

# Transcriptional and epigenomic landscapes of CNS and non-CNS vascular endothelial cells

Mark F Sabbagh<sup>1,2</sup>, Jacob S Heng<sup>1,2</sup>, Chongyuan Luo<sup>3,4</sup>, Rosa G Castanon<sup>3</sup>, Joseph R Nery<sup>3</sup>, Amir Rattner<sup>1</sup>, Loyal A Goff<sup>2,5</sup>, Joseph R Ecker<sup>3,4</sup>, Jeremy Nathans<sup>1,2,6,7\*</sup>

<sup>1</sup>Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, United States; <sup>2</sup>Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, United States; <sup>3</sup>Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, United States; <sup>4</sup>Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, United States; <sup>5</sup>Institute for Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, United States; <sup>6</sup>Department of Ophthalmology, Johns Hopkins University School of Medicine, Baltimore, United States; <sup>7</sup>Howard Hughes Medical Institute, Johns Hopkins University School of Medicine, Baltimore, United States

**Abstract** Vascular endothelial cell (EC) function depends on appropriate organ-specific molecular and cellular specializations. To explore genomic mechanisms that control this specialization, we have analyzed and compared the transcriptome, accessible chromatin, and DNA methylome landscapes from mouse brain, liver, lung, and kidney ECs. Analysis of transcription factor (TF) gene expression and TF motifs at candidate *cis*-regulatory elements reveals both shared and organ-specific EC regulatory networks. In the embryo, only those ECs that are adjacent to or within the central nervous system (CNS) exhibit canonical Wnt signaling, which correlates precisely with blood-brain barrier (BBB) differentiation and *Zic3* expression. In the early postnatal brain, single-cell RNA-seq of purified ECs reveals (1) close relationships between veins and mitotic cells and between arteries and tip cells, (2) a division of capillary ECs into vein-like and artery-like classes, and (3) new endothelial subtype markers, including new validated tip cell markers.

DOI: <https://doi.org/10.7554/eLife.36187.001>

\*For correspondence:  
jnathans@jhmi.edu

Competing interest: See  
[page 34](#)

Funding: See [page 34](#)

Received: 23 February 2018

Accepted: 21 August 2018

Published: 06 September 2018

Reviewing editor: Elisabetta Dejana, FIRC Institute of Molecular Oncology, Italy

© Copyright Sabbagh et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Introduction

In vertebrates, the vascular and immune systems constitute the only classes of cells that come into close proximity to nearly all other cell types in the body. These systems also share the property that a small number of basic cell types are further divided into functionally and molecularly distinct subtypes, some of which are specialized for the tissues in which they reside. For example, in the immune system, phagocytic cells are represented by monocytes in the blood, Kupffer cells in the liver, microglia in the central nervous system (CNS), Langerhans cells in the skin, and osteoclasts in bone. In the vascular system, organ-specific molecular and cellular specialization is typically at the level of microvascular (capillary) endothelial cells (ECs) (Aird, 2007a; Aird, 2007b; Potente and Mäkinen, 2017). Additional specializations, which apply across all tissues, distinguish arteries, veins, and capillaries.

Liver and brain capillaries exemplify two extremes of EC specialization. The highly permeable sinusoidal vascular endothelium of the liver acts as a sieve in which ECs form a discontinuous lining to allow efficient diffusional exchange between serum and hepatocytes. Hepatic capillaries also

exhibit one of the body's highest capacities for endocytosis through the expression of various classes of scavenger receptors (Aird, 2007b; Knolle and Wohleber, 2016; Poisson et al., 2017). In contrast, the vast majority of capillaries in the CNS have extremely low permeability. CNS ECs are characterized by reduced transcytosis, an absence of fenestrations, tight junctions between neighboring cells, and a variety of small molecule transporters and efflux pumps (Zhao et al., 2015). Together, these properties constitute the blood-brain barrier (BBB).

The widespread functional, morphological, and molecular heterogeneity of ECs must ultimately reflect different gene expression programs. The BBB program represents a particularly interesting case, as it involves the coordinated regulation of a large number of genes (Daneman et al., 2010). Early transplantation studies showed that the BBB properties of CNS ECs are induced by developing neural tissue (Stewart and Wiley, 1981), implicating an instructive signal from surrounding parenchyma. More recent work has shown that WNT and NORRIN ligands produced by neurons and glia activate receptors on ECs, and the resulting canonical Wnt signal initiates and maintains the BBB state (Liebner et al., 2008; Stenman et al., 2008; Daneman et al., 2009; Wang et al., 2012; Zhou et al., 2014; Cho et al., 2017). Interestingly, the same canonical Wnt signal acts earlier in vascular development to direct CNS angiogenesis.

Most molecular studies of EC heterogeneity have focused on defining transcriptome differences between EC subtypes (Chi et al., 2003; Daneman et al., 2010; Nolan et al., 2013; Hupe et al., 2017). Several recent studies have also investigated the enhancer landscape of ECs (Ernst et al., 2011; Hogan et al., 2017; Quillien et al., 2017; Zhou et al., 2017). However, there have been no genome-wide analyses that integrate and compare the transcriptome, chromatin structure, and DNA methylation landscapes of ECs. Accessible chromatin and genomic regions with low levels of CG methylation are of particular interest as they mark candidate *cis*-regulatory elements (CREs) (Stadler et al., 2011; Thurman et al., 2012; Stergachis et al., 2013; Ziller et al., 2013). Importantly, CG hypomethylation can also provide a record of a cell's developmental history, offering an opportunity to identify lineage-specific regulators (Hon et al., 2013; Mo et al., 2015; Mo et al., 2016).

Here, we explore transcriptome, chromatin, and DNA methylation differences among four populations of murine ECs to determine the factors that are associated with EC heterogeneity. We uncover unique sets of candidate CREs that define different EC subtypes, and we identify candidate transcription factor (TF) regulators using motif enrichment analysis of these regions. We provide evidence for precisely spaced and oriented ETS and ZIC TF motifs that likely play a role in defining the enhancer landscape of brain ECs. To explore intra-tissue EC heterogeneity, we apply RNA profiling of brain ECs at the single-cell level, leading to a molecular definition of the relationship among the major CNS EC subtypes. Together, these studies provide an integrated view of the genetic and epigenetic mechanisms underlying the unique tissue-specific programs of EC differentiation.

## Results

### Tissue-specific endothelial cell gene expression

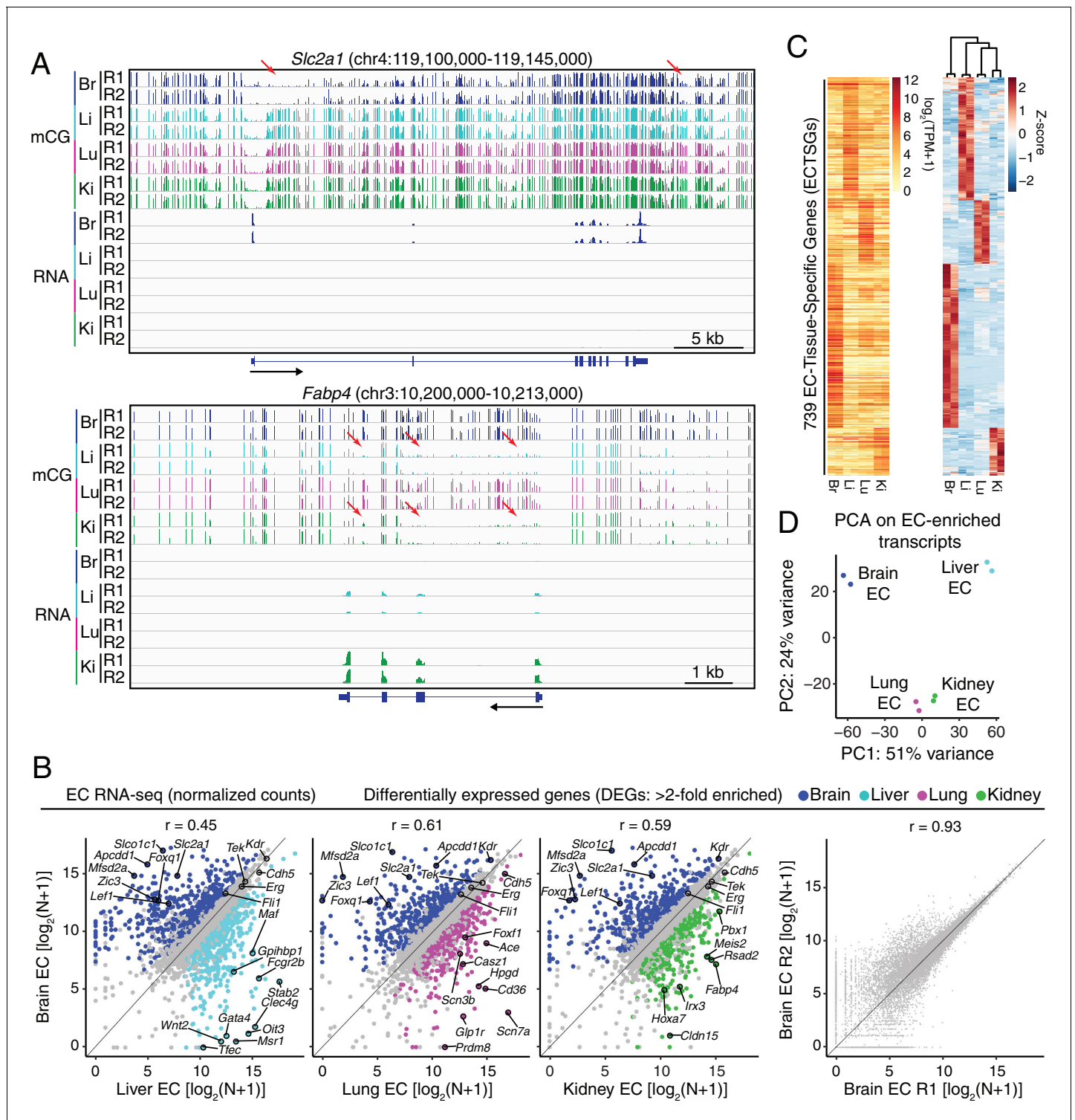
To characterize the genetic basis of inter-tissue heterogeneity among ECs, we set out to acquire a broad view of their genomic and epigenomic differences. Toward this end, we isolated ECs from four different murine tissues - brain, liver, lung, and kidney - at postnatal day 7 (P7) and performed RNA-seq, MethylC-seq (Lister et al., 2008), and ATAC-seq (Buenrostro et al., 2013) to generate an atlas of transcript abundances, single-base resolution DNA methylation, and accessible chromatin, respectively (Supplementary file 1). The data are available for public exploration as the Vascular Endothelial Cell Trans-omics Resource Database (VECTRDB) at <https://marksabbagh.shinyapps.io/vectrdb/> and [https://github.com/mfsabbagh/EC\\_Genomics](https://github.com/mfsabbagh/EC_Genomics). To visualize the data on a genome browser, see [http://neomorph.salk.edu/Endothelial\\_cell\\_methylome.php](http://neomorph.salk.edu/Endothelial_cell_methylome.php). At P7, the vasculature is still growing but it has also, presumably, acquired tissue-specific properties to support organ function. To isolate ECs, we harvested organs from *Tie2-GFP* (the *Tie2* gene is also known as *Tek*) mice, which specifically express green fluorescent protein (GFP) in ECs (Motoike et al., 2000), enzymatically dissociated the tissues into single-cell suspensions, and then purified ECs using fluorescence-activated cell sorting (FACS) (Figure 1—figure supplement 1A). This approach to EC isolation was optimized and validated by Daneman et al., 2010). We note that *Tie2-GFP* is not expressed in all

ECs. For example, we could not detect GFP in the capillaries of renal glomeruli by immunostaining (yellow arrows in **Figure 1—figure supplement 1B**). Each high-throughput sequencing experiment was performed on two or more independent biological replicates, and these exhibited high pairwise correlations (**Figure 1A** and far right panel of **Figure 1B**; **Figure 1—figure supplements 2, 3 and 4A**; **Figure 2A**; **Figure 2—figure supplement 1A**).

To filter out non-EC transcripts that derive from lysed parenchymal cells, we also conducted RNA-seq on GFP-negative FACS-sorted cells, and, for completeness, total dissociated tissue. By comparing the expression profiles of GFP-positive cells to GFP-negative cells, we determined a union set of 4117 protein-coding and non-coding EC-enriched transcripts from each tissue for further analysis (**Supplementary file 2**; see the RNA-seq section under Materials and methods for the three criteria applied in this comparison). Importantly, this list contains many known EC markers, including *Pecam1*, *Kdr*, *Cdh5*, and *Tek* (also known as *Tie2*). Notably absent from the list are the pan-leukocyte marker *Ptprc/Cd45*, the macrophage/microglia marker *Csf1r*, the vascular smooth muscle cell marker *Acta2*, the pericyte marker *Cspg4*, and various abundant parenchyma-specific markers such as *Syt11* (brain), *Alb* (liver), *Sftpc* (lung), and *Umod* (kidney) (**Figure 1—figure supplement 2B**). The GFP-negative fraction of dissociated liver was unexpectedly depleted for hepatocyte-specific transcripts suggesting that hepatocytes were lost during cell dissociation and/or flow cytometry. This fraction was enriched for transcripts associated with Kupffer cells (*Cd68*, *Cd80*, and *Cd86*) and cholangiocytes (*Epcam*, *Pkhd11*, and *St14*; *Li et al., 2017*) (**Figure 1—figure supplement 3A**). Based on the retention of known EC-specific transcripts and the removal of known parenchyma-specific transcripts, we conclude that the EC-enrichment criteria selected for bona fide EC-specific transcripts.

Focusing only on the 3853 protein-coding EC-enriched transcripts, scatter plots of between-sample normalized RNA-seq read counts for ECs from one tissue versus the three other EC subtypes show both previously identified and novel tissue-specific EC markers (**Figure 1B**; **Figure 1—figure supplement 4B**). In this analysis, the relatively low correlation between different tissue EC transcriptomes reflects the exclusion of housekeeping genes. Brain ECs specifically express mRNAs coding for transporters such as *Mfsd2a*, *Slc2a1*, and *Slco1c1* (**Figure 1B**, first three panels), whereas liver sinusoidal ECs specifically express mRNAs coding for scavenger receptors such as *Fcgr2b*, *Stab2*, and *Clec4g* (**Figure 1B**, first panel). Similarly, mRNA coding for angiotensin-converting enzyme (*Ace*), a known lung EC marker, was specifically enriched in lung ECs (**Figure 1B**, second panel). The differential expression analysis also identified two novel lung EC markers, *Scn7a* and *Scn3b* (**Figure 1B**; **Figure 1—figure supplements 3B and 4B**). The former is a sodium channel referred to as Na(x) that is responsive to extracellular levels of sodium and that was previously reported to be expressed in lung epithelium (*Watanabe et al., 2000*; *Watanabe et al., 2002*; *Xu et al., 2015*), while the latter is an auxiliary beta subunit that can modify the conductance of sodium channel alpha subunits (*Cusdin et al., 2010*). *Ctfr*, the cystic fibrosis transmembrane conductance regulator, and surfactant proteins *Sftpa1*, *Sftpb*, *Sftpc*, and *Sftpd* exhibit enrichment in the lung GFP-negative fraction, consistent with a known lung epithelial role for these genes (**Figure 1—figure supplement 3B**). In contrast, both *Scn7a* and *Scn3b* show a clear enrichment in lung ECs on par with *Ace* enrichment, implying that they function in ECs rather than in the epithelium (**Figure 1—figure supplement 3B**). We speculate that expression of *Scn7a* and *Scn3b* may allow lung ECs to directly sense serum sodium concentration, perhaps to regulate *Ace* expression and thereby provide feedback control of the renin-angiotensin-aldosterone system.

Transcripts from 2392 genes exhibited >2 fold differential abundance between one or more pairs of EC subtypes, with 739 exhibiting >2 fold differential expression in a comparison of ECs from one tissue versus ECs from each of the other three tissues (**Figure 1C**; **Supplementary file 2**; see Materials and methods). We will refer to these 739 genes as 'EC-tissue-specific genes' or 'ECTSGs'. For reference, this abbreviation and those that follow are summarized and defined in **Table 1**. Of the 739 ECTSGs, 685 are protein-coding and 54 are non-coding. Gene ontology (GO) enrichment analysis (*Ashburner et al., 2000*; *The Gene Ontology Consortium, 2017*; *The Gene Ontology Consortium, 2017*) of the four sets of ECTSGs was consistent with tissue-specific specializations of each vascular bed. For example, brain ECTSGs were enriched for GO terms such as 'neutral amino acid transport' and 'drug transmembrane transport' whereas liver ECTSGs were enriched for GO terms such as 'scavenger receptor activity' and 'vesicle-mediated transport' (**Supplementary file 3**). Consistent with the known role of Wnt signaling in brain endothelium development, brain ECTSGs were also



**Figure 1.** RNA-seq reveals inter-tissue EC heterogeneity. (A) Genome browser images showing CG methylation (top) and RNA expression (bottom) for two genes: *Slc2a1*, a glucose transporter expressed in brain ECs, and *Fabp4*, a fatty acid binding protein expressed in liver and kidney ECs. For DNA methylation, the mCG/CG ratio is shown, with the height of each bar indicating the fractional methylation (range: 0 to 1). For RNA-seq, histograms of the number of aligned reads are shown. For this and all other genome browser images, the heights of all the tracks of a given sequencing experiment are the same across samples. For both genes, tissue-specific gene expression is associated with tissue-specific hypomethylation near the TSS. Red arrows indicate illustrative examples of differential hypomethylation. Br, brain; Li, liver; Lu, lung; Ki, kidney. R1 and R2, biological replicates. Black arrows beneath this and all other genome browser images indicate the direction of transcription. (B) Scatter plots comparing cross-sample normalized RNA-seq counts for Brain ECs (y-axis) and Liver ECs (x-axis), Lung ECs (x-axis), and Kidney ECs (x-axis). The correlation coefficients (r) are shown above each plot. (C) Heatmap showing the expression of 739 EC-tissue-specific genes (ECTSGs) across four tissues (Brain, Liver, Lung, Kidney). The color scale represents the Z-score of the expression. (D) PCA plot showing the first two principal components (PC1 and PC2) of the EC-enriched transcripts. The variance explained by PC1 and PC2 is 51% and 24%, respectively. The plot shows that Brain ECs and Liver ECs cluster together, while Lung ECs and Kidney ECs cluster together.

Figure 1 continued on next page

Figure 1 continued

seq read counts of EC-expressed protein-coding genes from brain versus liver, lung, and kidney, showing only those transcripts with TPM >10 for each of the two RNA-seq replicates. Colored symbols indicate transcripts with FDR < 0.05 and enrichment >2 fold for the indicated tissue comparison. Right, comparison of cross-sample normalized RNA-seq read counts for protein-coding genes between the two brain EC replicates. Values depicted are the log<sub>2</sub> transformation of cross-sample normalized counts + 1. (C) Heatmaps depicting transcript abundances for 739 Endothelial Cell Tissue-Specific Genes (ECTSGs). Left, log<sub>2</sub> transformation of TPM + 1. Right, z-scores for the TPMs. (D) Principal component analysis of all EC-enriched transcripts from brain, liver, lung, and kidney. The two symbols for each sample represent biological replicates. In this and all other figures, tissue type is indicated by color: Br, brain, blue; Li, liver, cyan; Lu, lung, magenta; Ki, kidney, green.

DOI: <https://doi.org/10.7554/eLife.36187.002>

The following figure supplements are available for figure 1:

**Figure supplement 1.** *Tie2-GFP* transgenic mouse enables isolation of ECs.

DOI: <https://doi.org/10.7554/eLife.36187.003>

**Figure supplement 2.** GFP-positive FACS-sorted cells from P7 *Tie2-GFP* mice represent pure populations of ECs.

DOI: <https://doi.org/10.7554/eLife.36187.004>

**Figure supplement 3.** Expression of cell-type-specific transcripts in liver and lung samples.

DOI: <https://doi.org/10.7554/eLife.36187.005>

**Figure supplement 4.** RNA-seq comparisons between peripheral ECs.

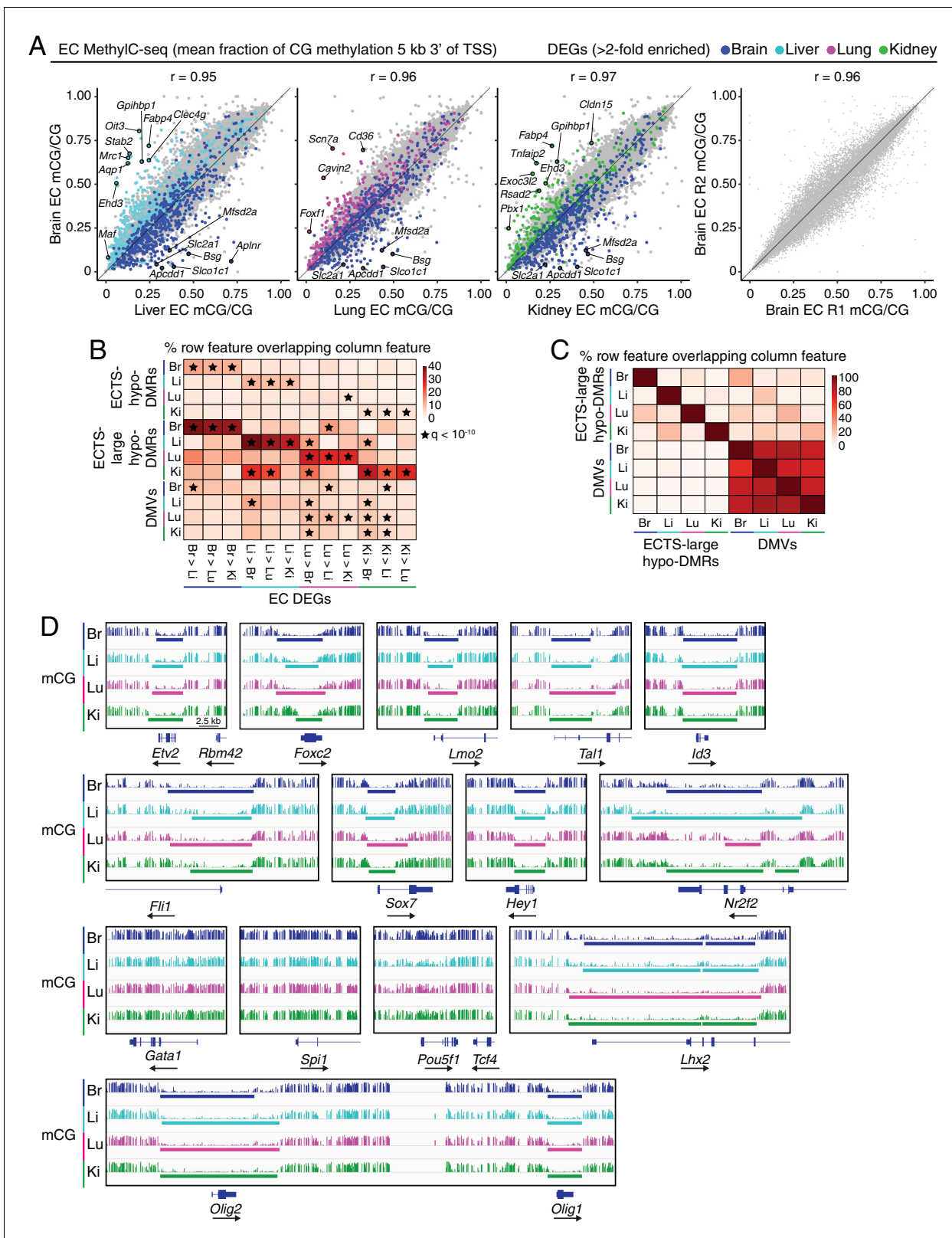
DOI: <https://doi.org/10.7554/eLife.36187.006>

enriched for the GO term 'negative regulation of canonical Wnt signaling pathway' (**Supplementary file 3**). Principal component analysis (PCA) of the four EC transcriptomes showed that the transcriptional signatures of brain ECs and liver ECs are substantially distinct from each other and from those of lung ECs and kidney ECs (**Figure 1D**), broadly consistent with earlier studies of embryonic and adult ECs (**Daneman et al., 2010; Hupe et al., 2017**).

## Identification of distinct classes of hypomethylated regions in endothelial cells

Low cytosine DNA methylation (mCG) immediately downstream of transcriptional start sites (TSSs) typically correlates with increased gene expression (**Schultz et al., 2015**). In visually comparing EC methylation profiles to transcription profiles, we find that differentially expressed genes are often associated with tissue-specific hypomethylation starting at the TSS and extending into the gene body (**Figure 1A**). As a result, plotting the mean fraction of mCG methylation in a five kilobase (kb) window downstream of the TSS for all protein-coding genes reveals an inverse relationship between methylation and gene expression among differentially expressed EC-enriched genes (**Figure 2A; Figure 2—figure supplement 1B**). However, these differences in methylation at differentially expressed genes occur in the context of highly similar global methylation patterns, as seen by correlation coefficients between different EC subtypes (0.95–0.97) that are nearly indistinguishable from the correlation coefficients between biological replicates (0.95–0.98; **Figure 2A; Figure 2—figure supplement 1A**).

The distinctive tissue-specific patterns of ECTSGs imply that ECs regulate transcription with a corresponding set of tissue-specific CREs. [Although the use of the term 'CRE' implies that these non-coding genomic regions influence nearby genes (i.e. act in cis), we recognize that some of these genetic elements may act between chromosomes (i.e. in trans).] To identify candidate EC CREs based on patterns of DNA methylation, we first parsed individual EC DNA methylomes into unmethylated regions (UMRs; hypomethylated, CG-rich) and low-methylated regions (LMRs; hypomethylated, CG-poor). These criteria were based on those of **Stadler et al. (2011)** and **Burger et al., 2013**, with some modifications (e.g. methylation threshold  $m < 0.3$ ; see Materials and methods) to account for a relatively low level of global CG methylation (mean of 67%). We identified 16,158 brain EC, 16,506 liver EC, 16,171 lung EC, and 15,979 kidney EC UMRs (median length = 1738 bp; median mCG = 10%), and 72,061 brain EC, 60,212 liver EC, 57,594 lung EC, and 71,343 kidney EC LMRs (median length = 490 bp; median mCG = 15%) (**Supplementary file 4**). 72% of the UMRs (12,581 of 17,406) are within 2 kb of a TSS whereas only 5.7% of the LMRs (6682 of 116,273) are within 2 kb of a TSS (**Figure 2—figure supplement 1C**), consistent with an earlier study showing that UMRs are enriched at promoters and LMRs are enriched at distal regulatory elements (**Stadler et al., 2011**).



