Title: The epigenomic landscape of single vascular cells reflects developmental origin and identifies disease risk loci

Authors: Chad S. Weldy, M.D., Ph.D.[1,2,#], Soumya Kundu, Ph.D.[3,4], João Monteiro, Ph.D.[1], Wenduo Gu, Ph.D.[1], Albert J. Pedroza, M.D.[6], Alex R. Dalal, M.D.[6], Matthew D. Worssam, Ph.D.[1], Daniel Li, M.D.[1], Brian Palmisano, M.D., Ph.D.[1], Quanyi Zhao, Ph.D.[1], Disha Sharma, Ph.D.[1], Trieu Nguyen, B.S.[1], Ramendra Kundu, Ph.D.[1], Michael P. Fischbein, M.D., Ph.D.[5,6], Jesse Engreitz, Ph.D.[4], Anshul B. Kundaje, Ph.D.[3,4], Paul P. Cheng, M.D., Ph.D.[1,6], Thomas Quertermous, M.D.[1,6,#]

[1]Department of Medicine, Division of Cardiovascular Medicine, Stanford University
[2]Stanford Center for Inherited Cardiovascular Disease
[3]Department of Computer Science, Stanford University
[4]Department of Genetics, Stanford University
[5]Department of Cardiothoracic Surgery, Stanford University
[6]Stanford Cardiovascular Institute, Stanford University

Running title: Epigenomic landscape of single vascular cells

[#]Co-correspondance:
Chad S. Weldy, MD., PhD
Falk CVRC, Rm 138
300 Pasteur Drive
Stanford, CA 94305
Email: weldyc@stanford.edu

Thomas Quertermous, M.D.
Falk CVRC, Rm 265
300 Pasteur Drive
Stanford, CA 94305
Email: tomq1@stanford.edu

Short title: Epigenomic landscape of single vascular cells

One Sentence Summary: Epigenomic landscape are not only cell type, but vascular site specific, with important implications in regulating vascular site-specific disease risk.

Abstract

Vascular sites have distinct susceptibility to atherosclerosis and aneurysm, yet the biological underpinning of vascular site-specific disease risk is largely unknown. Vascular tissues have different developmental origins that may influence global chromatin accessibility, and understanding differential chromatin accessibility, gene expression profiles, and gene regulatory networks (GRN) on single cell resolution may give key insight into vascular site-specific disease risk. Here, we performed single cell chromatin accessibility (scATACseq) and gene expression profiling (scRNAseq) of healthy adult mouse vascular tissue from three vascular sites, 1) aortic root and ascending aorta, 2) brachiocephalic and carotid artery, and 3) descending thoracic aorta. Through a comprehensive analysis at single cell resolution, we discovered key regulatory enhancers to not only be cell type, but vascular site specific in vascular smooth muscle (SMC), fibroblasts, and endothelial cells. We identified epigenetic markers of embryonic origin with differential chromatin accessibility of key developmental transcription factors such as *Tbx20*, *Hand2*, *Gata4*, and *Hoxb* family members and discovered transcription factor motif accessibility to be cell type and vascular site specific. Notably, we found ascending fibroblasts to have distinct epigenomic patterns, highlighting SMAD2/3 function to suggest a differential susceptibility to TGFβ, a finding we confirmed through *in vitro* culture of primary adventitial fibroblasts. Finally, to understand how vascular site-specific enhancers may regulate human genetic risk for disease, we integrated genome wide association study (GWAS) data for ascending and descending aortic dimension, and through using a distinct base resolution deep learning model to predict variant effect on chromatin

accessibility, ChromBPNet, to predict variant effects in SMC, Fibroblasts, and Endothelial cells within ascending aorta, carotid, and descending aorta sites of origin. We reveal that although cell type remains a primary influence on variant effects, vascular site modifies cell type transcription and highlights genomic regions that are enriched for specific TF motif footprints — including MEF2A, SMAD3, and HAND2. This work supports a paradigm that the epigenomic and transcriptomic landscape of vascular cells are cell type and vascular site-specific and that site-specific enhancers govern complex genetic drivers of disease risk.

Introduction

The risk to develop vascular disease is vascular site-specific — a clinical observation that has been well described for over a half century[1-3], yet the biological mechanisms by which this occurs remain poorly understood. Smooth muscle cells (SMC), fibroblasts, and endothelial cells, which make up vascular tissues, have distinct developmental origins[4-10] and this embryonic lineage diversity has been hypothesized to contribute to vascular site specific susceptibility to disease[5]. Genome wide association studies (GWAS) of vascular disease have revealed unique loci for disease in individual vascular sites, highlighting differential genetic and molecular etiologies for disease in differing vascular tissues[11-17]. Understanding the fundamental mechanisms that mediate this differential susceptibility to disease is critical to understand mechanisms of vascular disease and discover novel therapeutic targets.

Genomic structure and chromatin accessibility within the cell is intricately linked to function and governs transcriptional programs that drive disease risk[18]. This context is important when considering that heritability of complex disease is determined by common genetic variation that largely mediates disease risk through modifying genomic enhancer function within regions of open chromatin[19]. The recent advancement of single cell global chromatin accessibility profiling through *A*ssay for *T*ransposase *A*ccessible *C*hromatin *seq*uencing (scATACseq) in combination with single cell transcriptomic analysis (scRNAseq) has further defined that genetic risk of cardiac[20,21] and vascular disease[22-24] is driven in part through modification of cell type specific enhancers regulating expression of disease modifying genes in a cell type specific manner.

During vasculogenesis in early life, the carotid arteries and great vessels of the heart form from the pharyngeal arch arteries which appear in a craniocaudal sequence and then regress or remodel to form a definitive vascular pattern[25-28]. The adult murine aorta is composed of vascular cells from diverse embryonic origin including the secondary heart field, neural crest, and somitic mesoderm[5] and reside in spatially distinct domains[8]. Further, analysis of smooth muscle cells from adult healthy vascular tissue in mouse has shown a differential gene expression profile across vascular sites, which suggests relevance to vascular site-specific disease risk[29,30].

Epigenomic landscape of tissues may be influenced by embryonic origin, yet cell type specific enhancers of vascular tissues between vascular sites have yet to be mapped. By performing scATACseq combined with scRNAseq of healthy adult mouse aortic tissue from three vascular sites, 1) aortic root and ascending aorta, 2) brachiocephalic and carotid artery, and 3) descending thoracic aorta — representing the secondary heart field and neural crest, neural crest, and somitic mesoderm, respectively — we report that cell type specific enhancers of vascular cells are vascular site specific, suggest a developmental “memory”, and are related to disease risk genes. Through a comprehensive approach utilizing gene regulatory networks (GRN), *in vitro* culture of primary adventitial fibroblasts, and a novel technique to develop machine learning neural networks (ChromBPNet) to predict human variant effect within cell types across vascular sites, we not only reveal key differential chromatin accessibility of transcription factor (TF) motifs between vascular sites but provide data to support the hypothesis that genetic variants regulate cell function in not only a cell type, but vascular site-specific mechanism. This work provides important insights into disease risk across vascular

tissues and supports the concept that the epigenomic landscape of vascular cells is vascular site-specific and regulates disease risk.

Results

Single cell analysis of vascular tissue reveals cell type and vascular site specific epigenomic profiles

We performed microdissections to collect the vascular tissues from three vascular sites in 16 healthy adult C57Bl/6 male mice (14-16 weeks of age) including the 1) ascending aorta and aortic root, 2) brachiocephalic and right common carotid arteries, and 3) descending thoracic aorta (**Fig 1A**). Tissues were pooled by vascular site, underwent enzymatic digestion and mechanical dissociation, live cells were FACS sorted, and cells were then partitioned for either scRNAseq or for subsequent nuclei isolation and scATACseq using 10X genomics platform as previously described[22,31,32]. Following quality control and removal of low-quality cells, data was normalized and underwent linear dimensional reduction following standard protocols in R packages Seurat and Signac[33]. A total of 37,394 cells and 40,275 cells were of high data quality and were profiled with scRNAseq and scATACseq respectively with roughly 10,000-15,000 cells per vascular site and sequencing modality. Visualization with UMAP methodology for both scRNAseq and scATACseq datasets was performed (**Fig 1B&C**). UMAP visualization of scRNAseq and scATACseq data reveals a specific pattern where vascular site origin distinguishes cell clusters (**Fig 1B&C**).

Given the interrelationship between RNA expression and global chromatin accessibility, these datasets were then integrated together as previously described[22,33].

Integrated scRNAseq and scATACseq datasets underwent additional non-linear dimensional reduction and UMAP visualization (**Fig 1D**). The integrated UMAP representing both transcriptomic and chromatin accessibility visually reveals significant differences between cell types by vascular site (**Fig 1D**). Analysis of RNA expression of canonical genes to differentiate cell type shows strong correlation to overall gene chromatin activity (e.g. *Myh11* — smooth muscle, **Fig 1E&F**). Integrated cell clusters were then assigned to the major cell types of the vascular wall, vascular smooth muscle cells (SMCs), fibroblasts, endothelial cells, and macrophages (**Fig 1G**) using known lineage markers. To further focus on SMC and fibroblast cell populations, a subsetted analysis of SMC and fibroblasts was performed with UMAP visualization of scRNAseq, scATACseq, and integrated datasets, demonstrating clear separation of cells by original vascular site, demonstrating a unique vascular site-specific chromatin and RNA profile by cell type (**Fig 1H&I**).

Epigenomic patterns of vascular SMCs are cell type and vascular site specific

UMAP clustering of SMCs from RNA, ATAC, and integrated datasets demonstrate a pattern where SMCs from the ascending aorta, carotid artery, and descending thoracic vascular sites are distinct from one another (**Fig 1H**). Chromatin peak accessibility analysis between ascending and descending thoracic aorta SMCs identifies 4,805 peaks which are differentially accessible (**Fig 2A**). Peak locations were analyzed using Genomic Regions Enrichment of Annotations Tool (GREAT)[34]. GREAT performs genomic region-gene associations by assigning a regulatory domain for each gene, and then each genomic region is associated with all genes whose regulatory

domain it overlaps. Of these peaks, the majority (3,971/4,805; 83%) are associated with 2 genes within a region (**Fig 2B**, **Table S1**). A distribution of region-gene association distances to TSS is observed where most peaks lie downstream of the TSS (**Fig 2C**). Roughly 12% of region-gene associations are within 5kb of the TSS of a gene (**Fig 2C**) and approximately 43% (4,496/8,789) of region-gene associations are within 50kb of a TSS (**Fig 2C**).

When comparing peak accessibility between ascending and descending aorta, specific chromatin regions are identified as having marked differential accessibility. Among the top differentially accessible peaks in the ascending SMC population lie at chr8-57320451-57321235 and chr14-63244897-63245811, 140bp and 83bp upstream from the TSSs for *Hand2* and *Gata4* (**Fig 2D&E**). Coverage plots of chromatin accessibility reveal *Hand2* regulatory regions to be open in ascending and carotid SMCs while closed in descending SMCs (**Fig 2D**). This contrasts with regulatory regions of *Gata4* that appear to be in an open state in the ascending SMC however in a largely closed state in the carotid and descending SMCs (**Fig 2E**). Consistent with the role of Hox genes during the development of somites, top peaks that have increased accessibility in the descending aortic SMC include genomic regions that lie near the *Hoxb* family of transcription factors (i.e. chr11-9686379-96287301 at *Hoxb7*, **Fig 2F**, **Table S1**). Featureplots of peak accessibility for *Hand2* and *Hoxb7* peaks (**Fig 2G-I**) highlight distinct differential accessibility dependent on SMC vascular site of origin. Consistent with epigenetic marks, RNA expression analysis reveals *Hand2* expression in the ascending and carotid SMC while absent expression in the descending SMC (**Fig 2J**). Top differential peak analysis identifies other cardiac and vascular development

genes that have higher chromatin accessibility within the ascending aorta include *Tbx20*, *Tbx2*, *Gata4*, and *Wnt16*, while homeobox genes such as the *Hoxa/b/c* family of genes, *Smad2*, and are identified as having higher chromatin accessibility within the descending aorta (**Tables S2&S3**).

Evaluation of biological processes of differentially accessible chromatin regions using GREAT identifies key development pathways that characterize the ascending aorta SMC such as 'outflow tract morphogenesis', 'cardiac septum development', and 'embryonic heart tube development' as well as other pathways involving cellular signaling such as 'transmembrane receptor protein signaling', 'regulation of Ras protein' (**Fig S1A**). Biological processes that characterize the descending aorta SMC include developmental programs such as 'anterior/posterior pattern specification', however notable processes regulating cytoskeletal organization such as 'actin cytoskeleton organization', 'actomyosin structure organization', and 'regulation of SMC differentiation', as well as processes involved in the negative regulation of TGFβ signaling such as 'negative regulation of TGFβ receptor signaling' (**Fig S1B**).

We then performed transcription factor motif accessibility analysis using ChromVAR as previously described[35]. In comparing ascending versus descending aorta SMCs, we identified that ascending SMCs have notable transcription factor motif enrichment for AP1 factors (i.e. FOS:JUNB 'TGAGTCA', **Fig 2K**), neural crest and secondary heart field transcription factors such as *HAND2*, *GATA4*, and *TCF21* (i.e. *HAND2*, **Fig 2L**), as well as TFs with known roles in SMC differentiation such as *TWIST1* and *TEAD* factors (**Table S4**). Top differentially accessible motifs in the descending SMCs highlight GC rich KLF motifs (i.e. *KLF15*, **Fig 2M**), numerous AT rich

212 homeobox transcription factor motifs including HOX factors/Cdx1 (caudal type
213 homeobox 1)/Lhx1 (LIM homeobox 1), and MEF2 factors (i.e. *MEF2A*, **Fig 2N**, **Table**
214 **S5**). Motif enrichment analysis of the carotid SMCs suggest similarities to ascending
215 SMCs with notable enrichment for AP1 and *HAND2* motifs (**Table S6**), however we
216 identified additional motif enrichment that appeared to be largely distinct to the carotid
217 SMC including distal-less homeobox (DLX) motifs (*DLX5*, **Fig 2O**) as well as
218 neurodevelopmental E-box TFs such as *NEUROD1* and *NEUROG2* (*NEUROD1*, **Fig**
219 **2P**).

220 Gene expression analysis using FindMarker from scRNAseq data of SMCs
221 reveals 252 genes which have differential expression among vascular sites (147
222 ascending, 77 carotid, 28 descending) (**Table S7**) that largely match genes identified by
223 differential peak accessibility analysis (**Fig 2Q**). Top differentially expressed genes
224 within the ascending SMC population includes expected developmental genes such as
225 *Tbx20* and *Hand2* (**Fig 2Q**), secondary heart field marker *Tnnt2* (**Fig 2Q&R**), but also
226 numerous genes with previously demonstrated roles in atherosclerosis suggesting
227 vascular site-specific disease risk mechanisms (**Fig 2Q**), including *Ccn3* (cellular
228 communication network factor 3) which has previously reported vascular protective
229 effects by inhibiting neointimal formation and plaque development[36,37] and *Dcn* (decorin)
230 that has been shown to be protective in atherosclerosis[38]. Notably, we identify specific
231 vascular development and inflammation genes as having higher expression within
232 SMCs isolated from the carotid artery, such as *Klf4*, *Fosb*, *Jun*, and *Atf3* (**Fig 2Q&R**).
233 *Fosb* and *Jun* have specific roles in mediating cellular and arterial contractility[39], *Atf3* is
234 a CAD GWAS gene and has recently been identified to have vascular protective effect

in atherosclerosis[23], and *Klf4* has critical roles in mediating SMC phenotypic modulation and promotes atherosclerosis[40]. Specific genes which have higher expression within the descending thoracic SMCs include signal transduction genes such as *Rgs5* (Regulator of G protein signaling 5) (**Fig 2Q**, **Table S7**), *Fn1* (fibronectin 1), and *Ccdc3* (coiled-coil domain containing 3). *Rgs5* has been reported to be a specific marker of peripheral arterial smooth muscle, is downregulated in atherosclerosis, and acts to inhibit SMC proliferation and neointimal formation[41,42], *Fn1* is a putative CAD GWAS gene thought to be protective in CAD and loss of function has been implicated in thoracic aortic aneurysm[43,44], and *Ccdc3* has been found to inhibit TNFα mediated vascular inflammation[45].

