nature cardiovascular research

Article

https://doi.org/10.1038/s44161-025-00692-4

# Region-specific gene expression and sex inform about disease susceptibility in the aorta

Received: 30 September 2021

Accepted: 15 July 2025

Published online: 21 August 2025

Check for updates

Milagros C. Romay[1], Feiyang Ma[1], Ana Mompeón[1], Michele Silvestro[2,3], Gloria E. Hernandez[1,4], Jocelynda Salvador[1], Andrew L. Wang[5], Marie Vandestienne[6], Nathalie Bardin[7], Marcel Blot-Chabaud[7], Aurelie S. Leroyer[7], Hafid Ait-Oufella[6,8], Bhama Ramkhelawon[2,3] & M. Luisa Iruela-Arispe[1] ✉

Pathology in large vessels frequently develops at specific locations, implying that local stressors and spatially restricted gene expression are likely contributors to disease susceptibility. Here we perform single-cell transcriptomics in the carotids, the aortic arch and the thoracic and abdominal aorta to identify site- and sex-specific differences that could inform about vulnerability. Our findings revealed (1) regionally defined transcriptional profiles, (2) signatures associated with embryonic origins and (3) differential contributions of sex-specific effectors. Furthermore, cross-referencing regional-specific signatures with available genome-wide association study and expression quantitative trait loci databases identified 339 disease candidates associated with aorta distensibility, stiffness index and blood pressure. *CPNE8* and *SORBS2* were further evaluated and highlighted as strong causal candidates. Sex differences were predominantly observed in the thoracic and abdominal aorta. *MCAM* (CD146), a transcript with sex-skewed expression and lower in male mice and men, had significantly reduced expression in human aortic aneurysms. The findings reveal underlying diversity within vascular smooth muscle cell populations relevant to understanding site-specific and sex-specific variation of vascular pathologies.

The development of aortic aneurysms is multifactorial, shaped by host-intrinsic features and environmental risk factors that influence disease severity and progression. Efforts to identify genetic determinants have focused on the contributions of vascular smooth muscle cells (vSMCs), as many familial forms of thoracic aortic aneurysms arise from variants in key vSMC contractile proteins such as those encoded by *MYH11*, *ACTA2* and *MYLK*[1]. While several causal genes have been identified, many remain unknown, and, despite numerous genome-wide association studies (GWASs), progress has been limited. Interestingly, the location of aneurysms appears to be associated with distinct etiologies and histories[2,3]. For instance, abdominal aortic aneurysms (AAAs) are more prevalent than nonhereditary thoracic aneurysms (TAAs)[4] despite the overlap of risk factors[5]. Sex also impacts incidence and is a key determinant of severity. Females are relatively protected early in life but may be more susceptible to lethal dissections later[6]. The sex-related factors influencing incidence and severity remain unclear, although studies in individuals with Turner's syndrome (45X,O) support the contribution of sex chromosomes to aneurysm risk[7].

A full list of affiliations appears at the end of the paper. ✉e-mail: arispe@northwestern.edu

Embryologic ancestry also contributes to aortic disease susceptibility[8–11]. Lineage tracing has revealed the diverse embryological origins of vSMCs[12,13], which include the neural crest (NC), the second heart field (SHF) and somites[13,14]. These cells remain spatially segregated in the adult aorta[15], and their developmental origin influences their response to stimuli. In mouse models of Loeys–Dietz syndrome, responses to transforming growth factor (TGF)-β can either activate or repress SMAD2–SMAD3 signaling depending on vSMC developmental origin[16]. These findings suggest that vSMCs retain ancestral memory influencing disease susceptibility and therapeutic responses. Furthermore, vSMCs, like most cells, are endowed with positional identity[17] that is retained in the adult, as per expression of HOX family members[18–20]. Interestingly, altering the topographical expression of HOX leads to changes in gene expression and vascular remodeling. In sum, these results suggest that region-specific features and embryonic ancestry may influence disease susceptibility in addition to sex and hereditary factors.

Here we applied single-cell transcriptomics to profile 3,124 adult vSMCs in the carotids and in three different regions of the aorta: arch, thoracic and abdominal. We also used lineage tracing to elucidate expression patterns associated with embryonic origins. We identified a small (7.2%) subset of transcripts with site-enriched expression. Further curation of these site-enriched genes showed a statistically significant overlay with GWAS loci, linking regional skewed expression to disease susceptibility. We also found sex-specific differences in vSMCs, particularly in the thoracic and abdominal aorta, highlighting disease predilection. These findings reveal diversity in aortic vSMCs relevant to understanding the site- and sex-specific incidence of vascular pathologies.