**Figure 2.** MethylC-seq reveals distinct classes of hypomethylated regions in ECs. (A) Scatter plots comparing the mean fraction of CG methylation in a 5 kb window immediately 3' of the TSS for protein-coding genes from brain ECs versus liver, lung, and kidney ECs. Colored symbols indicate transcripts with FDR < 0.05 and enrichment >2 fold for the indicated tissue comparison. These plots show a positive correlation between EC tissue-specific gene expression and tissue-specific hypo-methylation. Right, comparison of CG methylation between the two brain EC biological replicates. (B) Heatmap

Figure 2 continued on next page

Figure 2 continued

indicating the percentage of each row that overlaps with differentially expressed genes. A significant proportion of ECTS-large hypo-DMRs overlap differentially expressed genes for the same EC-subtype relative to the other EC subtypes. Black stars indicate statistical significance at  $q < 1 \times 10^{-10}$ . (C) Heatmap indicating the percentage of each row that overlaps with either ECTS-large hypo-DMRs or DMVs. DMVs exhibit more overlap between EC subtypes than large hypo-DMRs. (D) Genome browser images showing methylation as in **Figure 1A** at various TF genes. Colored bars indicate DMVs. Each genome browser image is at the same scale.

DOI: <https://doi.org/10.7554/eLife.36187.007>

The following figure supplements are available for figure 2:

**Figure supplement 1.** Distinct classes of hypomethylated features in ECs.

DOI: <https://doi.org/10.7554/eLife.36187.008>

**Figure supplement 2.** DMVs at differentially expressed TF genes exhibit differential methylation.

DOI: <https://doi.org/10.7554/eLife.36187.009>

86% of UMRs (15,036 of 17,406) are shared by all four EC subtypes, whereas only 25% of LMRs (29,053 of 116,273) are shared by all four EC subtypes (**Figure 2—figure supplement 1D**).

We next searched for regions of the genome that were differentially hypomethylated in one or more pairwise comparisons among the four EC subtypes. Unlike the enumeration of UMRs/LMRs, which is an intrinsic property of each DNA methylome, this analysis relies on a comparison between DNA methylomes. Using a conservative statistical approach (*Lister et al., 2013*); see Materials and methods), we discovered 105,050 discrete differentially hypomethylated regions ('hypo-DMRs'; i.e., a % mC difference above threshold in one or more pairwise comparisons between EC tissue types; median length = 228 bp) with the largest number present in liver ECs (42,682), followed by lung ECs (32,004), brain ECs (30,471), and kidney ECs (17,027). A smaller number of hypo-DMRs were uniquely hypomethylated in one EC tissue type (25,612 in liver ECs, 13,862 in lung ECs, 16,540 in brain ECs, and 7132 in kidney ECs). We will refer to these regions as 'EC-tissue-specific hypo-DMRs' or 'ECTS-hypo-DMRs' (**Supplementary file 4**). Almost all hypo-DMRs (94.5%) are found >2 kb from a TSS, suggesting they reflect distal regulatory elements (**Figure 2—figure supplement 1C**). This pattern of EC tissue-type diversity among hypo-DMRs and LMRs implies differential utilization of the enhancer landscape.

**Table 1.** Abbreviations used throughout the text

Abbreviation	Definition
ATAC	Assay for Transposase Accessible Chromatin
BBB	Blood-brain barrier
CRE	Cis-regulatory element
DEG	Differentially expressed gene
DMR	Differentially methylated region
DMV	DNA methylation valley (UMR > 3 kb)
EC	Endothelial cell
ECTS-hypo-DMR	Endothelial cell tissue-specific hypo-DMR
ECTSAP	Endothelial cell tissue-specific ATAC-seq peak
ECTSG	Endothelial cell tissue-specific gene
GO	Gene ontology
LMR	Low-methylated region (hypomethylated, CG-poor)
PCA	Principal component analysis
scRNA-seq	Single-cell RNA-seq
TF	Transcription factor
TSS	Transcription start site
UMR	Unmethylated region (hypomethylated, CG-rich)

DOI: <https://doi.org/10.7554/eLife.36187.010>

A previous study in neurons has shown that many closely spaced hypo-DMRs are part of larger hypomethylated regions that correlate with the current transcriptional state of a cell and with higher levels of histone modifications associated with active enhancers (Mo et al., 2015). Therefore, we expanded our analysis to assess regions of low DNA methylation that span multiple kilobases. Using an approach similar to the one described by Mo et al. (2015), we merged closely spaced (separation <1 kb) hypo-DMRs into larger hypomethylated regions; we refer to such regions as ‘large hypo-DMRs’ if they are >2 kb in size (Supplementary file 4). ECTS-large hypo-DMRs are enriched at differentially expressed EC-enriched genes (q-value <10<sup>-10</sup>; Figure 2B; see Supplementary file 5).

DNA methylation valleys (DMVs; also referred to as DNA methylation canyons) are a distinct category of large hypomethylated regions that exhibit especially high enrichment at genes coding for important regulators of embryonic development including many TFs (Xie et al., 2013; Jeong et al., 2014; Mo et al., 2015; Mo et al., 2016). Defining DMVs as UMRs > 3 kb, we identified 2313 brain EC, 2695 liver EC, 2403 lung EC, and 2208 kidney EC DMVs (Supplementary file 4). We found very little overlap between large hypo-DMRs and DMVs, implying they represent distinct hypomethylated features (Figure 2C). These two categories of hypomethylation are further distinguished by the extent to which they are shared between EC types: 1602 DMVs are shared by all four EC subtypes and additional DMVs are shared between subsets of EC subtypes, whereas the majority of large hypo-DMRs are specific to a single EC subtype (Figure 2C). We used the Genomic Regions Enrichment of Annotations Tool (GREAT) to identify gene ontologies associated with EC DMVs (McLean et al., 2010). Consistent with previous studies in other tissues (Xie et al., 2013; Jeong et al., 2014; Mo et al., 2015; Mo et al., 2016), EC DMVs are highly enriched for genes encoding TFs (Figure 2—figure supplement 1E).

Previously identified DMVs in embryonic stem cells (ESCs) remain hypomethylated after cellular differentiation, with only a small subset demonstrating DNA methylation variations across cell types (Xie et al., 2013). These large regions of hypomethylation are maintained by Polycomb complexes that deposit repressive H3K27me3 histone modifications and are thought to support developmental plasticity, as DNA methylation may reflect a more stable repressive mechanism (Li et al., 2018). Interestingly, a subset of H3K27me3-positive DMVs found at lineage-specific expressed genes exhibit increased DNA methylation outside of the promoter region with a corresponding loss of H3K27me3, providing a snapshot of the developmental history of gene expression (Li et al., 2018). In all four subtypes of ECs, DMVs overlap *Etv2*, *Foxc2*, *Tal1*, *Lmo2*, *Id3*, and *Fli1*, genes coding for TFs that are important for establishing the EC lineage (Figure 2D). DMVs are also present at *Sox7*, *Hey1*, and *Nr2f2*, TFs important for venous or arterial differentiation. An exception to this pattern is seen with the early embryonic TF gene *Pou5f1/Oct4*, which lacks a DMV, consistent with a previous report that this locus exhibits increased DNA methylation upon cell differentiation (Xie et al., 2013). No EC DMVs overlap *Gata1* and *Spi1*, which are important for establishing the erythroid and myeloid/lymphoid lineages, respectively. Given that ECs and hematopoietic cells share the same developmental lineage, these observations suggest that DMVs occur at TFs that determine very early progenitor cell lineage commitments. Curiously, EC DMVs overlap some genes coding for lineage-determining TFs in completely unrelated cell types, including oligodendrocytes (*Olig1* and *Olig2*) and Muller glia (*Lhx2*) (Figure 2D).

At genes coding for TFs that are differentially expressed among EC subtypes, DMVs show EC subtype-specific differences in boundary locations and degree of methylation, with a broad trend of increased methylation 5’ of the TSS in the EC subtype(s) that express the TF gene as seen for *Foxf1*, *Foxf2*, and *Zic3* in brain; *Tbx3* in brain, lung, and kidney; *Tbx2*, *Gata6*, *Meis1*, and *Meis2* in liver, lung, and kidney; *Gata4* in liver; and *Irx3* in kidney (Figure 2—figure supplement 2A–C). This observation is consistent with a proposed model in which cell-type-specific expression of important TF regulators of lineage identity results in nearby hypermethylation (Li et al., 2018). We note that FOXF1, a lung EC-enriched TF implicated in pulmonary vascular development (Kalinichenko et al., 2001; Cai et al., 2016; Dharmadhikari et al., 2016), represents an exception to this methylation pattern, due to demethylation at the adjacent lung EC-expressed lncRNA gene, *Fendrr* (Figure 2—figure supplement 2A).



## HOX gene expression and DNA methylation: correlation with anterior-posterior position

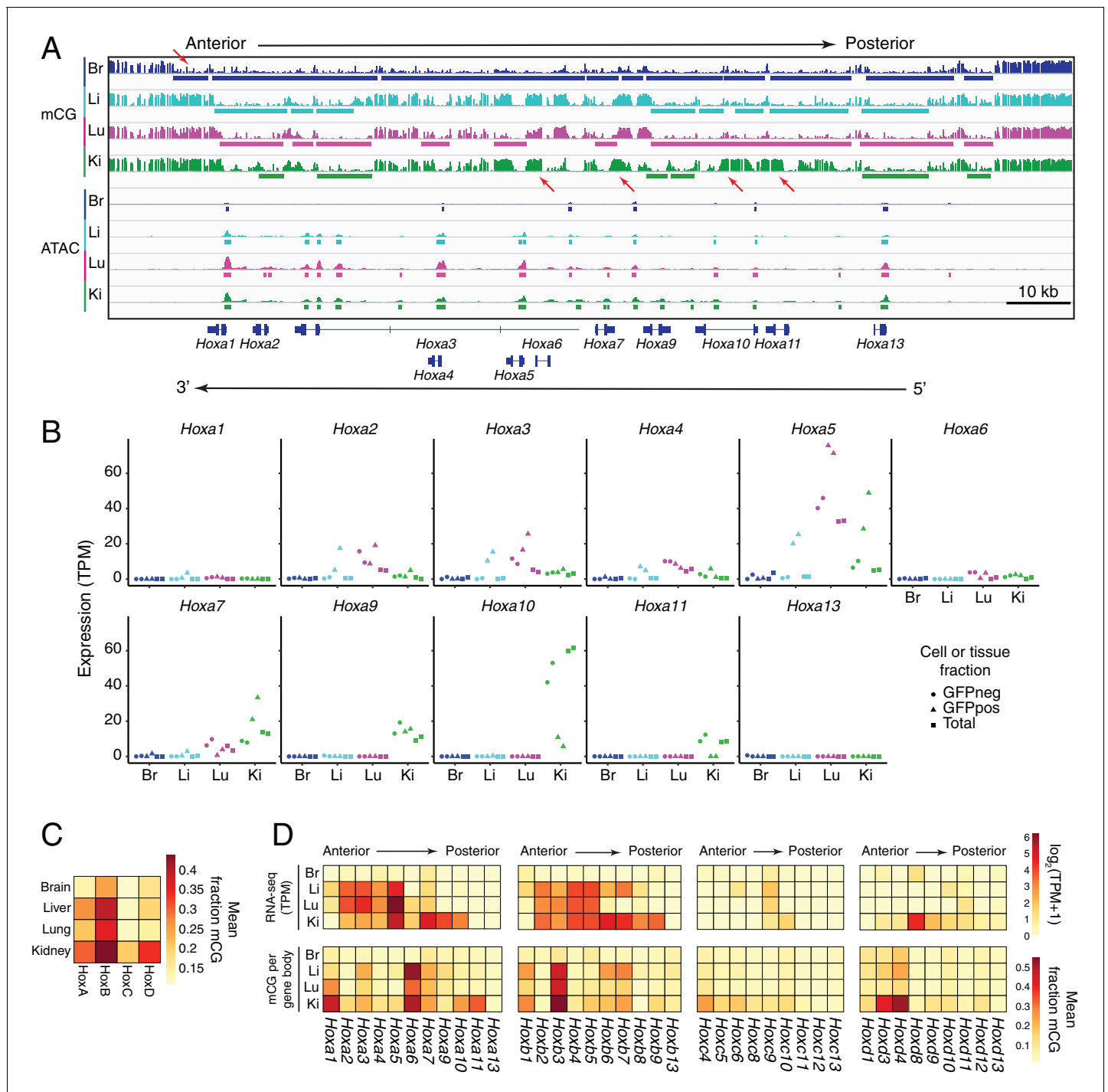
Some of the largest regions of hypomethylation are found in the HOX gene clusters. HOX genes do not exhibit the classic relationship between methylation and gene expression (Laurent *et al.*, 2010; Bock *et al.*, 2012; Xie *et al.*, 2013). Among the four EC subtypes, HOX cluster DNA methylation is lowest in brain ECs, intermediate in liver and lung ECs, and highest in kidney ECs (Figure 3A and C; Figure 3—figure supplement 1). This pattern of methylation correlates with the anterior/posterior position of the brain, liver, lung, and kidney – as well as with their embryonic primordia – along the body axis: the brain is most anterior, the lungs and liver are intermediate (and are derived from the primitive foregut), and the kidneys are the most posterior (and are derived from the ureteric bud and surrounding mesenchyme). For all four HOX clusters, the ATAC-seq patterns are similar for liver, lung, and kidney ECs, but the peak intensities are greatly reduced in brain ECs (Figure 3A; Figure 3—figure supplement 1).

In the HOX clusters, there is a roughly co-linear relationship between each gene's position along the chromosome and the anterior/posterior territory along the body axis where that HOX gene is expressed: genes that are named with low numbers (e.g. *Hoxa1*) are expressed more anteriorly and genes that are named with high numbers (e.g. *Hoxa13*) are expressed more posteriorly (Lewis, 1978; reviewed in Papageorgiou, 2012). Based on the EC subtype-specific RNA-seq data, we found that patterns of HOX gene expression correlate with the anterior/posterior position of each organ and its corresponding EC subtype (Figure 3B and D). In the HOX-A cluster, brain ECs express none of the *Hoxa* genes; liver ECs express *Hoxa2–Hoxa5*; lung ECs express *Hoxa2–Hoxa5* and *Hoxa7*; and kidney ECs express *Hoxa5* and *Hoxa7–Hoxa10*, with lower levels of expression in *Hoxa2–Hoxa4*. (Figure 3B and D). Interestingly, there does not appear to be a strong relationship between gene expression and gene body methylation (compare top and bottom heatmaps in Figure 3D). Rather, like other developmentally important TFs, the trend is that expression of a HOX gene correlates with increased methylation in the neighborhood of the gene body. For example, *Hoxa5* which is expressed in liver, lung, and kidney ECs but not brain ECs exhibits a hypomethylated gene body in all ECs but a hypermethylated promoter region only in peripheral ECs (Figure 3A). Similarly, the regions in and adjacent to *Hoxa10* are hypermethylated only in kidney ECs, which are the only ECs in this set that express *Hoxa10* (Figure 3A). For each of the four organs, the patterns of EC and parenchymal cell expression of *Hoxa* genes are closely matched (Figure 3B). These data confirm and extend two earlier studies: (1) a microarray transcriptome meta-analysis showing a correlation between patterns of HOX gene expression and the tissue-of-origin for a collection of human EC lines (Toshner *et al.*, 2014) and (2) a transgenic mouse study showing that *Hoxa3* and *Hoxc11* reporter expression in murine vasculature roughly corresponds to the positional identity specified by the HOX code (Pruett *et al.*, 2008). Taken together, the patterns of HOX gene expression and DNA methylation suggest that ECs in different organs 'know' their anterior-posterior position within the body.

## Comparisons between accessible chromatin and hypomethylation landscapes

As a second and complimentary approach to defining candidate EC CREs, we analyzed regions of accessible chromatin. In our initial analysis, we used the full range of ATAC-seq fragment lengths and identified a total of 51,740 discrete regions of accessible chromatin (i.e. ATAC-seq peaks; median length 656 bp) among the four EC subtypes (Supplementary file 6). Across replicates, the percent of mapped reads found in called peaks ranged between 16 and 34% (Figure 4—figure supplement 1A). Calling peaks on subsamples of total mapped reads for each ATAC-seq replicate indicates that, with this sequencing depth, the data is approaching an asymptote (Figure 4—figure supplement 1B). In deriving a consensus EC-tissue-specific peakset, we adopted a conservative approach by requiring each peak to be called in 2/2 or 2/3 replicates (Figure 4—figure supplement 1C).

For a quantitative assessment of open chromatin, we restricted our analyses to ATAC-seq fragments of <100 bp, as previous work suggested that these correspond to nucleosome-free regions of the genome (Buenrostro *et al.*, 2013). Using <100 bp fragments, we identified a total of 39,987 ATAC-seq peaks (median length 508 bp) among the four EC subtypes (Supplementary file 6).



**Figure 3.** In ECs, patterns of methylation and gene expression at HOX gene clusters correlate with anterior/posterior position. (A) Genome browser image showing CG methylation (top) and accessible chromatin (bottom) at the HOX-A gene cluster. HOX genes in this cluster are expressed in an anterior-posterior gradient corresponding to their position in the cluster, with genes near the 3' end of the cluster expressed more anteriorly and genes near the 5' end expressed more posteriorly. The degree of EC methylation is: brain < liver < lung < kidney. The degree of EC accessible chromatin is: brain < liver < lung < kidney. Colored bars indicate DMVs or ATAC-seq peaks. Red arrows pointing down indicate illustrative examples of differential hypomethylation. Red arrows pointing up indicate illustrative examples of differential hypermethylation. (B) Expression levels (TPMs) based on RNA-seq for each gene in the HOX-A cluster. (C) Heatmap depicting mean fraction of methylated CG across each HOX cluster for each EC subtype. (D) Heatmaps depicting  $\log_2(\text{TPM}+1)$  or mean fraction of methylated CG in the gene body for genes within each of the four HOX clusters.

DOI: <https://doi.org/10.7554/eLife.36187.011>

The following figure supplement is available for figure 3:

Figure 3 continued on next page

Figure 3 continued

**Figure supplement 1.** Methylation patterns and accessible chromatin at HOX-B, HOX-C, and HOX-D clusters in ECs.

DOI: <https://doi.org/10.7554/eLife.36187.012>

[Hereafter, the phrase 'ATAC-seq peaks' refers to <100 bp fragments unless otherwise stated.] Of the 31,358 brain EC, 18,926 liver EC, 24,970 lung EC, and 21,694 kidney EC ATAC-seq peaks, 35% are shared across all four EC subtypes, but only 1–16% are EC subtype-specific (6381 brain, 632 liver, 755 lung, and 397 kidney EC-tissue-specific ATAC-seq peaks, referred to hereafter as 'ECTSAPs') (**Figure 4—figure supplement 2B**).

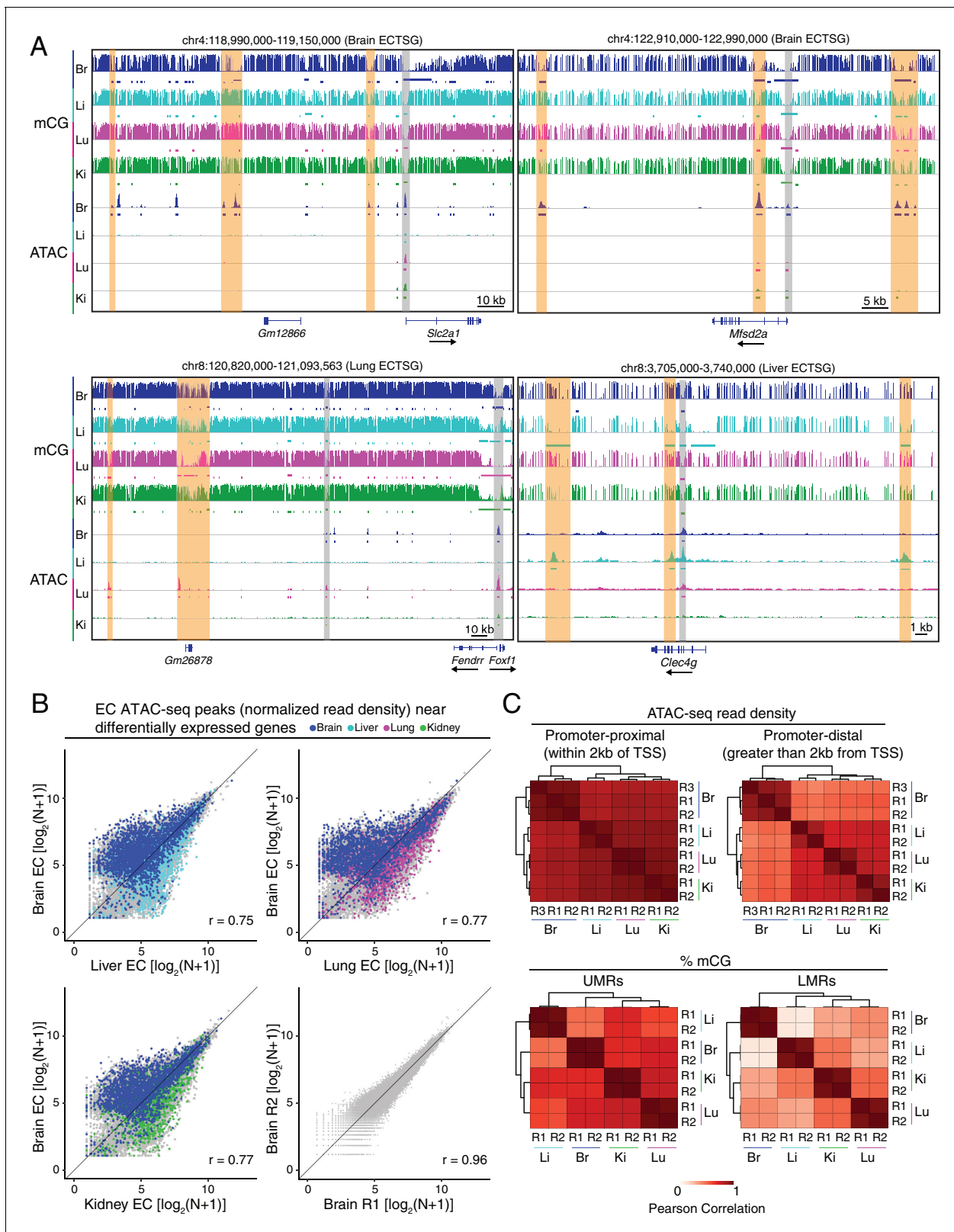
We observed multiple ECTSAPs near ECTSGs, as illustrated in **Figure 4A** using the full range of ATAC-seq fragment lengths. As indicated by the translucent orange bars in **Figure 4A**, ECTS-hypo-DMRs generally co-localize with ECTSAPs. For example, near brain EC-specific genes *Slc2a1* and *Mfsd2a*, multiple ECTSAPs and ECTS-hypo-DMRs co-localize (**Figure 4A**, upper two panels; **Figure 4—figure supplement 2A**). Similarly, the genomic region encompassing *Clec4g*, a liver EC-specific gene, contains co-localized ECTSAPs and ECTS-hypo-DMRs upstream and downstream of the gene body (**Figure 4A**, right bottom panel; **Figure 4—figure supplement 2A**). A region ~200 kb upstream of *Foxf1* contains two ECTSAPs that colocalize with ECTS-hypo-DMRs (**Figure 4A**, left bottom panel; **Figure 4—figure supplement 2A**), suggestive of a long-range enhancer. This region also includes a lncRNA gene, *Gm26878*, that is expressed specifically in lung ECs (**Figure 4—figure supplement 2A**). In **Figure 4A**, regions of chromatin accessibility and hypomethylation that are shared across more than one EC subtype are indicated by translucent gray bars.

In scatter plots of normalized ATAC-seq read densities, ECTSAPs for each tissue type are enriched near differentially expressed EC-enriched genes for that tissue (**Figure 4B**), and broad differences are apparent between EC subtypes ( $r = 0.75\text{--}0.92$ ), whereas biological replicates of the same subtype are highly similar ( $r = 0.93\text{--}0.97$ , **Figure 4B**; **Figure 4—figure supplement 3**). Additionally, a larger proportion of ECTSAPs and ECTS-hypo-DMRs are within 100 kb of ECTSGs that are expressed in the corresponding EC subtype relative to ECTSGs expressed in the other EC subtypes, a difference that is highly significant statistically (**Figure 4—figure supplement 2C**). This enrichment implies a role for these candidate CREs in EC subtype-specific gene expression.

Of the 39,987 total ATAC-seq peaks among all four EC subtypes, 64% are >2 kb from a TSS (i.e. are promoter-distal). Similar to the distributions described above for UMRs and LMRs, a majority of promoter-proximal (<2 kb from a TSS) ATAC-seq peaks are shared by all EC subtypes (10,336 of 14,474; 71%), whereas only a small minority of promoter-distal ATAC-seq peaks are shared by all EC subtypes (3476 of 25,517; 14%) (**Figure 4C**, upper two panels; **Figure 4—figure supplement 2D**). This trend is even more striking for methylation at UMRs and LMRs (**Figure 4C**, lower two panels). Chromatin accessibility at UMRs and LMRs and DNA methylation at ATAC-seq peaks similarly shows greater divergence among EC subtypes at distal candidate CREs (**Figure 4—figure supplement 2E**). The ATAC-seq comparisons also reveal a greater degree of similarity in chromatin accessibility among ECs derived from the three peripheral tissue (liver, lung, and kidney) relative to brain ECs (**Figure 4C**). Taken together, these comparisons suggest that the heterogeneity in gene expression that gives rise to tissue-specific EC differentiation may largely reflect differences in distal rather than proximal CREs. As illustrated by the hierarchical clustering analysis shown in **Figure 4C**, these differences are most pronounced in the case of brain ECs, a finding that potentially reflects the specialized and extensive program of gene expression associated with the BBB.