Fibroblast cell subset analysis reveals vascular site-specific gene programs and disease risk genes

Evaluation of fibroblast subset population similarly reveals differences in chromatin accessibility and gene expression across vascular sites (**Fig 1I**). Peak accessibility analysis between fibroblasts from the ascending and descending aorta reveals 7,008 peaks that are differentially accessible (**Fig 3A**, **Table S8**). Top peaks that have higher accessibility in the ascending aorta include peaks located at chr9-24773921-24774837 and chr14-63244897-63245811 that lie 95bp and 85bp upstream from the TSSs of *Tbx20* and *Gata4* respectively, while peak chr15-57985530-57986444 is increased in both ascending and carotid and lies 568bp downstream from the TSS of *Fam83a* (**Fig 3B-D**). Top peaks with higher accessibility in the descending aorta fibroblast population includes peaks near Hox related family members (i.e. chr11-

96298662-96299576 at *Hoxb6*) (**Fig 3E**), however additional peaks include chr17-35590469-35591345 in a gene desert 23,457bp downstream of noncoding RNA 2300002M23Rik, as well as chr8-114679333-114680184 that lies within an intron and 240,104bp downstream of the TSS of *Wwox* (**Fig 3F&G**). Featureplots of peak accessibility to peaks that correspond to *Fam83a* and *Wwox* (chr15-57985530-57986444 and chr8-114679333-114680184 respectively) highlight distinct vascular site-specific chromatin accessibility (**Fig 3H&I**).

Peak to gene association analysis with GREAT reveals top genes with higher peak accessibility in the ascending aorta to include expected key developmental genes such as *Tbx20*, *Gata4*, *Tcf21*, and *Hand2*, as well as signaling genes such as *Fam83a*, *Cul4b*, and *Cdkn2b* (**Table S9**). Peaks with higher accessibility within the descending aorta include peaks that lie near the *Hoxa/b/c* family of transcription factors as well as other developmental transcription factors such as *Pax1* and *Foxd1* (**Table S10**). Gene ontology pathway analysis using GREAT identifies top pathways to be enriched in the ascending aorta to include developmental pathways such as 'cardiac septum morphogenesis' and 'embryonic heart tube development', but also pathways that suggest a distinct response to TGFβ such as 'regulation of TGFβ receptor signaling' and 'regulation of cellular response to TGFβ stimulus' (**Fig S2A**). Pathways enriched within the descending aorta fibroblast include an enrichment of pathways involved in extracellular structure organization such as 'extracellular matrix organization', 'collagen fibril organization', and 'negative regulation of cell-cell adhesion' (**Fig S2B**).

Motif accessibility analysis using ChromVar within fibroblast populations reveals ascending fibroblasts to have a marked increase in accessibility of *TCF21* and other

bHLH family of transcription factors, similar to what was observed in the SMC ascending population (**Fig 3J&K**, **Table S11**). However, motif enrichment appears to further identify regulators of TGFβ signaling including the SMAD2:3:4 and TGIF1 motifs (**Fig 3L**). *SMAD3* is a member of the TGFβ superfamily and has causal roles in aneurysm[46] and coronary artery disease[47]. Like that observed in the SMC, descending aortic fibroblast appear to have increased accessibility of motifs that relate to homeobox transcription factors such as HOX family of TFs (i.e. *HOXA13*, **Fig 3M**), *MEIS1/2*, GC rich KLF family members, as well as *SOX9* and other SOX TFs further appear enriched in the descending fibroblasts (**Fig 3N**, **Table S12**). Motif enrichment for carotid fibroblast appear to show a distinct increase in accessibility to developmental TF *NKX6-1* (**Fig 3O**) as well as related TFs with similar sequence (i.e. *MEOX1/2*)(**Table S13**).

RNA expression analysis between vascular sites demonstrates significant concordance with peak accessibility analysis. Fibroblast specific gene expression analysis reveals 597 genes which have differential expression between vascular sites (314 ascending, 149 carotid, 134 descending)(**Fig 3P**, **Table S14**). The top 10 genes which differentiate vascular sites include expected development genes such as *Tbx20*, *Tcf21* (ascending), *Igfbp4*, *Sod3* (carotid), and *Hox*a7, *Col1a2* (descending)(**Fig 3Q**). A notable observation is that descending fibroblasts appear to be enriched for extracellular structure organization pathways by scATACseq (**Fig S2B**) and by RNA we observe a pattern of increased expression of collagen extracellular matrix genes (**Fig 3Q&R**). These genes overlap with known causal genes for hereditary thoracic aortopathy (HTAD) and loci associated with aortic dimension and aortic dissection risk[48,49]. Evaluation of these genes reveal a pattern where TGFβ related HTAD genes in

304 which gain of function leads to disease (i.e. *TGFBR1/2*, *TGFB2*, *SMAD3*) show
305 increased expression in the ascending fibroblasts, while collagen and matrix related
306 HTAD and other collagen genes where loss of function leads to disease (i.e. *FBN1*,
307 *COL3A1*, *LOX*) there is decreased expression (**Fig 3R**). This suggests distinct
308 regulation of extracellular matrix and TGFβ signaling genes in fibroblasts in a vascular
309 site-specific manner that may suggest a relationship to risk of thoracic aortic disease.
310
311 ***Endothelial cell subset analysis shows distinct vascular site-specific chromatin***
312 ***accessibility***
313 Single cell analysis of endothelial cell subset population reveals vascular site-
314 specific chromatin accessibility and gene expression programs. Visualization with
315 UMAP demonstrates separation of endothelial cells from the ascending aorta with
316 somewhat less pronounced differences between endothelial cells from the carotid artery
317 and descending aorta based on RNA, ATAC, and integrated datasets (**Fig 4A-C**). Peak
318 accessibility analysis between endothelial cells from the ascending and descending
319 aorta identifies 935 peaks with differential accessibility meeting significance using an
320 adjusted P value of 0.05 (**Fig 4D**) with 14,388 peaks meeting an unadjusted P value of
321 0.05 (**Table S15**). Top peaks that have increased accessibility in the ascending aorta
322 includes peak chr6-134981212-134982145 which lies 39bp upstream of the TSS of
323 *Apold1* (**Fig 4E&F**). *Apold1* (apolipoprotein L domain containing 1, aka VERGE) has
324 been previously reported as an endothelial cell specific stress response gene and
325 protects against vascular thrombosis[50,51]. Peak analysis further highlights peaks related
326 to Wnt signaling including peak chr6-18031351-18032250, which lies 1216bp upstream

of the TSS for *Wnt2* (**Fig 4G**). Coverage plot of this genomic region demonstrates that chromatin accessibility is nearly entirely closed in the descending aorta with only minor accessibility in the carotid (**Fig 4G**). Notably, the descending aorta endothelial cells have increased accessibility of Hox related genes (i.e. peak chr6-52225707-52226553, *Hoxa9*)(**Fig 4H&I**), as well as peaks related to genes involved in Notch signaling such as chr2-137109516-137110365 which lies 6707bp downstream of the TSS for *Jag1* (**Fig 4J**) where coverage plot of this genomic region reveals differential peak accessibility with carotid and descending endothelial cells having similar accessibility (**Fig 4J**). *Jag1* (jagged 1) and associated Notch pathway signaling has important roles in suppressing vascular smooth muscle chondrogenic fate and in the formation of the atherosclerotic fibrous cap[52,53].

Motif accessibility analysis was performed comparing ascending and descending aortic endothelial cell populations. Top differentially accessible motifs in the ascending endothelial cells highlight *LEF1* (**Fig 4K**, **Table S16**). *LEF1* plays a critical role in Wnt β-catenin signaling and regulates endothelial cell fate specification[54]. Other motifs enriched in the ascending population including *Hand2* as well as multiple ETS factors (**Table S16**). ETS factors play important roles in cellular growth and differentiation and regulate vascular inflammation and remodeling where inhibition of ETS transcription factors promotes vessel regression[55,56]. Top motifs enriched in the descending aortic endothelial cells highlights *NFIC* (**Fig 4L**, **Table S17**) where motif accessibility appears to be enriched in both the carotid and descending endothelial cells. Other motifs include multiple Hox family members, TEAD family members, and other nuclear factors such as

NFIX (**Table S17**). Motif enrichment in carotid endothelial cells highlights nuclear factors such as *NFIA/C/X* as well as AP1 factors (**Table S18**).

Differential gene expression analysis reveals specific patterns differentiating endothelial cell programs by vascular site (**Fig 4M**) with FindMarker analysis identifying 397 differentially expressed genes between vascular sites (174 ascending, 127 carotid, 96 descending) (**Table S19**), where heatmap analysis suggests greater similarity between carotid and descending endothelial cells compared to ascending (**Fig 4M**). Ascending aortic endothelial cells have increased expression of VEGF receptors *Kdr* (kinase insert domain receptor, aka VEGFR), *Flt1* (fms related receptor tyrosine kinase 1, aka VEGFR1), and *Flt4* (fms related receptor tyrosine kinase 4, aka VEGFR3) (Figure 4N&O). Descending aortic endothelial cells have notable increase in expression of *Edn1* as well as other BMP factors (i.e. *Bmp4*) as well as interestingly xenobiotic transformation gene *Cyp1b1* (cytochrome p450, 1b1) (**Fig 4P&Q**). Vascular site-specific expression of *Cyp1b1* is particularly notable as it contributes to abdominal aortic aneurysm and suggests a vascular site-specific mechanism of aneurysm[57,58]. Carotid endothelial cells have higher expression of specific genes that appear to play unique roles within the cerebral vasculature such as *Efemp1* (EGF containing fibulin extracellular matrix protein 1) (**Fig 4R**) an extracellular matrix glycoprotein which regulates vessel development and has been associated with intracranial vascular disease and white matter density through large genome wide association studies[59].

Macrophage cells have minimal epigenomic and transcriptional variation across vascular sites

Given the significant vascular site-specific epigenomic and RNA transcriptional profiles observed within vascular SMCs, fibroblasts, and endothelial cells, we next evaluated if resident macrophages within healthy vascular tissues harbor specific epigenomic and transcriptional profiles. Macrophage cells from integrated scATACseq/scRNAseq data were subsetted and subsequently analyzed. Importantly, UMAP visualization of macrophage cells with RNA, ATAC, and integrated datasets reveals no significant differences in macrophage cells across vascular sites (**Fig S3A-C**). Differential peak accessibility analysis between ascending and descending aorta macrophages do not reveal any peaks which meet statistical significance as differentially accessible (**Fig S3D**). Differential gene expression analysis with scRNAseq data reveal very minimal differences in gene expression (**Fig S3E**). This lack of vascular site-specific macrophage diversity across the aorta is notable given the contrasting finding within SMCs, fibroblasts, and endothelial cells, and that significant leukocyte diversity develops within the aorta during atherosclerosis[60].

Cell type and vascular site-specific gene regulatory networks highlight distinct ascending fibroblast regulatory networks

These data indicate that epigenomic landscape is not only cell type but vascular site-specific with key differences in transcription factor motif accessibility. To further understand how TF regulatory elements may further control gene expression on a cell type and vascular site-specific basis, we utilized this multi-modal single-cell data to infer gene regulatory networks (GRNs) with Pando[61]. By integrating scRNA and scATAC datasets and subsetting across vascular cell type and vascular site of origin, we utilized

Pando to infer GRNs through modeling gene expression with understanding the interaction of TF expression with TF binding sites and gene targets (**Fig 5A**). With Pando, we initiated GRN analysis by scanning candidate genomic regions and identified TF binding motifs to then infer GRN modules. To understand how regulatory networks may differ within cell type based on vascular site, we identified TF modules for aortic SMCs and Fibroblasts in both ascending and descending aortic sites. In these modules, we identified GRN TF and TF gene targets. In the SMC population, we identified 37 and 38 TF modules in the ascending and descending cells, respectively. In the fibroblast population, we identified 56 and 40 TF modules in the ascending and descending cells, respectively (**Fig 5B&C**). In both SMC and fibroblast cell populations, the majority of TF modules appear to be shared between ascending and descending cells (**Fig 5B&C**), with key TFs including *Atf3*, *Creb5*, *Fosb*, *Klf2/4/6*. However, although SMC TF modules show relatively equal proportion of TF modules that are distinct to ascending and descending, ascending fibroblasts appear to have a notable increase in distinct TF modules (20 vs 4) (**Fig 5C**). Distinct ascending fibroblast GRN TF modules include *Meox1*, which has recently been described as a master regulator of fibroblast activation in cardiac fibroblasts through epigenetic mechanisms[20]. GRN TF gene targets appear to have less overlap between ascending and descending cell populations in both SMCs and fibroblasts (**Fig 5D&E**), however consistent with GRN TF modules, there is a greater proportion of ascending fibroblast GRN gene targets compared to descending (108 vs 43) (**Fig 5D&E**). Visualization of ascending fibroblast GRN highlights the proximity of *Meox1* to other developmental TFs *Tbx20* and *Gata4* (**Fig 5F**). Other

distinct fibroblast ascending GRN TF modules include *Tcf21*, *Irf7/8*, *Sox4/7/17*, and *E2f8*.

Vascular site-specific Meox1 activation implicates epigenetic 'priming' for fibroblast activation in ascending aortic fibroblasts

Our findings implicate a cell type and vascular site-specific 'epigenetic memory' of important regulatory enhancer elements with functional effect on disease relevant pathways. A notable observation from evaluation of GRNs is that ascending fibroblasts have increased proportion of distinct regulatory networks from descending fibroblasts highlighting the TF *Meox1*. We interrogated our scRNAseq and scATACseq fibroblast integrated dataset where we see that *Meox1* has increased RNA expression in ascending fibroblasts, but this expression is largely limited to a specific cell population that represents valvular fibroblasts (**Fig 5G&H**). Importantly, we see that *Meox1* chromatin accessibility is much more expansive and extends throughout all ascending and includes carotid fibroblasts, with increased total chromatin accessibility in the ascending/carotid fibroblast populations (**Fig 5I**). This discrepancy of chromatin accessibility and RNA expression for non-valvular ascending fibroblasts may suggest evidence to support an epigenetic 'priming'.

Primary adventitial fibroblast response to TGFβ is dependent on vascular site and implicates functional developmental epigenomic memory

We have shown that vascular SMCs, fibroblasts, and endothelial cells have transcriptional and epigenomic features that are distinct to vascular site. Further, TF

motif accessibility analysis in fibroblasts reveals an increased accessibility of *AP1*, and *SMAD2:3:4*, and *TCF21* motifs in the ascending fibroblast population compared to carotid and descending fibroblasts (**Fig 3**), with further gene regulatory network analysis suggesting distinct regulatory TF module activity in ascending fibroblasts highlighting master regulator of fibroblast activation *Meox1* (**Fig 5**). This differential chromatin accessibility would suggest the potential for heightened biological response to TGFβ in ascending fibroblasts. To evaluate the functional effect of this differential chromatin accessibility and to identify if differential gene expression is retained following removal from vascular site flow conditions, we isolated and cultured primary adventitial fibroblasts from healthy 14 week old C57BL/6 mice from the ascending and descending aorta, passaged cells 3X allowing for separation from hemodynamic effects, and stimulated them with control or TGFβ (10ng/mL, 48hrs) and performed bulk RNA sequencing (n = 3 per condition) (**Fig 6A**).

Principal component analysis (PCA) of RNAseq data reveals distinct separation of samples based on origin of vascular site and stimulation with TGFβ (**Fig 6B**). Upon evaluation of differential gene expression by vascular site, we observe 1342 differentially expressed genes between ascending and descending fibroblasts (**Fig 6C**, **Table S20**). Ascending fibroblasts have increased expression of genes including *Hand2*, *Tbx20*, *Pcolce2*, *Tcf21*, and *Gata4* (**Fig 6C**, **Fig S4A**), while descending fibroblasts have increased expression of *Hox* family of genes (i.e. *Hoxb9*, *Hoxa7*, *Hoxc8*), *Col6a3*, and *Pax1* (**Fig 6C**, **Fig S4B**). When evaluating response to TGFβ stimulation, comparing all samples by TGFβ treatment, we identify 4,234 DE genes, with top upregulated genes including *Ada*, *Snx30*, *Wisp1*, *Cthrc1*, *Loxl2*, and *Akap5*,

while top downregulated genes include *Adm*, *Gstm1*, *Mcc*, *Rras2*, *Il6ra*, and *Vegfd* (**Fig 6D, Table S21**). To evaluate the effect of TGFβ by vascular site, we compared the transcriptomic effect of TGFβ stimulation within ascending and descending fibroblast samples. We identify TGFβ stimulation to induce a markedly higher transcriptional effect in ascending fibroblasts, with 6,733 DE genes (**Fig 6E**, **Table S22**) while only 2,250 DE genes within descending fibroblasts (**Fig 6F**, **Table S23**).