## Results

### Transcriptional characterization of vSMCs

To identify regional-specific transcripts in vSMC populations, we performed single-cell RNA sequencing (scRNA-seq) in four distinct sites: carotid arteries, the aortic arch, thoracic aorta and abdominal aorta (Fig. 1a,b). The rationale for selecting these sites was the combination of embryonic origin, hemodynamics and disease emergence. In total, we sequenced between 4,930 and 7,929 total cells per site from four males and four females using eight different libraries (Extended Data Fig. 1 and Supplementary Tables 1–3). Importantly, libraries were sequenced together, and Harmony was applied to minimize confounding issues related to comparisons between libraries. We identified 17 distinct cell clusters across the four sites (Extended Data Figs. 1 and 2 and Supplementary Table 2). The molecular identity of the cells residing in the 17 cell clusters was determined by expression of classical vascular cell type markers including endothelial cells (*Pecam1*), fibroblasts (*Pdgfra*) and vSMCs (*Myh11*). This protocol permitted resolution of seven cell types, data consistent with previous publications[21] (Fig. 1c and Extended Data Figs. 1h–j).

Following identification, vSMCs were computationally extracted from the other cell types for further analysis (Fig. 1d). Visualization of vSMCs using uniform manifold approximation and projection (UMAP) did not reveal strong intra-cell type diversity of vSMCs per anatomical region, consistent with their core primary identity despite distinct embryological origins (Fig. 1d and Supplementary Tables 1 and 3). From the vSMC population, 18,267 transcripts were analyzed for site-enriched expression, identifying 1,315 transcripts (7.2%) with location-enriched expression (adjusted $P$ value ($P_{adj}$) ≤ 0.05) (Fig. 1e and Supplementary Table 1). A heatmap of the top five location-enriched transcripts revealed the stark contrast between locations (Fig. 1f). Note that the carotids showed stronger similarity to the aortic arch, despite their different flow patterns, and more consistent with their common embryological origin. Importantly, while these anatomically defined vSMC populations contained regional-enriched expression, their core identity, as per the expression of classical vSMC markers (*Myh11*, *Acta2* and *Tagln*), was unchanged (Fig. 1f). Analysis of HOX gene clusters in vSMCs showed significant association between expression and anatomic location (Fig. 1g). These findings are consistent with previous reports showing maintained expression of HOX genes in vascular cells and other tissues dependent on anatomic location[19].

### Unique carotid vSMC signature

The carotids represent the major branch point from the aortic arch responsible for supplying blood to the brain (Fig. 2a,b), and their tunica media is populated by vSMCs derived from the NC[12]. A total of 405 transcripts (2.2% of total transcripts) distinguishes carotid arteries (carotid vSMCs; caSMCs) from their counterpart vSMCs in the aortic arch and the thoracic and abdominal aorta (Fig. 1e and Supplementary Tables 1 and 3). Gene ontology analysis of the top 50 carotid signature genes showed a significant enrichment for proteins involved with receptor serine–threonine kinases, NABA matrisome-associated proteins and WNT signaling (Fig. 2c and Supplementary Table 4). From these pathways, candidate gene validation using immunohistochemistry was performed for the genes *Rspo3 and Wif1*.

*Rspo3* is a gene that encodes a member of the R-spondin family of proteins and ranks within the top 25 most-enriched genes in the caSMC signature (Fig. 2d). Immunofluorescence (IF) staining of RSPO3 showed significantly increased expression in all layers of the carotid's tunica media relative to the other arterial beds (Fig. 2e,f). *Wif1*, a gene encoding a WNT signaling regulatory factor, ranked sixth within all caSMC signature genes (Fig. 2g). WIF1 IF staining displayed significantly increased WIF1 expression in all layers of carotid vascular smooth muscle relative to other arterial beds (Fig. 2h,i).