In all four EC subtypes, >80% of ATAC-seq peaks overlap with UMRs or LMRs (**Figure 4—figure supplement 4A**), in agreement with previous findings in neurons and photoreceptors (**Mo et al., 2015**; **Mo et al., 2016**). ATAC-seq peaks from all four EC subtypes exhibit similarly reduced levels of DNA methylation compared to random size-matched genomic regions. Moreover, ECTSAPs exhibit levels of DNA methylation that are preferentially reduced in their respective EC DNA methylomes (**Figure 4—figure supplement 4B**). Reciprocally, there were preferential increases in the EC subtype-specific ATAC-seq signal in the respective ECTS-hypo-DMRs (**Figure 4—figure supplement 4C**).

Interestingly, there are many more LMRs than ATAC-seq peaks, and therefore only a minority of LMRs overlap with ATAC-seq peaks:<10% of liver EC and kidney EC LMRs and <20% of brain and lung EC LMRs overlap with ATAC-seq peaks (**Figure 4—figure supplement 4A**). In contrast, >75%



**Figure 4.** Accessible chromatin and hypomethylated regions reveal candidate EC regulatory elements. (A) Genome browser images showing CG methylation (top) and accessible chromatin (bottom) around ECTSGs. Colored bars under mCG tracks mark UMRs (upper row) and LMRs (lower row). For accessible chromatin, histograms of ATAC-seq reads are shown. Colored bars under ATAC tracks indicate the called ATAC-seq peaks. Vertical orange bars highlight co-localizing cell-type-specific open chromatin and differentially hypomethylated DNA. Vertical gray bars highlight shared regions

Figure 4 continued on next page

Figure 4 continued

of open chromatin in at least three EC subtypes. The ATAC-seq reads shown here represent the full range of ATAC-seq fragment sizes. (B) Scatter plots of normalized ATAC-seq read density (N) within ATAC-seq peaks called from <100 bp fragments. Values shown are  $\log_2(N + 1)$ . Colored symbols correspond to peaks for which the closest annotated TSS is a differentially expressed EC-enriched gene from the indicated tissue. Lower right, comparison between two brain EC ATAC-seq biological replicates. (C) Pairwise Pearson correlation heatmaps for ATAC-seq read density within ATAC-seq peaks at promoter-proximal (<2 kb from TSS; upper left) or promoter-distal (>2 kb from TSS; upper right) and percent CG methylation in UMRs (lower left) and LMRs (lower right).

DOI: <https://doi.org/10.7554/eLife.36187.013>

The following figure supplements are available for figure 4:

**Figure supplement 1.** Quality control for ATAC-seq.

DOI: <https://doi.org/10.7554/eLife.36187.014>

**Figure supplement 2.** EC subtype differences in distal epigenetic features.

DOI: <https://doi.org/10.7554/eLife.36187.015>

**Figure supplement 3.** Differentially accessible chromatin between peripheral ECs.

DOI: <https://doi.org/10.7554/eLife.36187.016>

**Figure supplement 4.** Relationship between accessible and hypomethylated regions of the EC genome.

DOI: <https://doi.org/10.7554/eLife.36187.017>

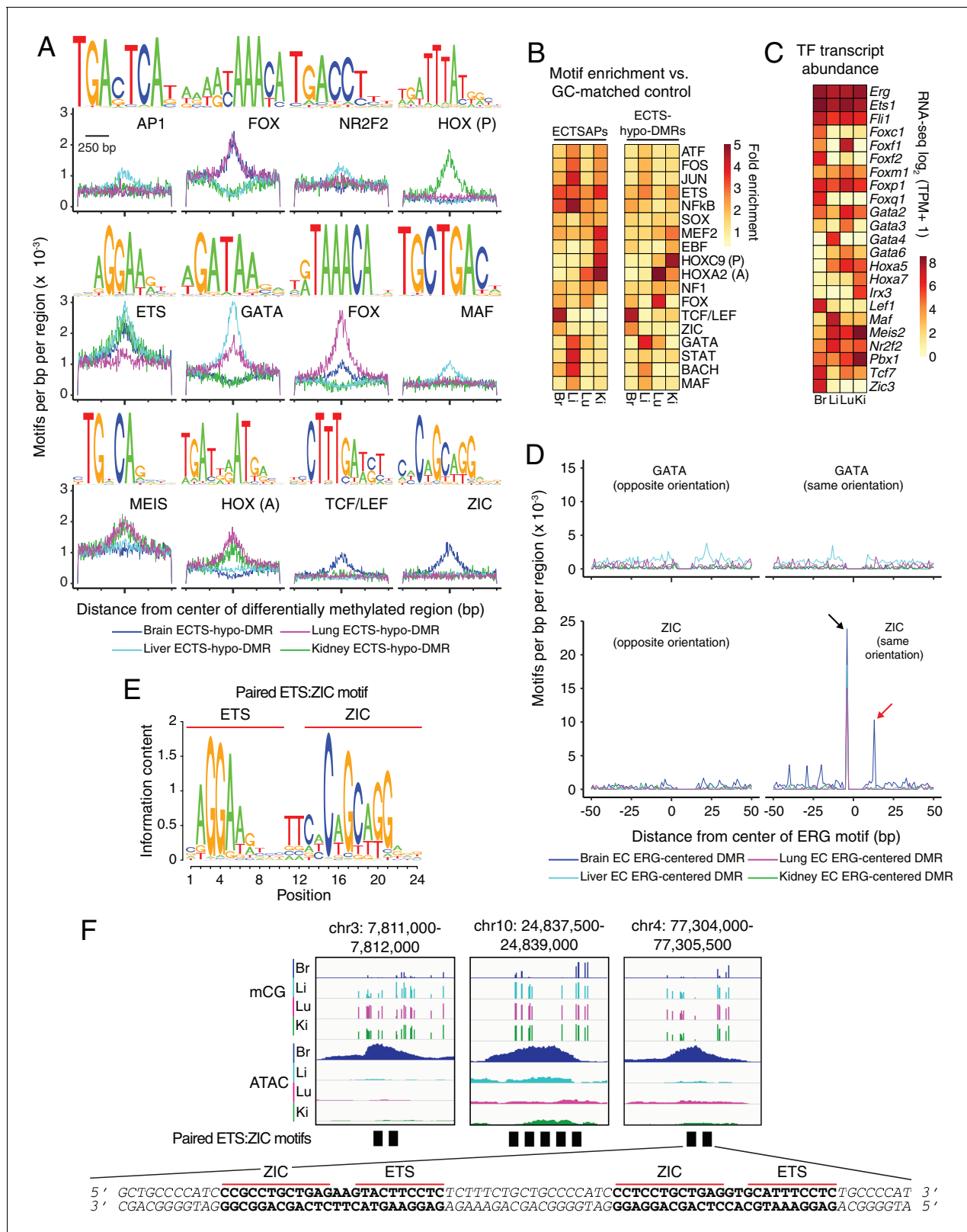
of UMRs overlap with ATAC-seq peaks in all four EC subtypes. In comparing ECTS-hypo-DMRs and ECTSAPs, <5% of liver EC, lung EC, and kidney EC hypo-DMRs overlap with their respective ATAC-seq peaks, whereas 13% of brain EC hypo-DMRs overlap with brain ATAC-seq peaks. As two explanations for this numerical difference, we suggest either (1) that many distal enhancer regions may not exhibit a sufficient degree of chromatin accessibility to permit insertion by two transposon complexes, the requisite event for detection of DNA accessibility by ATAC-seq, or (2) that distal enhancer regions may have greater variance in their occupancy across the population of cells, resulting in a general decrease in average accessibility.

### Tissue-specific EC transcription factors and their DNA targets

If LMRs and ATAC-seq peaks mark CREs, then they should be enriched for the DNA binding motifs of the TFs that control the distinct transcriptional, chromatin, and DNA methylation landscapes observed in ECs from different tissues. Current models of transcriptional regulation posit that TFs act in a combinatorial fashion, co-localizing at CREs to give rise to cell-type-specific transcription (Heinz and Glass, 2011; Heinz et al., 2015). These models predict that CREs in different EC subtypes should (1) share motifs corresponding to TFs that orchestrate patterns of gene expression common to all ECs, and (2) exhibit subtype-specific motifs corresponding to TFs that orchestrate patterns of gene expression distinctive to each EC subtype.

To investigate these models, we used Hypergeometric Optimization of Motif EnRichment (HOMER), a suite of tools for discovering enriched motifs in genomic sequences (Heinz et al., 2010). The HOMER de novo motif detection strategy parses genomic sequences under consideration into all possible k-mers of a desired motif length and searches for enrichment of each k-mer, a strategy that is not dependent on existing biochemical definitions of TF motifs (Figure 5A). The results of the k-mer analysis are then compared to a compendium of known TF motifs. HOMER also directly searches the genomic sequences under consideration for known TF motifs (Figure 5B). In general, we found the two approaches to be in good agreement.

To address the first prediction, we first identified LMRs and ATAC-seq peaks that are not present in photoreceptors and brain neurons (Mo et al., 2015; Mo et al., 2016). This should filter out CREs associated with 'housekeeping' genes, thereby enriching for CREs that reflect the EC regulatory landscape. This procedure identified 56,755 LMRs and 12,200 ATAC-seq peaks that are relatively EC-specific (Supplementary files 4 and 6). From this subset, 8035 LMRs and 899 ATAC-seq peaks were common to all four EC subtypes. Analyzing this common subset using HOMER revealed a highly significant (LMRs:  $p\text{-value}=10^{-1691}$ ; ATAC-seq peaks:  $p\text{-value}=10^{-148}$ ) enrichment over random GC-matched genomic regions for a sequence that resembles the binding motif of class 1 members of the ETS family of TFs, which includes ERG, FLI1, ETV2, and ETS1 (Wei et al., 2010; Hollenhorst et al., 2011); Figure 5—figure supplement 1A). The finding of a common enriched class 1 ETS family motif is not unexpected given the large body of data supporting a role for class 1



**Figure 5.** Motif enrichment analysis identifies candidate TF regulators of tissue-specific EC development and function. (A) HOMER-identified enriched motifs in ECTS-hypo-DMRs. Frequency of the indicated motif as a function of distance from the center of ECTS-hypo-DMRs. Shown above each individual plot is the position weight matrix (PWM) of the enriched nucleotide sequence. The TF family that most closely matches the motif is indicated below the PWM. (B) Heatmap showing the fold enrichment for the indicated TF motifs (% ECTSAPs or ECTS-hypo-DMRs containing the motif divided by % GC-matched control). (C) Heatmap showing TF transcript abundance (RNA-seq  $\log_2(\text{TPM} + 1)$ ) for the indicated TF motifs across brain, liver, lung, and kidney. (D) Motif frequency plots showing motifs per bp per region as a function of distance from the center of ERG motifs. Plots are shown for GATA (opposite orientation) and GATA (same orientation), and ZIC (opposite orientation) and ZIC (same orientation). (E) Information content plot showing information content as a function of position for a paired ETS:ZIC motif. (F) Genomic tracks showing mCG and ATAC tracks for three regions: chr3: 7,811,000-7,812,000; chr10: 24,837,500-24,839,000; and chr4: 77,304,000-77,305,500. Below the tracks are paired ETS:ZIC motifs and a sequence alignment.

Figure 5 continued

by % GC-matched background genomic regions containing the motif). Representative members of TF families that exhibited a significant enrichment (FDR < 0.001) are shown. (C) Heatmap showing TPMs for the four EC subtypes for a subset of TFs with the motifs shown in (A). Values shown are  $\log_2(\text{TPM} + 1)$ . (D) ECTS-hypo-DMRs were centered on the motif for ERG, a member of the ETS family, and the frequencies of the indicated motifs were plotted as a function of distance from the ERG motif with a bin size of 1 bp. Black arrow: the ZIC motif ends with the sequence AGG and the ERG motif begins with the sequence AGG, thereby generating a frequency spike in all four EC subtypes at position -5 bp that represents the overlap of the two sites. Red arrow: the frequency spike for the ZIC motif at +11 bp is only present in one orientation and only in brain ECTS-hypo-DMRs centered on the ERG motif. (E) PWM for the consensus sequence of the paired ETS:ZIC motif. (F) Representative instances of the paired ETS:ZIC motif from brain ECTS-hypo-DMRs. Each instance is represented by a black rectangle. The bottom strand of the sequence in (F) matches the consensus sequence shown in (E). DOI: <https://doi.org/10.7554/eLife.36187.018>

The following figure supplements are available for figure 5:

**Figure supplement 1.** Candidate TF regulators of EC gene expression.

DOI: <https://doi.org/10.7554/eLife.36187.019>

**Figure supplement 2.** TF binding motifs in candidate CREs near ECTSGs.

DOI: <https://doi.org/10.7554/eLife.36187.020>

ETS TFs in EC development and function (McLaughlin et al., 2001; Pham et al., 2007; Meadows et al., 2011; Oh et al., 2015; Sumanas and Choi, 2016). We also detected k-mers matching the SOX (p-value= $10^{-101}$ ) and GATA (p-value= $10^{-94}$ ) binding motifs in shared EC-only LMRs and - unexpectedly - a k-mer matching the telomere sequence TTAGGG (p-value= $10^{-123}$ ) in shared EC-only ATAC-seq peaks (Figure 5—figure supplement 1A).

To address the second prediction, we used HOMER to search for TF motifs in sequences present in the four sets of ECTSAPs and the four sets of ECTS-hypo-DMRs. In addition to the EC subtype-specific enriched motifs described in the paragraphs below, this analysis revealed the class 1 ETS motif in both ECTSAPs (p-value= $10^{-45}$ – $10^{-718}$ ) and ECTS-hypo-DMRs (p-value= $10^{-30}$ – $10^{-1169}$ ) (Figure 5A and B). This observation suggests that ETS family TFs not only act in shared CREs to determine and maintain common aspects of the EC phenotype, but they also cooperate with EC subtype-specific TFs in EC subtype-specific CREs. As determined from the EC RNA-seq datasets, all four subtypes of ECs express members of the ETS family of TFs, including ERG, FLI1, and ETS1 (Figure 5C). As seen in Figure 5B, the motif analysis using ECTSAPs was in good agreement with the motif analysis using ECTS-hypo-DMRs, but the latter exhibited a higher signal-to-noise ratio, most likely due to the larger number of ECTS-hypo-DMRs.

Significantly, HOMER identified several EC subtype-specific enriched motifs (Figure 5A and B). As different members of a given TF family generally recognize the same or nearly the same motif, we determined which family members are likely to be relevant for tissue-specific EC transcriptional regulation by assessing their transcript abundances in the EC RNA-seq datasets (Figure 5C). A summary of motifs and TFs for each of the four EC subtypes follows (see also Supplementary file 7).

- In brain ECTS-hypo-DMRs, we identified enriched k-mers that closely match the motifs of the Forkhead box (FOX), TCF/LEF, and ZIC TF families (Figure 5A). Transcripts corresponding to three Fox family members (*Foxq1*, *Foxc1*, and *Foxf2*), two Tcf/Lef family members (*Tcf7* and *Lef1*), and the Zic family member *Zic3* were highly enriched in brain ECs (Figure 5C). LEF/TCF family members are the effectors of canonical Wnt signaling in the nucleus (Cadigan and Waterman, 2012). The specific enrichment of LEF/TCF motifs in brain ECTS-hypo-DMRs and ECTSAPs is consistent with previous evidence that canonical Wnt signaling in ECs plays a central role in CNS angiogenesis and BBB development and maintenance, but that it has little or no role in vascular development outside of the CNS (Liebner et al., 2008; Stenman et al., 2008; Daneman et al., 2009; Wang et al., 2012; Zhou et al., 2014; Cho et al., 2017). The brain EC-specific enrichment of ZIC motifs is consistent with previous work showing that the accumulation of ZIC3 in the developing retinal vasculature depends on NORRIN/FRIZZLED4 (i.e. canonical Wnt) signaling (Wang et al., 2012). Analysis of candidate CREs near the *Zic3* gene in all four EC subtypes revealed the presence of a brain ECTSAP ~50 kb downstream of the gene that contains ETS and TCF/LEF binding sites, suggestive of a nearby Wnt-responsive enhancer element (translucent orange bar in Figure 5—figure supplement 2A and B).

- In liver ECTS-hypo-DMRs, we identified enriched k-mers that closely match the motifs recognized by members of the GATA and MAF TF families (Figure 5A). RNA-seq across the four EC

populations showed that *Gata4* and *Maf* have the greatest enrichment among members of their respective families in liver ECs (**Figure 5C**). Consistent with these observations, a recent study found that GATA4 controls liver EC development and function (*Géraud et al., 2017*). In liver ECTS-hypo-DMRs, we also detected enrichment for a k-mer that corresponds to the NR2F2/COUP-TFII motif (**Figure 5A**). As NR2F2/COUP-TFII is essential for establishing venous EC identity (*You et al., 2005*), enrichment of this motif could reflect enrichment of venous ECs due to inclusion of the portal vein and its tributaries or a bona fide role for NR2F2 in liver sinusoid development. Like *Gata4* and *Maf*, *Nr2f2* transcripts are enriched in liver ECs (**Figure 5C**).

- In lung ECTS-hypo-DMRs, we identified enriched k-mers that closely match the motifs of the FOX, GATA, and homeobox TF families (**Figure 5A**). *Foxp1*, *Foxm1*, *Foxf1*, *Gata2*, *Gata3*, *Gata6*, and *Hoxa5* transcripts are enriched in lung ECs (**Figure 5C**).

- In kidney ECTS-hypo-DMRs, we identified enriched k-mers that closely match the homeobox motif (**Figure 5A**). RNA-seq shows strong kidney-specific enrichment for homeobox transcripts *Ir3*, *Hoxa7*, *Pbx1*, and *Meis2* (**Figure 5C**). A motif corresponding to anterior HOX proteins (e.g. HOXA2; **Figure 5B**) was enriched in lung and kidney whereas a motif corresponding to posterior HOX proteins (e.g. HOXC9; **Figure 5B**) showed enrichment primarily in kidney, consistent with the *Hox* gene expression patterns described above (**Figure 3** and **Figure 3—figure supplement 1**).

If the enriched motifs play a role in regulating EC subtype-specific gene expression, then one might expect an enrichment for the same motifs in candidate EC subtype-specific CREs near ECTSGs in one EC subtype relative to candidate CREs near ECTSGs in other EC subtypes. Consistent with this expectation, ECTS-hypo-DMRs found within 100 kb of ECTSGs exhibit statistically significant enrichment for most of the motifs in **Figure 5A** (**Figure 5—figure supplement 1B**). ECTSAPs within 100 kb of ECTSGs were also significantly enriched for several of these motifs (**Figure 5—figure supplement 1C**). **Figure 5—figure supplement 2** shows five examples of individual ECTSGs and their nearby candidate CREs with tissue-specific TF and ETS motifs marked in the regions encompassed by ECTS-hypo-DMRs and ECTSAPs (**Figure 5—figure supplement 2C–F**; data are shown for *Zic3* and for the four genes shown in **Figure 4A**).

## Identification of a paired ETS and ZIC motif with precise spacing and orientation

If ETS TFs, which are expressed in all ECs, cooperate with tissue-specific EC TFs to control expression of ECTSGs, then there might be a non-random spatial relationship between ETS motifs and motifs for tissue-specific TFs. To test this idea, we compiled all of the ECTS-hypo-DMRs containing an ERG (i.e. ETS family) motif, aligned these ECTS-hypo-DMRs on the ERG motif, and then generated, for each of the four EC subtypes, motif frequency histograms with a 1 bp bin size for the lung EC-enriched FOXF1 motif, the liver EC-enriched GATA4 motif, the kidney EC-enriched HOXC9 motif, the brain EC-enriched ZIC3 and TCF/LEF motifs as a function of distance from the center of the ERG motif (**Figure 5D**; **Figure 5—figure supplement 1D**).

From this analysis, two general patterns emerged. One pattern is exemplified by the frequency of GATA4 motifs in liver ECTS-hypo-DMRs, which are modestly enriched within ~40 bp of the ERG motif. This suggests that in liver ECs GATA4 may collaborate with ERG and/or other class 1 ETS factors but without forming a unique ETS-GATA4-DNA ternary complex. The second pattern is exemplified by the spatial relationship between the ERG and ZIC3 motifs. Here, we observed a spike in ZIC3 motif frequency precisely 2 bp from the end of and in the same orientation as the ERG motif only in brain ECTS-hypo-DMRs (red arrow in **Figure 5D** lower right panel). A partial overlap of the ERG and ZIC3 motifs (the nucleotide sequence AGG occurs at the start of the ERG motif and at the end of the ZIC3 motif) generates a second spike in ZIC3 motif frequency that is present in all four EC subtypes but is unlikely to be of biological significance (black arrow in **Figure 5D** lower right panel). Additional correlations are observed between the ERG motif and FOXF1 and HOXC9 motifs in lung and kidney ECTS-hypo-DMRs, respectively (red arrows in **Figure 5—figure supplement 1D**). The paired FOX:ETS motif in lung ECTS-hypo-DMRs matches a composite FOX:ETS motif found in several EC-specific enhancers that is known to be bound and synergistically activated by FOX and ETS factors (*De Val et al., 2008*; *Robinson et al., 2014*). The paired ETS:HOX motif in kidney ECTS-hypo-DMRs suggests a similar interaction between ETS factors and HOX factors; however, we found no previous characterization of such an interaction in the literature. A partial overlap of the nucleotide sequence AGG between the TCF/LEF motif and the ERG motif generates a spike in frequency



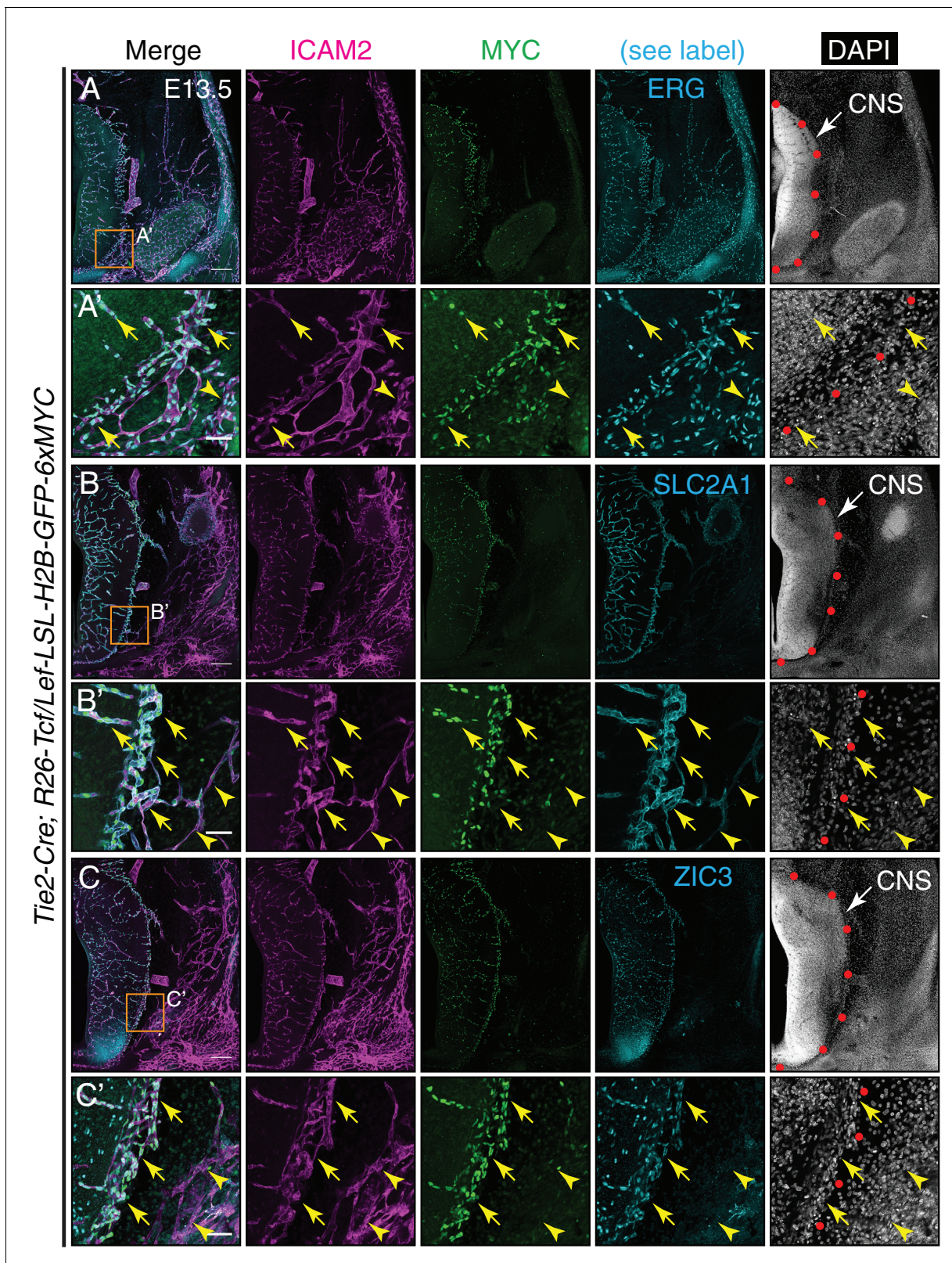
that is present in all four EC subtypes but is unlikely to be of biological significance (black arrow in **Figure 5—figure supplement 1D**).

The proximity, precise spacing, and orientation of the ERG and ZIC3 motifs in brain ECTS-hypo-DMRs suggests that these two transcription factors form a ternary complex by binding to this larger DNA element. To determine a consensus for this element, we repeated the HOMER analysis on brain ECTS-hypo-DMRs with longer motifs. This analysis revealed a consensus motif that we will refer to as the paired ETS:ZIC motif (**Figure 5E**; **Supplementary file 7**). Strikingly, identically spaced ETS and ZIC motifs lead to robust enhancer activity in the ascidian *Ciona*, implying an ancient and conserved role for the cooperation of ETS factors and ZIC factors in transcriptional regulation (**Farley et al., 2016**). Genome-wide, there are only 1944 instances of the paired ETS:ZIC motif, and this paired motif is significantly enriched in candidate brain EC CREs (both hypo-DMRs and ATAC-seq peaks) relative to random GC-matched genomic regions (DMR p-value= $10^{-45}$ ; ATAC peak p-value= $10^{-17}$ ) and to peripheral EC CREs (DMR p-value= $10^{-74}$ ; ATAC peak p-value= $10^{-7}$ ; **Figure 5F**; **Figure 5—figure supplement 1E**). Interestingly, 81% of the brain EC CREs that contain one or more paired ETS:ZIC motifs overlap the long terminal repeats (LTRs) of the mouse-specific endogenous retrovirus (ERV) family RLTR45/ERV4\_2. Transposable elements, especially LTR retrotransposons, provide a rich genomic substrate for the evolution of non-coding regulatory elements (**Emera and Wagner, 2012**; **Chuong et al., 2013**; **Chuong et al., 2016**); reviewed in **Thompson et al., 2016**). We speculate that RLTR45/ERV4\_2 may represent an example of a transposable element family that is being co-opted into host gene regulatory networks that utilize ETS and ZIC factors.