By performing an interaction analysis, we reveal key genes with differential response to TGFβ by vascular site. In this analysis, 422 genes meet FDR cutoff for interaction significance (**Fig S5**, **Table S24**). When plotting the log2FC_interaction (log2FC > 0 represents greater effect TGFβ in ascending; log2FC < 0 represents greater effect TGFβ in descending), we identify that matrix gene *Eln* is the top interacting gene that has much lower expression in descending fibroblasts but is rapidly upregulated in response to TGFβ (**Fig S5A&B**). Here, we identify that *Meox1* has a marked interaction in response to TGFβ with greater effect in ascending fibroblasts, where although low expression of *Meox1* is present in ascending fibroblasts, *Meox1* expression significantly increases with TGFβ treatment, an effect not seen in descending fibroblasts (**Fig S5C**). This relationship also includes other TGFβ response genes such as *Gdf6* that are more highly induced in ascending fibroblasts, while *Tgfb1* and *Tgfb3* are more induced in descending fibroblasts (**Fig S5C**, **Table S24**). Numerous collagen genes have differential response to TGFβ by vascular site, where *Col8a1* and *Col5a2* are upregulated by TGFβ to a greater extent in descending fibroblasts (**Fig S5D**). Alternatively, *Col28a1* and *Col6a1* are downregulated by TGFβ to a greater extent in descending fibroblasts (**Fig S5D**). TFs such as *Klf4* and *Ptx3* are more highly

expressed in ascending fibroblasts and then downregulated by TGFβ to a greater extent than descending fibroblasts (**Fig S5E**). However, *Ahr* is downregulated by TGFβ in descending fibroblasts only while *Hoxb5* shows differing effects by vascular site, where it is modestly upregulated in ascending while downregulated in descending fibroblasts (**Fig S5E**).

By evaluating the 200bp DNA sequence upstream of the TSS of differentially expressed RNA transcripts in ascending versus descending fibroblast comparison, we evaluated TF motif enrichment using HOMER[62]. In control samples, DE genes upregulated in ascending vs descending fibroblasts reveal enrichment of key TFs including SMAD3 (TWGTCTGV), KLFs including KLF4 (GCCACACCCA), FOS (NDATGASTCAYN), and HIF1α (TACGTGCV)(p-values 1.0E-10, 1.0E-13, 1.0E-07, and 1.0E-05, respectively). DE genes upregulated in descending fibroblasts reveal enrichment of TFs including HOXA11 (TTTTATGGCM), HOXA9 (RGCAATNAAA), HOXC6 (GGCCATAAATCA), (p-values 1.0E-07, 1.0E-06, 1.0E-03, respectively) (**Fig 6G**). We then similarly evaluated the motif enrichment of the 200bp sequence upstream of the TSS of upregulated genes in ascending fibroblasts in response to TGFβ and then for descending fibroblasts in response to TGFβ (**Fig 6H&I**). In this comparison, although motif enrichment to TGFβ has similarities between ascending and descending fibroblasts with GC rich motifs such as SP2/5 and KLF factors similarly enriched (**Fig 6H&I**), distinct differences emerge. Ascending fibroblast response to TGFβ motif enrichment highlights AP1 factors, SMAD3, as well as MEF2 factors (**Fig 6H**). In the descending fibroblast response to TGFβ highlights ETS factors and other NRF factors and to a lesser extent AP1 factors (**Fig 6I**). However SMAD3 and MEF2 factors are not

509 identified as being enriched. The differential enrichment of these TFs in differentially
510 expressed genes between vascular site and in response to TGFβ suggests this
511 differential chromatin accessibility to have functional effect on global transcriptome.
512
513 ***Aortic dimension GWAS identifies genes with vascular site-specific expression***
514 Our data indicates that epigenomic landscape are not only cell type, but vascular
515 site-specific. However, it is unclear if these vascular site-specific enhancers and gene
516 programs relate to human genetic evidence of disease. To evaluate this question, we
517 leveraged the data from a recently performed GWAS to understand the genetic
518 determinants of ascending versus descending aortic dimension[17]. Through an algorithm
519 based method to evaluate cardiac MRI imaging of UK Biobank participants, Pirrucello et
520 al. (2022)[17] identified 82 and 47 genomic loci which met genome wide significance for
521 ascending and descending aortic dimensions, respectively. We hypothesized that if
522 vascular site-specific regulatory enhancers influence genetic disease risk, that we would
523 observe aortic dimension GWAS genes to be differentially expressed in a vascular site-
524 specific manner. By taking the nearest gene for each locus, we evaluated if these
525 GWAS genes are enriched in the ascending versus descending differentially expressed
526 RNA gene lists for each cell type (SMC, Fibro, Endo). Following performing a differential
527 gene expression analysis between ascending and descending aortic SMC, Fibro, and
528 Endo, cell types, we evaluated the proportion of GWAS aortic dimension genes to be
529 differentially expressed. We observed that these GWAS genes are enriched in the
530 differential gene analysis within SMC, Fibroblasts, and Endothelial cells. In SMCs, 25%
531 (21/82) and 21% (10/47) of ascending and descending aortic dimension GWAS genes

532 are differentially expressed (*p=5.4E-6* and *p=2.5E-4*, compared to random gene
533 set)(**Fig 7A-C**). In Fibroblasts, 32% (26/82) and 34% (16/47) of ascending and
534 descending aortic dimension GWAS genes are differentially expressed (*p=9.9E-5* and
535 *p=1.8E-4*, compared to random gene set)(**Fig 7D-F**). Whereas, in endothelial cells, 15%
536 (12/82) and 17% (8/47) of ascending and descending aortic dimension GWAS genes
537 are differentially expressed (*p=0.024* and *p=0.019*, compared to random gene set)(**Fig**
538 **7G-I**). These data highlight that a large proportion of genes that regulate aortic
539 dimension are differentially expressed across vascular sites within a cell type and
540 possibly suggests fibroblasts to be a primary cell type.
541
542 ***ChromBPnet predicts human genotype effect of chromatin accessibility in a cell***
543 ***type and vascular site-specific manner***
544 These data indicate that genes for which variants regulate aortic dimension have
545 not only cell type but vascular site-specific expression and epigenomic patterns. This
546 may suggest that a gene variant can have a differing effect on TF motif binding,
547 chromatin accessibility, and gene regulation in not only a cell type but a vascular site-
548 specific context. To better understand this vascular site-specific epigenomic regulation,
549 we aimed to predict how a gene variant that influences aortic dimension (Pirruccello et
550 al., 2022)[17] may affect chromatin accessibility in a cell type and vascular site specific
551 manner using ChromBPNet[63]. ChromBPNet is a novel, bias factorized, base-resolution
552 deep learning model of chromatin accessibility[63]. Prior training of the ChromBPNet
553 model has been predicated on the concept that gene variants affect TF binding of cis-
554 regulatory elements (cREs) in a cell context-specific manner[63]. Training of a

convolutional neural network (CNN) for ChromBPNet has utilized chromatin accessibility data across 5 ENCODE Tier 1 cell lines[63], however, although this takes into account how a variant may affect TF binding and chromatin accessibility within cell type, it does not allow for the investigation of a variant effect within cell type but across vascular sites. To overcome this barrier, we utilized our multiomic single cell RNA and ATAC sequencing data and trained the ChromBPNet model in 9 different groups (Ascending aorta, Carotid, and Descending aorta, within SMC, Fibroblasts, and Endothelial cells) in our mouse dataset. Given this 'genome-agnostic' approach, this model can predict the fundamental effect of base pair change on chromatin accessibility. These individual cell type and vascular site-specific models can then be applied to the human genome. We leveraged the variants identified from the prior GWAS on aortic dimension[17], performed linkage disequilibrium (LD) expansion to identify SNPs in LD with lead SNPs, and used each of the 9 models to score variant effect for 27,556 variants based on cell type and vascular site. We identified 469 high scoring aorta diameter LD-expanded GWAS variants across all of the vascular site datasets (example, rs2959350, in LD with lead SNP rs55736442 at *ANGPT1* locus, chr8:107367260:G:A, Ascending Endothelial cells, **Fig 8A**). In this model, the variant allele has a pronounced activating effect on chromatin accessibility across a ~300bp region.

For each variant, we scored the absolute log fold change (abs_logFC) from variant effect as well as the p-value of the abs_logFC (abs_logFC.pval). To understand how predicted variant effect may vary within cell type across vascular site, we first evaluated the abs_logFC.pval and compared between ascending and descending SMC models (**Fig 8B**). This reveals significant variation in p-value of logFC for a given SNP

between vascular sites (**Fig 8B**). However, upon evaluation of the variation of abs_logFC for each SNP within SMC across vascular site, variation is significantly reduced (**Fig 8C**). A similar level of variation of variant effect is seen when comparing ascending and descending fibroblast and endothelial cell models (**Fig 8D&E**). To understand the predicted variant effect between cell types, we then compared the predicted abs_logFC in ascending SMC vs Fibro, SMC vs Endo, and Fibro vs Endo (**Fig 8F-H**). We observe a consistent pattern of a greater degree of variation in comparing between cell types as opposed to within cell type across vascular site, consistent with an overall understanding that variant effect is primarily influenced by cell type, however within cell type, vascular site continues to have significant variant effect.

Following linear regression analysis of the abs_logFC within cell type between ascending and descending models (i.e. Asc SMC vs Desc SMC), we identified the top 1% of deviant SNPs (furthest away from linear regression line). These top 1% deviant SNPs highlight specific variants that appear to have vascular site-specific effect. For example, rs11677932 is the lead SNP (chr2:237315312:G:A) at the *COL6A3* locus, a gene we have previously highlighted as having differential expression within cell type across vascular site. The variant effect is predicted to lead to epigenetic silencing; however, this effect is greatest in SMC, compared to fibroblast and endothelial cells, and within SMCs, this effect is greatest in descending SMC compared to ascending SMC (**Fig 8I**).

The identification of variants that have vascular site-specific effects on chromatin accessibility may suggest that variants can change pioneer TF binding to motifs that influence vascular site-specific epigenetic patterns. To better understand these variants

and the potential TF motifs that they regulate, we took the top 1% of deviant SNPs away from the linear regression, and from these SNPs, we selected a 200bp genomic window (+/- 100bp around each SNP) and performed unbiased Motif Discovery function with MEME Suite[64]. Known motif analysis of top discovered motifs was then performed using TOMTOM (HOCOMOCOv11)[65]. This analysis was performed for top deviant SNPs in SMC, fibroblast, and endothelial cell populations, and here, we identified novel motifs enriched in the genomic regions of these SNPs with differential effect on vascular site. TOMTOM analysis of discovered motifs revealed numerous TF motifs that we have similarly identified as having differential motif accessibility between vascular sites. For example, analysis of top discovered motif from SMC data reveals motif enrichment for *MEF2A/C*, consistent with our prior finding of enriched *MEF2A/C* motif in descending SMC (**Fig 8J**). Similarly, in analysis of the top discovered motif from fibroblast data, we observe an enrichment for multiple zinc finger motifs including *IKZF1* and *ZN250*, however we also see an enrichment for *TEAD4* (**Fig 8K**), a motif we have previously observed to have differential accessibility. Further, in the second top motif in this fibroblast data, we observe an enrichment for a *SMAD3* motif (**Fig 8L**), consistent with our previously observed increased in *SMAD3* motif accessibility in ascending aortic fibroblasts. Finally, evaluation of these SNPs from endothelial data further reveals a discovered motif that is notably enriched for a *HAND2* motif (**Fig 8M**), a motif we had similarly seen enriched in ascending aortic cell types. These data suggest that variants that have differential effect on chromatin accessibility between vascular sites lie in genomic regions enriched for motifs that have differential accessibility between vascular

623 sites. These data further support our hypothesis that vascular site-specific epigenomic
624 patterns influence human genetic determinants of vascular disease risk.
625
626 **Discussion**
627 Chromatin architecture and cis-regulatory elements such as enhancers and
628 promoters are critical in mediating cellular gene programs in a cell type specific
629 manner[19]. Disease associated gene variants identified through GWAS are now
630 increasingly being recognized to influence disease risk through modification of these
631 regulatory regions of the genome[66]. Additionally, there is now growing experimental
632 evidence that vascular disease associated gene loci influence disease risk in cell type
633 specific mechanisms, where common human genetic variation can modify the function
634 of cell type specific enhancers[22,23]. Risk of vascular diseases, such as atherosclerosis,
635 aneurysm, or autoinflammatory vasculitides, are vascular site-specific, with disease
636 associated genetic loci suggesting differing genetic mechanisms of disease. These
637 observations raise the question as to what extent do cell type specific enhancers and
638 gene programs vary by vascular site, and do these variations contribute to disease risk?
639 Here, we evaluated cell type and vascular site-specific enhancer and gene
640 expression profiles of healthy vascular tissue in adult mice from three disease relevant
641 vascular sites, 1) ascending aorta and aortic root, 2) brachiocephalic and right common
642 carotid arteries, and 3) descending thoracic aorta (**Fig 1A**). These vascular sites
643 represent the developmental diversity that makes up the aorta, with these regions
644 arising from the secondary heart field and neural crest, neural crest, and somitic
645 mesoderm, respectively. This work has revealed thousands of differentially accessible

enhancers within vascular smooth muscle, fibroblasts, and endothelial cells. Through use of this single cell epigenomic and transcriptomic multi-omic data, computational analysis of gene regulatory networks, *in vitro* culture of primary aortic fibroblast cells, and by using a novel machine learning approach to train and predict gene variant effect on chromatin accessibility across vascular sites in the human genome (ChromBPNet), we have defined an important observation — that epigenomic patterns are not only cell type but vascular site specific. This data supports our understanding that gene variants appear to regulate chromatin accessibility through influencing pioneer TF motif binding of key TF (i.e. *SMAD3*, *MEF2A/C*, *HAND2*, *TEAD4*) in a vascular site-specific mechanism. These data have important implications for our understanding of vascular site-specific disease risk and may give insight into novel mechanisms of disease.

The activation of key developmental transcription factors is crucial in the coordinated cellular development of the fetal and adult vasculature. Although expression of genes such as *Hand2*, *Tbx20*, *Gata4*, *Wnt* and those encoding related WNT signaling molecules, and *Hoxa/b/c* family of transcription factors have previously been known to mediate vascular development[5,8,25-28], the residual chromatin accessibility and gene expression of these developmental transcription factors in the adult vasculature has been less well characterized. A notable observation of our study is that key regulatory enhancers of development genes *Tbx20*, *Hand2*, and *Gata4* have increased chromatin accessibility in the ascending aorta and carotid artery in comparison to the descending aorta in smooth muscle and fibroblasts (**Figs 2&3**). Whereas *Hoxb* and *Hoxa/Hoxc* family of transcription factors have increased chromatin accessibility in the descending aorta (**Figs 2&3**). Although it is possible that

669 hemodynamic factors may be contributing to these epigenetic differences across
670 vascular sites, our data from culturing primary adventitial fibroblasts *in vitro* suggests
671 that these transcriptomic patterns are retained even when culturing in a dish and across
672 3 passages (**Fig 6**). These findings suggest that these vascular cells retain an
673 epigenetic 'memory' of their developmental program, suggesting these cells are poised
674 to turn on these gene programs.

675 This finding further raises the question as to the role of these development genes
676 in mediating vascular disease risk in adulthood. Although genetic variation in *TBX20*
677 has been associated with a spectrum of congenital cardiac lesions relating to cardiac
678 and vascular development[67], there have been increasing observations that genetic loci
679 near *TBX20* meet genome wide significance for disease in adulthood including coronary
680 artery disease[68], myocardial infarction[69], blood pressure[70], and aortic dimension and
681 distensibility[17,71]. Similarly, variants near *GATA4* and members of the *HOXB* family of
682 transcription factors such as *HOXB7* have been associated with hypertension[72,73], while
683 *HAND2* variants are associated with aortic dimension and atrial fibrillation[17,74]. The
684 vascular site and cell type-specific chromatin accessibility observed in our study of
685 these developmental transcription factors may suggest their ongoing role in mediating
686 vascular disease in adulthood in a vascular site-specific mechanism. This hypothesis of
687 developmental TFs involved in vascular disease pathogenesis has been similarly
688 supported by work from our lab on developmental TFs *TCF21* and *ZEB2*[22,32]. These
689 TFs — which have important roles in vascular development, regulating cell state
690 transitions and endothelial to mesenchymal transition[75,76] — have been identified
691 through GWAS as having additional roles in the development of CAD[12,77]. SMC specific

692 deletion of *Tcf21* and *Zeb2* in the mouse revealed a significant effect on transcriptional
693 regulation, epigenetic landscape, and plaque characteristic, further defining their roles
694 as causal CAD genes[22,32].