### Unique aortic arch vSMC signature

The collected aortic arch region extended from the aortic root to the left subclavian branch, in concordance with the limits of NC origin[5]. A total of 192 genes (1% of all transcripts) distinguished aortic arch smooth muscle cells (SMCs) from their counterparts in the thoraco-abdominal aorta and carotid arteries (Figs. 1e and 2j,k and Supplementary Tables 1 and 3). Gene ontology analysis for this region showed significant enrichment for transcripts involved in heart development, blood circulation and striated muscle tissue development (Fig. 2l and Supplementary Table 4). From these identified signature genes, validation was performed for aggrecan (*Acan*) and angiotensin-converting enzyme I (*Ace*). ACAN IF showed significant expression in all layers of aortic arch vascular smooth muscle relative to smooth muscle from the carotids and the thoracic and abdominal aorta (Fig. 2m–o). *Ace*, encoding the enzyme responsible for the conversion of angiotensin I to vasoconstricting angiotensin II (AngII), was also identified as significantly enriched in aortic arch vSMCs (Fig. 2p–r).

### Contribution of developmental ancestry to the adult aortic arch signature

In addition to the aortic arch's unique biomechanical properties, the vSMCs that comprise the arch emerge from two distinct developmental origins: the NC and the SHF, which is predicted to contribute to the emergence of site-specific disease in this region[10]. Lineage tracing in combination with single-cell sequencing was performed to determine whether distinct developmental origins impact adult expression patterns. Thus, the *Mef2c*$^{Cre}$ and *Wnt1*$^{Cre}$ mouse lines, in combination with reporter genes (enhanced green fluorescent protein (eGFP) and tdTomato), were used to lineage trace adult descendants of SHF and NC cells, respectively (Fig. 3a). The efficiency of the lineage-tracing strategy is shown in Fig. 3b, where cells derived from the NC are shown in green by virtue of *Wnt1*$^{Cre}$-mediated expression of eGFP. Here, the *Wnt1*$^{Cre}$ line was crossed with mT/mG mice to visualize recombined and unrecombined cells (Fig. 3b). Importantly, while the reporter was excised during development, adult cells remained in place with little migration or intermixing of vSMCs from other cell populations. Interdigitation