### Canonical Wnt signaling in CNS ECs versus non-CNS ECs

At present, it is not known whether CNS EC-specific effects of canonical Wnt signaling reflect a spatial restriction on Wnt signaling - perhaps reflecting the spatial distribution of the relevant ligands - or whether canonical Wnt signaling occurs throughout the vasculature but has different downstream consequences in CNS and non-CNS tissues. To address this question, we visualized canonical Wnt signaling at cellular resolution in CNS and non-CNS vasculature using an EC-specific Cre recombinase (*Tie2-Cre*) and a Cre-dependent Wnt reporter (*R26-Tcf/Lef-LSL-H2B-GFP-6xMYC*). With this combination of genetic elements, Cre-mediated excision of a *loxP*-flanked transcription stop cassette allows the multimerized LEF/TCF motifs, together with a minimal promoter, to report canonical Wnt signaling by controlling production of a nuclear-localized histone H2B-GFP-6xMYC fusion protein specifically in ECs (**Cho et al., 2017**).

In coronal sections of E13.5 embryos, the pan-EC TF ERG and the pan-EC plasma membrane protein ICAM2 mark all ECs (**Figure 6A and A'**). At this stage, BBB development is already underway, as judged by the accumulation of the glucose transporter GLUT1/SLC2A1, a BBB marker, specifically in CNS and perineural ECs (**Figure 6B and B'**). At E13.5, ZIC3 accumulation in the vasculature is also limited to CNS and perineural ECs (**Figure 6C and C'**; **Figure 6—figure supplement 1**), a distribution that closely matches that of GLUT1/SLC2A1. Outside of the vasculature, ZIC3 accumulates in developing neurons in the ventral CNS. Similarly, vascular accumulation of LEF1, which is both a mediator and a marker of canonical Wnt signaling, is limited to CNS and perineural ECs (**Figure 6—figure supplement 2A and A'**). Visualizing the accumulation of the Wnt reporter, H2B-GFP-6xMYC, with anti-MYC immunostaining shows a neural and perineural distribution, closely matching that of SLC2A1, ZIC3, and LEF1 (**Figure 6**; **Figure 6—figure supplement 1**). We note that cells positive for the Wnt reporter occur outside of the CNS, yet their rounded morphology and lack of association with blood vessels suggests that these cells are macrophages or other immune cells, consistent with the specificity of *Tie2-Cre*. At P7, H2B-GFP-6xMYC accumulates in CNS ECs, as well as ECs in the renal medulla (**Figure 6—figure supplement 2B–E**), and ZIC3 and SLC2A1 are restricted to the CNS vasculature (data not shown). These protein distributions are consistent with the CNS EC-specific expression of *Lef1* and *Zic3* found by RNA-seq (**Figure 5C**) and the CNS EC-enrichment of TCF/LEF and ZIC3 motifs in ECTS-hypo-DMRs (**Figure 5A and B**). Taken together, these data imply that, among ECs, canonical Wnt signaling is largely confined to the CNS, with a corresponding restriction in the expression of TFs that mediate (LEF1) and respond to (LEF1 and ZIC3) canonical Wnt signaling.



**Figure 6.** Canonical Wnt signaling in CNS but not peripheral ECs at E13.5. (A–C) Coronal sections of E13.5 *Tie2-Cre;R26-Tcf/Lef-LSL-H2B-GFP-6xMYC* embryos near the cephalic flexure. The markers are: ICAM2 (pan-EC membrane protein), MYC (the canonical Wnt reporter), ERG (pan-EC TF), SLC2A1 (the glucose transporter GLUT1; a BBB marker), ZIC3, and DAPI. (A'–C') Higher magnification of the boxed regions in (A–C). The boundary between CNS and peripheral tissue is marked on the DAPI image with red circles. The nuclear MYC signal reveals canonical Wnt signaling in CNS ECs (yellow arrowheads). *Figure 6 continued on next page*

Figure 6 continued

arrows) but not in peripheral ECs (yellow arrowheads). ZIC3 is present in CNS ECs (yellow arrows) but not in peripheral ECs (yellow arrowheads), and in developing neurons in the ventral CNS. Scale bars in A, B, and C: 200  $\mu\text{m}$ . Scale bars in A', B', and C': 50  $\mu\text{m}$ .

DOI: <https://doi.org/10.7554/eLife.36187.021>

The following figure supplements are available for figure 6:

**Figure supplement 1.** Quantification of immunostaining in **Figure 6**.

DOI: <https://doi.org/10.7554/eLife.36187.022>

**Figure supplement 2.** LEF1 in ECs in the E13.5 CNS, and canonical Wnt signaling in the P7 brain, liver, lung, and kidney.

DOI: <https://doi.org/10.7554/eLife.36187.023>

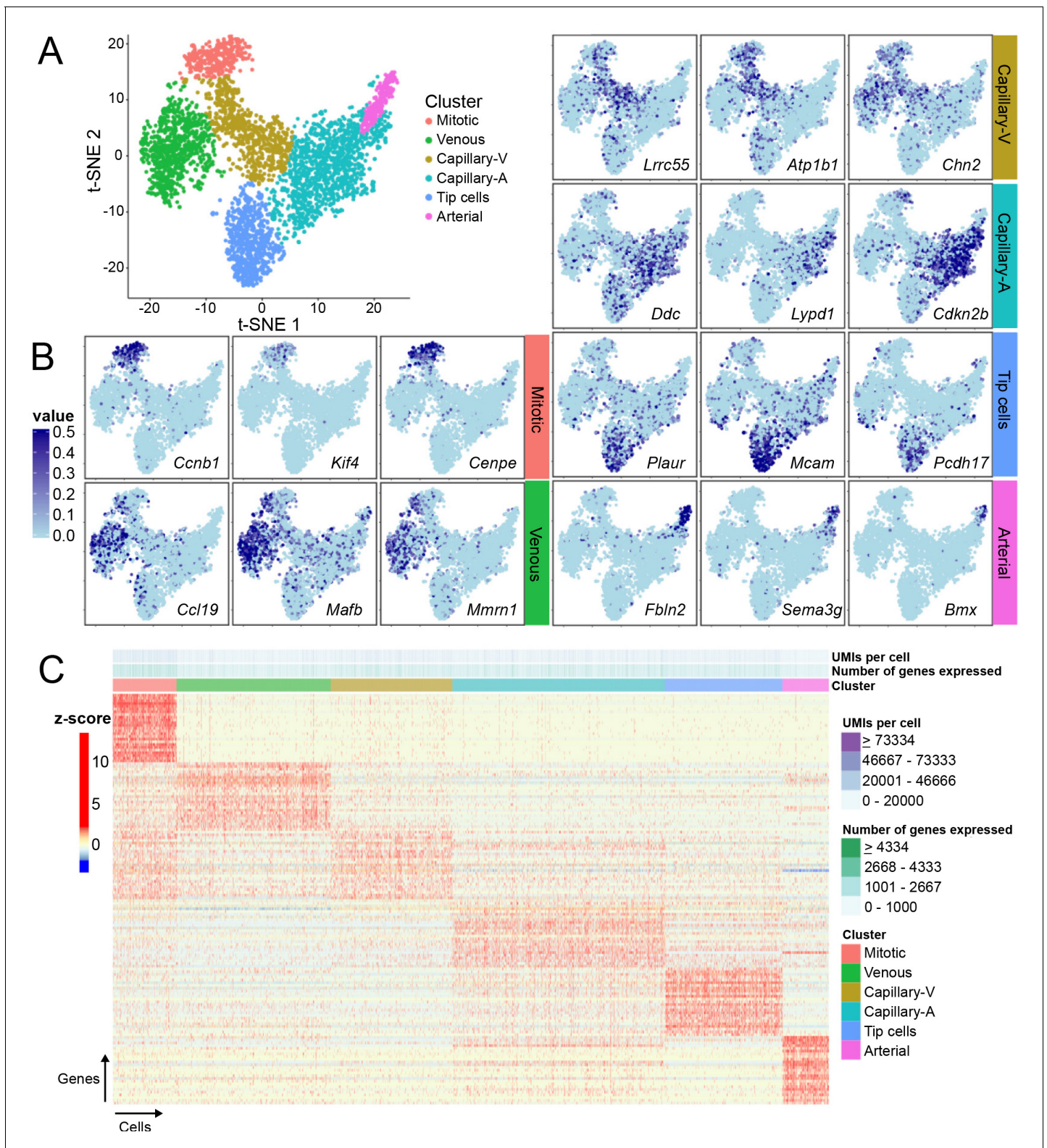
## Single-cell RNA-seq reveals heterogeneity among brain ECs

Up to this point, the focus has been on EC heterogeneity between tissues. Here, we address the question of intra-tissue EC heterogeneity. It is well established that arterial, venous, and capillary ECs represent distinct subtypes based on morphology, function, and gene expression profiles (Aird, 2007a; Potente and Mäkinen, 2017), but it is not clear whether or to what extent these EC subtypes can be further subdivided. It is also not clear how developing EC subtypes, such as tip cells and proliferating cells, are related to the more mature EC subtypes. To address these questions, we assessed EC gene expression in the developing CNS at single-cell resolution by performing single-cell RNA-seq (scRNA-seq) on 3,946 FACS-purified GFP-positive ECs from a P7 *Tie2-GFP* mouse brain.

scRNA-Seq data were processed and analyzed using a modified dpFeature approach as described in the Monocle R/Bioconductor package (Trapnell et al., 2014). t-Distributed Stochastic Neighbor Embedding (t-SNE) of cells revealed a single interconnected distribution of gene expression profiles suggesting a continuity of transcriptional identity across brain ECs at this developmental time point consistent with the observed zonation of EC subtypes along the vasculature. Subsequent clustering analysis identified six EC clusters (Figure 7A) that could be assigned as tip, mitotic, venous, or arterial cells (one cluster each) or capillary cells (two adjacent clusters) based on known EC subtype marker gene expression (Figure 7B). One of the capillary clusters is enriched for cells expressing *DOPA Decarboxylase (Ddc)*, a gene known to be expressed in ECs in bovine aorta (Sorriente et al., 2012). The ECs in the high-*Ddc* cluster also express multiple genes that are characteristic of arterial ECs and this cluster was therefore named 'capillary-A' (capillary-arterial). The second capillary cluster was relatively deficient in *Ddc*-expressing cells, and it was enriched for cells that express genes characteristic of venous ECs. The second cluster was therefore named 'capillary-V' (capillary-venous). Mitotic ECs express multiple genes that are also expressed in capillary-V and venous ECs (Figure 7C), which is consistent with the observation that proliferating cells arise from veins during sprouting angiogenesis (Ehling et al., 2013; Xu et al., 2014; Hasan et al., 2017; Pitulescu et al., 2017). Capillary-A ECs express multiple genes that are also expressed in tip cells and arterial ECs, but tip cells and arterial ECs show less mutual similarity (Figure 7C).

The distribution of DDC was examined by immunostaining of P7 WT mouse brain (Figure 7—figure supplement 1). DDC was localized to puncta, most likely corresponding to axon terminals of dopaminergic neurons; this pattern was particularly concentrated in the striatum, a brain region receiving dense dopaminergic innervation (Figure 7—figure supplement 1C and D). Most interestingly, we detected both fine- and coarse-grained heterogeneity in DDC immunostaining in GS Lectin-positive capillaries in different anatomic structures within the same brain sections, as seen by comparing the intensity of the general EC marker GS-Lectin and anti-DDC immunoreactivity. Larger vessels showed minimal DDC immunoreactivity (yellow arrows in Figure 7—figure supplement 1C). The heterogeneity in DDC immunoreactivity among capillary ECs likely corresponds to the differential capillary-A versus capillary-V gene expression patterns defined by scRNA-seq.

Although ECs within the six clusters express many genes in common, there were multiple genes expressed more or less selectively by each cluster (Figure 7C; Supplementary file 8 lists the top 25 genes for each cluster). The identification of multiple new markers for tip cells is especially interesting as it substantially extends earlier screens for tip cell markers based on the overproduction of tip cells in *Dll4<sup>+/-</sup>* mouse retinas and on laser capture microdissection (del Toro et al., 2010; Strasser et al., 2010). Five genes that were previously validated as tip cell markers in the retina —



**Figure 7.** Single-cell RNA-seq of P7 brain ECs reveals intra-tissue heterogeneity. (A) t-SNE plot of 3946 P7 brain ECs showing six clusters corresponding to tip cells, and mitotic, venous, capillary-venous (Capillary-V), capillary-arterial (Capillary-A), and arterial ECs. (B) The t-SNE plot from (A) showing expression of three marker genes with enriched expression for each EC cluster. Cells with no RNA-seq reads are shown in light blue; darker blue represents greater number of reads. (C) Heatmap showing scaled expression (z-scores) for the 25 most enriched marker genes for each EC cluster. **Supplementary file 8** lists the genes plotted in (C). Rows represent genes and columns represent cells.

Figure 7 continued on next page

Figure 7 continued

DOI: <https://doi.org/10.7554/eLife.36187.024>

The following figure supplement is available for figure 7:

**Figure supplement 1.** DOPA decarboxylase is expressed non-uniformly in the CNS vasculature.

DOI: <https://doi.org/10.7554/eLife.36187.025>

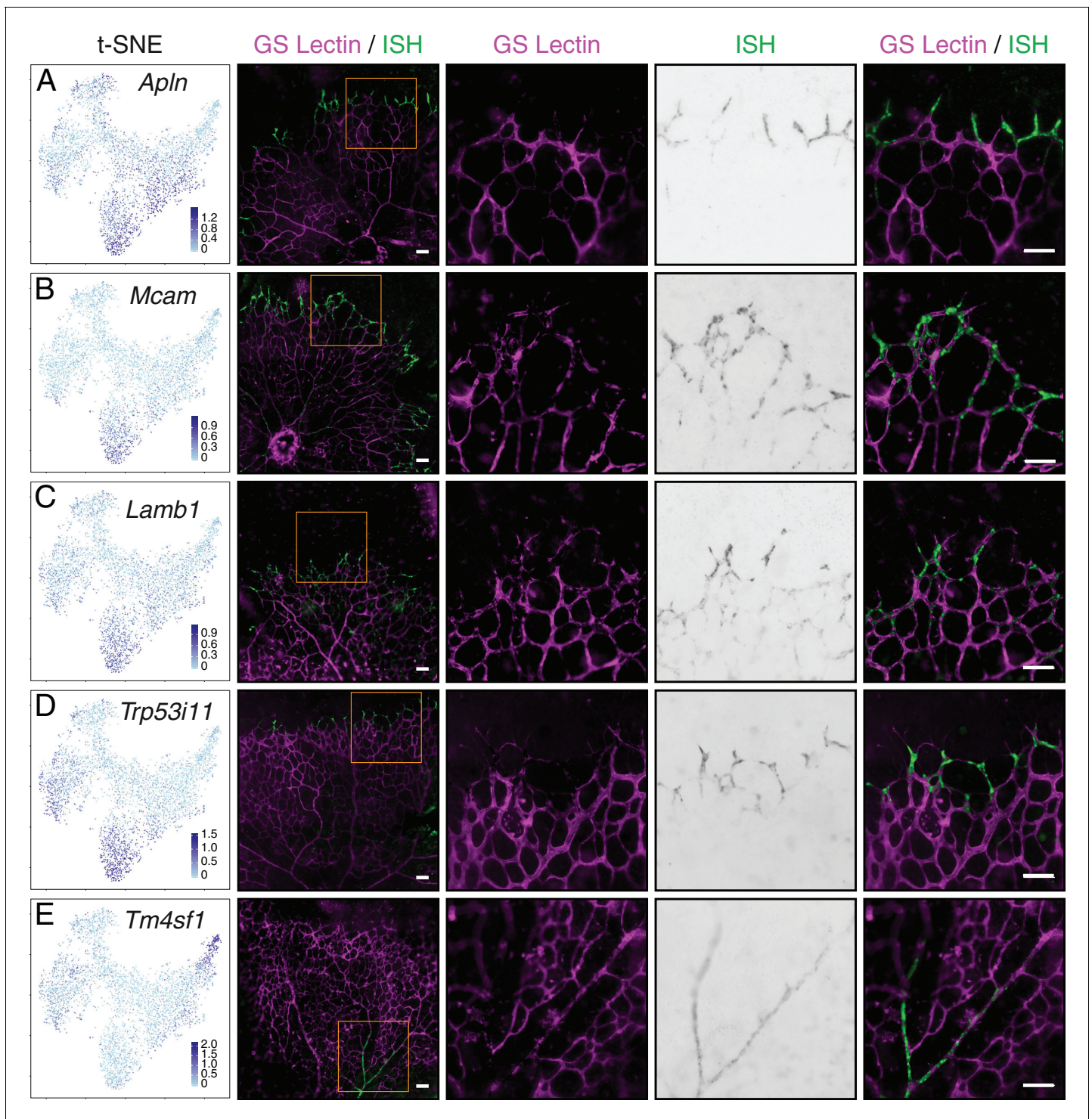
*Plaur*, *Angpt2*, *Lcp2*, *Cxcr4*, *Apln*, and *Kcne3* — are enriched in the scRNAseq tip cell cluster (*del Toro et al., 2010; Strasser et al., 2010; Rattner et al., 2013*). Interestingly, 19 of the 25 most enriched tip cell transcripts from the scRNA-seq analysis code for membrane or secreted proteins, many of which are known or suspected regulators of information flow. These include ion channels (the mechanosensor PIEZO2, voltage-gated sodium channel beta subunit SCN1B, and potassium channels KCNE3 and KCNA5), cell adhesion proteins (protocadherin PCDH17, cell adhesion molecules MADCAM1 and MCAM, and lectin CLEC1A), receptor tyrosine kinase signaling regulator SIRPA, G-protein-coupled receptor CMKLR1, plasminogen regulator SERPINE1/PAI, urokinase receptor PLAUR, ligands adrenomedullin (ADM) and ANGPT2, and ECM protein SMOC2. Among the six intracellular proteins in the list of top 25 tip cell markers, most are involved in signaling, including cyclic nucleotide phosphodiesterase 4B (PDE4B), serine/threonine phosphatase PPM1J/PP2C-zeta, NADPH oxidase organizer-1 (NOXO1), and the ubiquitin ligase HECW2, which regulates EC-EC junction stability via its substrate angiomin-like-1 (*Choi et al., 2016*).

To independently assess the expression of candidate tip cell genes, we performed whole mount in situ hybridization (ISH) on P5-7 retina (*Figure 8*). At this developmental stage, the vascular plexus on the vitreal face of the retina is growing outward from the optic disc and tip cells are localized to the growing vascular front. Consistent with the scRNA-seq analysis, *Apln*, *Mcam*, *Lamb1*, and *Trp53i11* exhibited enriched expression at the angiogenic front of the superficial vascular plexus (*Figure 8A–D*), indicating tip cell enrichment of these genes. As a specificity control, *Tm4sf1* exhibited only arterial expression by both scRNA-seq analysis and whole mount retina ISH (*Figure 8E*).

To describe the relationships between the six EC clusters, the scRNA-seq data were used to construct a cell trajectory map (*Figure 9A*). This analysis produced a map with two branch points and four terminal states [mitotic (M), venous (V), arterial (A), and tip cell (T)]. Capillary-V and capillary-A ECs populate the regions adjacent to the first and second branch points, respectively. The first branch point represents a choice between the venous state and the arterial and tip-cell states, implying that, at the transcriptome level, tip cells are more closely related to arterial ECs than to venous ECs. As examples of gene expression patterns that support this model, *Hey1* is more highly expressed in arterial ECs and tip cells than in venous ECs, whereas *Nr2f2* is more highly expressed in mitotic and venous ECs than in arterial ECs or tip-cells (*Figure 9B*). The second branch point represents a choice between the arterial state and the tip-cell state. *Fbln2* and *Plaur* are examples of genes that are more highly expressed in arterial ECs and tip-cells, respectively (*Figure 9B*).

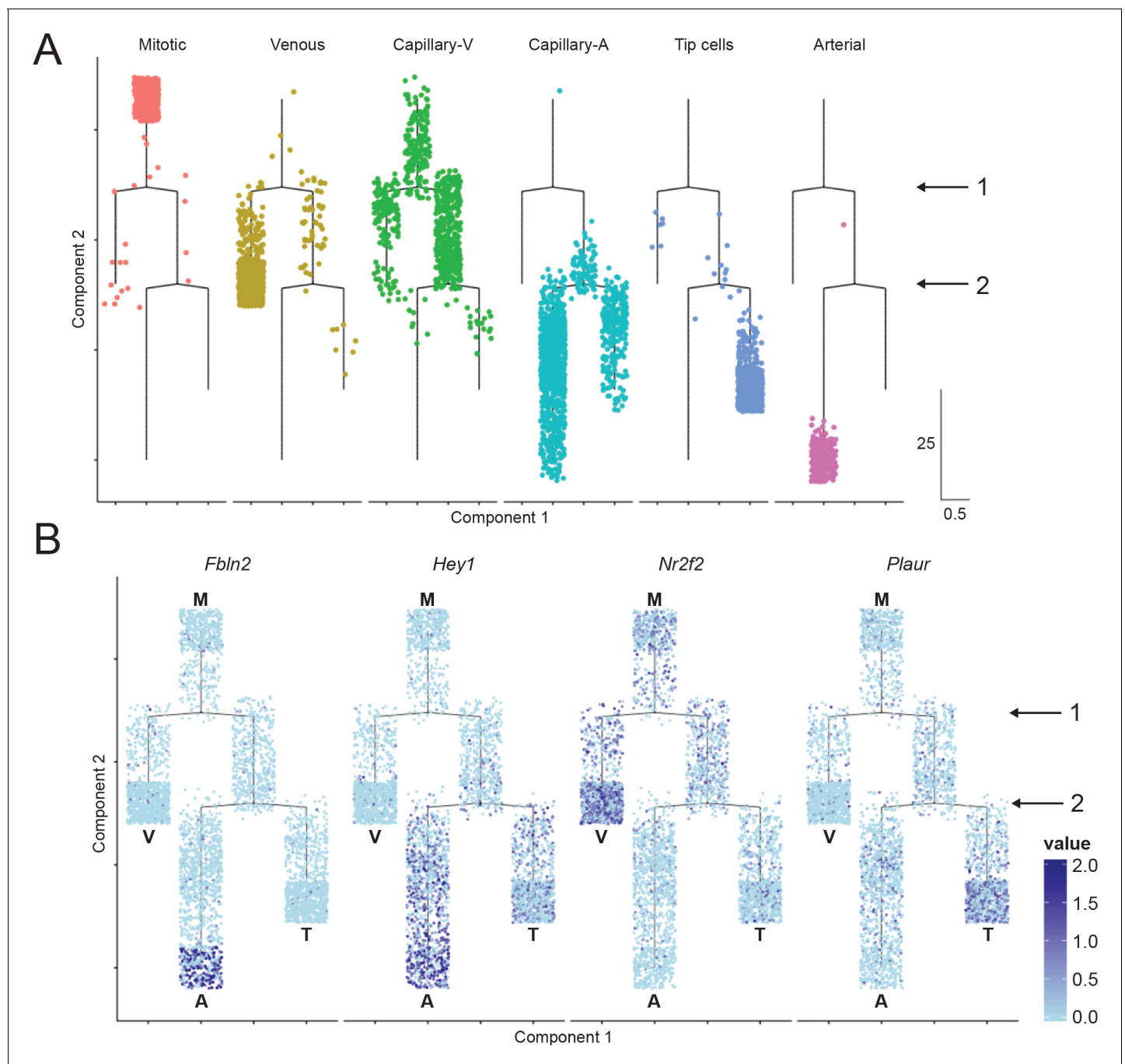
Among the transcripts with intriguing differences between EC clusters were several coding for cyclin-dependent kinase (CDK) inhibitors. In particular, *Cdkn1c/p57* expression was enriched in artery and vein ECs; *Cdkn1a/p21* expression was enriched in mitotic cells, tip cells, and capillary-A ECs; and *Cdkn2b/p15* expression was enriched in capillary-A ECs (*Figure 9—figure supplement 1A*). *Cdkn1a/p21* expression suggests that these ECs have exited the cell cycle (*Spencer et al., 2013*). Interestingly, TGF-beta2 (*Tgfb2*) was widely expressed but Latent TGF-beta-binding protein 4 (*Ltbp4*) expression was concentrated in arterial ECs, which utilize TGF-beta signaling to coordinate vascular smooth muscle cell development (*ten Dijke and Arthur, 2007; Figure 9—figure supplement 1A*). Transcripts coding for SMAD6 and SMAD7, two inhibitory SMADs induced by TGF-beta signaling (*Afrakhte et al., 1998*), were enriched in the capillary-A and arterial EC clusters (*Figure 9—figure supplement 1A*). Both CDKN1A and CDKN2B are known to mediate TGF-beta-induced cell cycle arrest (*Hannon and Beach, 1994; Datto et al., 1995*), consistent with a role for TGF-beta signaling in CNS EC development.

We also observed enrichment of transcripts coding for the chemokine receptor CXCR4 in tip cell, capillary-A, and arterial EC clusters, consistent with the role of this receptor in mediating angiogenic sprouting and artery formation (*Strasser et al., 2010; Busmann et al., 2011; Ehling et al., 2013; Xu et al., 2014; Hasan et al., 2017; Pitulescu et al., 2017; Figure 9—figure supplement 1A*).