695 TF motif accessibility analysis identified an important observation that
696 *SMAD2:SMAD3* motif is enriched in ascending aortic fibroblasts, highlighting a potential
697 vascular site-specific response to TGFβ signaling. It was first identified nearly 30 years
698 ago in chick embryos that vascular SMCs derived from the neural crest have enhanced
699 response to TGFβ signaling compared to ectoderm derived SMCs[78,79]. More recently,
700 transcriptomic differences across vascular sites and organ specific fibroblast
701 transcriptomic identity has been reported[30,80]. However, epigenomic landscape and
702 differential motif accessibility across vascular sites has not been reported. By isolating
703 and culturing primary adventitial fibroblasts from ascending and descending aorta, *in*
704 *vitro* experiments with TGFβ stimulation with bulk RNA sequencing confirmed that
705 differential gene expression by vascular site is retained following isolation and *in vitro*
706 culture and revealed ascending fibroblasts to have a markedly greater response to
707 TGFβ than descending fibroblasts (**Fig 6**), and motif accessibility analysis of
708 differentially expressed genes further suggests distinct response to AP1 factors such as
709 JUN, SMAD2/3, and MEF2 factors (**Fig 6H**). These are two important observations that
710 reveals 1) DE genes by vascular site in fibroblasts are independent of differential
711 laminar and turbulent flow and 2) implicates differential chromatin accessibility and
712 relevant TF motif accessibility to have causal biological and disease relevant function.

713 Evaluation of GRNs on a cell type and vascular site-specific basis highlights the
714 ascending aortic fibroblasts to have a notable activation of *Meox1* as a key GRN TF

(**Fig 5**). *MEOX1* is a master regulator of fibroblast activation *Meox1*[20] and we further identified an epigenetic 'priming' of key disease relevant genes. Primary fibroblast culture shows similar baseline RNA expression of *Meox1* between vascular sites, but *Meox1* is upregulated to a much greater extent in ascending fibroblasts in response to TGFβ (**Fig S5C**), consistent with epigenetic priming and activation of key vascular site-specific GRN.

Prior work has discovered that genomic variants identified through GWAS have cell type specific effects, however, understanding if gene variants may have not only a cell type but vascular site-specific effect is challenging to determine. Our data here highlights that many of the GWAS genes identified from a recent aortic dimension GWAS[17] are differentially expressed when comparing ascending versus descending aortic SMC, fibroblasts, and endothelial cells, with a notable enrichment of these genes in the fibroblast DE gene analysis (**Fig 7**). However, to understand how these variants may regulate chromatin accessibility on a base pair resolution, we applied a unique model, where by training a ChromBPNet model in each cell type and vascular site, we employed 9 distinct models to predict variant effect of aortic dimension on a cell type and vascular site-specific context. This work revealed that although cell type remains the primary influence on chromatin accessibility, vascular site is an important influence. Our work further highlights that variants that appear to modify effect based on vascular site are enriched for genomic positions that have key TF motifs, suggesting that these variants may modify the binding of TFs such as *SMAD3*, *MEF2A* and others and regulate disease risk in a vascular site-specific mechanism (**Fig 8**).

738 **Limitations**

739 In this study, our aim was to evaluate single cell enhancer and transcriptional
740 profiles across vascular sites in healthy tissue. A limitation of this study is that we did
741 not evaluate the epigenomic profiles in a disease state across vascular sites. However,
742 by isolating and culturing primary adventitial fibroblasts and evaluating differential
743 response to TGFβ, we gained further insight into this mechanism. We anticipate that
744 there are dynamic changes in chromatin accessibility and gene expression programs in
745 disease, as has been previously observed[22,23], and future studies will be aimed at
746 understanding how these dynamic chromatin accessibility changes occur across
747 vascular sites. Similarly, recent work has demonstrated changes in chromatin
748 accessibility in vascular tissue with aging[81]. Our study selected for adult mice that are
749 14-16 weeks of age, which represents a young adult. We anticipate that chromatin
750 accessibility dynamically changes with aging and future studies to evaluate how this
751 change differs across vascular sites may be particularly insightful.
752
753 **Conclusions**

754 By performing combined scRNAseq and scATACseq on vascular tissue across
755 three vascular sites, we reveal for the first time that the epigenomic landscape and
756 transcriptional profiles of vascular smooth muscle, fibroblasts, and endothelial cells, are
757 specific to anatomic origin, that genomic regions and genes that differentiate cells by
758 vascular site are weighted towards developmental genes, and that vascular cells have
759 an epigenetic 'memory' of their developmental program. We discovered that differential
760 chromatin accessibility appears to 'prime' vascular site-specific gene regulatory

761 networks in disease relevant mechanisms, and finally, that genetic variants that
762 influence aortic dimension appear to regulate chromatin accessibility in not only a cell
763 type, but vascular site-specific mechanism. This work supports the paradigm that
764 genetic mechanisms of disease influence disease risk in a vascular site-specific
765 manner, gives unique insight into vascular site-specific transcriptional and epigenetic
766 programs, and further creates a valuable single cell atlas for the vascular biology
767 community.
768
769 **Methods**
770 **Data Availability**
771 All scRNAseq, scATACseq, and bulk RNAseq data has been deposited to the
772 National Center for Biotechnology Information Gene Expression Omnibus under GEO
773 accession numbers GSE296197 (scRNA/ATACseq) and GSE296074 (bulk RNAseq).
774
775 **Mice and Micro-Dissections**
776 Male 14-16-week-old C57Bl/6 mice were purchased from Jackson Laboratory
777 (Bar Harbor, ME). The animal study protocol was approved by the Administrative Panel
778 on Laboratory Animal Care at Stanford University and procedures were followed in
779 accordance with institutional guidelines.
780 Mice were anesthetized with isoflurane and sacrificed with cervical dislocation
781 technique. Vascular tissue was flushed with injection of 5mL of phosphate buffered
782 saline (PBS) into the left ventricle after an incision was made at the right atrium. Aortic
783 tissue was dissected including the aortic root and ascending aorta up to the take-off of

784 the brachiocephalic artery. The brachiocephalic artery and its extension into the right
785 common carotid artery were carefully dissected under stereoscope. The descending
786 thoracic aorta was then isolated from past the left subclavian artery down to the renal
787 arteries.
788
789 **Vascular Tissue Dissociation, Cell Capture, and Sequencing**
790 Tissues were collected and dissociated for single cell capture as previously
791 described[22,31,32]. Briefly, vascular tissue was washed three times in PBS, tissues were
792 then placed into an enzymatic dissociation cocktail (2 U ml−1 Liberase (5401127001;
793 Sigma–Aldrich) and 2 U ml−1 elastase (LS002279; Worthington) in Hank's Balanced
794 Salt Solution (HBSS)) for 45min at 37 °C. Tissues were then gently minced and
795 dissociated with pipette. The cell suspension was strained and then pelleted by
796 centrifugation at 500g for 5 min. The enzyme solution was then discarded, and cells
797 were resuspended in fresh HBSS. To increase biological replication, two sets of 8 mice
798 (16 mice total) were used to obtain two single-cell suspensions for each vascular tissue
799 (aortic root/ascending aorta, brachiocephalic/carotid, and descending thoracic aorta).
800 Cells were FACS sorted and live cells were identified as previously described[32]. Cells
801 were sorted on a Sony SH800S cell sorter, where cells were gated on forward/side
802 scatter parameters to exclude small debris and then gated on forward scatter height
803 versus forward scatter area to exclude obvious doublet events. Approximately 100-
804 150,000 live cells were sorted for each vascular site for each capture, where a portion of
805 cells were taken directly to scRNAseq capture. For single cell ATAC, cells were

collected in BSA-coated tubes, and nuclei isolated per 10X recommended protocol, and captured on the 10X scATAC platform.

All single-cell capture and library preparation was performed in the Quertermous Lab. Cells were loaded into a 10x Genomics microfluidics chip and encapsulated with barcoded oligo-dT-containing gel beads using the 10x Genomics Chromium controller according to the manufacturer's instructions. Single-cell libraries were then constructed according to the manufacturer's instructions (Illumina). Libraries from individual samples were multiplexed into one lane before sequencing on an Illumina platform with targeted depth of 50,000 reads per cell for RNA and 75,000 reads/cell for ATAC. Sequencing was performed by MedGenome (Foster City, CA). Post filtering for non-cells, mean number of reads within peaks per cell in scATAC data was 18,000-20,000 as was seen in our prior report[22].

Analysis of Single-Cell Sequencing Data

scRNAseq

Fastq files from each vascular site (6 total RNA captures) were aligned to the reference genome (mm10) individually using CellRanger Software (10x Genomics). Dataset was then analyzed and captures were integrated using the R package Seurat[33]. The dataset was trimmed of cells expressing fewer than 1000 genes, and genes expressed in fewer than 50 cells. The number of genes, number of unique molecular identifiers and percentage of mitochondrial genes were examined to identify outliers. As an unusually high number of genes can result from a 'doublet' event, in which two different cell types are captured together with the same barcoded bead, cells with >7500

genes were discarded. Cells containing >7.5% mitochondrial genes were presumed to be of poor quality and were also discarded. QC of nFeature_RNA, nCount_RNA, and percent.mt are included in **Figure S6A-C**. The gene expression values then underwent library-size normalization and normalized using established Single-Cell Transform function in Seurat. Principal component analysis was used for dimensionality reduction, followed by clustering in principal component analysis space using a graph-based clustering approach via Louvain algorithm. Batch correction was performed with reciprocal PCA (RPCA). UMAP was then used for two-dimensional visualization of the resulting clusters. Analysis, visualization and quantification of gene expression and generation of gene module scores were performed using Seurat's built-in function such as "FeaturePlot", "VlnPlot", and "FindMarker."

scATACseq

Fastq files from each vascular site (6 total ATAC captures) were aligned to the reference ATAC genome (mm10) individually using CellRanger Software (10x Genomics). Individual datasets were aggregated and peak calling was performed using the CellRanger aggr command without subsampling normalization. The aggregated dataset was then analyzed using the R package Signac[33]. The dataset was first trimmed of cells containing fewer than 1000 peaks, and peaks found in fewer than 10 cells. The subsequent cells were then again filtered based off TSS enrichment, nucleosome signal, and percent of reads that lies within peaks found within the larger dataset. Cells with greater than 20,000 reads within peaks or fewer than 3000 peaks, <2 transcription start site (TSS) enrichment, or <15% reads within peaks were removed as they are

likely poor-quality nuclei. QC of pct_reads_in_peaks, peak_region_fragments, TSS.enrichment, and nucleosome_signal are included in **Figure S6D-G**. Fragment histograms based on high and low nucleosome group (NS >4, NS <4) and TSS Plot based on high and low TSS enrichment (TSS >2, TSS <2) are included in **Figure S6H&I**. The remaining cells were then processed using RunTFIDF(), RunSVD() functions from Signac to allow for latent semantic indexing (LSI) of the peaks[82,83], which was then used to create UMAPs. Batch correction for UMAP visualization was performed with Harmony. Differentially accessible peaks between different populations of cells were found using FindMarker function, using number of peaks as latent variable to correct for depth. Motif matrix was obtained from JASPAR 2020, aligned onto BSgenome.Mmusculus.UCSC.mm10. Accessibility analysis around transcription factor motifs was performed using ChromVar[35]. Merging of scRNA and ATAC data was performed using Pseudo-expression of each gene created using scATACseq GeneActivity function to assign peaks to nearest genes expressed in the scRNA dataset, and mapped onto each other using Canonical Correlation Analysis. For gene regulatory network (GRN) analysis, RNA and ATAC integrated datasets were subsampled based on cell type and GRNs were inferred using Pando[61]. In this method, variable features are identified from RNA dataset and GRN is initiated with 'initiate_grn()' function. Candidate regions for TF binding motifs are identified and GRNs are inferred using 'infer_grn()' function. GRN modules are then identified using 'find_modules()' function for each cell type and vascular site reported. Network graphs are then visualized using 'plot_network_graph()' function and further analysis was performed following standard Pando workflow.

ChromBPNet

Processing of single-cell ATAC-seq data

Methods for ChromBPNet model has been recently described in detail[63,84]. For each cell type and vascular site cluster (i.e. SMC/Fibro/Endo, Asc/Car/Desc) from the dataset, we created pooled all of the fragments that belonged to each cluster based on barcode to cell type maps to create pseudobulk fragment files. We then splitting each fragment in half to obtain one read per strand and converted each fragment file to a tagalign file. We then used these files as input to the ENCODE ATAC-seq pipeline v1.10.0 (available at https://github.com/ENCODE-DCC/atac-seq-pipeline), which generated peak calls using MACS2 (available at https://github.com/macs3-project/MACS). For downstream analyses involving ChromBPNet, we used the overlap peak set, which only includes peaks called from the original sample that overlap with peak calls from both of the pseudo-replicates created by the pipeline by randomly allocating the reads from the original sample into each synthetic replicate.

Training ChromBPNet models

To model the cell type and vascular–specific chromatin accessibility in each vascular bed cell type, we trained ChromBPNet models on their cell type–resolved pseudobulk ATAC-seq profiles (ChromBPNet is available at https://github.com/kundajelab/ChromBPNet). For each sample, we used as input the peak and tagalign files generated by the ENCODE ATAC-seq pipeline, along with a Tn5-bias model trained on the descending aorta SMC. We trained the ChromBPNet models using a 5-fold cross-validation scheme — ensuring that each chromosome

appeared in the test set of at least one cross-validation fold. **Table S25** outlines the chromosomes included in the train, validation, and test splits for each fold in mouse.

Upon completion of training, any model that continued to respond to the tn5 motif's consensus sequences after unplugging the bias model, indicating an unsuccessful bias-correction procedure, was re-trained until the model showed a limited response to the tn5 motif's consensus sequences embedded in random genomic backgrounds. Specifically, this retraining was performed if the maximum prediction from the profile head for any base-pair from the examples containing the tn5 motif exceeded 0.002.

Creating the Variant List

To create a comprehensive set of variants associated with aorta diameter, we first obtained all of the genome-wide significant GWAS variants from Pirruccello et al.[17] (p-value < 5e-8). Next, we used Plink v1.9 to find all variants from the AFR, AMR, EAS, EUR, and SAS populations in 1000 genomes that are in high LD (r2 > 0.8) with the genome-wide significant GWAS variants, and we used TopLD to do the same for the African, East Asian, European, and South Asian populations in TOPMed. The final variant list included all of the genome-wide significant variants and any variants in high LD with them from any of the above populations in either 1000 genomes or TOPMed.

Predicting variant effects using ChromBPNet models

To score each variant in this study, we used the ChromBPNet models for each cell type to predict the base-resolution scATAC-seq coverage profiles for the 1 kb genomic sequence centered at each variant and containing the reference and alternate allele. We then estimated the variant's effect size using two measures: (1) the log2 fold

change in total predicted coverage (total counts) within each 1 kb window for the alternate versus reference allele and (2) the Jensen–Shannon distance (JSD) between the base-resolution predicted probability profiles for the reference and alternate allele (capturing changes in profile shape).

We assessed statistical significance for these scores using empirical null distributions constructed by shuffling the 2114 bp sequence around each variant multiple times while preserving dinucleotide frequency. Next, each shuffled sequence was duplicated, and the variant's reference or alternate allele was inserted at the center, resulting in a total of one million null variants for each set of observed variants scored. Each null variant was scored with the same procedure as the observed variants. For each observed variant, we then computed the proportion of null variants with an equally high or higher (more extreme) score to derive empirical *P*-values for the log2 fold change and JSD scores. The code base for scoring variants is at https://github.com/kundajelab/variant-scorer.

