4.3×10^4 FFU units/well
- Nipah virus: 2.2×10^4 PFU units/well (9.97×10^5 $TCID_{50}$ units/well)
- Sendai virus: 8 HA units/well

After 1 hour of viral infection, cells were washed twice in EGM2 medium, and then artery and vein ECs were cultured in their respective growth media (EGM2 for artery ECs and EGM2 + SB505124 + RO4929097 + Forskolin for vein ECs) in standard incubator conditions (37°C and 5% CO_2). After viral infection, ECs were profiled 6, 12, 24, or 48 hours later to quantify viral

replication, gene expression (by RNA-seq), and cytokine secretion (by ELISA), as described below.

Extensive cell death was observed in Nipah-infected artery ECs at 24 and 48 hours post-infection, as indicated in **Fig. 7B**, which was taken into consideration when interpreting experimental results from those timepoints.

Quantification of viral replication in hPSC-derived ECs

To quantify viral genomes present in the media at 6, 12, 24, and 48 hours post-infection of hPSC-derived artery and vein ECs, 140 µL of the culture media (i.e., supernatant) was placed in 560 µL of lysis buffer AVL (Qiagen QIAamp Viral RNA extraction kit, Qiagen, 52904), in addition to 560 µL of 100% ethanol, in order to inactivate it for transfer out of the BSL4 laboratory. Then, RNA was extracted following the manufacturer's instructions. Subsequently, qRT-PCR was performed using primers directed against the viral genome and the AgPath-ID One-Step RT-PCR Kit (Thermo Fisher, 4387391).

To quantify intracellular viral genomes present within the cells at 6, 12, 24, and 48 hours post-infection of hPSC-derived artery and vein ECs, the cells within each well were lysed with 350 µL RLT buffer (Qiagen RNeasy Mini Kit, Qiagen, 74106) containing 1% β-Mercaptoethanol, in addition to 350 µL of 70% ethanol, in order to inactivate materials for transfer out of the BSL4 laboratory. Afterwards, RNA was extracted following the manufacturer's instructions. Subsequently, qRT-PCR was performed using primers and FAM-labeled probes directed against the viral genome and the AgPath-ID One-Step RT-PCR Kit (Thermo Fisher, 4387391). Following PCR conditions were used: reverse transcription was performed at 45°C for 15 min, followed by initial cDNA denaturation at 95°C for 10 minutes, followed by 43 cycles with denaturation at 95°C for 15 seconds, and annealing at 60°C for 1 minute.

Andes virus	
Forward primer	AAGGCAGTGGAGGTGGAC
Reverse primer	CCCTGTTGGATCAACTGGTT
Probe	ACGGGCAGCTGTGTCTACATTGGA 6Fam-BBQ
Ebola virus	
Forward primer	GTTCGTCTCCATCCTCTTGCA
Reverse primer	TGAGGGAAAAGACCATGCTCA
Probe	TGCTCCTTTCGCCCGACTTTTGAACC 6Fam-BBQ
Nipah virus	
Forward primer	GTTCAGGCTAGAGAGGCAAAATTT
Reverse primer	CCCCTTCATCGATATCTTGATCA
Probe	CTGCAGGAGGTGTGCTCATTGGAGG 6Fam-BBQ

Quantification of cytokine secretion by virally-infected hPSC-derived ECs

At 24 or 48 hours post-infection of hPSC-derived artery and vein ECs, supernatants (culture media) were harvested, and stored at -80 °C. Media was then thawed, and the concentration of secreted interferon-β (IFNβ) protein in the 2:1 diluted media was quantified using the Human IFNβ Quantikine QuicKit ELISA Kit (R&D Systems, QK410), as described previously[49] and as per the manufacturer's instructions. Absorbance was quantified using a TECAN Sunrise plate

reader (Tecan Trading AG, Switzerland) with a spectral filter. The concentration of IFNβ in the samples was calculated by comparing the absorbance to that of a standard curve generated with human IFNβ, after subtracting background absorbance, using TECAN Magellan software. Data shown represent the average of results from three technical replicates using H1 hPSCs.

RNA-seq of virally-infected hPSC-derived ECs

At 6, 12, 24, or 48 hours post-infection of hPSC-derived artery and vein ECs, cells were lysed in 350 µL of buffer RLT from the RNeasy Micro Kit (Qiagen, 74004) and then triturated with 350 µL of 70% ethanol to inactivate infectious material for transfer out of the BSL4 laboratory. RNA was then extracted as per the manufacturer's instructions. Bulk-population RNA-seq libraries were constructed and sequenced by Novogene.