**Figure 8.** Single-cell RNA-seq of P7 brain ECs identifies novel tip cell markers. (A–E) Whole mount retina in situ hybridization (ISH) for known tip cell marker *Apln* (A); novel tip cell markers *Mcam* (B), *Lamb1* (C), and *Trp53i11* (D); and a novel arterial marker *Tm4sf1* (E). The first two columns, from left to right, are the t-SNE plot from **Figure 7A** showing expression of each marker gene and a low-magnification merged image of flatmount retina with blood vessels marked by GS Lectin (magenta) and ISH signal (green). The third to fifth columns showed the boxed region in column two at higher magnification, with separate signals (columns three and four) and the merged signals (column five). All scale bars are 50  $\mu$ m.

DOI: <https://doi.org/10.7554/eLife.36187.026>



**Figure 9.** Cell trajectory analysis of brain ECs based on single-cell RNA-seq. (A) Plot showing the position of cells in each EC cluster on the constructed cell trajectory. (B) Summary of the two branch points (labeled '1' and '2') in the cell trajectory analysis. M, mitotic; V, venous; A, arterial; T, tip cells. Examples of markers that show differential expression as cells diverge from branch points 1 and 2. Cells with no RNA-seq reads are shown in light blue; dark blue represents greater number of reads. Hey1 is enriched in arterial and tip cells (i.e. the products of the right side of branchpoint 1); Nr2f2 is enriched in venous ECs; Fbln2 is enriched in arterial ECs; and Plaur is enriched in tip cells.

DOI: <https://doi.org/10.7554/eLife.36187.027>

The following figure supplement is available for figure 9:

**Figure supplement 1.** Brain ECs exhibit heterogeneous expression of cyclin-dependent kinase inhibitors, components of TGF-beta signaling, and components of CXCR4 signaling.

DOI: <https://doi.org/10.7554/eLife.36187.028>

Transcripts coding for CXCL12, the ligand for CXCR4, were highly enriched in the arterial EC cluster, suggesting that tip cells in the developing brain vasculature may receive a CXCL12 signal from arterial ECs. CXCL12 is expressed in pulmonary and coronary arterial ECs, where it mediates patterning of arterial vasculature (Chang *et al.*, 2017; Kim *et al.*, 2017). Cell trajectory representations of CDK inhibitor, TGF-beta signaling, and CXCR4/CXCL12 transcripts are shown in **Figure 9—figure supplement 1B**.

## Discussion

The experiments and analyses presented here define the transcriptome, accessible chromatin, and DNA methylome landscapes for ECs from brain, liver, lung, and kidney, and they relate differences in those landscapes to the regulatory programs that control tissue-specific EC heterogeneity. The resulting genome-wide analysis of candidate CREs, TF motifs, and TF expression reveal the distinctive gene regulatory architecture of each EC subtype. Using scRNA-seq, we define the relationships among developing and mature EC subtypes within the brain based on an unbiased clustering of gene expression profiles. In comparing CNS versus non-CNS ECs, the most prominent differences in TF motifs and TF expression include CNS EC-specific utilization of TCF/LEF and ZIC3 motifs and CNS EC-specific expression of TCF/LEF factors and ZIC3, implying a prominent role for canonical Wnt signaling and ZIC3 DNA binding in CNS EC-specific gene expression.

## The DNA methylation landscape as a guide to current and past gene expression

The relationship between patterns of mammalian DNA methylation and gene expression has been an object of inquiry for more than 40 years (Riggs, 1975; Razin and Riggs, 1980; Ambrosi *et al.*, 2017). The discovery of diverse methyl CG binding proteins (Shimbo and Wade, 2016) and the more recent characterization of DNA methylation effects on the binding affinities of diverse TFs (Hu *et al.*, 2013; Kribelbauer *et al.*, 2017; Yin *et al.*, 2017) imply that DNA methylation/demethylation can play a causal role in the control of chromatin state and the utilization of enhancers and promoters via its effects on protein binding. Reciprocally, changes in chromatin state and gene transcription can alter the accessibility of DNA to the enzymatic machinery of methylation and demethylation, leading to changes in DNA methylation (Stroud *et al.*, 2017).

Comparative analyses of DNA methylation are especially powerful because technical developments now permit the genome-wide determination of DNA methylation at single-base resolution starting with small numbers of cells. Using this approach, we observe EC subtype-specific patterns of CG methylation that correlate with current patterns of gene expression and that also appear to reflect patterns of gene expression at earlier developmental points, especially at TF genes. As recognized in several previous studies, DNA methylation and gene expression do not exhibit a simple monotonic relationship (Mo *et al.*, 2015; Amabile *et al.*, 2016; Liu *et al.*, 2016; Neri *et al.*, 2017). Many expressed genes exhibit promoter demethylation, with a variable degree of demethylation extending into the gene body. Demethylation can also reflect earlier epochs in which the gene was expressed, thus serving as a mnemonic epigenetic feature (Hon *et al.*, 2013; Mo *et al.*, 2015). Paradoxically, in the HOX cluster, methylation appears to correlate inversely with current gene expression: it is lowest among CNS ECs, which show little or no HOX gene expression, and it is highest among kidney ECs, which show the highest levels of HOX gene expression. This inverse correlation could reflect a loss (or retention) of the association of Polycomb repressive complexes with HOX clusters (Eskeland *et al.*, 2010; Kundu *et al.*, 2017; Lau *et al.*, 2017), a resulting loss (or retention) of chromatin compaction (Chambeyron *et al.*, 2005; Fraser *et al.*, 2009; Ferraiuolo *et al.*, 2010; Noordermeer *et al.*, 2011), and a corresponding increase (or decrease) in access to DNA methyltransferases.

Previous work has linked DNA demethylation of promoter-distal regions with enhancers (Stadler *et al.*, 2011; Schultz *et al.*, 2015). In the present study, LMRs and ECTS-hypo-DMRs have been used to identify candidate CREs and enriched TF motifs. As many promoter-distal hypomethylated regions are differentially hypomethylated in one or more pairwise comparisons between EC subtypes, the availability of the complete DNA methylomes for brain, liver, lung, and kidney ECs provides a powerful resource for identifying candidate tissue-specific EC CREs. While promoter-distal ATAC-seq peaks generally co-localize with LMRs and/or ECTS-hypo-DMRs, promoter-distal



hypomethylated regions are ~3 fold more numerous than promoter-distal regions of chromatin accessibility, and, as a result, TF motif enrichment analyses using hypomethylated regions provided greater statistical significance than the comparable analysis with accessible chromatin.

Rod photoreceptors exhibit a similar discrepancy in the number of promoter-distal hypomethylated features and accessible chromatin, and *Mo et al., 2016* suggested that the degree of chromatin compaction associated with the extremely small nuclei of rods could restrict access of DNA methyltransferases to DNA. As they mature, EC nuclei flatten and elongate (*Cao et al., 1998; Merks et al., 2006; Parsa et al., 2011*), which enhances chromatin compaction (*Versaevel et al., 2012*). Hence, it is not inconceivable that one of the explanations for the EC discordance between hypomethylated features and accessible chromatin may be EC nuclear structure.

### Organ-specific regulatory circuits for EC gene expression

Current evidence suggests that all ECs differentiate from hemangioblasts, a common precursor to both the vascular EC and hematopoietic lineages. Restricted expression of lineage determining TFs directs a population of mesodermal cells away from the blood cell fate and toward the EC fate, giving rise to EC progenitors that express the vascular endothelial growth factor (VEGF) receptor KDR (*Marcelo et al., 2013; Park et al., 2013*). In particular, the ETS factor ETV2 and the forkhead box factor FOXC2 cooperate to establish KDR +hemangioblasts (*De Val et al., 2008; Sumanas and Lin, 2006; Sumanas et al., 2008; Park et al., 2013*). Continued expression of ETV2 and other ETS factors, including ERG, FLI1, and ETS1, together with GATA2 and FOXC2, commits these progenitors to the EC lineage (*Pimanda et al., 2007; Liu and Patient, 2008*).

Arterial-venous differentiation represents the earliest observable differentiation of blood vessels in embryonic development. This process occurs through extracellular cues that activate specific transcriptional regulatory networks (reviewed in *Potente and Mäkinen, 2017*). In particular, VEGF and Notch signaling drive arterial differentiation and gene expression through SOX factors (*Fischer et al., 2004; Corada et al., 2013; Hermkens et al., 2015*), whereas NR2F2-mediated suppression of Notch signaling establishes venous endothelium (*You et al., 2005*). Although these more global EC transcriptional networks are well established, there remains a large gap in our understanding regarding the signaling pathways and transcriptional regulators that drive organ-specific EC identities.

In the present study, we provide transcriptional and epigenomic evidence for TFs that likely control tissue-specific specialization of ECs. For example, liver sinusoidal ECs exhibit differential expression of *Gata4* and *Maf* with a corresponding enrichment of GATA and MAF binding motifs in candidate CREs. Liver EC-specific deletion of *Gata4* leads to a conversion of sinusoidal endothelium to continuous endothelium similar to that found in brain vasculature (*Géraud et al., 2017*), and deletion of *Maf* globally results in a defective erythropoietic microenvironment in the liver (*Kusakabe et al., 2011*). Kidney ECs appear to utilize transcriptional networks that depend on homeobox TFs as seen by the enrichment of a HOX motif in kidney EC-specific CREs. Our observation that kidney ECs differentially express *Hoxa7*, *Pbx1*, and *Meis2* is consistent with a role for PBX and MEIS proteins as cofactors for HOX proteins (*Choe et al., 2014; Longobardi et al., 2014; De Kumar et al., 2017*). PBX1 is also essential in the developing renal mesenchyme (*Schnabel et al., 2003*), and its deletion in renal vascular mural cells leads to disrupted renal vascular development (*Hurtado et al., 2015*). The data presented here imply that PBX1 also plays a role in kidney ECs. Similarly, *lrx3* - another kidney EC-specific homeobox TF - was previously identified as a regulator of nephron segment identity in *Xenopus* (*Reggiani et al., 2007*), and the data presented here imply that it also plays a role in kidney ECs. These observations suggest that shared expression of TFs in parenchymal cells and ECs in the same organ may be a not uncommon pattern.

### Inter- and intra-tissue heterogeneity of vascular endothelial cells

The inter-tissue heterogeneity of EC structure and function has long been recognized as a central component of organ specialization (*Aird, 2007a; Aird, 2007b*). For example, in secondary lymphoid organs, leukocyte trafficking from the intravascular compartment to the surrounding parenchyma is mediated by leukocyte adhesion molecules on the luminal face of ECs in high endothelial venules (HEVs); in the kidney, serum filtration is controlled by ECs in renal glomeruli with a high density of fenestrations and a luminal glycocalyx that functions as a charge filter; in the liver, efficient

detoxification of xenobiotics requires rapid equilibration of serum contents with hepatocytes, a process that is facilitated by ECs with large trans-cellular fenestrae; and in the lung, regulation of systemic blood pressure is controlled by EC expression of angiotensin-converting enzyme (*Bakhle and Vane, 1974*).

For many tissues, culturing the resident ECs leads to a loss of tissue-specific EC markers, implicating paracrine signaling from the parenchyma in tissue-specific EC differentiation and maintenance. For example, HEV ECs isolated from human tonsils lose multiple tissue-specific markers after 2 days in culture (*Lacorre et al., 2004*). In some systems, co-culture with parenchymal cells leads to retention of organ-specific markers: co-culture of cardiac ECs with cardiac myocytes induces expression of markers characteristic of cardiac microvascular ECs (*Aird et al., 1997*), and co-culture of hepatic sinusoidal ECs with hepatocytes induces expression of markers characteristic of hepatic ECs (*Edwards et al., 2005*). With the exception of CNS ECs (discussed below), the identities of the organ-specific paracrine signals remain unknown.

The present genomic approach complements these foundational observations. Analyses of the transcriptome, accessible chromatin, and DNA methylome landscapes can be used to more precisely define organ-specific EC differentiation in vivo and in vitro, providing genomic snapshots of the different stages of EC differentiation, the effects of tissue-specific paracrine signals, and the response of ECs to disease-associated perturbations. With respect to the last point, this approach should prove especially interesting in dissecting the chromatin and epigenomic architecture of ECs within tumor vasculature, where a distinctive transcriptome signature has been defined (*St Croix et al., 2000*).

Until recently, the study of intra-organ EC heterogeneity has generally been limited to a comparison of large artery, large vein, and microvascular ECs, as these distinctions correspond to the vessel classes that can be physically separated (*Chi et al., 2003*). This technical limitation is circumvented by scRNA-seq, which provides a broad sampling of the EC population that is independent of vessel size and structure, enabling an unbiased clustering of cell classes based only on gene expression profiles. A recent scRNA-seq study of adult mouse brain vasculature provided the first molecular profile of vascular zonation along the arteriovenous axis (*Vanlandewijck et al., 2018*). The data and analyses presented here both corroborate the concept of an EC continuum and extend the findings of Vanlandewijck and colleagues to the early postnatal brain vasculature. In particular, the present work provides strong support for a model in which proliferative ECs are more closely related to venous ECs and tip cells are more closely related to arterial ECs. The data also reveal an unexpected heterogeneity among capillary ECs, with distinct subpopulations that reflect a more venous-like or more arterial-like transcriptional signature. The molecular mechanisms that orchestrate the development of these EC properties remain largely unknown.

## A genomic view of canonical Wnt signaling and CNS vascular development

The earliest evidence for parenchymal signaling to CNS ECs came from embryonic transplantation experiments in which coelomic cavity ECs acquired BBB characteristics when they invaded pieces of transplanted brain, and brain ECs lost BBB characteristics when they invaded pieces of transplanted mesoderm (*Stewart and Wiley, 1981; Risau et al., 1986*). Subsequent co-culture experiments showed that astrocytes, which send end-foot processes to contact ECs and pericytes, can induce BBB properties in non-CNS ECs and can maintain BBB properties in long-term cultures of CNS-derived ECs (*Hayashi et al., 1997*); reviewed in *Helms et al., 2016*).

At present, canonical Wnt signaling is the only well-characterized cell-cell signaling pathway that promotes BBB formation and maintenance. In the developing and adult CNS, canonical Wnt ligands are secreted by glia (and possibly neurons) and activate receptors on ECs. In mice, EC-specific loss of beta-catenin leads to abortive CNS angiogenesis and loss of BBB integrity and BBB markers; similarly, double knockout of *Wnt7a* and *Wnt7b* in the embryonic CNS inhibits CNS angiogenesis and suppresses the expression of BBB markers (*Liebner et al., 2008; Stenman et al., 2008; Daneman et al., 2009; Zhou et al., 2014*). Nearly identical phenotypes are observed with mutations in the genes coding for GPR124 and RECK, essential receptor cofactors for WNT7A/WNT7B signaling (*Kuhnert et al., 2010; Anderson et al., 2011; Cullen et al., 2011; Zhou and Nathans, 2014; Posokhova et al., 2015; Cho et al., 2017*). In the retina, analogous phenotypes are produced by mutations in the genes coding for NORRIN, FRIZZLED4, and TSPAN12 - the ligand, receptor, and

co-activator components of a canonical Wnt signaling system that operates in parallel with the WNT7A/WNT7B system (Xu et al., 2004; Junge et al., 2009; Ye et al., 2009; Wang et al., 2012).

The present work extends our understanding of canonical Wnt signaling and BBB development in several ways. First, the location of LEF/TCF motifs in candidate CNS EC-specific CREs permits an initial assessment of which CNS EC genes might be direct targets of canonical Wnt-regulation. Second, the expression of the canonical Wnt reporter exclusively in ECs within and on the surface of the embryonic CNS argues that canonical Wnt signaling does not function as a widely distributed and permissive signal for CNS EC development but may, by virtue of its spatial localization, determine the spatial distribution of the CNS EC fate. The simplest explanation for the observed spatial distribution in canonical Wnt signaling is that it corresponds to the distribution of the relevant ligands: WNT7A, WNT7B, and NORRIN. Third, the nearly identical CNS-specific distribution of ZIC3-expressing ECs - together with (1) the presence of LEF/TCF motifs in candidate CREs near the Zic3 gene, (2) the enrichment of ZIC3 motifs in candidate CNS EC-specific CREs, and (3) the dependence of ZIC3 accumulation on NORRIN/FRIZZLED4 signaling in the developing retinal vasculature (Wang et al., 2012) - strongly suggests that ZIC3 is a downstream effector of the canonical Wnt-controlled EC gene expression program.

In summary, we have dissected from a set of candidate CREs common to many cell types, a subset of candidate CREs that are specific to vascular ECs, and an even smaller subset of candidate CREs that are distinctive for tissue-specific EC subtypes and that are likely to control tissue-specific EC gene expression. The genomic resources and analytical approaches developed here should help in identifying the full network of TFs and TF-binding sites that comprise the downstream effectors of parenchyma-derived signals for EC differentiation and specialization.

## Materials and methods

### Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Genetic reagent ( <i>Mus musculus</i> ; Male)	Tie2-GFP	The Jackson Laboratory	Stock No: 003658; RRID:IMSR_JAX:003658	
Genetic reagent ( <i>M. musculus</i> ; Male)	Tie2-Cre	The Jackson Laboratory	Stock No: 008863	
Genetic reagent ( <i>M. musculus</i> ; Male)	R26-Tcf/Lef-LSL-H2B-GFP-6xMYC	PMID: 28803732		
Antibody	anti-CD11b BV421 (rat monoclonal)	Biolegends	Cat No: 101235; RRID:AB_10897942	1:1000
Antibody	anti-ICAM2 (rat monoclonal)	BD Biosciences	Cat No: 553326	1:300
Antibody	anti-6xMYC (chicken polyclonal)	PMID: 24411735		1:10000
Antibody	anti-GLUT1 (rabbit polyclonal)	Thermo Fisher Scientific	RB-9052	1:500
Antibody	anti-ERG (rabbit polyclonal)	Cell Signaling	A7L1G	1:500
Antibody	anti-dopa decarboxylase (goat polyclonal)	R and D Systems	AF3564	1:500
Antibody	anti-ZIC3 (rabbit polyclonal)	PMID: 23217714		1:50
Other	Alexa Fluor 594-conjugated GS Lectin	Thermo Fisher Scientific	L21416	1:200
Peptide, recombinant protein	Tn5 transposase	Illumina	FC-121-1030	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Commercial assay or kit	Worthington Papain Dissociation Kit	Worthington Biochemical Corporation	LK003160	
Commercial assay or kit	Rneasy Micro Plus Kit	Qiagen	74034	
Commercial assay or kit	Dneasy Blood and Tissue Kit	Qiagen	69504	
Commercial assay or kit	MinElute GelExtraction Kit	Qiagen	28604	
Commercial assay or kit	Agencourt AMPure XP beads	Beckman Coulter	A63880	
Commercial assay or kit	EZ DNA Methylation-Direct Kit	Zymo	D5021	
Software, algorithm	RSEM	PMID: 21816040	RRID:SCR_013027	
Software, algorithm	DESeq2	PMID: 25516281	RRID:SCR_015687	
Software, algorithm	Bowtie2	PMID: 22388286		
Software, algorithm	MACS2	PMID: 18798982		
Software, algorithm	DiffBind	PMID: 22217937	RRID:SCR_012918	
Software, algorithm	deepTools	PMID: 27079975		
Software, algorithm	HOMER	PMID: 20513432		
Software, algorithm	Monocle	PMID: 24658644		
Software, algorithm	Methylpy	PMID: 26030523		
Software, algorithm	BEDTools	PMID: 20110278	RRID:SCR_006646	

## Mice

P7 homozygous *Tie2-GFP* mice (**Motoike et al., 2000**); JAX 003658) were used for isolation of tissue-specific ECs. To visualize Wnt activity in ECs, homozygous *Tie2-Cre* mice (**Kisanuki et al., 2001**); JAX 008863) were crossed to homozygous *R26-Tcf/Lef-LSL-H2B-GFP-6xMYC* mice (**Cho et al., 2017**). To control for the possibility of sex-dependent differences, male mice were used for RNA-seq, ATAC-seq, and MethylC-seq. All mice were housed and handled according to the approved Institutional Animal Care and Use Committee (IACUC) protocol MO16M367 of the Johns Hopkins Medical Institutions.

## EC isolation

ECs were isolated as previously described with some modifications (**Daneman et al., 2010**; **Zhang et al., 2014**). Reagents used included the Worthington Papain Dissociation System (LK003160, Worthington Biochemical Corporation, Lakewood, NJ), Dulbecco's PBS (DPBS, 14287072, Thermo Fisher Scientific, Waltham, MA), trehalose (T0167-10G, MilliporeSigma, Burlington, MA), bovine serum albumin (A7906, Sigma-Aldrich, Burlington, MA), 20 µm cell strainers (Cell-Trics, Sysmex Partec, Gortitz, Germany), and anti-CD11b BV421 (101235, Biolegends, San Diego, CA). For liver, lung, and kidney, a single P7 *Tie2-GFP* mouse typically yielded enough tissue for FACS sorting. For P7, two mice were used. Mice were euthanized and tissues of interest were rapidly

placed into Dulbecco's PBS (DPBS), minced using a razor blade, and enzymatically dissociated in 2.5 ml of papain solution (20 units papain per ml; 1 mM L-cysteine; 0.5 mM EDTA; 5% trehalose; 100 units of DNase per ml) for 1 hr at 37° Celsius. Following this incubation, ovomucoid inhibitor/BSA solution reconstituted at 10 mg per ml was added at 1:10 dilution. The dissociated tissue was gently triturated to form a single-cell suspension. This suspension was then layered on top of a solution consisting of 5 mg per ml ovomucoid inhibitor/BSA and 5% trehalose and centrifuged at 70 x g for 5 min at room temperature. The cell pellet was resuspended in 0.02% BSA and 5% trehalose solution (FACS buffer) and subsequently filtered through a 20 um mesh filter. After another round of centrifugation at 300 x g for 5 min, the cells were resuspended in FACS buffer. Cells were incubated with anti-CD11b BV421 (a macrophage marker) at 1:1000 dilution for 30 min at room temperature. Following antibody staining, cells were pelleted at 300 x g and washed twice with 1 ml of FACS buffer.

Cells were sorted using a MoFlo XDP Sorter (Beckman Coulter, Brea, CA). Viable ECs were defined as GFP positive, propidium iodide negative, and CD11b BV421 negative. For RNA-seq, viable ECs (GFP-positive) and viable parenchymal cells (GFP-negative) were sorted directly into QIAGEN Buffer RLT Plus. For single-cell RNA-seq, ATAC-seq, and MethylC-seq, viable ECs were sorted into FACS buffer for downstream sample preparation.

### Immunohistochemistry

E13.5 embryos were immersion fixed in 4% paraformaldehyde (PFA) overnight at 4° Celsius. For P7 brain, liver, lung, and kidneys, mice were deeply anesthetized and then the tissues were fixed by intracardiac perfusion with PBS followed by 4% PFA in PBS. Tissues were further fixed by immersion in 4% PFA in PBS overnight at 4°C. Embryos and tissues were embedded in 3% low-melting point agarose and vibratome sectioned at a thickness of 120 um. The following reagents were used: DAPI, rat anti-ICAM2 (1:300; 553326, BD BioSciences, San Jose, CA), chicken anti-6xMYC (1:10,000; *Wu et al., 2014*), rabbit anti-GLUT1 (1:500; RB-9052, Thermo Fisher Scientific), rabbit anti-ERG (1:500; A7L1G, Cell Signaling Technology, Danvers, MA), goat anti-dopa decarboxylase (1:500; AF3564, R and D Systems, Minneapolis, MN), Alexa Fluor 594-conjugated GS Lectin (1:200; L21416; Thermo Fisher Scientific), rabbit anti-ZIC3 (1:50; *Wang et al., 2012*). For the images in **Figure 6A and C**, the unit for quantification was a rectangle 410 um high x 355 um wide, and quantifications were performed using three adjacent rectangles overlying the CNS and three adjacent rectangles overlying the mesoderm (i.e. non-CNS) region. Within each rectangle, the number of MYC+, ERG+, or ZIC3 +nuclei were counted and the vasculature was traced in Adobe Illustrator and the lengths of the traced lines were quantified by using ImageJ to count the resulting pixels.

### Whole mount retina in situ hybridization

Flat-mount retina ISH was performed as previously described (*Rattner et al., 2013*). Briefly, enucleated eyes were fixed in 2% PFA for 5 min at room temperature and then transferred to hypertonic PBS (2x PBS) for 5 min at room temperature. Retinas were dissected and four radial incisions were made. With the vitreal surface facing up, buffer was slowly removed until retinas were flat. In order to fix the flattened retinas, methanol cooled to -20°C was added dropwise. For ISH, fixed retinas were first transferred to 4% PFA for 5 min at room temperature, then washed three times with RNAase-free PBS with 0.1% Tween-20 (PBST), digested with 20 ug/mL Proteinase K (P2308, Sigma-Aldrich) for 15 min at room temperature, fixed in 0.2% glutaraldehyde with 4% PFA, washed three times with PBST, and finally transferred to hybridization solution (50% formamide, 5x saline-sodium citrate buffer, 5 mM EDTA, 2% Blocking Reagent (11096176001, Sigma-Aldrich), 50 ug/mL yeast RNA, 100 ug/mL heparin, 0.5% CHAPS, and 0.1% Triton X-100) at 65°C. After 2 hr of prehybridization, digoxigenin-labeled riboprobes were added, and hybridization proceeded overnight at 65°C. After a series of washes (*Rattner et al., 2013*), retinas were blocked for 2 hr at room temperature in maleic acid buffer with 0.1% Tween-20 (MABT), 2% Blocking Reagent, and 10% normal goat serum. Then, retinas were incubated with anti-digoxigenin-AP Fab fragments (1:1000; 11093274910, Sigma-Aldrich) overnight at 4°C. The following day, retinas were washed four times for 1 hr each at room temperature and then overnight at 4°C in MABT. Retinas were washed three times in AP buffer with 0.1% Tween-20 for 5 min at room temperature, and then they were stained for 1–4 hr with AP buffer and nitro-blue tetrazolium (NBT; 350 ug/mL) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP; 175 ug/mL). After staining, retinas were washed twice with PBST, post-fixed for 5 min with 4% PFA,

washed twice with PBST, and stained with Alexa Fluor 594-conjugated GS Lectin (1:100) overnight at 4°C. For imaging, retinas were dehydrated by transferring the retinas between increasing concentrations of ethanol (25% in PBST, 50% in PBST, 75% in PBST, and 100%). Finally, retinas were cleared in benzyl benzoate:benzyl alcohol (2:1). For presentation, the purple ISH signal was converted to green, and the GS Lectin signal was converted to magenta.