Motif enrichment analysis from scored ChromBPNet variants

To identify variants that have differential effect based on vascular site (i.e. between ascending and descending aortic cell types), we performed linear regression analysis for absolute log fold change between all scored variants between ascending and descending SMC, fibroblasts, and endothelial cell ChromBPNet models. We identified the top 1% of variants that 'deviate' from the linear regression line that represent the top 1% of variants that have a differing effect based on vascular site within cell type. The genomic position of this top 1% variant list was expanded 100bp upstream and downstream to have a 200bp window that was converted to a BED file.

These genomic positions were then used unbiased Motif Discovery function with MEME Suite[64]. Known motif analysis of top discovered motifs was then performed using TOMTOM (HOCOMOCOv11)[65].

Primary adventitial fibroblast culture and bulk RNA sequencing

We isolated the aortic root and ascending aorta from 6, 12-week-old male C57/BL6 mice as well as the descending thoracic aorta to the level of the renal arteries and carefully separated the media from the adventitia by gentle traction with forceps. Adventitia was cut into small pieces with microscissors and plated into 12 well plates with DMEM media with 5% FBS. Following cell attachment to the plate, cells were cultured at 37C and expanded for 48 hours. Cells were washed with PBS and passaged 1x and plated into a 12 well plate at 65,000 cells/mL. Upon cell culture reaching ~80% confluence, cells were treated with either solute control or mouse TGFβ (invitae) (10ng/mL) for 48 hours. Following 48hr treatment, RNA was isolated using Qiagen RNeasy kits and RNA/cDNA library was sequenced with 250M paired end reads. Fastq files were aligned to MM10 murine genome and raw read count and expression normalization was performed with featureCounts[85] and DEseq2[86], respectively. Using DEseq2 standard pipeline, we performed differential gene expression analysis between groups (i.e. Asc-control versus Desc-control; Asc-TGFβ versus Desc-TGFβ; Asc-control versus Asc-TGFβ; Desc-control versus Desc-TGFβ). Further interaction analysis was performed where DESeq2 pipeline evaluates genes whose response to TGFβ is not the same in both regions — genes that are regulated by TGFβ in one aortic region but not the other, or that change in opposite directions, helping identify region-specific

967 responses to TGFβ. Differential predicted transcription factor motif accessibility from
968 bulk RNAseq analysis of the 200bp upstream of the transcription start site (TSS) from
969 differentially expressed genes was performed with Hypergeometric Optimization of Motif
970 EnRichment (HOMER)[62].

971

972 **Acknowledgements**

973 **Funding:**
974
975 National Institutes of Health grant F32HL160067 (CW)
976 National Institutes of Health grant L30HL159413 (CW)
977 National Institutes of Health grant K08HL167699 (CW)
978 National Institutes of Health grant K08HL153798 (PC)
979 National Institutes of Health grant R01HL134817 (TQ)
980 National Institutes of Health grant R01HL139478 (TQ)
981 National Institutes of Health grant R01HL156846 (TQ)
982 National Institutes of Health grant R01HL151535 (TQ)
983 National Institutes of Health grant R01HL145708 (TQ)
984 National Institutes of Health grant UM1 HG011972 (TQ)
985 American Heart Association grant 20CDA35310303 (PC)
986 American Heart Association grant 23CDA1042900 (CW)
987 American Heart Association grant 23POST1018991 (WG).
988

989 **Disclosures**

990 TQ is a scientific advisor for Amgen. CSW is a consultant and advisory board member
991 for Avidity Biosciences and AIRNA bio. All other authors have no disclosures.

992

993 **References**
994
995 1. Haimovici H, Maier N. Fate of Aortic Homografts in Canine Atherosclerosis. 3.
996 Study of Fresh Abdominal and Thoracic Aortic Implants into Thoracic Aorta: Role
997 of Tissue Susceptibility in Atherogenesis. *Arch Surg*. 1964;89:961-969. doi:
998 10.1001/archsurg.1964.01320060029006
999 2. Haimovici H, Maier N, Strauss L. Fate of aortic homografts in experimental
1000 canine atherosclerosis; study of fresh thoracic implants into abdominal aorta.
1001 *AMA Arch Surg*. 1958;76:282-288. doi: 10.1001/archsurg.1958.01280200104012

3. Haimovici H, Maier N, Strauss L. Fate of aortic homografts in experimental canine atherosclerosis. II. Study of fresh abdominal aortic implants into abdominal aorta. *AMA Arch Surg*. 1959;78:239-245. doi: 10.1001/archsurg.1959.04320020061010
4. Katz TC, Singh MK, Degenhardt K, Rivera-Feliciano J, Johnson RL, Epstein JA, Tabin CJ. Distinct compartments of the proepicardial organ give rise to coronary vascular endothelial cells. *Dev Cell*. 2012;22:639-650. doi: 10.1016/j.devcel.2012.01.012
5. Majesky MW. Developmental basis of vascular smooth muscle diversity. *Arterioscler Thromb Vasc Biol*. 2007;27:1248-1258. doi: 10.1161/ATVBAHA.107.141069
6. Mikawa T, Gourdie RG. Pericardial mesoderm generates a population of coronary smooth muscle cells migrating into the heart along with ingrowth of the epicardial organ. *Dev Biol*. 1996;174:221-232. doi: 10.1006/dbio.1996.0068
7. Red-Horse K, Ueno H, Weissman IL, Krasnow MA. Coronary arteries form by developmental reprogramming of venous cells. *Nature*. 2010;464:549-553. doi: 10.1038/nature08873
8. Sawada H, Rateri DL, Moorleghen JJ, Majesky MW, Daugherty A. Smooth Muscle Cells Derived From Second Heart Field and Cardiac Neural Crest Reside in Spatially Distinct Domains in the Media of the Ascending Aorta-Brief Report. *Arterioscler Thromb Vasc Biol*. 2017;37:1722-1726. doi: 10.1161/ATVBAHA.117.309599
9. Tian X, Hu T, Zhang H, He L, Huang X, Liu Q, Yu W, He L, Yang Z, Zhang Z, et al. Subepicardial endothelial cells invade the embryonic ventricle wall to form coronary arteries. *Cell Res*. 2013;23:1075-1090. doi: 10.1038/cr.2013.83
10. Volz KS, Jacobs AH, Chen HI, Poduri A, McKay AS, Riordan DP, Kofler N, Kitajewski J, Weissman I, Red-Horse K. Pericytes are progenitors for coronary artery smooth muscle. *Elife*. 2015;4. doi: 10.7554/eLife.10036
11. Jones GT, Tromp G, Kuivaniemi H, Gretarsdottir S, Baas AF, Giusti B, Strauss E, Van't Hof FN, Webb TR, Erdman R, et al. Meta-Analysis of Genome-Wide Association Studies for Abdominal Aortic Aneurysm Identifies Four New Disease-Specific Risk Loci. *Circ Res*. 2017;120:341-353. doi: 10.1161/CIRCRESAHA.116.308765
12. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47:1121-1130. doi: 10.1038/ng.3396
13. Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet*. 2011;43:333-338. doi: 10.1038/ng.784
14. van der Harst P, Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ Res*. 2018;122:433-443. doi: 10.1161/CIRCRESAHA.117.312086
15. Strawbridge RJ, Ward J, Bailey MES, Cullen B, Ferguson A, Graham N, Johnston KJA, Lyall LM, Pearsall R, Pell J, et al. Carotid Intima-Media Thickness:

Novel Loci, Sex-Specific Effects, and Genetic Correlations With Obesity and Glucometabolic Traits in UK Biobank. *Arterioscler Thromb Vasc Biol*. 2020;40:446-461. doi: 10.1161/ATVBAHA.119.313226

16. Franceschini N, Giambartolomei C, de Vries PS, Finan C, Bis JC, Huntley RP, Lovering RC, Tajuddin SM, Winkler TW, Graff M, et al. GWAS and colocalization analyses implicate carotid intima-media thickness and carotid plaque loci in cardiovascular outcomes. *Nat Commun*. 2018;9:5141. doi: 10.1038/s41467-018-07340-5
17. Pirruccello JP, Chaffin MD, Chou EL, Fleming SJ, Lin H, Nekoui M, Khurshid S, Friedman SF, Bick AG, Arduini A, et al. Deep learning enables genetic analysis of the human thoracic aorta. *Nat Genet*. 2022;54:40-51. doi: 10.1038/s41588-021-00962-4
18. Oudelaar AM, Higgs DR. The relationship between genome structure and function. *Nat Rev Genet*. 2021;22:154-168. doi: 10.1038/s41576-020-00303-x
19. Anene-Nzelu CG, Lee MCJ, Tan WLW, Dashi A, Foo RSY. Genomic enhancers in cardiac development and disease. *Nat Rev Cardiol*. 2022;19:7-25. doi: 10.1038/s41569-021-00597-2
20. Alexanian M, Przytycki PF, Micheletti R, Padmanabhan A, Ye L, Travers JG, Gonzalez-Teran B, Silva AC, Duan Q, Ranade SS, et al. A transcriptional switch governs fibroblast activation in heart disease. *Nature*. 2021;595:438-443. doi: 10.1038/s41586-021-03674-1
21. Hocker JD, Poirion OB, Zhu F, Buchanan J, Zhang K, Chiou J, Wang TM, Zhang Q, Hou X, Li YE, et al. Cardiac cell type-specific gene regulatory programs and disease risk association. *Sci Adv*. 2021;7. doi: 10.1126/sciadv.abf1444
22. Cheng P, Wirka RC, Clarke LS, Zhao Q, Kundu R, Nguyen T, Nair S, Sharma D, Kim HJ, Shi H, et al. ZEB2 Shapes the Epigenetic Landscape of Atherosclerosis. *Circulation*. 2022. doi: 10.1161/CIRCULATIONAHA.121.057789
23. Wang Y, Gao H, Wang F, Ye Z, Mokry M, Turner AW, Ye J, Koplev S, Luo L, Alsaigh T, et al. Dynamic changes in chromatin accessibility are associated with the atherogenic transitioning of vascular smooth muscle cells. *Cardiovasc Res*. 2021. doi: 10.1093/cvr/cvab347
24. Depuydt MAC, Prange KHM, Slenders L, Ord T, Elbersen D, Boltjes A, de Jager SCA, Asselbergs FW, de Borst GJ, Aavik E, et al. Microanatomy of the Human Atherosclerotic Plaque by Single-Cell Transcriptomics. *Circ Res*. 2020;127:1437-1455. doi: 10.1161/CIRCRESAHA.120.316770
25. Aquino JB, Sierra R, Montaldo LA. Diverse cellular origins of adult blood vascular endothelial cells. *Dev Biol*. 2021;477:117-132. doi: 10.1016/j.ydbio.2021.05.010
26. Anderson MJ, Pham VN, Vogel AM, Weinstein BM, Roman BL. Loss of unc45a precipitates arteriovenous shunting in the aortic arches. *Dev Biol*. 2008;318:258-267. doi: 10.1016/j.ydbio.2008.03.022
27. Li P, Pashmforoush M, Sucov HM. Mesodermal retinoic acid signaling regulates endothelial cell coalescence in caudal pharyngeal arch artery vasculogenesis. *Dev Biol*. 2012;361:116-124. doi: 10.1016/j.ydbio.2011.10.018
28. Paffett-Lugassy N, Singh R, Nevis KR, Guner-Ataman B, O'Loughlin E, Jahangiri L, Harvey RP, Burns CG, Burns CE. Heart field origin of great vessel precursors

relies on nkx2.5-mediated vasculogenesis. *Nat Cell Biol*. 2013;15:1362-1369. doi: 10.1038/ncb2862

29. Dobnikar L, Taylor AL, Chappell J, Oldach P, Harman JL, Oerton E, Dzierzak E, Bennett MR, Spivakov M, Jorgensen HF. Disease-relevant transcriptional signatures identified in individual smooth muscle cells from healthy mouse vessels. *Nat Commun*. 2018;9:4567. doi: 10.1038/s41467-018-06891-x
30. Yu L, Zhang J, Gao A, Zhang M, Wang Z, Yu F, Guo X, Su G, Zhang Y, Zhang M, et al. An intersegmental single-cell profile reveals aortic heterogeneity and identifies a novel Malat1(+) vascular smooth muscle subtype involved in abdominal aortic aneurysm formation. *Signal Transduct Target Ther*. 2022;7:125. doi: 10.1038/s41392-022-00943-x
31. Kim JB, Zhao Q, Nguyen T, Pjanic M, Cheng P, Wirka R, Travisano S, Nagao M, Kundu R, Quertermous T. Environment-Sensing Aryl Hydrocarbon Receptor Inhibits the Chondrogenic Fate of Modulated Smooth Muscle Cells in Atherosclerotic Lesions. *Circulation*. 2020;142:575-590. doi: 10.1161/CIRCULATIONAHA.120.045981
32. Wirka RC, Wagh D, Paik DT, Pjanic M, Nguyen T, Miller CL, Kundu R, Nagao M, Coller J, Koyano TK, et al. Atheroprotective roles of smooth muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell analysis. *Nat Med*. 2019;25:1280-1289. doi: 10.1038/s41591-019-0512-5
33. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, Hao Y, Stoeckius M, Smibert P, Satija R. Comprehensive Integration of Single-Cell Data. *Cell*. 2019;177:1888-1902 e1821. doi: 10.1016/j.cell.2019.05.031
34. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*. 2010;28:495-501. doi: 10.1038/nbt.1630
35. Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods*. 2017;14:975-978. doi: 10.1038/nmeth.4401
36. Liu J, Ren Y, Kang L, Zhang L. Overexpression of CCN3 inhibits inflammation and progression of atherosclerosis in apolipoprotein E-deficient mice. *PLoS One*. 2014;9:e94912. doi: 10.1371/journal.pone.0094912
37. Shimoyama T, Hiraoka S, Takemoto M, Koshizaka M, Tokuyama H, Tokuyama T, Watanabe A, Fujimoto M, Kawamura H, Sato S, et al. CCN3 inhibits neointimal hyperplasia through modulation of smooth muscle cell growth and migration. *Arterioscler Thromb Vasc Biol*. 2010;30:675-682. doi: 10.1161/ATVBAHA.110.203356
38. Al Haj Zen A, Caligiuri G, Sainz J, Lemitre M, Demerens C, Lafont A. Decorin overexpression reduces atherosclerosis development in apolipoprotein E-deficient mice. *Atherosclerosis*. 2006;187:31-39. doi: 10.1016/j.atherosclerosis.2005.08.023
39. Licht AH, Nubel T, Feldner A, Jurisch-Yaksi N, Marcello M, Demicheva E, Hu JH, Hartenstein B, Augustin HG, Hecker M, et al. Junb regulates arterial contraction capacity, cellular contractility, and motility via its target Myl9 in mice. *J Clin Invest*. 2010;120:2307-2318. doi: 10.1172/JCI41749