Computational analysis of RNA-seq data from virally-infected hPSC-derived ECs

Quality control of RNA-seq reads was performed using FastQC[155] and Trim Galore[156], as described in the above section "Bulk-population RNA-seq of uninfected cells". RNA-seq reads from virus-infected cells contained both human and viral reads, and two parallel workflows were used to analyze human or viral gene expression:

- **Human gene expression analysis**: Human gene expression was analyzed as described in the above section "Bulk-population RNA-seq of uninfected cells". In brief, read pseudoalignment to human reference genome hg38 and quantification of transcript-level abundances was performed using Kallisto[157]. Ensembl transcript-to-gene ID mapping (EnsDb.Hsapiens.v86[151]), gene-level count summarization (tximport[169]), and differential gene expression analysis (DESeq2[152]) was then performed.
- **Viral gene expression analysis**: To determine the overall percentage of viral reads within a given RNA-seq library, Kallisto[157] was used to pseudoalign reads to the reference genomes of Ebola virus, Yambuku variant, Mayinga isolate (NCBI accession number NC_002549.1) or Sendai virus (NCBI accession number NC_001552). This enabled the detection of all viral reads within a given sample. Then the number of total viral counts was divided over the number of total human + viral counts to estimate what proportion of the cellular transcriptome comprised viral transcripts. Viral read alignment to the Andes virus genome was not performed. We performed RNA-seq of polyadenylated mRNAs and therefore did not expect to capture Andes virus mRNAs, which are not polyadenylated.

Further RNA-seq analysis of human gene expression was conducted as follows:

- **Heatmaps**: The rlog function of DESeq2[152] was used to perform regularized log transformation. Transformed counts were used to plot heatmaps of gene expression using pheatmap[172].
- **Visualization of timecourse gene expression**: Expression counts normalized by the DESeq function of DESeq2[152] were used to create line graphs of timecourse gene expression.
- **Differential gene expression analysis**: Genes that were differentially expressed between samples were determined using the lmFit linear modeling tool provided by limma[160].
- **Artery- vs. vein-specific transcriptional responses to viral infection**: Genes with (1) $\log_2 > 0$ expression and (2) adjusted $P<0.05$ in virus-infected artery ECs relative to virus-infected vein ECs, and (3) no statistically significant (adjusted $P>0.05$) expression between uninfected artery vs. vein ECs were categorized as genes with artery-specific induction upon

viral infection. The reciprocal was used to define genes with vein-specific induction upon viral infection.

- **Conserved artery- vs. vein-specific transcriptional responses to viral infection**: Finally, we determined genes that were consistently induced by viral infection in artery ECs, but not vein ECs, and *vice versa*. From the above list of genes with artery-specific induction upon viral infection (described in the above section, "Artery- vs. vein-specific transcriptional responses to viral infection"), we selected genes with greater than $\log_2$ 1.5-fold (i.e., $2^{1.5}$-fold) higher expression in infected vs. uninfected cells at 2 days post-infection. We then determined overlapping genes, e.g., genes that were consistently upregulated in artery ECs by Ebola, Sendai, and Andes infection. The same analysis was performed on vein ECs, to discover genes that were consistently upregulated in vein ECs by Ebola, Sendai, and Andes infection. These genes were enumerated in Venn diagrams. Nipah was not included in this analysis owing to extensive cell death within 24-48 hours of viral infection.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical tests are specified in the figure legend accompanying each experimental result. Unpaired t-tests were used to test for statistical significance in qPCR and ELISA data, and P values were reported. Wilcoxon rank sum tests were performed to test for statistical significance in module score differences quantified by scRNAseq, and P values were reported. The following convention was used in figures: not significant (n.s., $P>0.05$), $*P<0.05$, and $**P<0.01$.

Q-values were reported for differentially-expressed genes detected by bulk-population RNA-seq, differentially-accessible genomic loci detected by Omni-ATAC-seq, and transcription factor motif enrichments detected by Omni-ATAC-seq. The q-value represents the P value adjusted for multiple hypothesis testing using the Benjamini-Hochberg method to control the false discovery rate.

Quantification of mouse embryo images was performed as described in the above section, "Quantification of mouse embryo images".