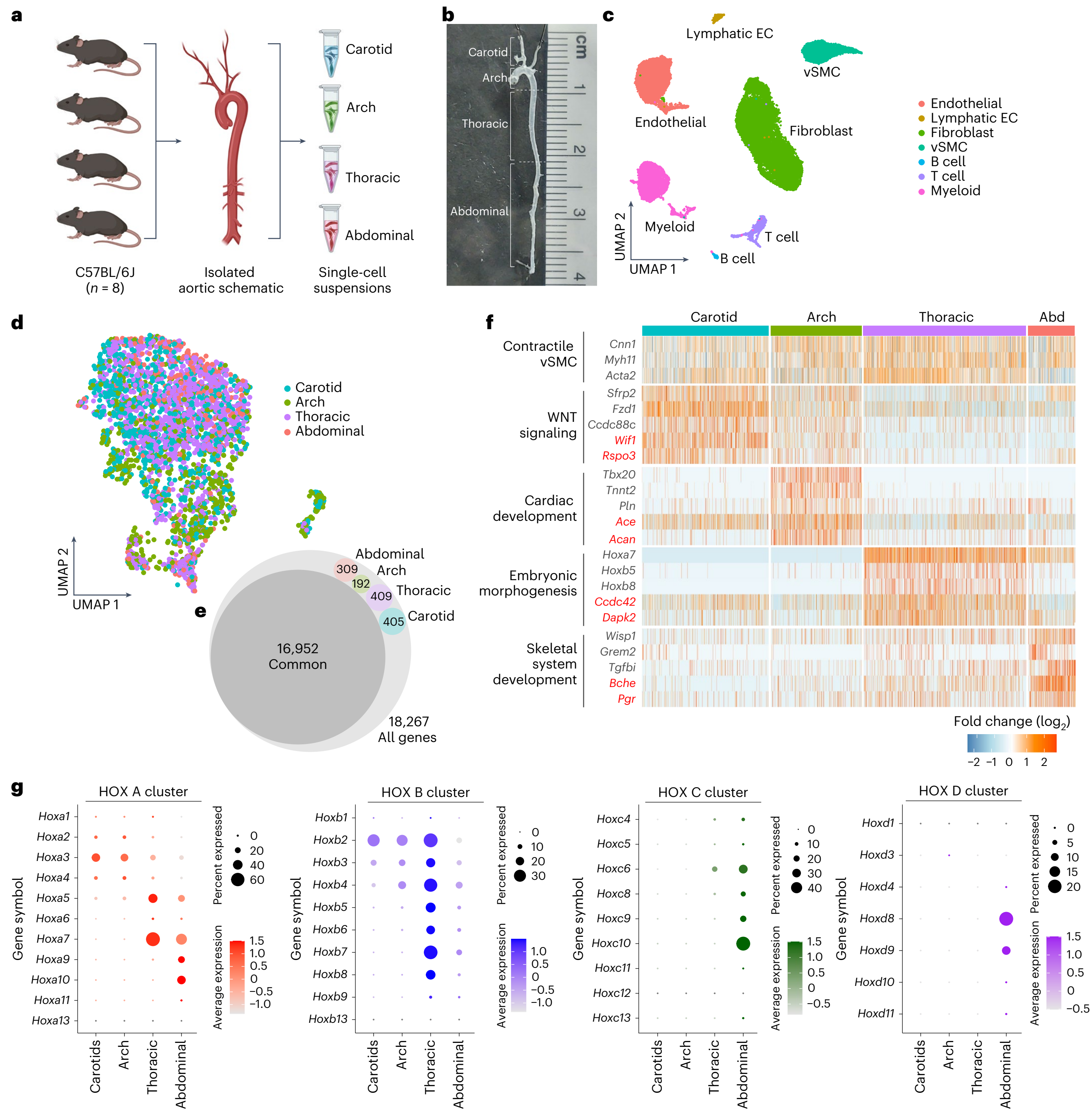



**Fig. 1 | Region-specific vSMC transcriptomics. a**, Experimental design. Twelve-week-old C57BL/6J mice (*n* = 8; four males and four females) were sacrificed, and the aorta and carotid arteries were dissected based on anatomical landmarks and digested into four single-cell preparations per sex for generation of eight scRNA-seq libraries. **b**, Representative mouse aorta used in single-cell preparations with brackets indicating the four anatomic sites used for scRNA-seq. **c**, UMAP of the cell types identified in the eight libraries combined. EC, endothelial cell. **d**, UMAP of vSMCs from the four anatomic locations. **e**, Circular packing diagram showing the relative contribution of site-enriched vSMC signature genes to the total vSMC transcriptome. **f**, Heatmap expression of classical vSMC cell type markers relative to regional signature genes (top five) and their related gene ontology categories. Markers in red indicate transcripts validated by immunohistochemistry (shown later). Abd, abdominal; *Wisp1*, *Ccn4* (gene symbol). **g**, Dot plot visualization of the expression of the four HOX clusters in vSMCs by anatomic location. Panel **a** created using BioRender.com.

between the regions of the ascending aorta, the subclavian and the descending aorta is shown on the right (Fig. 3b). To minimize the risk of cell death, we avoided flow cytometry sorting and instead identified eGFP- and tdTomato-expressing cells bioinformatically from scRNA-seq analysis of these tissues. Despite distinct levels of eGFP and TdTomato (Fig. 3c,d) both populations expressed equivalent and high levels of the vSMC markers *Myh11* and *Mylk* (Fig. 3e,f). A combined UMAP of NC-derived (NC-vSMC) and SHF-derived (SHF-vSMC) cells shows their respective distributions (Fig. 3g). Expression of the top 50 differentially expressed genes between NC-vSMC and SHF-vSMC populations highlighted unique features (Fig. 3h). Gene ontology of NC-vSMC and SHF-vSMC genes with the adult aortic arch showed contributions to classical cardiac-associated genes, including *Tnnt2* (cardiac troponin) and *Tbx20* in the SHF population with the heart development gene ontology category; whereas NC-enriched *Fgf2* was associated with vascular development (Fig. 3i). Importantly, violin plots from a subset of top differentially expressed genes showed that memory from past developmental origin clearly impacts gene expression in the adult (Fig. 3j–o).