40. Yoshida T, Kaestner KH, Owens GK. Conditional deletion of Kruppel-like factor 4 delays downregulation of smooth muscle cell differentiation markers but accelerates neointimal formation following vascular injury. *Circ Res*. 2008;102:1548-1557. doi: 10.1161/CIRCRESAHA.108.176974
41. Li J, Adams LD, Wang X, Pabon L, Schwartz SM, Sane DC, Geary RL. Regulator of G protein signaling 5 marks peripheral arterial smooth muscle cells and is downregulated in atherosclerotic plaque. *J Vasc Surg*. 2004;40:519-528. doi: 10.1016/j.jvs.2004.06.021
42. Daniel JM, Prock A, Dutzmann J, Sonnenschein K, Thum T, Bauersachs J, Sedding DG. Regulator of G-Protein Signaling 5 Prevents Smooth Muscle Cell Proliferation and Attenuates Neointima Formation. *Arterioscler Thromb Vasc Biol*. 2016;36:317-327. doi: 10.1161/ATVBAHA.115.305974
43. Paloschi V, Kurtovic S, Folkersen L, Gomez D, Wagsater D, Roy J, Petrini J, Eriksson MJ, Caidahl K, Hamsten A, et al. Impaired splicing of fibronectin is associated with thoracic aortic aneurysm formation in patients with bicuspid aortic valve. *Arterioscler Thromb Vasc Biol*. 2011;31:691-697. doi: 10.1161/ATVBAHA.110.218461
44. Soubeyrand S, Lau P, Nikpay M, Dang AT, McPherson R. Common Polymorphism That Protects From Cardiovascular Disease Increases Fibronectin Processing and Secretion. *Circ Genom Precis Med*. 2022:CIRCGEN121003428. doi: 10.1161/CIRCGEN.121.003428
45. Azad AK, Chakrabarti S, Xu Z, Davidge ST, Fu Y. Coiled-coil domain containing 3 (CCDC3) represses tumor necrosis factor-alpha/nuclear factor kappaB-induced endothelial inflammation. *Cell Signal*. 2014;26:2793-2800. doi: 10.1016/j.cellsig.2014.08.025
46. van de Laar IM, Oldenburg RA, Pals G, Roos-Hesselink JW, de Graaf BM, Verhagen JM, Hoedemaekers YM, Willemsen R, Severijnen LA, Venselaar H, et al. Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. *Nat Genet*. 2011;43:121-126. doi: 10.1038/ng.744
47. Iyer D, Zhao Q, Wirka R, Naravane A, Nguyen T, Liu B, Nagao M, Cheng P, Miller CL, Kim JB, et al. Coronary artery disease genes SMAD3 and TCF21 promote opposing interactive genetic programs that regulate smooth muscle cell differentiation and disease risk. *PLoS Genet*. 2018;14:e1007681. doi: 10.1371/journal.pgen.1007681
48. Renard M, Francis C, Ghosh R, Scott AF, Witmer PD, Ades LC, Andelfinger GU, Arnaud P, Boileau C, Callewaert BL, et al. Clinical Validity of Genes for Heritable Thoracic Aortic Aneurysm and Dissection. *J Am Coll Cardiol*. 2018;72:605-615. doi: 10.1016/j.jacc.2018.04.089
49. Chou E, Pirruccello JP, Ellinor PT, Lindsay ME. Genetics and mechanisms of thoracic aortic disease. *Nat Rev Cardiol*. 2023;20:168-180. doi: 10.1038/s41569-022-00763-0
50. Regard JB, Scheek S, Borbiev T, Lanahan AA, Schneider A, Demetriades AM, Hiemisch H, Barnes CA, Verin AD, Worley PF. Verge: a novel vascular early response gene. *J Neurosci*. 2004;24:4092-4103. doi: 10.1523/JNEUROSCI.4252-03.2004

51. Diaz-Canestro C, Bonetti NR, Wust P, Nageswaran V, Liberale L, Beer JH, Montecucco F, Luscher TF, Bohacek J, Camici GG. Apold1 deficiency associates with increased arterial thrombosis in vivo. *Eur J Clin Invest*. 2020;50:e13191. doi: 10.1111/eci.13191
52. Briot A, Jaroszewicz A, Warren CM, Lu J, Touma M, Rudat C, Hofmann JJ, Airik R, Weinmaster G, Lyons K, et al. Repression of Sox9 by Jag1 is continuously required to suppress the default chondrogenic fate of vascular smooth muscle cells. *Dev Cell*. 2014;31:707-721. doi: 10.1016/j.devcel.2014.11.023
53. Martos-Rodriguez CJ, Albarran-Juarez J, Morales-Cano D, Caballero A, MacGrogan D, de la Pompa JL, Carramolino L, Bentzon JF. Fibrous Caps in Atherosclerosis Form by Notch-Dependent Mechanisms Common to Arterial Media Development. *Arterioscler Thromb Vasc Biol*. 2021;41:e427-e439. doi: 10.1161/ATVBAHA.120.315627
54. Hubner K, Grassme KS, Rao J, Wenke NK, Zimmer CL, Korte L, Muller K, Sumanas S, Greber B, Herzog W. Wnt signaling positively regulates endothelial cell fate specification in the Fli1a-positive progenitor population via Lef1. *Dev Biol*. 2017;430:142-155. doi: 10.1016/j.ydbio.2017.08.004
55. Oettgen P. Regulation of vascular inflammation and remodeling by ETS factors. *Circ Res*. 2006;99:1159-1166. doi: 10.1161/01.RES.0000251056.85990.db
56. Schafer CM, Gurley JM, Kurylowicz K, Lin PK, Chen W, Elliott MH, Davis GE, Bhatti F, Griffin CT. An inhibitor of endothelial ETS transcription factors promotes physiologic and therapeutic vessel regression. *Proc Natl Acad Sci U S A*. 2020;117:26494-26502. doi: 10.1073/pnas.2015980117
57. Mukherjee K, Pingili AK, Singh P, Dhodi AN, Dutta SR, Gonzalez FJ, Malik KU. Testosterone Metabolite 6beta-Hydroxytestosterone Contributes to Angiotensin II-Induced Abdominal Aortic Aneurysms in Apoe(-/-) Male Mice. *J Am Heart Assoc*. 2021;10:e018536. doi: 10.1161/JAHA.120.018536
58. Thirunavukkarasu S, Khan NS, Song CY, Ghafoor HU, Brand DD, Gonzalez FJ, Malik KU. Cytochrome P450 1B1 Contributes to the Development of Angiotensin II-Induced Aortic Aneurysm in Male Apoe(-/-) Mice. *Am J Pathol*. 2016;186:2204-2219. doi: 10.1016/j.ajpath.2016.04.005
59. Traylor M, Tozer DJ, Croall ID, Lisiecka-Ford DM, Olorunda AO, Boncoraglio G, Dichgans M, Lemmens R, Rosand J, Rost NS, et al. Genetic variation in PLEKHG1 is associated with white matter hyperintensities (n = 11,226). *Neurology*. 2019;92:e749-e757. doi: 10.1212/WNL.0000000000006952
60. Zernecke A, Winkels H, Cochain C, Williams JW, Wolf D, Soehnlein O, Robbins CS, Monaco C, Park I, McNamara CA, et al. Meta-Analysis of Leukocyte Diversity in Atherosclerotic Mouse Aortas. *Circ Res*. 2020;127:402-426. doi: 10.1161/CIRCRESAHA.120.316903
61. Fleck JS, Jansen SMJ, Wollny D, Zenk F, Seimiya M, Jain A, Okamoto R, Santel M, He Z, Camp JG, et al. Inferring and perturbing cell fate regulomes in human brain organoids. *Nature*. 2023;621:365-372. doi: 10.1038/s41586-022-05279-8
62. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38:576-589. doi: 10.1016/j.molcel.2010.05.004

63. Pampari A, Shcherbina A, Kvon EZ, Kosicki M, Nair S, Kundu S, Kathiria AS, Risca VI, Kuningas K, Alasoo K, et al. ChromBPNet: bias factorized, base-resolution deep learning models of chromatin accessibility reveal cis-regulatory sequence syntax, transcription factor footprints and regulatory variants. *bioRxiv*. 2025. doi: 10.1101/2024.12.25.630221
64. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. *Nucleic Acids Res*. 2015;43:W39-49. doi: 10.1093/nar/gkv416
65. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. *Genome Biol*. 2007;8:R24. doi: 10.1186/gb-2007-8-2-r24
66. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science*. 2012;337:1190-1195. doi: 10.1126/science.1222794
67. Chen Y, Xiao D, Zhang L, Cai CL, Li BY, Liu Y. The Role of Tbx20 in Cardiovascular Development and Function. *Front Cell Dev Biol*. 2021;9:638542. doi: 10.3389/fcell.2021.638542
68. Koyama S, Ito K, Terao C, Akiyama M, Horikoshi M, Momozawa Y, Matsunaga H, Ieki H, Ozaki K, Onouchi Y, et al. Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. *Nat Genet*. 2020;52:1169-1177. doi: 10.1038/s41588-020-0705-3
69. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, Narita A, Konuma T, Yamamoto K, Akiyama M, et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet*. 2021;53:1415-1424. doi: 10.1038/s41588-021-00931-x
70. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, Ntalla I, Surendran P, Liu C, Cook JP, et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet*. 2017;49:403-415. doi: 10.1038/ng.3768
71. Benjamins JW, Yeung MW, van de Vegte YJ, Said MA, van der Linden T, Ties D, Juarez-Orozco LE, Verweij N, van der Harst P. Genomic insights in ascending aortic size and distensibility. *EBioMedicine*. 2022;75:103783. doi: 10.1016/j.ebiom.2021.103783
72. Takeuchi F, Akiyama M, Matoba N, Katsuya T, Nakatochi M, Tabara Y, Narita A, Saw WY, Moon S, Spracklen CN, et al. Interethnic analyses of blood pressure loci in populations of East Asian and European descent. *Nat Commun*. 2018;9:5052. doi: 10.1038/s41467-018-07345-0
73. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, Schoech A, Pasaniuc B, Price AL. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am J Hum Genet*. 2019;104:65-75. doi: 10.1016/j.ajhg.2018.11.008
74. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, Almgren P, Alonso A, Anderson CD, Aragam KG, et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet*. 2018;50:1225-1233. doi: 10.1038/s41588-018-0133-9
75. Acharya A, Baek ST, Huang G, Eskiocak B, Goetsch S, Sung CY, Banfi S, Sauer MF, Olsen GS, Duffield JS, et al. The bHLH transcription factor Tcf21 is required

for lineage-specific EMT of cardiac fibroblast progenitors. *Development*. 2012;139:2139-2149. doi: 10.1242/dev.079970

76. Fardi M, Alivand M, Baradaran B, Farshdousti Hagh M, Solali S. The crucial role of ZEB2: From development to epithelial-to-mesenchymal transition and cancer complexity. *J Cell Physiol*. 2019. doi: 10.1002/jcp.28277
77. Erdmann J, Kessler T, Munoz Venegas L, Schunkert H. A decade of genome-wide association studies for coronary artery disease: the challenges ahead. *Cardiovasc Res*. 2018;114:1241-1257. doi: 10.1093/cvr/cvy084
78. Gadson PF, Jr., Dalton ML, Patterson E, Svoboda DD, Hutchinson L, Schram D, Rosenquist TH. Differential response of mesoderm- and neural crest-derived smooth muscle to TGF-beta1: regulation of c-myb and alpha1 (I) procollagen genes. *Exp Cell Res*. 1997;230:169-180. doi: 10.1006/excr.1996.3398
79. Topouzis S, Majesky MW. Smooth muscle lineage diversity in the chick embryo. Two types of aortic smooth muscle cell differ in growth and receptor-mediated transcriptional responses to transforming growth factor-beta. *Dev Biol*. 1996;178:430-445. doi: 10.1006/dbio.1996.0229
80. Forte E, Ramialison M, Nim HT, Mara M, Li JY, Cohn R, Daigle SL, Boyd S, Stanley EG, Elefanty AG, et al. Adult mouse fibroblasts retain organ-specific transcriptomic identity. *Elife*. 2022;11. doi: 10.7554/eLife.71008
81. Xie W, Ke Y, You Q, Li J, Chen L, Li D, Fang J, Chen X, Zhou Y, Chen L, et al. Single-Cell RNA Sequencing and Assay for Transposase-Accessible Chromatin Using Sequencing Reveals Cellular and Molecular Dynamics of Aortic Aging in Mice. *Arterioscler Thromb Vasc Biol*. 2022;42:156-171. doi: 10.1161/ATVBAHA.121.316883
82. Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, Steemers FJ, Trapnell C, Shendure J. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science*. 2015;348:910-914. doi: 10.1126/science.aab1601
83. Cusanovich DA, Reddington JP, Garfield DA, Daza RM, Aghamirzaie D, Marco-Ferreres R, Pliner HA, Christiansen L, Qiu X, Steemers FJ, et al. The cis-regulatory dynamics of embryonic development at single-cell resolution. *Nature*. 2018;555:538-542. doi: 10.1038/nature25981
84. Marderstein AR, Kundu S, Padhi EM, Deshpande S, Wang A, Robb E, Sun Y, Yun CM, Pomales-Matos D, Xie Y, et al. Mapping the regulatory effects of common and rare non-coding variants across cellular and developmental contexts in the brain and heart. *bioRxiv*. 2025. doi: 10.1101/2025.02.18.638922
85. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30:923-930. doi: 10.1093/bioinformatics/btt656
86. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550. doi: 10.1186/s13059-014-0550-8

Figures and Figure Legends

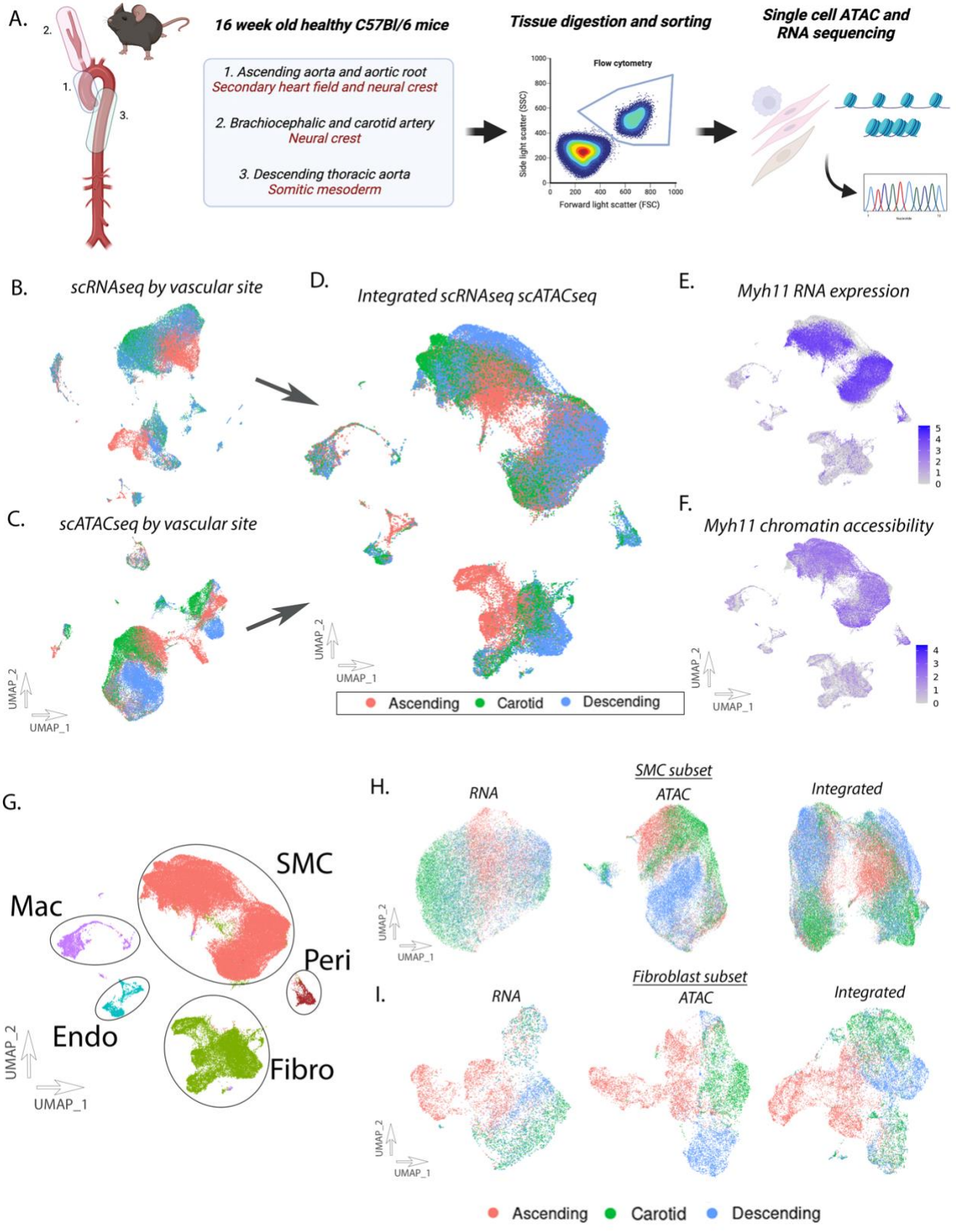


Figure 1. Transcriptomic and epigenomic landscape of single vascular cells. Single cell RNA seq (scRNAseq) and single cell ATAC seq (scATACseq) was performed on vascular tissue in adult healthy mice C57Bl/6) in three vascular sites (aortic root/ascending aorta, brachiocephalic/carotid artery, descending thoracic aorta)(A). UMAP of scRNAseq data (B), scATACseq data (C), and integrated datasets (D). RNA expression *Myh11* in integrated dataset (E). Chromatin accessibility of *Myh11* in integrated dataset (F). Differentiation of major cell populations of the vascular wall (G). Subset analysis for scRNA, scATAC, and integrated datasets in SMC (H) and Fibroblast (I).