## Microscopy

Confocal images were captured with an LSM700 (Zeiss, Jena, Germany) microscope. ISH images were captured with a Zeiss Imager Z1 Microscope using Zeiss AxioVision 4.6 software. Image processing was performed using FIJI (*Schindelin et al., 2012*).

## RNA and DNA sample preparation

RNA was extracted from an aliquot of dissociated tissue prior to filtering and antibody staining and from GFP-positive and GFP-negative FACS sorted cells using the RNeasy Micro Plus kit (74034, QIAGEN, Venlo, Netherlands). DNA for MethylC-seq was prepared from GFP-positive FACS-sorted cells using the DNeasy Blood and Tissue kit (69504, QIAGEN). For ATAC-seq, ~50,000 GFP-positive FACS-sorted cells were resuspended in ice-cold lysis buffer (0.25 M sucrose, 25 mM KCl, 5 mM MgCl<sub>2</sub>, 20 mM Tricine-KOH, 0.1% Igepal CA-630) and immediately centrifuged at 500 x g for 10 min at 4°C to prepare nuclei. The resulting nuclear pellet was resuspended in a 50 ul reaction volume in Tn5 transposase and transposase reaction buffer (FC-121–1030, Illumina Inc., San Diego, CA) and the tagmentation reaction was incubated at 37°C for 30 min.

## Library preparation and sequencing

Each RNA-seq, ATAC-seq, and MethylC-seq analysis was conducted on two biological replicates except for brain EC ATAC-seq, which was conducted on three biological replicates. Single-cell RNA-seq was conducted with a single sample. Libraries for RNA-seq and ATAC-seq were prepared as previously described, with minor modifications (*Buenrostro et al., 2015; Lister et al., 2013; Mo et al., 2015; Mo et al., 2016*). For RNA-seq, total RNA was converted to cDNA and amplified (Ovation Ultralow System V2-32, 0342HV, NuGEN Technologies, San Carlos, CA). Amplified cDNA was fragmented, end-repaired, linker-adapted, and single-end sequenced for 75 cycles on a NextSeq500 (Illumina Inc.). Tagmented DNA was purified using QIAGEN MinElute GelExtraction kit (28604, Qiagen). ATAC-seq libraries were PCR amplified for 11 cycles. Agencourt AMPure XP beads (A63880, Beckman Coulter) were used to purify ATAC-seq libraries, which were then paired-end sequenced for 36 cycles on a NextSeq500. MethylC-seq library preparation discussed below.

## Data analysis

Most data analysis was performed as previously described (*Mo et al., 2016*). For basic data processing, exploration, and visualization, we used deepTools (*Ramírez et al., 2016*), BEDTools (*Quinlan and Hall, 2010*), RStudio (*Team RS, 2016*), the tidyverse collection of R packages (*Wickham, 2017*), ggplot2 (*Wickham, 2009*), pheatmap (*Kolde, 2015*), and custom scripts. Reads were aligned to the mm10 genome using Bowtie2 (*Langmead and Salzberg, 2012*). The Integrative Genomics Viewer (IGV) was used to visualize bigwig files generated by deepTools bamcoverage (*Robinson et al., 2011; Thorvaldsdóttir et al., 2013*).

## RNA-seq

Three different sets of RNA-seq data were generated from papain dissociated P7 *Tie2-GFP* brain, liver, lung, and kidney: (1) total tissue; (2) GFP-negative FACS-sorted cells; and (3) GFP-positive FACS-sorted cells. RSEM version 1.3.0 (*Li and Dewey, 2011*) was used to align reads with Bowtie2 to the mm10/GRCm38 genome and calculate transcript expression using Ensembl annotation (*rsem-calculate-expression* –bowtie2 –estimate-rspd –append-names –output-genome-bam –sort-bam-by-coordinate). Differentially expressed genes were identified using EBSeq version 1.18.0 (*Leng and Kendziorski, 2015*). To filter out background transcripts from surrounding parenchymal cells, we determined a set of EC-enriched transcripts for each tissue by comparing GFP-positive and GFP-negative sorted cells. A transcript was considered EC-enriched if it met the following three criteria: (1) a minimum two-fold enrichment in GFP-positive compared to GFP-negative samples; (2) a

posterior probability of differential expression (PPDE) greater than or equal to 0.95 [PPDE = (1 - false discovery rate)]; (3) relative expression greater than or equal to 10 transcripts per million (TPM) in both biological replicates. A gene was considered to be differentially expressed between EC subtypes if it met the following criteria: (1) EC-enriched with minimum of two-fold enrichment between one subtype and all three other subtypes; (2) a PPDE greater than or equal to 0.95; (3) a TPM value greater than or equal to 10 in both biological replicates. Principal component analysis was performed on 'regularized log'-transformed data using the DESeq2 *rlog* and *plotPCA* function (Love et al., 2014). Gene ontology enrichment analysis was performed on EC tissue-specific genes using the online enrichment analysis tool (Mi et al., 2017) from the Gene Ontology Consortium.

## ATAC-seq

We aligned ATAC-seq data using Bowtie2 (Version 2.3.2 t -X 2000 -no-mixed -no-discordant) and then removed duplicate reads (*picard MarkDuplicates*). Peaks were called using MACS2 (Version 2.1.1.20160309 *callpeak* -nomodel -keep-dup all -shift -100 -extsize 200 -call-summits) (Zhang et al., 2008). Peaks were then filtered for fold-change >2 and -log(qvalue)>2. ATAC-seq peaks visualized on the browser were called using all fragment sizes. To determine the fraction of mapped reads in called peaks, we used deepTools *multiBamSummary* (BED-file -outRawCounts -samFlagInclude 64). Samtools *view* (-s) was used to subsample reads. For all downstream analyses, ATAC-seq peaks were called using paired-end sequencing fragments < 100 bp in length, which are enriched for nucleosome-free regions.

To identify differential ATAC-seq peaks between ECs isolated from brain, liver, lung, and kidney, we used DiffBind (Stark and Brown, 2011; Ross-Innes et al., 2012) with both DESEQ2 and EdgeR (Robinson et al., 2010; McCarthy et al., 2012) methods. DiffBind was used to develop a set of consensus peaks between replicates using the requirement that peaks must be in at least two thirds of the replicates (minOverlap = 0.66). From this new set of peaks, a consensus peakset was generated corresponding to the discrete set of peaks identified between all samples being compared (minOverlap = 1). Log<sub>2</sub> transformed normalized read counts in each sample are then tabulated over each peak region and plotted. To retrieve a set of high-confidence, cell type-enriched peaks, differential peak analysis was carried out using DESEQ2 and EdgeR. From both sets of analysis, we filtered for peaks with an absolute fold difference >2 and FDR < 0.05. In R, the union of these two sets of peaks was used to generate a final set of differentially enriched peaks for each sample. Principal component analysis was performed on 'regularized log'-transformed data using the DESeq2 *rlog* and *plotPCA* function.

## Methylome sequencing

DNA methylome libraries were prepared using a modified snmC-seq protocol adapted for bulk DNA samples (Luo et al., 2017). 20 ng of purified genomic DNA with 0.5% unmethylated lambda DNA spike-in (D1521, Promega) was bisulfite converted using the EZ DNA Methylation-Direct Kit (D5021, Zymo) following the product manual and was eluted in 10 µl M-Elution buffer. 9 µl converted DNA was mixed with 1 µl P5L\_random oligo (5 µM, /5SpC3/TTCCCTACACGACGCTC TTCCGATCTNNNNNNNNNN, IDT) followed by heat denaturing at 95°C for 3 min using a thermocycler. The samples were chilled on ice for 2 min and mixed with 10 µl enzyme mix containing 2 µl of Blue Buffer (B0110, Enzymatics), 1 µl of 10 mM dNTP (N0447L, NEB), 1 µl of Klenow exo- (50 U/µl, P7010-HC-L, Enzymatics) and 6 µl H<sub>2</sub>O. The samples were incubated using a thermocycler with the following program: 4°C for 5 min, ramp up to 25°C at 0.1°C/s, 25°C for 5 min, ramp up to 37°C at 0.1°C/s, 37°C for 60 min, 4°C. A 3 µl enzyme mix containing 2 µl of Exonuclease 1 (20U/µl, X8010L, Enzymatics) and 1 µl of Shrimp Alkaline Phosphatase (rSAP, M0371L, NEB) was added to the samples. The samples were incubated at 37°C for 30 min using a thermocycler. Samples were cleaned-up using 0.8x SPRI beads and eluted with 10 µl M-Elution buffer. Eluted samples were heated at 95°C for 3 min using a thermocycler and were chilled on ice for 2 min. Samples were mixed with 10.5 µl Adaptase master mix containing 2 µl Buffer G1, 2 µl Regent G2, 1.25 µl Reagent G3, 0.5 µl Enzyme G4, 0.5 µl Enzyme G5 and 4.25 µl M-Elution buffer (Accel-NGS Adaptase Module for Single Cell Methyl-Seq Library Preparation, 33096, Swift Biosciences). Samples were incubated at 37°C for 30 min, 95°C for 2 min and 4°C using a thermocycler. 1 µl 30 µM P5 indexing primer and 5 µl 10 µM P7 indexing primer and 25 µl KAPA HiFi HotStart ReadyMix, (KK2602, KAPA BIOSYSTEMS) were

added into each sample followed by thermocycling using the following program - 95°C for 2 min, 98°C for 30 s, 10 cycles of (98°C for 15 s, 64°C for 30 s, 72°C for 2 min), 72°C for 5 min and then 4°C. PCR products were cleaned up using 0.8x SPRI beads twice. Methylome libraries were sequenced using an Illumina HiSeq 4000 instrument with 5% PhiX DNA spike-in.

Adaptor sequences were trimmed from sequencing reads using Cutadapt 1.11 with parameters `-f fastq -q 20 -m 50 -a AGATCGGAAGAGCACACGTCTGAAAC -A AGATCGGAAGAGCGTCGTG TAGGGA`. We further trimmed 10 bp from both 5'- and 3'- ends of both R1 and R2 reads with parameters `-f fastq -u 10 -m 30`. Trimmed R1 and R2 reads were mapped to the mm10 reference genome as single-end reads using bismark 0.14.4 with parameter `-bowtie2`. `-pbat` option was used for mapping R1 reads. Mapped reads were sorted using Samtools 1.3 followed by removing duplicate reads using Picard 1.141 MarkDuplicates with parameter `REMOVE_DUPLICATES = true`. Reads were further filtered by `MAPQ > 20` using samtools view with parameter `-bhq20`. Methylation level at each cytosine was summarized using `call_methylated_sites` from Methylpy package (<https://github.com/yupenghe/methylpy>).

### Methylome analysis and identification of ECTS-hypo-DMRs, UMRs, LMRs, and large hypomethylation features

DMRs were identified as previously described in (Mo et al., 2015) using *DMRfind* from the Methylpy package with parameters `num_sims = 3000, num_sig_tests = 100, use_mc_status = False, dmr_max_dist = 250, sig_cutoff = 0.01`. DMRs showing hypo-methylation in both biological replicates were used for further analyses.

UMRs and LMRs were identified using MethylSeekR (Burger et al., 2013) with `m` (methylation level)=0.3% and 5% FDR. DMVs were identified as UMRs > 3 kb. To identify large ECTS-hypo-DMRs, all ECTS hypo-DMRs with inter-DMR distances < 1 kb were merged (`bedtools merge -d 1000`). Large ECTS-hypo-DMRs were defined as merged ECTS-hypo-DMRs>2 kb.

For MethylC-seq scatterplots in **Figure 2A** and **Figure 2—figure supplement 1**, *multiBigwigSummary* (BED-file `-outRawCounts`) of deepTools was used to compute the mean fraction of mCG in a 5 kb window 3' of the TSS of protein-coding genes. Gene ontology analysis of shared DMVs was performed using Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean et al., 2010). To compare the degree of methylation between EC subtypes at DMVs overlapping differentially expressed TF genes, we applied the following steps: (1) within each EC subtype, all DMVs in a 20 kb window 5' to the TSS of the gene of interest were joined into a single contiguous block, (2) overlapping DMVs between EC subtypes were merged to generate a genomic interval that spanned the greatest extent of all the DMVs, (3) within the merged interval, the fraction mCG methylation was calculated for each EC subtype. Computation was performed using *multiBigwigSummary*. In **Figure 4—figure supplement 3B and C**, *computeMatrix* (reference-point `-referencePoint center`) was used to compute either the degree of methylation at ATAC-seq peaks or the ATAC-seq signal at UMRs, LMRs, and DMRs as a function of distance from the center of regions. Then the deepTools function *plotProfile* was used to generate output data in table form for plotting in R. The ATAC-seq signal at hypomethylated features was upper quartile normalized using the *uqua* function of the R package NOISeq (Tarazona et al., 2011; Tarazona et al., 2015). For **Figure 4—figure supplement 3C**, each line was translated vertically so that the value at position -1500 bp is 0.

### Feature overlap analysis

*Bedtools intersect* (`-u`) was used to determine features that overlapped with other features by greater than or equal to 1 bp. *Bedtools window* (`-w 100000`) was used to identify CREs within 100 kb of ECTSGs. For **Figure 2B**, statistically significant enrichment was tested by hypergeometric distribution using MATLAB *hygecdf* function as previously described in Mo et al., 2015. For **Figure 4—figure supplement 2C**, statistically significant enrichment was tested with *hygecdf(number of category i feature overlapping with category j DEG, sample size of either all hypo-DMRs and ECTS-hypo-DMRs or all differential APs and ECTSAPs, number of features (either DMRs or APs) that overlap with category j DEG, sample size of category i feature, 'upper')*.



## Transcription factor DNA binding motif enrichment analysis

TF binding motifs in candidate regulatory regions (ECTSAPs and ECTS-hypo-DMRs) were identified and analyzed using the HOMER suite of tools for motif discovery (Heinz *et al.*, 2010), in particular, *findMotifsGenome.pl* (-size given). HOMER searches for enriched motifs using a de novo strategy of looking for enriched k-mers in target regions compared to a set of random background genomic regions. HOMER also maintains a custom database of known motifs based on published ChIP-Seq data sets. From the set of known motifs, we filtered for motifs exhibiting a q-value <0.001 and a fold enrichment (defined as the % of target regions with the motif divided by the % of background regions with the motif) >2 for each sample. These were considered to be statistically significant. De novo-enriched motifs could typically be matched to a motif from the list of enriched known motifs. Because many members of a TF family share a similar core motif, we chose one representative member of each TF family motif to display in motif heatmaps (Figures 5B and F). *AnnotatePeaks.pl* (-m -size 1000 -hist 5) was used to generate histograms of enriched motifs, and *annotatePeaks.pl* (-center -size given -multi) was used to center candidate regions on instances of motifs. To identify the number of features containing a given motif, we used *annotatePeaks.pl* (-m -nomotifs -noann -nogene). To determine statistically significant enrichment of a motif (as in Figure 5—figure supplement 1B,C and E), we used *hygecdf*(number of category *i* feature overlapping with category *j* motif, sample size of either ECTS-hypo-DMRs or ECTSAPs within 100 kb of ECTSGs, number of features (either DMRs or APs) that contain category *j* motif, sample size of category *i* feature, 'upper').

## Droplet-based single-cell RNA-seq of brain ECs

FACS purified GFP-positive P7 brain ECs from a *Tie2-GFP* mouse were processed through the GemCode Single Cell Platform using the GemCode Gel Bead, Chip, and Library Kits (10X Genomics, Pleasanton, CA) as per the manufacturer's protocol. In brief, papain-dissociated ECs were sorted into 0.4% BSA-DPBS, and approximately 10,000 FACS-purified singlet brain ECs were added to the channel with a final recovery rate of 4,257 cells. ECs were then partitioned into gel beads in emulsion in the GemCode instrument for cell lysis and barcoded oligo-DT priming and reverse transcription of poly-adenylated RNA. This was followed by PCR amplification, shearing, and 5' adaptor and sample index attachment. Libraries were sequenced on an Illumina HiSeq 2500.

## Analyses of single EC transcriptomes

scRNA-seq data were pre-processed using the cellranger pipeline (v2.0; 10x Genomics) with default settings. The normalized geneXcell matrix was used as input for the Monocle R/Bioconductor package (Trapnell *et al.*, 2014). Log-normalized expression estimates (with a pseudocount of 1) were used as input for a modified dpFeature analysis (Monocle). Briefly, high-dispersion genes were identified as those genes with a residual greater than 1 \* the estimated mean-variance split using the 'estimateDispersions' method in Monocle. Non-endothelial-specific transcripts in the high-dispersion gene set, defined as transcripts enriched >2 fold by bulk RNA-seq in the FACS-purified GFP-negative fraction compared to the GFP-positive fraction, likely derived from high-abundance transcripts released and solubilized by lysis of non-ECs during the preparation of the single-cell suspension. 311 non-EC specific genes were excluded from the high-dispersion gene set and removed from subsequent analyses. For preliminary cell clustering, principal component analysis (PCA) was performed on the high-variance genes and components 2–10 were visualized in two dimensions using t-Distributed Stochastic Neighbor Embedding (t-SNE) (Van Der Maaten and Hinton, 2008; Krijthe, 2015; Macosko *et al.*, 2015). Component one was excluded as it was determined to primarily reflect capture efficiency (num\_genes\_expressed). Initial clusters were identified from the 2D embedding using Mclust. To refine the list of genes contributing to the initial clustering, we performed a Monocle differential gene test with respect to cluster assignment, accounting for total mRNAs in both full and reduced models as well. The top 1200 genes with significant differential expression across clusters were then selected and used as input for a second PCA and components 2–8 were reduced into two dimensions using tSNE. Final cluster assignments were determined using Mclust on the updated 2D embedding. Clusters were annotated based on the expression of canonical EC subtype markers and a differential gene test between the subtypes was performed using the R package Monocle. The final six EC clusters conformed well to the expression patterns of known markers for arterial, venous,

tip, capillary and mitotic cells. Cell trajectory analyses were carried out in Monocle on the same 1200 differentially expressed genes used for the second PCA.

## Acknowledgements

We thank Hao Zhang of the Johns Hopkins Bloomberg School of Public Health Flow Cytometry Core Facility, and Haiping Hao, Linda Orzolek, and Jasmeet Sethi of the Johns Hopkins Medical Institutions Deep Sequencing and Microarray Core Facility for technical assistance.

## Additional information

### Competing interests

Jeremy Nathans: Reviewing editor, *eLife*. The other authors declare that no competing interests exist.

### Funding

Funder	Grant reference number	Author
Howard Hughes Medical Institute		Joseph R Ecker Jeremy Nathans
National Eye Institute	R01EY018637	Jacob S Heng Amir Rattner Jeremy Nathans
Arnold and Mabel Beckman Foundation		Jacob S Heng Jeremy Nathans
Eunice Kennedy Shriver National Institute of Child Health and Human Development	F30HD088023	Mark F Sabbagh
National Science Foundation	IOS-1665692	Loyal A Goff
Maryland Stem Cell Research Fund	2016-MSCRFI-2805	Loyal A Goff
Johns Hopkins University	Science of Learning and Synergy Awards	Loyal A Goff

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

Mark F Sabbagh, Conceptualization, Data curation, Software, Formal analysis, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing; Jacob S Heng, Conceptualization, Data curation, Software, Formal analysis, Funding acquisition, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing; Chongyuan Luo, Data curation, Software, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing—review and editing; Rosa G Castanon, Joseph R Nery, Validation, Investigation; Amir Rattner, Conceptualization, Resources, Software, Supervision, Methodology, Project administration, Writing—review and editing; Loyal A Goff, Resources, Software, Supervision, Funding acquisition, Visualization, Writing—review and editing; Joseph R Ecker, Resources, Supervision, Funding acquisition, Writing—review and editing; Jeremy Nathans, Conceptualization, Resources, Supervision, Funding acquisition, Methodology, Writing—original draft, Project administration, Writing—review and editing

### Author ORCIDs

Mark F Sabbagh  <http://orcid.org/0000-0003-1996-5251>

Jacob S Heng  <http://orcid.org/0000-0001-6291-6688>

Chongyuan Luo  <http://orcid.org/0000-0002-8541-0695>

Loyal A Goff [id](http://orcid.org/0000-0003-2875-451X) <http://orcid.org/0000-0003-2875-451X>

Joseph R Ecker [id](http://orcid.org/0000-0001-5799-5895) <http://orcid.org/0000-0001-5799-5895>

Jeremy Nathans [id](http://orcid.org/0000-0001-8106-5460) <http://orcid.org/0000-0001-8106-5460>

### Ethics

Animal experimentation: This study was performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. All mice were housed and handled according to the approved Institutional Animal Care and Use Committee (IACUC) protocol MO16M367 of the Johns Hopkins Medical Institutions.

### Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.36187.041>

Author response <https://doi.org/10.7554/eLife.36187.042>

---

## Additional files

### Supplementary files

- Supplementary file 1. Characteristics of each sequencing sample. (A) RNA-seq. (B) ATAC-seq. (C) MethylC-seq

DOI: <https://doi.org/10.7554/eLife.36187.029>

- Supplementary file 2. Gene expression data. (A-C) Transcript abundances in raw expected read counts (A), cross-sample normalized read counts (B), and TPMs (C). (D) List of EC-enriched transcripts. (E) List of all differentially expressed genes between one or more pairs of EC subtypes. (F) List of ECTSGs

DOI: <https://doi.org/10.7554/eLife.36187.030>

- Supplementary file 3. Gene ontology (GO) enrichment analysis (A) GO enrichment analysis of brain ECTSGs. (B) GO enrichment analysis of liver ECTSGs. (C) GO enrichment analysis of lung ECTSGs. (D) GO enrichment analysis of kidney ECTSGs.

DOI: <https://doi.org/10.7554/eLife.36187.031>

- Supplementary file 4. Hypomethylated features in each EC subtype (A-D) UMRs for each EC subtype. (E-H) LMRs for each EC subtype. (I-L) DMRs for each EC subtype. (I'-L') ECTS-hypo-DMRs for each EC subtype. (M-P) DMVs for each EC subtype. (Q-T) ECTS-large hypo-DMRs for each EC subtype. (U) LMRs found only in ECs relative to photoreceptors and cortical neurons. (V) LMRs from (U) that are shared between EC subtypes

DOI: <https://doi.org/10.7554/eLife.36187.032>

- Supplementary file 5. Numbers underlying heatmaps shown throughout figures.

DOI: <https://doi.org/10.7554/eLife.36187.033>

- Supplementary file 6. Accessible chromatin peaks in each EC subtype. (A-D) ATAC-seq peaks for each EC subtype called using the full range of ATAC-seq fragment lengths. (E-H) ATAC-seq peaks for each EC subtype called using <100 nt ATAC-seq fragments. (I-L) Differential ATAC-seq peaks (<100 nt) for each EC subtype. (I'-L') ECTSAPs (<100 nt) for each EC subtype. (M) ATAC-seq peaks (<100 nt) found only in ECs relative to photoreceptors and cortical neurons. (N) ATAC-seq peaks from (M) that are shared between EC subtypes.

DOI: <https://doi.org/10.7554/eLife.36187.034>

- Supplementary file 7. HOMER motif files used in this study. (A) HOMER motif files for enriched k-mers identified in ECTS-hypo-DMRs and ECTSAPs. (B) HOMER motif file used for representative member of TF families. (C) HOMER motif file used for paired ETS:ZIC motif.

DOI: <https://doi.org/10.7554/eLife.36187.035>

- Supplementary file 8. Top 25 genes for each single-cell RNA-seq cluster. (A) Arterial cluster markers. (B) Capillary-A cluster markers. (C) Capillary-V cluster markers. (D) Mitotic cluster markers. (E) Tip cell cluster markers. (F) Venous cluster markers.

DOI: <https://doi.org/10.7554/eLife.36187.036>

- Transparent reporting form

DOI: <https://doi.org/10.7554/eLife.36187.037>

### Data availability

Sequencing data have been deposited in GEO under accession code GSE111839

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database, license, and accessibility information
Sabbagh MF, Heng JS, Luo C, Castanon RG, Nery JR, Rattner A, Goff LA, Ecker JR, Nathans J	2018	Transcriptional and Epigenomic Landscapes of CNS and non-CNS Vascular Endothelial Cells	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE111839">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE111839</a>	Publicly available at the NCBI Gene Expression Omnibus (accession no: GSE111839).