### Unique thoracic aorta vSMC signature

SMCs of the thoracic aorta are derived from somites[13]. In total, 409 genes (2.2% of total) compose the transcriptional signature for thoracic aorta vSMCs (taSMCs) (Figs. 1e and 4a,b and Supplementary Tables 1 and 3). Gene ontology on the top 50 genes of the thoracic aorta signature identified significant enrichment for genes involved in hematopoiesis, response to amphetamine and the DNA damage response (Fig. 4c and Supplementary Table 4). Two genes were selected to validate using immunohistochemistry: *Ccdc42* and *Dapk2*. *Ccdc42*, encoding a protein associated with centrosome assembly, was in the top ten enriched taSMC genes[22] (Fig. 4d). IF localization of CCDC42 identified preferential expression in the thoracic aorta, relative to the aortic arch, the abdominal aorta and carotid arteries (Fig. 4e,f). *Dapk2*, encoding a calcium–calmodulin-dependent kinase associated with cell death, was in the top 20 taSMC candidate genes (Fig. 4g). IF identification of DAPK2 showed predominant expression in the thoracic aorta relative to the other three arterial beds (Fig. 4h,i).

### Unique abdominal aorta vSMC signature

SMCs of the abdominal aorta extending from the diaphragm to the iliac bifurcation are also derived from somites[13]. In total, 309 genes (1.7% of total)) composed the molecular signature for abdominal vSMCs (Figs. 1e and 4j,k and Supplementary Tables 1 and 3). Gene ontology on the top 50 genes preferentially expressed by abdominal SMCs identified significant enrichment for genes involved in skeletal system development, urogenital system development and response to hormones (Fig. 4l and Supplementary Table 4). Two genes were selected to validate the signature using immunohistochemistry: *Pgr* and *Bche*. Progesterone receptor (PGR), a member of the steroid receptor family, which mediates the action of progesterone in target tissues, was significantly enriched in abdominal SMCs (Fig. 4m). PGR IF showed that the protein was preferentially localized in the nuclei of SMCs of the abdominal aorta, relative to the aortic arch, the abdominal aorta and carotid arteries (Fig. 4n,o). Butyrylcholinesterase (encoded by *Bche*) serves as one of two enzymes that metabolize acetylcholine, the primary neurotransmitter of the parasympathetic nervous system that plays a key role in muscle relaxation (Fig. 4p). BCHE IF showed predominant expression in vSMCs proximal to the lumen of the abdominal aorta relative to the other three arterial beds (Fig. 4q,r).

### Site-enriched signatures are associated with site-specific vascular pathologies

The identification of a cohort of vSMC genes enriched at specific vascular locations (7.2% of total vSMC expression) suggested potential relevance to site-specific pathologies. To test this prediction, we performed enrichment analysis with known GWAS candidate genes. In total, 3,135 genes representing significant and suggestive GWAS associations in cardiovascular (CV) traits that had known mouse orthologs were identified through manual curation (Fig. 5a). The relative ratio of GWAS candidate genes in site-enriched gene signatures was then compared to the ratio of GWAS candidate genes in the ubiquitous vSMC transcriptome using Fisher's exact test. Site-enriched signature genes were significantly more likely to be CV-associated GWAS candidate genes, with 339 genes, representing 30–52% relative to the total cohort of site-enriched vSMC genes per anatomic site (Fig. 5b and Supplementary Table 5).

Most genome-wide association risk loci fall in noncoding regions of the genome, complicating the inference of the molecular pathway from DNA variation to disease-associated phenotypes. One approach for prioritizing GWAS candidate genes involves using expression quantitative trait locus (eQTL) mapping that links DNA variation to a molecular outcome such as messenger RNA expression. To identify site-enriched GWAS candidate genes with evidence of molecular regulation by genetic variation in humans, we turned to the Genotype–Tissue Expression (GTEx) project. GTEx is a publicly available database of gene expression and matched genotype information from multiple tissues. We chose to focus our analysis solely on the site-enriched genes from the mouse aortic arch, as samples collected in GTEx are restricted to the human aortic arch (GTEx does not collect from the thoracic–abdominal aorta).