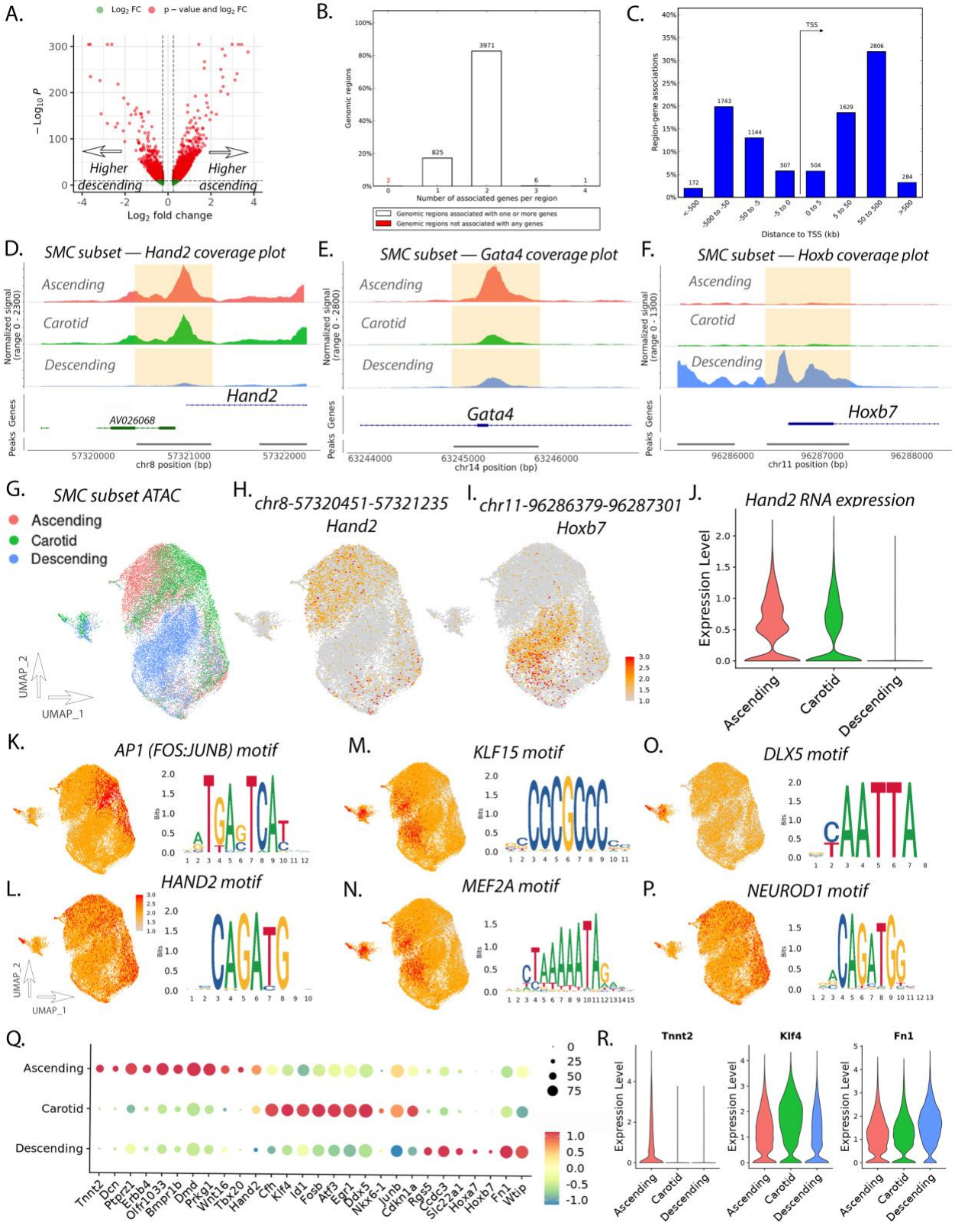


Figure 2. SMC subset analysis reveals differential chromatin peak accessibility, gene expression, and transcription factor motif accessibility between vascular sites. Volcano plot of differentially accessible peaks comparing ascending and descending SMC (A). Peak to gene analysis revealing number of associated genes per genomic region (B). Histogram of distance from peak to associated gene TSSs (C). Coverage plots of *Hand2* (D), *Gata4* (E), and *Hoxb* (F) for ascending, carotid, and descending SMC. UMAP of scATACseq data for SMC subset (G). Featureplot of *Hand2* (H) and *Hoxb7* (I) peaks. Violin plot of *Hand2* RNA expression between vascular sites (J). Featureplot of motif accessibility and motif sequence for *AP1* (K), *Hand2* (L), *Klf15* (M), *Mef2a* (N), *Dlx5* (O), and *Neurod1* (P). Dotplot of RNA expression for top differentially expressed transcripts (Q). Violinplot of SMC RNA expression across vascular sites for *Tnnt2*, *Klf4*, and *Fn1* (R).

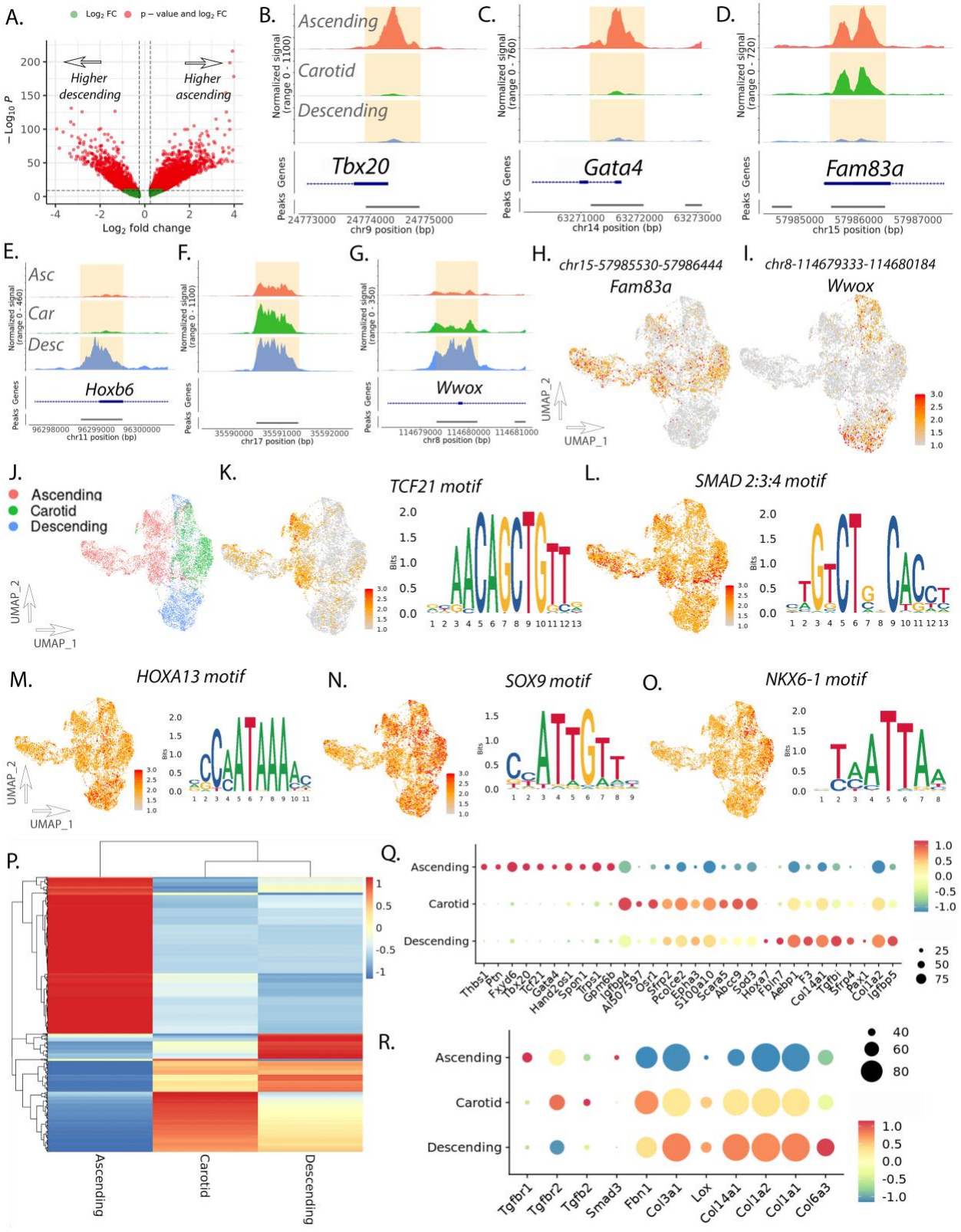
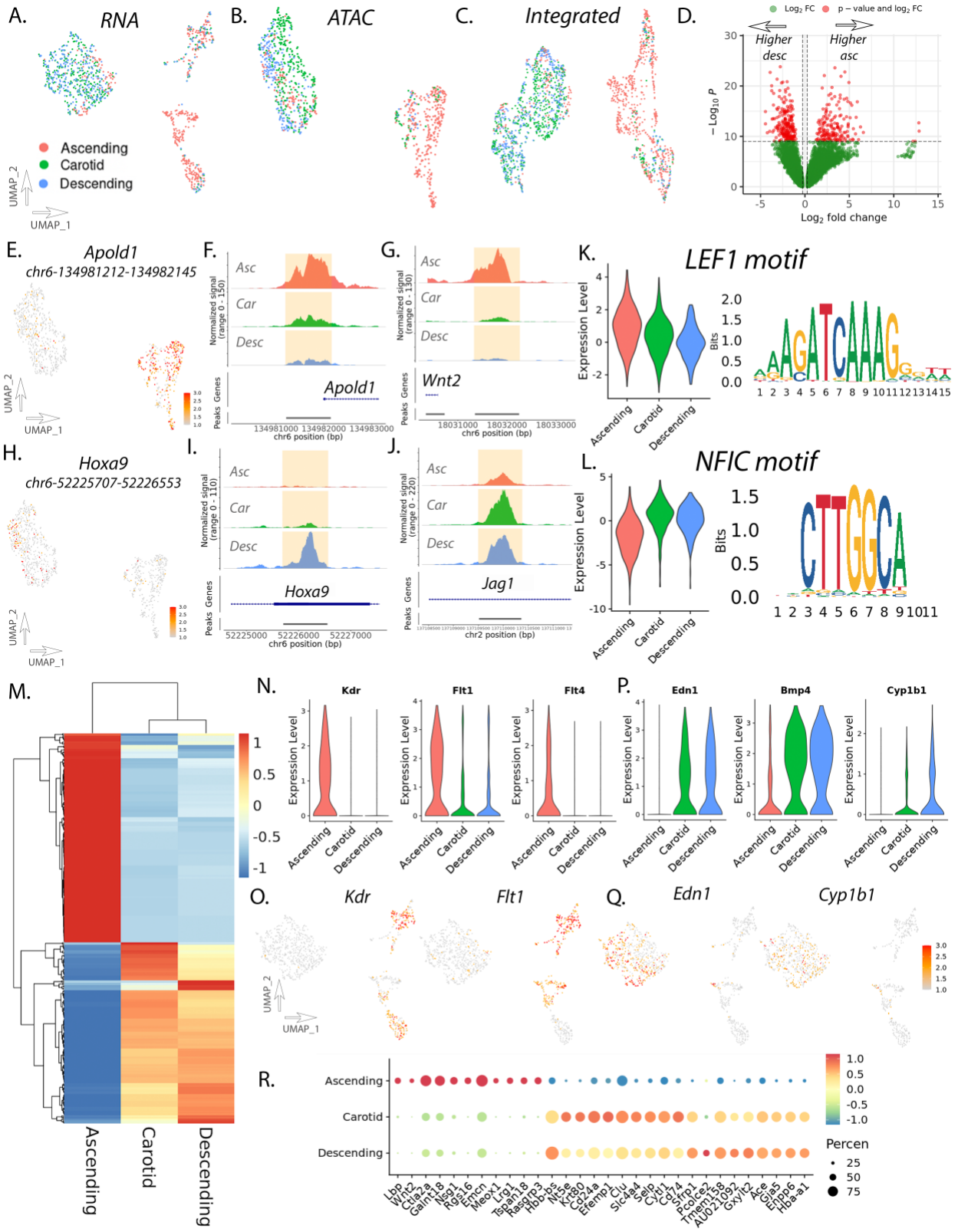


Figure 3. Fibroblast subset analysis reveals vascular site-specific enhancers for developmental and disease genes. Volcano plot of differentially accessible peaks between ascending and descending aortic fibroblasts (A). Coverage plot for ascending, carotid, and descending aortic fibroblasts at *Tbx20* (B), *Gata4* (C), *Fam83a* (D), *Hoxb6* (E), *chr17* gene desert (F), and *Wwox* (G). Featureplot of scATACseq data of fibroblast subset for peaks near *Fam83a* (H) and *Wwox* (I). UMAP of scATACseq data from fibroblast subset by vascular site (J). Featureplot of scATACseq for motif accessibility and motif sequence for *Tcf21* (K), *Smad2:3:4* (L), *Hoxa13* (M), *Sox9* (N), and *Nkx6-1* (O). Heatmap of scRNAseq differentially expressed genes in ascending, carotid, and descending fibroblast subset (P). Dotplot of scRNAseq transcript expression for top differentially expressed genes in fibroblast subset (Q). Dotplot of scRNAseq transcript expression for hereditary aortopathy genes (R).



1359

Figure 4. Endothelial cell subset analysis identifies differentially accessible chromatin peaks and differential transcription factor (TF) motif accessibility. UMAP of RNA (A), ATAC (B), and integrated (C) datasets for endothelial cell subset. Volcano plot of differentially accessible peaks between ascending and descending endothelial cells (D). Featureplot of peak accessibility of *Apold1* peak (E). Coverage plots for *Apold1* (F) and *Wnt2* (G) genomic regions in endothelial cell subset. Featureplot of peak accessibility of *Hoxa9* peak (H). Coverage plots for *Hoxa9* (I) and *Jag1* (J) genomic regions in endothelial cell subset. Violinplot of motif accessibility and motif sequence for *LEF1* (K) and *NFIC* (L). Heatmap of differential RNA expression by endothelial cell vascular site (M). Violinplots of RNA expression in endothelial cells for *Kdr*, *Flt1*, and *Flt4* (N). Featureplots of RNA expression for *Kdr* and *Flt1* (O). Violinplots of RNA expression in endothelial cells for *Edn1*, *Bmp4*, and *Cyp1b1* (P). Featureplots of RNA expression for *End1* and *Cyp1b1* (Q). Dotplot of scRNAseq transcript expression for top differentially expressed genes in endothelial cell subset (R).