## References

- Afrakhte M, Morén A, Jossan S, Itoh S, Sampath K, Westermark B, Heldin CH, Heldin NE, ten Dijke P. 1998. Induction of inhibitory Smad6 and Smad7 mRNA by TGF-beta family members. *Biochemical and Biophysical Research Communications* **249**:505–511. DOI: <https://doi.org/10.1006/bbrc.1998.9170>, PMID: 9712726
- Aird WC, Edelberg JM, Weiler-Guettler H, Simmons WW, Smith TW, Rosenberg RD. 1997. Vascular bed-specific expression of an endothelial cell gene is programmed by the tissue microenvironment. *The Journal of Cell Biology* **138**:1117–1124. DOI: <https://doi.org/10.1083/jcb.138.5.1117>, PMID: 9281588
- Aird WC. 2007a. Phenotypic heterogeneity of the endothelium: I. structure, function, and mechanisms. *Circulation Research* **100**:158–173. DOI: <https://doi.org/10.1161/01.RES.0000255691.76142.4a>, PMID: 17272818
- Aird WC. 2007b. Phenotypic heterogeneity of the endothelium: ii. representative vascular beds. *Circulation Research* **100**:174–190. DOI: <https://doi.org/10.1161/01.RES.0000255690.03436.ae>, PMID: 17272819
- Amabile A, Migliara A, Capasso P, Biffi M, Cittaro D, Naldini L, Lombardo A. 2016. Inheritable silencing of endogenous genes by Hit-and-Run targeted epigenetic editing. *Cell* **167**:219–232. DOI: <https://doi.org/10.1016/j.cell.2016.09.006>, PMID: 27662090
- Ambrosi C, Manzo M, Baubec T. 2017. Dynamics and context-dependent roles of DNA methylation. *Journal of Molecular Biology* **429**:1459–1475. DOI: <https://doi.org/10.1016/j.jmb.2017.02.008>, PMID: 28214512
- Anderson KD, Pan L, Yang XM, Hughes VC, Walls JR, Dominguez MG, Simmons MV, Burfeind P, Xue Y, Wei Y, Macdonald LE, Thurston G, Daly C, Lin HC, Economides AN, Valenzuela DM, Murphy AJ, Yancopoulos GD, Gale NW. 2011. Angiogenic sprouting into neural tissue requires Gpr124, an orphan G protein-coupled receptor. *PNAS* **108**:2807–2812. DOI: <https://doi.org/10.1073/pnas.1019761108>, PMID: 21282641
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: tool for the unification of biology. *Nature Genetics* **25**:25–29. DOI: <https://doi.org/10.1038/75556>
- Bakhtle YS, Vane JR. 1974. Pharmacokinetic function of the pulmonary circulation. *Physiological Reviews* **54**:1007–1045. DOI: <https://doi.org/10.1152/physrev.1974.54.4.1007>, PMID: 4371230
- Bock C, Beerman I, Lien WH, Smith ZD, Gu H, Boyle P, Gnirke A, Fuchs E, Rossi DJ, Meissner A. 2012. DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. *Molecular Cell* **47**:633–647. DOI: <https://doi.org/10.1016/j.molcel.2012.06.019>, PMID: 22841485
- Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods* **10**:1213–1218. DOI: <https://doi.org/10.1038/nmeth.2688>, PMID: 24097267
- Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. 2015. ATAC-seq: a method for assaying chromatin accessibility Genome-Wide. *Current Protocols in Molecular Biology* **109**. DOI: <https://doi.org/10.1002/0471142727.mb2129s109>, PMID: 25559105
- Burger L, Gaidatzis D, Schübeler D, Stadler MB. 2013. Identification of active regulatory regions from DNA methylation data. *Nucleic Acids Research* **41**:e155. DOI: <https://doi.org/10.1093/nar/gkt599>, PMID: 23828043
- Bussmann J, Wolfe SA, Siekmann AF. 2011. Arterial-venous network formation during brain vascularization involves hemodynamic regulation of chemokine signaling. *Development* **138**:1717–1726. DOI: <https://doi.org/10.1242/dev.059881>, PMID: 21429983
- Cadigan KM, Waterman ML. 2012. TCF/LEFs and wnt signaling in the nucleus. *Cold Spring Harbor Perspectives in Biology* **4**:a007906. DOI: <https://doi.org/10.1101/cshperspect.a007906>, PMID: 23024173
- Cai Y, Bolte C, Le T, Goda C, Xu Y, Kalin TV, Kalinichenko VV. 2016. FOXF1 maintains endothelial barrier function and prevents edema after lung injury. *Science Signaling* **9**:ra40. DOI: <https://doi.org/10.1126/scisignal.aad1899>, PMID: 27095594
- Cao Y, Linden P, Farnebo J, Cao R, Eriksson A, Kumar V, Qi JH, Claesson-Welsh L, Alitalo K. 1998. Vascular endothelial growth factor C induces angiogenesis in vivo. *PNAS* **95**:14389–14394. DOI: <https://doi.org/10.1073/pnas.95.24.14389>, PMID: 9826710

- Chambeyron S**, Da Silva NR, Lawson KA, Bickmore WA. 2005. Nuclear re-organisation of the *hoxb* complex during mouse embryonic development. *Development* **132**:2215–2223. DOI: <https://doi.org/10.1242/dev.01813>, PMID: 15829525
- Chang AH**, Raftrey BC, D'Amato G, Surya VN, Poduri A, Chen HI, Goldstone AB, Woo J, Fuller GG, Dunn AR, Red-Horse K. 2017. DACH1 stimulates shear stress-guided endothelial cell migration and coronary artery growth through the CXCL12-CXCR4 signaling Axis. *Genes & Development* **31**:1308–1324. DOI: <https://doi.org/10.1101/gad.301549.117>, PMID: 28779009
- Chi JT**, Chang HY, Haraldsen G, Jahnsen FL, Troyanskaya OG, Chang DS, Wang Z, Rockson SG, van de Rijn M, Botstein D, Brown PO. 2003. Endothelial cell diversity revealed by global expression profiling. *PNAS* **100**:10623–10628. DOI: <https://doi.org/10.1073/pnas.1434429100>, PMID: 12963823
- Cho C**, Smallwood PM, Nathans J. 2017. Reck and Gpr124 are essential receptor cofactors for Wnt7a/Wnt7b-specific signaling in mammalian CNS angiogenesis and Blood-Brain barrier regulation. *Neuron* **95**:1056–1073. DOI: <https://doi.org/10.1016/j.neuron.2017.07.031>, PMID: 28803732
- Choe SK**, Ladam F, Sagerström CG. 2014. TALE factors poise promoters for activation by *hox* proteins. *Developmental Cell* **28**:203–211. DOI: <https://doi.org/10.1016/j.devcel.2013.12.011>, PMID: 24480644
- Choi KS**, Choi HJ, Lee JK, Im S, Zhang H, Jeong Y, Park JA, Lee IK, Kim YM, Kwon YG. 2016. The endothelial E3 ligase HECW2 promotes endothelial cell junctions by increasing AMOTL1 protein stability via K63-linked ubiquitination. *Cellular Signalling* **28**:1642–1651. DOI: <https://doi.org/10.1016/j.cellsig.2016.07.015>, PMID: 27498087
- Chuong EB**, Elde NC, Feschotte C. 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* **351**:1083–1087. DOI: <https://doi.org/10.1126/science.aad5497>, PMID: 26941318
- Chuong EB**, Rumi MA, Soares MJ, Baker JC. 2013. Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nature Genetics* **45**:325–329. DOI: <https://doi.org/10.1038/ng.2553>, PMID: 23396136
- Corada M**, Orsenigo F, Morini MF, Pitulescu ME, Bhat G, Nyqvist D, Breviaro F, Conti V, Briot A, Iruela-Arispe ML, Adams RH, Dejana E. 2013. Sox17 is indispensable for acquisition and maintenance of arterial identity. *Nature Communications* **4**:2609. DOI: <https://doi.org/10.1038/ncomms3609>, PMID: 24153254
- Cullen M**, Elzarrad MK, Seaman S, Zudaire E, Stevens J, Yang MY, Li X, Chaudhary A, Xu L, Hilton MB, Logsdon D, Hsiao E, Stein EV, Cuttitta F, Haines DC, Nagashima K, Tessarollo L, St Croix B. 2011. GPR124, an orphan G protein-coupled receptor, is required for CNS-specific vascularization and establishment of the blood-brain barrier. *PNAS* **108**:5759–5764. DOI: <https://doi.org/10.1073/pnas.1017192108>, PMID: 21421844
- Cusdin FS**, Nietlispach D, Maman J, Dale TJ, Powell AJ, Clare JJ, Jackson AP. 2010. The sodium channel  $\beta$ 3-subunit induces multiphasic gating in NaV1.3 and affects fast inactivation via distinct intracellular regions. *Journal of Biological Chemistry* **285**:33404–33412. DOI: <https://doi.org/10.1074/jbc.M110.114058>, PMID: 20675377
- Daneman R**, Agalliu D, Zhou L, Kuhnert F, Kuo CJ, Barres BA. 2009. Wnt/ $\beta$ -catenin signaling is required for CNS, but not non-CNS, angiogenesis. *PNAS* **106**:641–646. DOI: <https://doi.org/10.1073/pnas.0805165106>, PMID: 19129494
- Daneman R**, Zhou L, Agalliu D, Cahoy JD, Kaushal A, Barres BA. 2010. The mouse blood-brain barrier transcriptome: a new resource for understanding the development and function of brain endothelial cells. *PLoS One* **5**:e13741. DOI: <https://doi.org/10.1371/journal.pone.0013741>, PMID: 21060791
- Datto MB**, Li Y, Panus JF, Howe DJ, Xiong Y, Wang XF. 1995. Transforming growth factor  $\beta$  induces the cyclin-dependent kinase inhibitor p21 through a p53-independent mechanism. *PNAS* **92**:5545–5549. DOI: <https://doi.org/10.1073/pnas.92.12.5545>, PMID: 7777546
- De Kumar B**, Parker HJ, Paulson A, Parrish ME, Pushel I, Singh NP, Zhang Y, Slaughter BD, Unruh JR, Florens L, Zeitlinger J, Krumlauf R. 2017. HOXA1 and TALE proteins display cross-regulatory interactions and form a combinatorial binding code on HOXA1 targets. *Genome Research* **27**:1501–1512. DOI: <https://doi.org/10.1101/gr.219386.116>, PMID: 28784834
- De Val S**, Chi NC, Meadows SM, Minovitsky S, Anderson JP, Harris IS, Ehlers ML, Agarwal P, Visel A, Xu SM, Pennacchio LA, Dubchak I, Krieg PA, Stainier DY, Black BL. 2008. Combinatorial regulation of endothelial gene expression by ets and forkhead transcription factors. *Cell* **135**:1053–1064. DOI: <https://doi.org/10.1016/j.cell.2008.10.049>, PMID: 19070576
- del Toro R**, Prahst C, Mathivet T, Siegfried G, Kaminker JS, Larrivee B, Breant C, Duarte A, Takakura N, Fukamizu A, Penninger J, Eichmann A. 2010. Identification and functional analysis of endothelial tip cell-enriched genes. *Blood* **116**:4025–4033. DOI: <https://doi.org/10.1182/blood-2010-02-270819>, PMID: 20705756
- Dharmadhikari AV**, Sun JJ, Gogolewski K, Carofino BL, Ustiyani V, Hill M, Majewski T, Szafranski P, Justice MJ, Ray RS, Dickinson ME, Kalinichenko VV, Gambin A, Stankiewicz P. 2016. Lethal lung hypoplasia and vascular defects in mice with conditional *Foxf1* overexpression. *Biology Open* **5**:1595–1606. DOI: <https://doi.org/10.1242/bio.019208>, PMID: 27638768
- Edwards S**, Lalor PF, Nash GB, Rainger GE, Adams DH. 2005. Lymphocyte traffic through sinusoidal endothelial cells is regulated by hepatocytes. *Hepatology* **41**:451–459. DOI: <https://doi.org/10.1002/hep.20585>, PMID: 15723297
- Ehling M**, Adams S, Benedito R, Adams RH. 2013. Notch controls retinal blood vessel maturation and quiescence. *Development* **140**:3051–3061. DOI: <https://doi.org/10.1242/dev.093351>, PMID: 23785053

- Emera D**, Wagner GP. 2012. Transformation of a transposon into a derived prolactin promoter with function during human pregnancy. *PNAS* **109**:11246–11251. DOI: <https://doi.org/10.1073/pnas.1118566109>, PMID: 22733751
- Ernst J**, Kheradpour P, Mikkelsen TS, Shores N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, Ku M, Durham T, Kellis M, Bernstein BE. 2011. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**:43–49. DOI: <https://doi.org/10.1038/nature09906>, PMID: 21441907
- Eskeland R**, Leeb M, Grimes GR, Kress C, Boyle S, Sproul D, Gilbert N, Fan Y, Skoutchi AI, Wutz A, Bickmore WA. 2010. Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Molecular Cell* **38**:452–464. DOI: <https://doi.org/10.1016/j.molcel.2010.02.032>, PMID: 20471950
- Farley EK**, Olson KM, Zhang W, Rokhsar DS, Levine MS. 2016. Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. *PNAS* **113**:6508–6513. DOI: <https://doi.org/10.1073/pnas.1605085113>, PMID: 27155014
- Ferraiuolo MA**, Rousseau M, Miyamoto C, Shenker S, Wang XQ, Nadler M, Blanchette M, Dostie J. 2010. The three-dimensional architecture of hox cluster silencing. *Nucleic Acids Research* **38**:7472–7484. DOI: <https://doi.org/10.1093/nar/gkq644>, PMID: 20660483
- Fischer A**, Schumacher N, Maier M, Sendtner M, Gessler M. 2004. The notch target genes Hey1 and Hey2 are required for embryonic vascular development. *Genes & Development* **18**:901–911. DOI: <https://doi.org/10.1101/gad.291004>, PMID: 15107403
- Fraser J**, Rousseau M, Shenker S, Ferraiuolo MA, Hayashizaki Y, Blanchette M, Dostie J. 2009. Chromatin conformation signatures of cellular differentiation. *Genome Biology* **10**:R37. DOI: <https://doi.org/10.1186/gb-2009-10-4-r37>, PMID: 19374771
- Géraud C**, Koch PS, Zierow J, Klapproth K, Busch K, Olsavszky V, Leibing T, Demory A, Ulbrich F, Dieltz M, Singh S, Sticht C, Breitkopf-Heinlein K, Richter K, Karppinen SM, Pihlajaniemi T, Arnold B, Rodewald HR, Augustin HG, Schledzewski K, et al. 2017. GATA4-dependent organ-specific endothelial differentiation controls liver development and embryonic hematopoiesis. *Journal of Clinical Investigation* **127**:1099–1114. DOI: <https://doi.org/10.1172/JCI90086>, PMID: 28218627
- Hannon GJ**, Beach D. 1994. p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. *Nature* **371**:257–261. DOI: <https://doi.org/10.1038/371257a0>, PMID: 8078588
- Hasan SS**, Tsaryk R, Lange M, Wisniewski L, Moore JC, Lawson ND, Wojciechowska K, Schnittler H, Siekmann AF. 2017. Endothelial notch signalling limits angiogenesis via control of artery formation. *Nature Cell Biology* **19**:928–940. DOI: <https://doi.org/10.1038/ncb3574>, PMID: 28714969
- Hayashi Y**, Nomura M, Yamagishi S, Harada S, Yamashita J, Yamamoto H. 1997. Induction of various blood-brain barrier properties in non-neural endothelial cells by close apposition to co-cultured astrocytes. *Glia* **19**:13–26. DOI: [https://doi.org/10.1002/\(SICI\)1098-1136\(199701\)19:1<13::AID-GLIA2>3.0.CO;2-B](https://doi.org/10.1002/(SICI)1098-1136(199701)19:1<13::AID-GLIA2>3.0.CO;2-B), PMID: 8989564
- Heinz S**, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular Cell* **38**:576–589. DOI: <https://doi.org/10.1016/j.molcel.2010.05.004>, PMID: 20513432
- Heinz S**, Glass CK. 2011. Roles of Lineage-Determining Transcription Factors in Establishing Open Chromatin: Lessons From High-Throughput Studies. In: *Current Topics in Microbiology and Immunology*. **356** p. 1–15. DOI: [https://doi.org/10.1007/82\\_2011\\_142](https://doi.org/10.1007/82_2011_142)
- Heinz S**, Romanoski CE, Benner C, Glass CK. 2015. The selection and function of cell type-specific enhancers. *Nature Reviews Molecular Cell Biology* **16**:144–154. DOI: <https://doi.org/10.1038/nrm3949>, PMID: 25650801
- Helms HC**, Abbott NJ, Burek M, Cecchelli R, Couraud PO, Deli MA, Förster C, Galla HJ, Romero IA, Shusta EV, Stebbins MJ, Vandenhaute E, Weksler B, Brodin B. 2016. In vitro models of the blood-brain barrier: An overview of commonly used brain endothelial cell culture models and guidelines for their use. *Journal of Cerebral Blood Flow & Metabolism* **36**:862–890. DOI: <https://doi.org/10.1177/0271678X16630991>, PMID: 26868179
- Hermkens DM**, van Impel A, Urasaki A, Bussmann J, Duckers HJ, Schulte-Merker S. 2015. Sox7 controls arterial specification in conjunction with hey2 and efnb2 function. *Development* **142**:1695–1704. DOI: <https://doi.org/10.1242/dev.117275>, PMID: 25834021
- Hogan NT**, Whalen MB, Stolze LK, Hadelik NK, Lam MT, Springstead JR, Glass CK, Romanoski CE. 2017. Transcriptional networks specifying homeostatic and inflammatory programs of gene expression in human aortic endothelial cells. *eLife* **6**:e22536. DOI: <https://doi.org/10.7554/eLife.22536>, PMID: 28585919
- Hollenhorst PC**, McIntosh LP, Graves BJ. 2011. Genomic and biochemical insights into the specificity of ETS transcription factors. *Annual Review of Biochemistry* **80**:437–471. DOI: <https://doi.org/10.1146/annurev-biochem.79.081507.103945>, PMID: 21548782
- Hon GC**, Rajagopal N, Shen Y, McCleary DF, Yue F, Dang MD, Ren B. 2013. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nature Genetics* **45**:1198–1206. DOI: <https://doi.org/10.1038/ng.2746>, PMID: 23995138
- Hu S**, Wan J, Su Y, Song Q, Zeng Y, Nguyen HN, Shin J, Cox E, Rho HS, Woodard C, Xia S, Liu S, Lyu H, Ming GL, Wade H, Song H, Qian J, Zhu H. 2013. DNA methylation presents distinct binding sites for human transcription factors. *eLife* **2**:e00726. DOI: <https://doi.org/10.7554/eLife.00726>, PMID: 24015356
- Hupe M**, Li MX, Kneitz S, Davydova D, Yokota C, Kele J, Hot B, Stenman JM, Gessler M. 2017. Gene expression profiles of brain endothelial cells during embryonic development at bulk and single-cell levels. *Science Signaling* **10**:eaag2476. DOI: <https://doi.org/10.1126/scisignal.aag2476>, PMID: 28698213

- Hurtado R**, Zewdu R, Mtui J, Liang C, Aho R, Kurylo C, Selleri L, Herzlinger D. 2015. Pbx1-dependent control of VMC differentiation kinetics underlies gross renal vascular patterning. *Development* **142**:2653–2664. DOI: <https://doi.org/10.1242/dev.124776>, PMID: 26138478
- Jeong M**, Sun D, Luo M, Huang Y, Challen GA, Rodriguez B, Zhang X, Chavez L, Wang H, Hannah R, Kim SB, Yang L, Ko M, Chen R, Göttgens B, Lee JS, Gunaratne P, Godley LA, Darlington GJ, Rao A, et al. 2014. Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nature Genetics* **46**:17–23. DOI: <https://doi.org/10.1038/ng.2836>, PMID: 24270360
- Junge HJ**, Yang S, Burton JB, Paes K, Shu X, French DM, Costa M, Rice DS, Ye W. 2009. TSPAN12 regulates retinal vascular development by promoting norrin- but not Wnt-induced FZD4/beta-catenin signaling. *Cell* **139**:299–311. DOI: <https://doi.org/10.1016/j.cell.2009.07.048>, PMID: 19837033
- Kalinichenko VV**, Lim L, Stolz DB, Shin B, Rausa FM, Clark J, Whitsett JA, Watkins SC, Costa RH. 2001. Defects in pulmonary vasculature and perinatal lung hemorrhage in mice heterozygous null for the forkhead box f1 transcription factor. *Developmental Biology* **235**:489–506. DOI: <https://doi.org/10.1006/dbio.2001.0322>, PMID: 11437453
- Kim BG**, Kim YH, Stanley EL, Garrido-Martin EM, Lee YJ, Oh SP. 2017. CXCL12-CXCR4 signalling plays an essential role in proper patterning of aortic arch and pulmonary arteries. *Cardiovascular Research* **113**:1677–1687. DOI: <https://doi.org/10.1093/cvr/cvx188>, PMID: 29016745
- Kisanuki YY**, Hammer RE, Miyazaki J, Williams SC, Richardson JA, Yanagisawa M. 2001. Tie2-Cre transgenic mice: a new model for endothelial cell-lineage analysis in vivo. *Developmental Biology* **230**:230–242. DOI: <https://doi.org/10.1006/dbio.2000.0106>, PMID: 11161575
- Knolle PA**, Wohleber D. 2016. Immunological functions of liver sinusoidal endothelial cells. *Cellular & Molecular Immunology* **13**:347–353. DOI: <https://doi.org/10.1038/cmi.2016.5>, PMID: 27041636
- Kolde R**. 2015. pheatmap: Pretty Heatmaps. R Package. 1.0.8. <https://CRAN.R-project.org/package=pheatmap>
- Kribelbauer JF**, Laptenko O, Chen S, Martini GD, Freed-Pastor WA, Prives C, Mann RS, Bussemaker HJ. 2017. Quantitative analysis of the DNA methylation sensitivity of transcription factor complexes. *Cell Reports* **19**:2383–2395. DOI: <https://doi.org/10.1016/j.celrep.2017.05.069>, PMID: 28614722
- Krijthe JH**. 2015. Rtsne: t-distributed stochastic neighbor embedding using a Barnes-Hut implementation. <https://github.com/jkrijthe/Rtsne>
- Kuhnert F**, Mancuso MR, Shamloo A, Wang HT, Choksi V, Florek M, Su H, Fruttiger M, Young WL, Heilshorn SC, Kuo CJ. 2010. Essential regulation of CNS angiogenesis by the orphan G protein-coupled receptor GPR124. *Science* **330**:985–989. DOI: <https://doi.org/10.1126/science.1196554>, PMID: 21071672
- Kundu S**, Ji F, Sunwoo H, Jain G, Lee JT, Sadreyev RI, Dekker J, Kingston RE. 2017. Polycomb repressive complex 1 generates discrete compacted domains that change during differentiation. *Molecular Cell* **65**:432–446. DOI: <https://doi.org/10.1016/j.molcel.2017.01.009>, PMID: 28157505
- Kusakabe M**, Hasegawa K, Hamada M, Nakamura M, Ohsumi T, Suzuki H, Tran MT, Kudo T, Uchida K, Ninomiya H, Chiba S, Takahashi S. 2011. c-Maf plays a crucial role for the definitive erythropoiesis that accompanies Erythroblastic Island formation in the fetal liver. *Blood* **118**:1374–1385. DOI: <https://doi.org/10.1182/blood-2010-08-300400>, PMID: 21628412
- Lacorre DA**, Baekkevold ES, Garrido I, Brandtzaeg P, Haraldsen G, Amalric F, Girard JP. 2004. Plasticity of endothelial cells: rapid dedifferentiation of freshly isolated high endothelial venule endothelial cells outside the lymphoid tissue microenvironment. *Blood* **103**:4164–4172. DOI: <https://doi.org/10.1182/blood-2003-10-3537>, PMID: 14976058
- Langmead B**, Salzberg SL. 2012. Fast gapped-read alignment with bowtie 2. *Nature Methods* **9**:357–359. DOI: <https://doi.org/10.1038/nmeth.1923>, PMID: 22388286
- Lau MS**, Schwartz MG, Kundu S, Savol AJ, Wang PI, Marr SK, Grau DJ, Schorderet P, Sadreyev RI, Tabin CJ, Kingston RE. 2017. Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. *Science* **355**:1081–1084. DOI: <https://doi.org/10.1126/science.aah5403>, PMID: 28280206
- Laurent L**, Wong E, Li G, Huynh T, Tsirigos A, Ong CT, Low HM, Kin Sung KW, Rigoutsos I, Loring J, Wei CL. 2010. Dynamic changes in the human methylome during differentiation. *Genome Research* **20**:320–331. DOI: <https://doi.org/10.1101/gr.101907.109>, PMID: 20133333
- Leng N**, Kendziorski C. 2015. EBSeq: An R package for gene and isoform differential expression analysis of RNA-seq data. R Package. 1.18.0.
- Lewis EB**. 1978. A gene complex controlling segmentation in Drosophila. *Nature* **276**:565–570. DOI: <https://doi.org/10.1038/276565a0>, PMID: 103000
- Li B**, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**:323. DOI: <https://doi.org/10.1186/1471-2105-12-323>, PMID: 21816040
- Li B**, Dorrell C, Canaday PS, Pelz C, Haft A, Finegold M, Grompe M. 2017. Adult mouse liver contains two distinct populations of cholangiocytes. *Stem Cell Reports* **9**:478–489. DOI: <https://doi.org/10.1016/j.stemcr.2017.06.003>, PMID: 28689996
- Li Y**, Zheng H, Wang Q, Zhou C, Wei L, Liu X, Zhang W, Zhang Y, Du Z, Wang X, Xie W. 2018. Genome-wide analyses reveal a role of polycomb in promoting hypomethylation of DNA methylation valleys. *Genome Biology* **19**:18. DOI: <https://doi.org/10.1186/s13059-018-1390-8>, PMID: 29422066
- Liebner S**, Corada M, Bangsow T, Babbage J, Taddei A, Czupalla CJ, Reis M, Felici A, Wolburg H, Fruttiger M, Taketo MM, von Melchner H, Plate KH, Gerhardt H, Dejana E. 2008. Wnt/beta-catenin signaling controls development of the blood-brain barrier. *The Journal of Cell Biology* **183**:409–417. DOI: <https://doi.org/10.1083/jcb.200806024>, PMID: 18955553