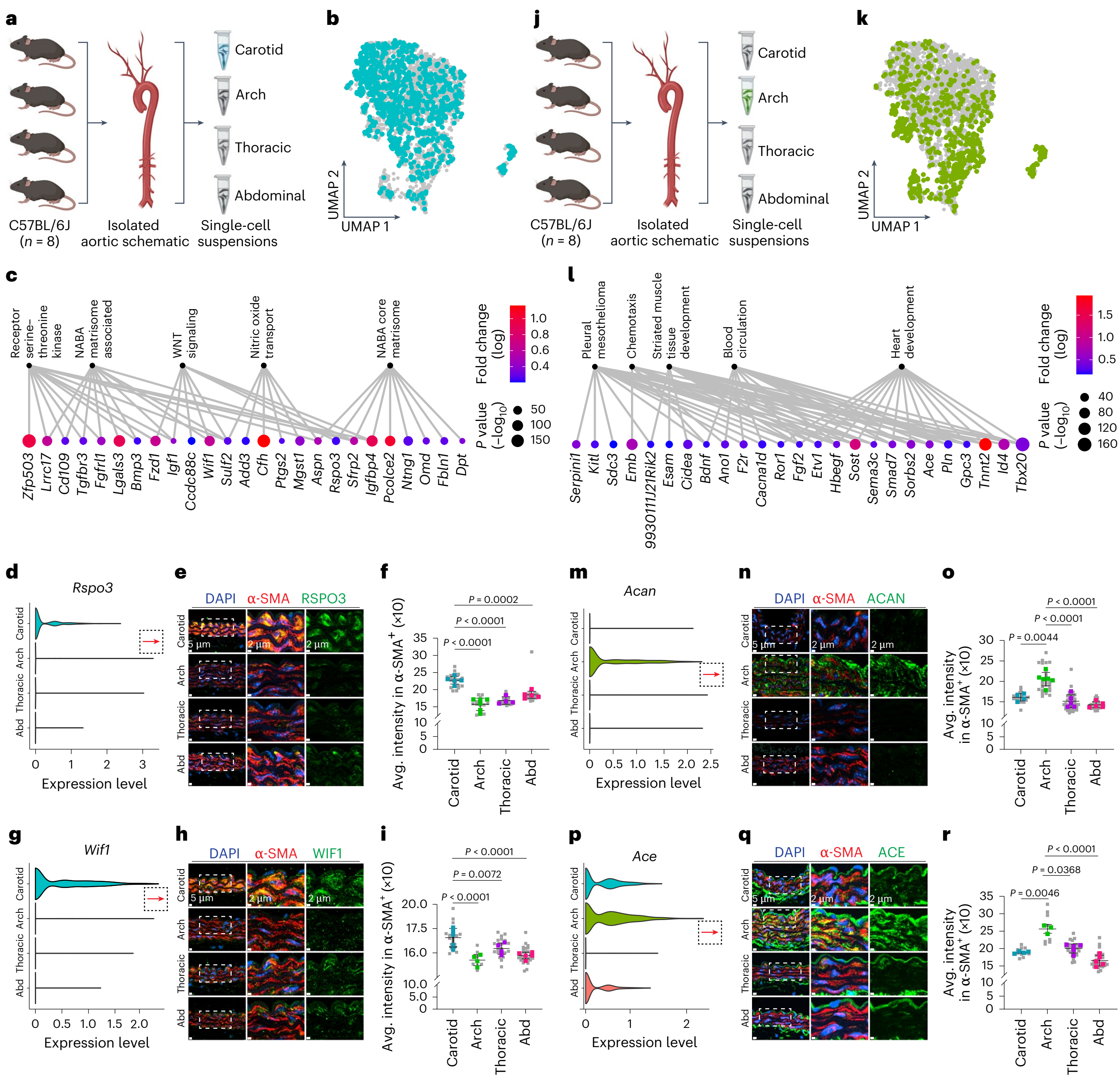
The relative ratio of GWAS candidate genes with eQTLs in the aortic arch gene signature was then compared to the ratio of GWAS