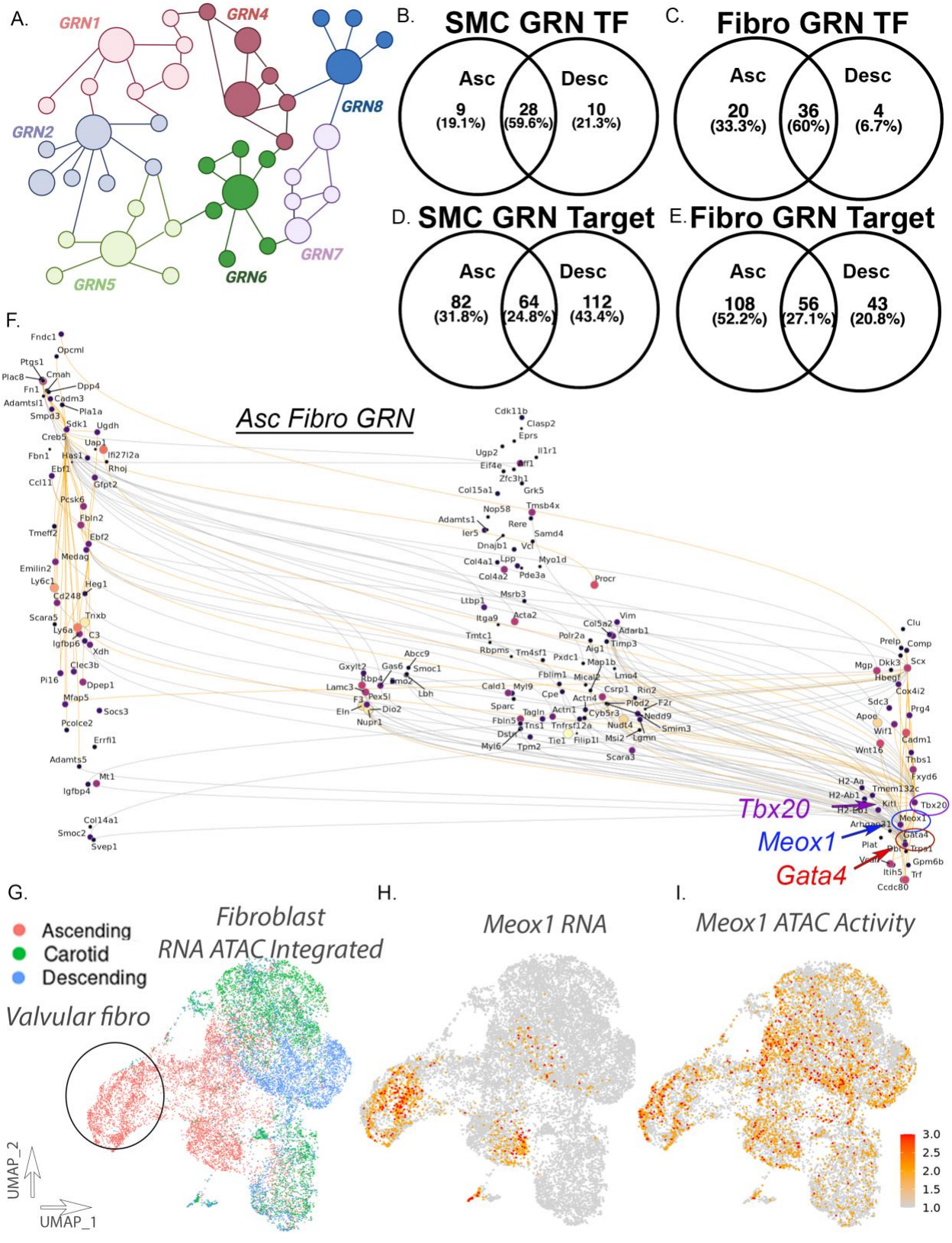


Figure 5. Gene regulatory network analysis reveals cell type and vascular site-specific regulatory networks within the vessel wall. Schematic diagram of gene regulatory networks (GRN)(A). Venn diagrams of GRN transcription factors comparing ascending and descending SMCs (B) and fibroblasts (C). Venn diagrams of GRN transcription factors gene targets comparing ascending and descending SMCs (D) and fibroblasts (E). GRN visualization for ascending fibroblasts network analysis (F). UMAP of scRNA/scATAC integrated fibroblast dataset by vascular site (G). Featureplot of *Meox1* RNA expression (H) and chromatin accessibility (I) in integrated fibroblast dataset.

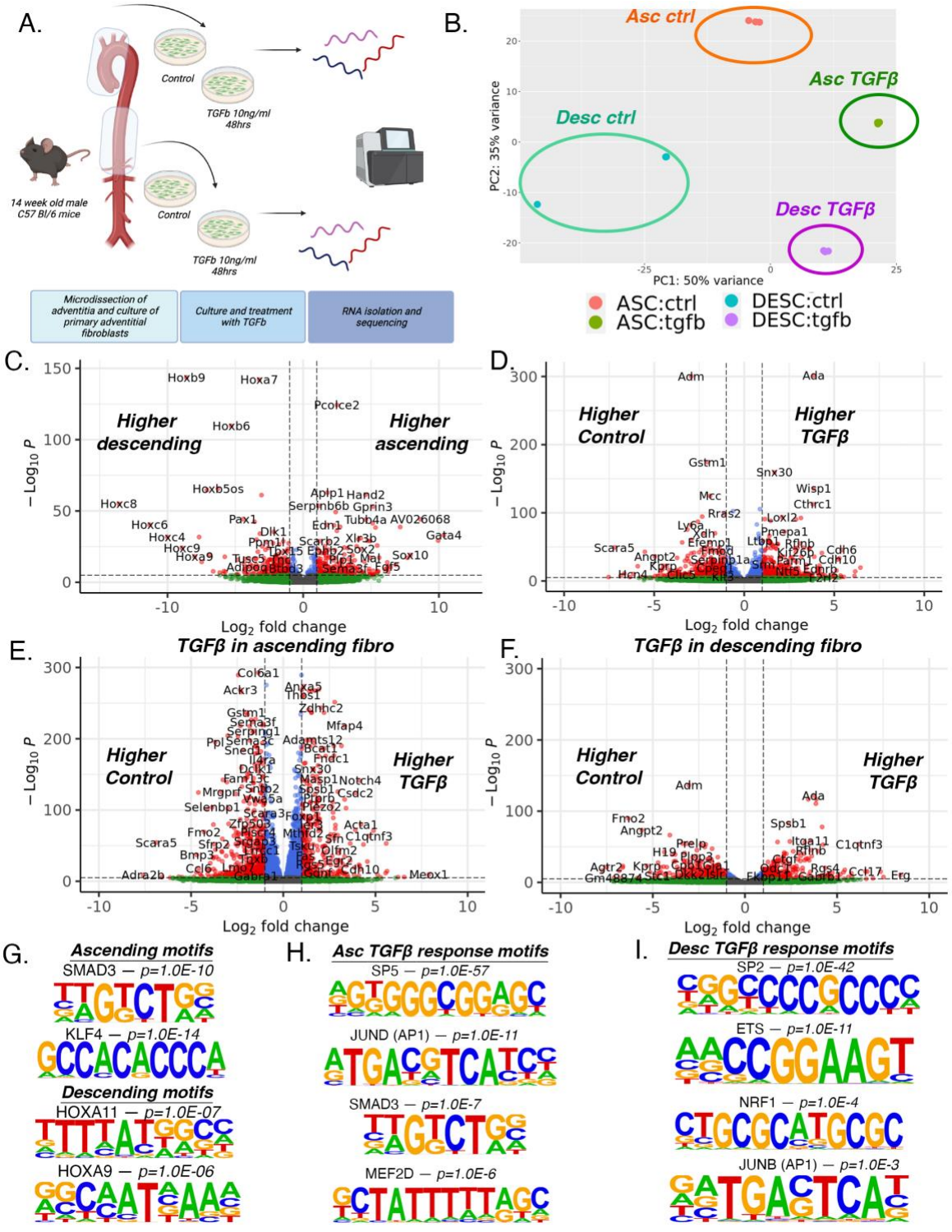


Figure 6. Primary adventitial fibroblasts retain vascular site-specific transcriptomic features following *in vitro* culture revealing vascular site-specific response to TGFβ. Schematic diagram of primary adventitial fibroblast culture,

1390 treatment, and bulk RNA sequencing (A). Principal component analysis (PCA) by
1391 vascular site and TGFβ treatment (B). Volcano plots showing differential gene
1392 expression analysis of all samples by vascular site (C) and by TGFβ treatment (D).
1393 Volcano plot of DE gene analysis in response to TGFβ in ascending (E) and descending
1394 (F) fibroblasts. Motif sequences of identified enriched transcription factors of 200bp
1395 sequences upstream of gene TSSs with increased expression in ascending and
1396 descending fibroblasts (G). Motif sequences of identified enriched transcription factors
1397 of 200bp sequences upstream of gene TSSs in of upregulated genes in ascending
1398 fibroblasts in response to TGFβ treatment (H) and descending fibroblasts in response to
1399 TGFβ treatment (I).
1400

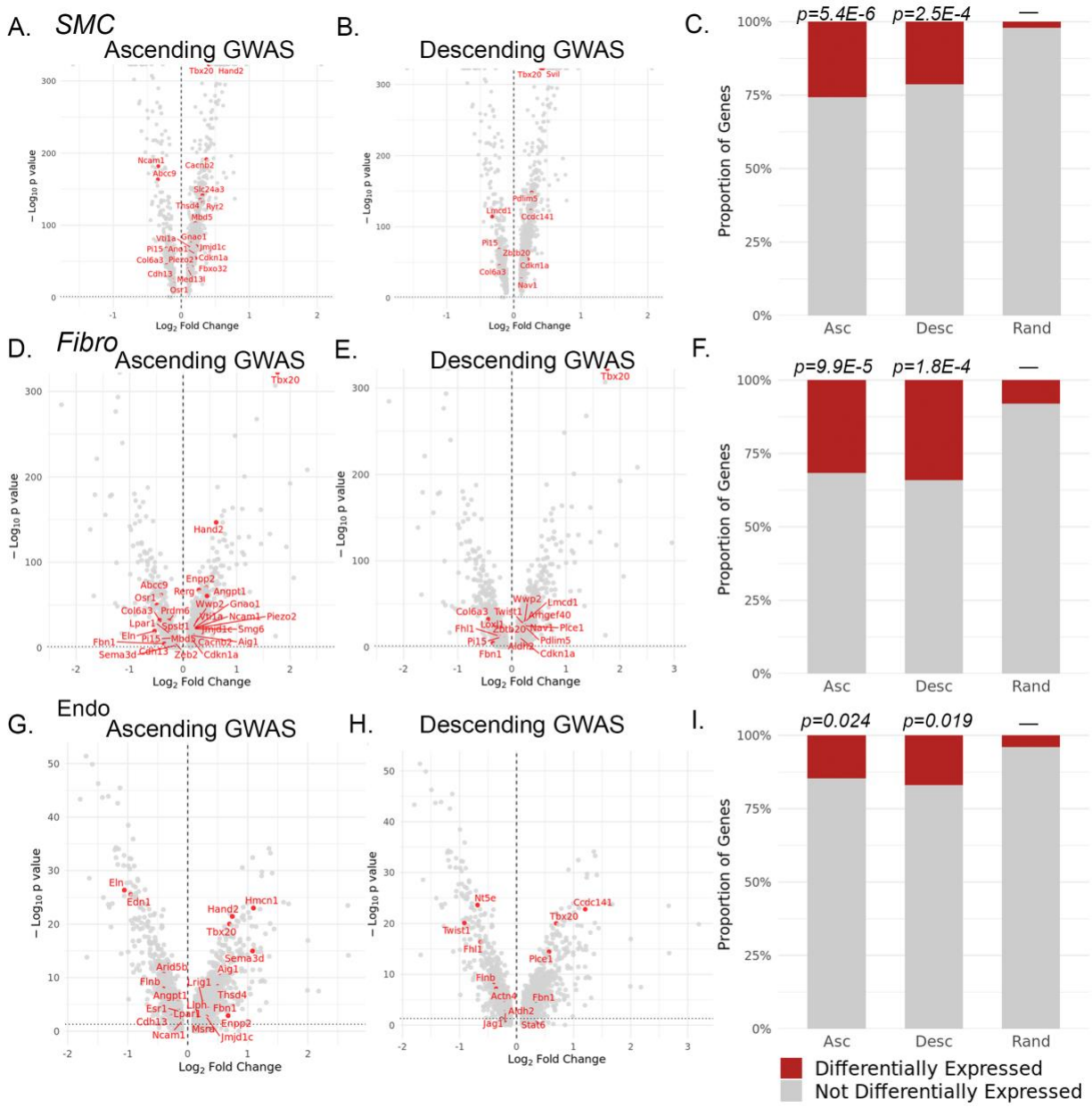
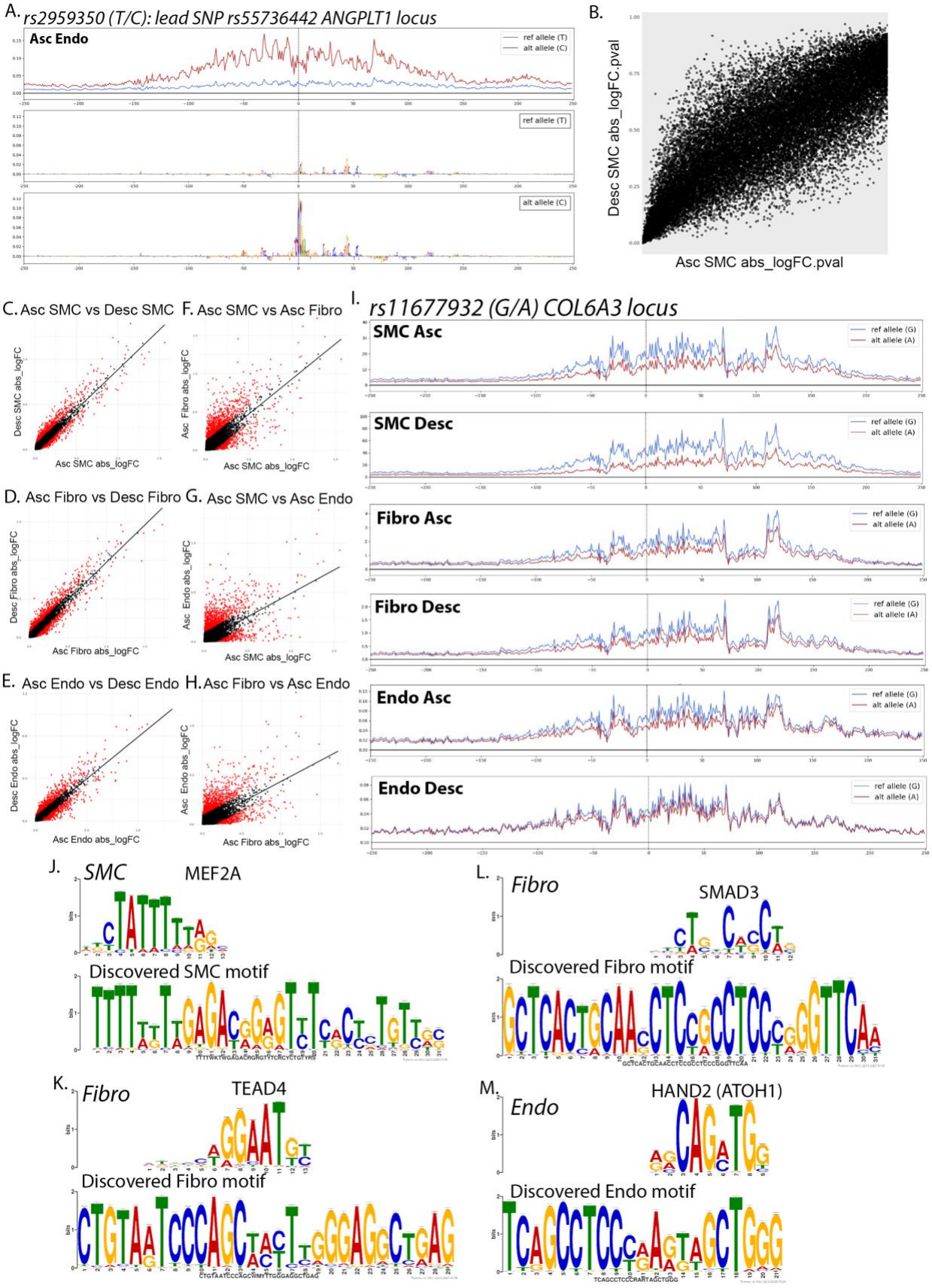


Figure 7. Aortic dimension GWAS genes are differentially expressed in a cell type and vascular site specific context. Volcano plots of ascending versus descending differential RNA expression and highlighting ascending and descending aortic dimension GWAS genes in SMC (A-B), fibroblast (D-E), and endothelial cells (G-H). Stacked barcharts showing the proportion of ascending/descending aortic dimension GWAS genes differentially expressed compared to random gene list in SMCs (C),

1408 fibroblasts (F), and endothelial cells (I). P values represent chi-quared statistical test for

1409 significance compared to random 100 gene list.

1410



1411

Figure 8. ChromBPNet machine learning models predict variant effect on chromatin accessibility to be cell type and vascular site-specific. Example locus of ChromBPNet prediction of chromatin accessibility in reference and alternative alleles for rs2959350 at *ANGPLT3* locus with predicted base pair contributing effects (A). Scatter plot of absolute log fold change (abs_logFC) p values for scored variants in ascending SMC versus descending SMC models (B). Scatter plots of abs_logFC for scored variants in ascending SMC versus descending SMC models (C), ascending fibro versus descending fibro models (D), ascending endothelial versus descending endothelial models (E), ascending SMC versus ascending fibroblast models (F), ascending SMC versus ascending endothelial models (G), and ascending fibroblast versus ascending endothelial models (H). Example locus of ChromBPNet prediction of chromatin accessibility in reference and alternative alleles for rs11677932 at *COL6A3* locus for ascending and descending SMCs, fibroblasts, and endothelial cell models (I). Examples of discovered and matched motifs in SMC – *MEF2A* (J), Fibroblasts – *TEAD4* (K), Fibroblasts – *SMAD3* (L), and Endothelial – *HAND2* (M).