- Lister R**, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, Yu M, Tonti-Filippini J, Heyn H, Hu S, Wu JC, Rao A, Esteller M, He C, Haghghi FG, Sejnowski TJ, et al. 2013. Global epigenomic reconfiguration during mammalian brain development. *Science* **341**:1237905. DOI: <https://doi.org/10.1126/science.1237905>, PMID: 23828890
- Lister R**, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. 2008. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* **133**:523–536. DOI: <https://doi.org/10.1016/j.cell.2008.03.029>, PMID: 18423832
- Liu F**, Patient R. 2008. Genome-wide analysis of the zebrafish ETS family identifies three genes required for hemangioblast differentiation or angiogenesis. *Circulation Research* **103**:1147–1154. DOI: <https://doi.org/10.1161/CIRCRESAHA.108.179713>, PMID: 18832752
- Liu XS**, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, Shu J, Dadon D, Young RA, Jaenisch R. 2016. Editing DNA methylation in the mammalian genome. *Cell* **167**:233–247. DOI: <https://doi.org/10.1016/j.cell.2016.08.056>, PMID: 27662091
- Longobardi E**, Penkov D, Mateos D, De Florian G, Torres M, Blasi F. 2014. Biochemistry of the tale transcription factors PREP, MEIS, and PBX in vertebrates. *Developmental Dynamics* **243**:59–75. DOI: <https://doi.org/10.1002/dvdy.24016>, PMID: 23873833
- Love MI**, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**:550. DOI: <https://doi.org/10.1186/s13059-014-0550-8>, PMID: 25516281
- Luo C**, Keown CL, Kurihara L, Zhou J, He Y, Li J, Castanon R, Lucero J, Nery JR, Sandoval JP, Bui B, Sejnowski TJ, Harkins TT, Mukamel EA, Behrens MM, Ecker JR. 2017. Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. *Science* **357**:600–604. DOI: <https://doi.org/10.1126/science.aan3351>, PMID: 28798132
- Macosko EZ**, Basu A, Satija R, Nemes J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA. 2015. Highly parallel Genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**:1202–1214. DOI: <https://doi.org/10.1016/j.cell.2015.05.002>, PMID: 26000488
- Marcelo KL**, Goldie LC, Hirschi KK. 2013. Regulation of endothelial cell differentiation and specification. *Circulation Research* **112**:1272–1287. DOI: <https://doi.org/10.1161/CIRCRESAHA.113.300506>, PMID: 23620236
- McCarthy DJ**, Chen Y, Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**:4288–4297. DOI: <https://doi.org/10.1093/nar/gks042>, PMID: 22287627
- McLaughlin F**, Ludbrook VJ, Cox J, von Carlowitz I, Brown S, Randi AM. 2001. Combined genomic and antisense analysis reveals that the transcription factor *erg* is implicated in endothelial cell differentiation. *Blood* **98**:3332–3339. DOI: <https://doi.org/10.1182/blood.V98.12.3332>, PMID: 11719371
- McLean CY**, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology* **28**:495–501. DOI: <https://doi.org/10.1038/nbt.1630>, PMID: 20436461
- Meadows SM**, Myers CT, Krieg PA. 2011. Regulation of endothelial cell development by ETS transcription factors. *Seminars in Cell & Developmental Biology* **22**:976–984. DOI: <https://doi.org/10.1016/j.semcdb.2011.09.009>, PMID: 21945894
- Merks RM**, Brodsky SV, Goligorsky MS, Newman SA, Glazier JA. 2006. Cell elongation is key to in silico replication of in vitro vasculogenesis and subsequent remodeling. *Developmental Biology* **289**:44–54. DOI: <https://doi.org/10.1016/j.ydbio.2005.10.003>, PMID: 16325173
- Mi H**, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD. 2017. PANTHER version 11: expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements. *Nucleic Acids Research* **45**:D183–D189. DOI: <https://doi.org/10.1093/nar/gkw1138>, PMID: 27899595
- Mo A**, Luo C, Davis FP, Mukamel EA, Henry GL, Nery JR, Urich MA, Picard S, Lister R, Eddy SR, Beer MA, Ecker JR, Nathans J. 2016. Epigenomic landscapes of retinal rods and cones. *eLife* **5**:e11613. DOI: <https://doi.org/10.7554/eLife.11613>, PMID: 26949250
- Mo A**, Mukamel EA, Davis FP, Luo C, Henry GL, Picard S, Urich MA, Nery JR, Sejnowski TJ, Lister R, Eddy SR, Ecker JR, Nathans J. 2015. Epigenomic signatures of neuronal diversity in the mammalian brain. *Neuron* **86**:1369–1384. DOI: <https://doi.org/10.1016/j.neuron.2015.05.018>, PMID: 26087164
- Motoike T**, Loughna S, Perens E, Roman BL, Liao W, Chau TC, Richardson CD, Kawate T, Kuno J, Weinstein BM, Stainier DY, Sato TN. 2000. Universal GFP reporter for the study of vascular development. *Genesis* **28**:75–81. DOI: [https://doi.org/10.1002/1526-968X\(200010\)28:2<75::AID-GENE50>3.0.CO;2-S](https://doi.org/10.1002/1526-968X(200010)28:2<75::AID-GENE50>3.0.CO;2-S), PMID: 11064424
- Neri F**, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, Maldotti M, Anselmi F, Oliviero S. 2017. Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**:72–77. DOI: <https://doi.org/10.1038/nature21373>, PMID: 28225755
- Nolan DJ**, Ginsberg M, Israely E, Palikuqi B, Poulos MG, James D, Ding BS, Schachterle W, Liu Y, Rosenwaks Z, Butler JM, Xiang J, Rafii A, Shido K, Rabbany SY, Elemento O, Rafii S. 2013. Molecular signatures of tissue-specific microvascular endothelial cell heterogeneity in organ maintenance and regeneration. *Developmental Cell* **26**:204–219. DOI: <https://doi.org/10.1016/j.devcel.2013.06.017>, PMID: 23871589
- Noordermeer D**, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D. 2011. The dynamic architecture of hox gene clusters. *Science* **334**:222–225. DOI: <https://doi.org/10.1126/science.1207194>, PMID: 21998387
- Oh SY**, Kim JY, Park C. 2015. The ETS factor, ETV2: a master regulator for vascular endothelial cell development. *Molecules and Cells* **38**:1029–1036. DOI: <https://doi.org/10.14348/molcells.2015.0331>, PMID: 26694034



- Papageorgiou S.** 2012. Comparison of models for the collinearity of hox genes in the developmental axes of vertebrates. *Current genomics* **13**:245–251. DOI: <https://doi.org/10.2174/138920212800543093>, PMID: 23115525
- Park C,** Kim TM, Malik AB. 2013. Transcriptional regulation of endothelial cell and vascular development. *Circulation Research* **112**:1380–1400. DOI: <https://doi.org/10.1161/CIRCRESAHA.113.301078>, PMID: 23661712
- Parsa H,** Upadhyay R, Sia SK. 2011. Uncovering the behaviors of individual cells within a multicellular microvascular community. *PNAS* **108**:5133–5138. DOI: <https://doi.org/10.1073/pnas.1007508108>, PMID: 21383144
- Pham VN,** Lawson ND, Mugford JW, Dye L, Castranova D, Lo B, Weinstein BM. 2007. Combinatorial function of ETS transcription factors in the developing vasculature. *Developmental Biology* **303**:772–783. DOI: <https://doi.org/10.1016/j.ydbio.2006.10.030>, PMID: 17125762
- Pimanda JE,** Ottersbach K, Knezevic K, Kinston S, Chan WY, Wilson NK, Landry JR, Wood AD, Kolb-Kokocinski A, Green AR, Tannahill D, Lacaud G, Kouskoff V, Göttgens B. 2007. Gata2, Flt1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. *PNAS* **104**:17692–17697. DOI: <https://doi.org/10.1073/pnas.0707045104>, PMID: 17962413
- Pitulescu ME,** Schmidt I, Giaimo BD, Antoine T, Berkenfeld F, Ferrante F, Park H, Ehling M, Biljes D, Rocha SF, Langen UH, Stehling M, Nagasawa T, Ferrara N, Borggrefe T, Adams RH. 2017. Dll4 and Notch signalling couples sprouting angiogenesis and artery formation. *Nature Cell Biology* **19**:915–927. DOI: <https://doi.org/10.1038/ncb3555>, PMID: 28714968
- Poisson J,** Lemoine S, Boulanger C, Durand F, Moreau R, Valla D, Rautou PE. 2017. Liver sinusoidal endothelial cells: Physiology and role in liver diseases. *Journal of Hepatology* **66**:212–227. DOI: <https://doi.org/10.1016/j.jhep.2016.07.009>, PMID: 27423426
- Posokhova E,** Shukla A, Seaman S, Volate S, Hilton MB, Wu B, Morris H, Swing DA, Zhou M, Zudaire E, Rubin JS, St Croix B. 2015. GPR124 functions as a WNT7-specific coactivator of canonical  $\beta$ -catenin signaling. *Cell Reports* **10**:123–130. DOI: <https://doi.org/10.1016/j.celrep.2014.12.020>, PMID: 25558062
- Potente M,** Mäkinen T. 2017. Vascular heterogeneity and specialization in development and disease. *Nature Reviews Molecular Cell Biology* **18**:477–494. DOI: <https://doi.org/10.1038/nrm.2017.36>, PMID: 28537573
- Pruett ND,** Visconti RP, Jacobs DF, Scholz D, McQuinn T, Sundberg JP, Awgulewitsch A. 2008. Evidence for Hox-specified positional identities in adult vasculature. *BMC Developmental Biology* **8**:93. DOI: <https://doi.org/10.1186/1471-213X-8-93>, PMID: 18826643
- Quillien A,** Abdalla M, Yu J, Ou J, Zhu LJ, Lawson ND. 2017. Robust identification of developmentally active endothelial enhancers in zebrafish using FANS-Assisted ATAC-Seq. *Cell Reports* **20**:709–720. DOI: <https://doi.org/10.1016/j.celrep.2017.06.070>, PMID: 28723572
- Quinlan AR,** Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**:841–842. DOI: <https://doi.org/10.1093/bioinformatics/btq033>, PMID: 20110278
- Ramírez F,** Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research* **44**:W160–W165. DOI: <https://doi.org/10.1093/nar/gkw257>, PMID: 27079975
- Rattner A,** Yu H, Williams J, Smallwood PM, Nathans J. 2013. Endothelin-2 signaling in the neural retina promotes the endothelial tip cell state and inhibits angiogenesis. *PNAS* **110**:E3830–E3839. DOI: <https://doi.org/10.1073/pnas.1315509110>, PMID: 24043815
- Razin A,** Riggs AD. 1980. DNA methylation and gene function. *Science* **210**:604–610. DOI: <https://doi.org/10.1126/science.6254144>, PMID: 6254144
- Reggiani L,** Raciti D, Airik R, Kispert A, Brändli AW. 2007. The prepattern transcription factor Irx3 directs nephron segment identity. *Genes & Development* **21**:2358–2370. DOI: <https://doi.org/10.1101/gad.450707>, PMID: 17875669
- Riggs AD.** 1975. X inactivation, differentiation, and DNA methylation. *Cytogenetic and Genome Research* **14**:9–25. DOI: <https://doi.org/10.1159/000130315>, PMID: 1093816
- Risau W,** Hallmann R, Albrecht U, Henke-Fahle S. 1986. Brain induces the expression of an early cell surface marker for blood-brain barrier-specific endothelium. *The EMBO Journal* **5**:3179–3183. PMID: 3545814
- Robinson AS,** Materna SC, Barnes RM, De Val S, Xu SM, Black BL. 2014. An arterial-specific enhancer of the human endothelin converting enzyme 1 (ECE1) gene is synergistically activated by Sox17, FoxC2, and Etv2. *Developmental Biology* **395**:379–389. DOI: <https://doi.org/10.1016/j.ydbio.2014.08.027>, PMID: 25179465
- Robinson JT,** Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nature Biotechnology* **29**:24–26. DOI: <https://doi.org/10.1038/nbt.1754>, PMID: 21221095
- Robinson MD,** McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**:139–140. DOI: <https://doi.org/10.1093/bioinformatics/btp616>, PMID: 19910308
- Ross-Innes CS,** Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, Ali S, Chin SF, Palmieri C, Caldas C, Carroll JS. 2012. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**:389–393. DOI: <https://doi.org/10.1038/nature10730>, PMID: 22217937
- Schindelin J,** Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**:676–682. DOI: <https://doi.org/10.1038/nmeth.2019>, PMID: 22743772

- Schnabel CA**, Godin RE, Cleary ML. 2003. Pbx1 regulates nephrogenesis and ureteric branching in the developing kidney. *Developmental Biology* **254**:262–276. DOI: [https://doi.org/10.1016/S0012-1606\(02\)00038-6](https://doi.org/10.1016/S0012-1606(02)00038-6), PMID: 12591246
- Schultz MD**, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, Rajagopal N, Nery JR, Urich MA, Chen H, Lin S, Lin Y, Jung I, Schmitt AD, Selvaraj S, Ren B, Sejnowski TJ, Wang W, Ecker JR. 2015. Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* **523**:212–216. DOI: <https://doi.org/10.1038/nature14465>, PMID: 26030523
- Shimbo T**, Wade PA. 2016. Proteins that read DNA methylation. *Advances in Experimental Medicine and Biology* **945**:303–320. DOI: [https://doi.org/10.1007/978-3-319-43624-1\\_13](https://doi.org/10.1007/978-3-319-43624-1_13), PMID: 27826844
- Sorriento D**, Santulli G, Del Giudice C, Anastasio A, Trimarco B, Iaccarino G. 2012. Endothelial cells are able to synthesize and release catecholamines both in vitro and in vivo. *Hypertension* **60**:129–136. DOI: <https://doi.org/10.1161/HYPERTENSIONAHA.111.189605>, PMID: 22665130
- Spencer SL**, Cappell SD, Tsai FC, Overton KW, Wang CL, Meyer T. 2013. The proliferation-quiescence decision is controlled by a bifurcation in CDK2 activity at mitotic exit. *Cell* **155**:369–383. DOI: <https://doi.org/10.1016/j.cell.2013.08.062>, PMID: 24075009
- St Croix B**, Rago C, Velculescu V, Traverso G, Romans KE, Montgomery E, Lal A, Riggins GJ, Lengauer C, Vogelstein B, Kinzler KW. 2000. Genes expressed in human tumor endothelium. *Science* **289**:1197–1202. DOI: <https://doi.org/10.1126/science.289.5482.1197>, PMID: 10947988
- Stadler MB**, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, Tiwari VK, Schübeler D. 2011. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**:490–495. DOI: <https://doi.org/10.1038/nature10716>, PMID: 22170606
- Stark R**, Brown G. 2011. *DiffBind: Differential Binding Analysis of ChIP-Seq Peak Data*.
- Stenman JM**, Rajagopal J, Carroll TJ, Ishibashi M, McMahon J, McMahon AP. 2008. Canonical Wnt signaling regulates organ-specific assembly and differentiation of CNS vasculature. *Science* **322**:1247–1250. DOI: <https://doi.org/10.1126/science.1164594>, PMID: 19023080
- Stergachis AB**, Neph S, Reynolds A, Humbert R, Miller B, Paige SL, Vernot B, Cheng JB, Thurman RE, Sandstrom R, Haugen E, Heimfeld S, Murry CE, Akey JM, Stamatoyannopoulos JA. 2013. Developmental fate and cellular maturity encoded in human regulatory DNA landscapes. *Cell* **154**:888–903. DOI: <https://doi.org/10.1016/j.cell.2013.07.020>, PMID: 23953118
- Stewart PA**, Wiley MJ. 1981. Developing nervous tissue induces formation of blood-brain barrier characteristics in invading endothelial cells: a study using quail–chick transplantation chimeras. *Developmental Biology* **84**:183–192. DOI: [https://doi.org/10.1016/0012-1606\(81\)90382-1](https://doi.org/10.1016/0012-1606(81)90382-1), PMID: 7250491
- Strasser GA**, Kaminker JS, Tessier-Lavigne M. 2010. Microarray analysis of retinal endothelial tip cells identifies CXCR4 as a mediator of tip cell morphology and branching. *Blood* **115**:5102–5110. DOI: <https://doi.org/10.1182/blood-2009-07-230284>, PMID: 20154215
- Stroud H**, Su SC, Hrvatin S, Greben AW, Renthall W, Boxer LD, Nagy MA, Hochbaum DR, Kinde B, Gabel HW, Greenberg ME. 2017. Early-Life gene expression in neurons modulates lasting epigenetic states. *Cell* **171**:1151–1164. DOI: <https://doi.org/10.1016/j.cell.2017.09.047>, PMID: 29056337
- Sumanas S**, Choi K. 2016. ETS transcription factor ETV2/ER71/Etsrp in hematopoietic and vascular development. *Current Topics in Developmental Biology* **118**:77–111. DOI: <https://doi.org/10.1016/bs.ctdb.2016.01.005>, PMID: 27137655
- Sumanas S**, Gomez G, Zhao Y, Park C, Choi K, Lin S. 2008. Interplay among Etsrp/ER71, Scl, and Alk8 signaling controls endothelial and myeloid cell formation. *Blood* **111**:4500–4510. DOI: <https://doi.org/10.1182/blood-2007-09-110569>, PMID: 18270322
- Sumanas S**, Lin S. 2006. Ets1-related protein is a key regulator of vasculogenesis in zebrafish. *PLoS Biology* **4**:e10. DOI: <https://doi.org/10.1371/journal.pbio.0040010>, PMID: 16336046
- Tarazona S**, Furió-Tarí P, Turrà D, Pietro AD, Nueda MJ, Ferrer A, Conesa A. 2015. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. *Nucleic Acids Research* **43**:e140. DOI: <https://doi.org/10.1093/nar/gkv711>, PMID: 26184878
- Tarazona S**, García-Alcalde F, Dopazo J, Ferrer A, Conesa A. 2011. Differential expression in RNA-seq: a matter of depth. *Genome Research* **21**:2213–2223. DOI: <https://doi.org/10.1101/gr.124321.111>, PMID: 21903743
- Team R Studio**. 2016. RStudio: Integrated Development for R. Boston, RStudio, Inc. <http://www.rstudio.com/>
- ten Dijke P**, Arthur HM. 2007. Extracellular control of TGFbeta signalling in vascular development and disease. *Nature Reviews Molecular Cell Biology* **8**:857–869. DOI: <https://doi.org/10.1038/nrm2262>, PMID: 17895899
- The Gene Ontology Consortium**. 2017. Expansion of the gene ontology knowledgebase and resources. *Nucleic Acids Research* **45**:D331–D338. DOI: <https://doi.org/10.1093/nar/gkw1108>, PMID: 27899567
- Thompson PJ**, Macfarlan TS, Lorincz MC. 2016. Long terminal repeats: from parasitic elements to building blocks of the transcriptional regulatory repertoire. *Molecular Cell* **62**:766–776. DOI: <https://doi.org/10.1016/j.molcel.2016.03.029>, PMID: 27259207
- Thorvaldsdóttir H**, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* **14**:178–192. DOI: <https://doi.org/10.1093/bib/bbs017>, PMID: 22517427
- Thurman RE**, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, Garg K, John S, Sandstrom R, Bates D, Boatman L, Canfield TK, Diegel M, Dunn D, Ebersol AK, Frum T, et al. 2012. The accessible chromatin landscape of the human genome. *Nature* **489**:75–82. DOI: <https://doi.org/10.1038/nature11232>, PMID: 22955617

- Toshner M**, Dunmore BJ, McKinney EF, Southwood M, Caruso P, Upton PD, Waters JP, Ormiston ML, Skepper JN, Nash G, Rana AA, Morrell NW. 2014. Transcript analysis reveals a specific HOX signature associated with positional identity of human endothelial cells. *PLoS One* **9**:e91334. DOI: <https://doi.org/10.1371/journal.pone.0091334>, PMID: 24651450
- Trapnell C**, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology* **32**:381–386. DOI: <https://doi.org/10.1038/nbt.2859>, PMID: 24658644
- Van Der Maaten L**, Hinton G. 2008. Visualizing data using t-SNE. *J Machine Learning Res* **9**:2579–2605.
- Vanlandewijck M**, He L, Mäe MA, Andrae J, Ando K, Del Gaudio F, Nahar K, Lebouvier T, Laviña B, Gouveia L, Sun Y, Raschperger E, Räsänen M, Zarb Y, Mochizuki N, Keller A, Lendahl U, Betscholtz C. 2018. A molecular atlas of cell types and zonation in the brain vasculature. *Nature* **554**:475–480. DOI: <https://doi.org/10.1038/nature25739>, PMID: 29443965
- Versaevel M**, Grevesse T, Gabriele S. 2012. Spatial coordination between cell and nuclear shape within micropatterned endothelial cells. *Nature Communications* **3**:671. DOI: <https://doi.org/10.1038/ncomms1668>, PMID: 22334074
- Wang Y**, Rattner A, Zhou Y, Williams J, Smallwood PM, Nathans J. 2012. Norrin/Frizzled4 signaling in retinal vascular development and blood brain barrier plasticity. *Cell* **151**:1332–1344. DOI: <https://doi.org/10.1016/j.cell.2012.10.042>, PMID: 23217714
- Watanabe E**, Fujikawa A, Matsunaga H, Yasoshima Y, Sako N, Yamamoto T, Saegusa C, Noda M. 2000. Nav2/NaG channel is involved in control of salt-intake behavior in the CNS. *The Journal of Neuroscience* **20**:7743–7751. DOI: <https://doi.org/10.1523/JNEUROSCI.20-20-07743.2000>, PMID: 11027237
- Watanabe E**, Hiyama TY, Kodama R, Noda M. 2002. NaX sodium channel is expressed in non-myelinating Schwann cells and alveolar type II cells in mice. *Neuroscience Letters* **330**:109–113. DOI: [https://doi.org/10.1016/S0304-3940\(02\)00708-5](https://doi.org/10.1016/S0304-3940(02)00708-5), PMID: 12213645
- Wei GH**, Badis G, Berger MF, Kivioja T, Palin K, Enge M, Bonke M, Jolma A, Varjosalo M, Gehrke AR, Yan J, Talukder S, Turunen M, Taipale M, Stunnenberg HG, Ukkonen E, Hughes TR, Bulky ML, Taipale J. 2010. Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *The EMBO Journal* **29**:2147–2160. DOI: <https://doi.org/10.1038/emboj.2010.106>, PMID: 20517297
- Wickham H**. 2009. *Ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag. DOI: <https://doi.org/10.1007/978-0-387-98141-3>
- Wickham H**. 2017. tidyverse: Easily Install and Load the 'Tidyverse'. *R Package*. 1.2.1. <https://CRAN.R-project.org/package=tidyverse>
- Wu H**, Luo J, Yu H, Rattner A, Mo A, Wang Y, Smallwood PM, Erlanger B, Wheelan SJ, Nathans J. 2014. Cellular resolution maps of X chromosome inactivation: implications for neural development, function, and disease. *Neuron* **81**:103–119. DOI: <https://doi.org/10.1016/j.neuron.2013.10.051>, PMID: 24411735
- Xie W**, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, Whitaker JW, Tian S, Hawkins RD, Leung D, Yang H, Wang T, Lee AY, Swanson SA, Zhang J, Zhu Y, Kim A, Nery JR, Urich MA, Kuan S, et al. 2013. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* **153**:1134–1148. DOI: <https://doi.org/10.1016/j.cell.2013.04.022>, PMID: 23664764
- Xu C**, Hasan SS, Schmidt I, Rocha SF, Pitulescu ME, Bussmann J, Meyen D, Raz E, Adams RH, Siekmann AF. 2014. Arteries are formed by vein-derived endothelial tip cells. *Nature Communications* **5**:5758. DOI: <https://doi.org/10.1038/ncomms6758>, PMID: 25502622
- Xu Q**, Wang Y, Dabdoub A, Smallwood PM, Williams J, Woods C, Kelley MW, Jiang L, Tasman W, Zhang K, Nathans J. 2004. Vascular development in the retina and inner ear: control by norrin and Frizzled-4, a high-affinity ligand-receptor pair. *Cell* **116**:883–895. DOI: [https://doi.org/10.1016/S0092-8674\(04\)00216-8](https://doi.org/10.1016/S0092-8674(04)00216-8), PMID: 15035989
- Xu W**, Hong SJ, Zhong A, Xie P, Jia S, Xie Z, Zeitchek M, Niknam-Bienias S, Zhao J, Porterfield DM, Surmeier DJ, Leung KP, Galiano RD, Mustoe TA. 2015. Sodium channel Nax is a regulator in epithelial sodium homeostasis. *Science translational medicine*. **7**:ra177. DOI: <https://doi.org/10.1126/scitranslmed.aad0286>, PMID: 26537257
- Ye X**, Wang Y, Cahill H, Yu M, Badea TC, Smallwood PM, Peachey NS, Nathans J. 2009. Norrin, frizzled-4, and Lrp5 signaling in endothelial cells controls a genetic program for retinal vascularization. *Cell* **139**:285–298. DOI: <https://doi.org/10.1016/j.cell.2009.07.047>, PMID: 19837032
- Yin Y**, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, Das PK, Kivioja T, Dave K, Zhong F, Nitta KR, Taipale M, Popov A, Ginno PA, Domcke S, Yan J, Schübeler D, Vinson C, Taipale J. 2017. Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* **356**:eaaj2239. DOI: <https://doi.org/10.1126/science.aaj2239>, PMID: 28473536
- You LR**, Lin FJ, Lee CT, DeMayo FJ, Tsai MJ, Tsai SY. 2005. Suppression of Notch signalling by the COUP-TFII transcription factor regulates vein identity. *Nature* **435**:98–104. DOI: <https://doi.org/10.1038/nature03511>, PMID: 15875024
- Zhang Y**, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keefe S, Phatnani HP, Guarnieri P, Caneda C, Ruderisch N, Deng S, Liddelow SA, Zhang C, Daneman R, Maniatis T, Barres BA, Wu JQ, Jq W. 2014. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *Journal of Neuroscience* **34**:11929–11947. DOI: <https://doi.org/10.1523/JNEUROSCI.1860-14.2014>, PMID: 25186741
- Zhang Y**, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biology* **9**:R137. DOI: <https://doi.org/10.1186/gb-2008-9-9-r137>, PMID: 18798982

- Zhao Z**, Nelson AR, Betsholtz C, Zlokovic BV. 2015. Establishment and Dysfunction of the Blood-Brain Barrier. *Cell* **163**:1064–1078. DOI: <https://doi.org/10.1016/j.cell.2015.10.067>, PMID: 26590417
- Zhou P**, Gu F, Zhang L, Akerberg BN, Ma Q, Li K, He A, Lin Z, Stevens SM, Zhou B, Pu WT, Wt P. 2017. Mapping cell type-specific transcriptional enhancers using high affinity, lineage-specific Ep300 bioChIP-seq. *eLife* **6**: e22039. DOI: <https://doi.org/10.7554/eLife.22039>, PMID: 28121289
- Zhou Y**, Nathans J. 2014. Gpr124 controls CNS angiogenesis and blood-brain barrier integrity by promoting ligand-specific canonical wnt signaling. *Developmental Cell* **31**:248–256. DOI: <https://doi.org/10.1016/j.devcel.2014.08.018>, PMID: 25373781
- Zhou Y**, Wang Y, Tischfield M, Williams J, Smallwood PM, Rattner A, Taketo MM, Nathans J. 2014. Canonical WNT signaling components in vascular development and barrier formation. *Journal of Clinical Investigation* **124**:3825–3846. DOI: <https://doi.org/10.1172/JCI76431>, PMID: 25083995
- Ziller MJ**, Gu H, Müller F, Donaghey J, Tsai LT, Kohlbacher O, De Jager PL, Rosen ED, Bennett DA, Bernstein BE, Gnirke A, Meissner A. 2013. Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**:477–481. DOI: <https://doi.org/10.1038/nature12433>, PMID: 23925113