**Fig. 2 | Transcriptional specificity of carotid and arch vSMCs.** **a**, Schematic of experimental design. **b**, UMAP of vSMCs highlighting the caSMC population. **c**, Visualization of the top five gene ontology categories from the top 50 unique caSMC-expressed genes calculated via hypergeometric distribution. *P* values for enrichment categories are as follows: NABA core matrisome, $3.13 \times 10^{-10}$; nitric oxide transport, $3.31 \times 10^{-8}$; WNT signaling, $3.93 \times 10^{-8}$; NABA matrisome associated, $1.69 \times 10^{-6}$; receptor serine–threonine kinase, $3.13 \times 10^{-10}$. **d**, Violin plot visualizing the expression pattern of *Rspo3*. **e**, Representative IF images of RSPO3 expression (green) in the anatomic sites. DAPI, 4,6-diamidino-2-phenylindole. **f**, Quantification of RSPO3 expression across sites: carotid, $n = 21$ images, 5 animals; arch, $n = 12$ images, 6 animals; thoracic, $n = 21$ images, 4 animals; abdominal, $n = 21$ images, 5 animals; 84 images in total. Kruskal–Wallis test, *P* values: $P < 0.0001$, carotid versus arch; $P < 0.0001$, carotid versus thoracic; $P = 0.0002$, carotid versus abdominal. Avg., average. **g**, Violin plot visualizing the expression pattern of *Wif1*. **h**, Representative IF images of WIF1 expression (green) in the four anatomic sites. **i**, Quantification of WIF1 expression in the four anatomic sites. Carotid, $n = 20$ images, 5 animals; arch, $n = 8$ images, 4 animals; thoracic, $n = 22$ images, 4 animals; abdominal, $n = 30$ images, 5 animals; 80 images in total. Brown–Forsythe and Welch's ANOVA, *P* values: $P < 0.0001$, carotid versus arch; $P = 0.0072$, carotid versus thoracic; $P < 0.0001$, carotid versus abdominal. **j**, Schematic of experimental design highlighting the arch-derived vSMC population. **k**, UMAP visualization of vSMCs from the four anatomic locations highlighting the distribution of the arch-derived vSMCs. **l**, Visualization of the top five gene ontology categories from the top 50 unique arch vSMC-expressed genes calculated via hypergeometric distribution. *P* values for enrichment categories are as follows: heart development, $2.21 \times 10^{-8}$; blood circulation, $2.31 \times 10^{-7}$; striated muscle tissue development, $7.50 \times 10^{-7}$; chemotaxis, $1.82 \times 10^{-6}$; pleural mesothelioma, $7.78 \times 10^{-8}$. **m**, Violin plot visualizing the expression pattern of *Acan*. **n**, Representative IF images of ACAN expression (green) in the four anatomic sites. **o**, Quantification of ACAN expression in the four anatomic sites. Carotid, $n = 21$ images, 4 animals; arch, $n = 26$ images, 6 animals; thoracic, $n = 43$ images, 5 animals; abdominal, $n = 44$ images, 5 animals; 134 images in total. Kruskal–Wallis test, *P* values: $P = 0.0044$, arch versus carotid; $P < 0.0001$, arch versus thoracic; $P < 0.0001$, arch versus abdominal. **p**, Violin plot visualizing the expression pattern of *Ace*. **q**, Representative IF images of ACE expression (green) in the four anatomic sites. **r**, Quantification of ACE expression in the four anatomic sites. Carotid, $n = 12$ images, 4 animals; arch, $n = 6$ images, 3 animals; thoracic, $n = 30$ images, 5 animals; abdominal, $n = 27$ images, 5 animals; 75 images in total. Kruskal–Wallis test, *P* values: $P = 0.0046$, arch versus carotid; $P = 0.0368$, arch versus thoracic; $P < 0.0001$, arch versus abdominal. For **c**,**l**, color indicates the degree of enrichment (log) of expression, while circle size indicates the *P* value ($-\log_{10}$). For **e**,**h**,**n**,**q**, the boxed red arrow indicates the anatomic site of interest. For **f**,**i**,**o**,**r**, $n = 8$–84 images per region (gray), 4–6 animals per location (colored). Error bars represent mean ± s.d. Two-sided Kruskal–Wallis test (*Rspo3*, *Acan* and *Ace*) and two-sided Brown–Forsythe and Welch's ANOVA (*Wif1*) with multiple-testing comparison. Panels **a** and **j** created using BioRender.com.

https://doi.org/10.1038/s44161-025-00692-4

candidate genes with eQTLs in the ubiquitous vSMC transcriptome (Fig. 5c and Supplementary Table 6). In fact, from the 192 arch-enriched transcripts (Figs. 1e and 2j–r), 109 of those genes were associated with human genetic variation, either by CV GWAS, eQTL mapping or both (Supplementary Table 6). While most of the GWAS–eQTL pairs were associated with blood pressure (including *ACE* and *SORBS2*), several were assigned to aneurysm and/or structural traits including aortic diameter (*THSD4*), distensibility (*CHSY1*) and aortic stiffness (*ACTR2*, *CPNE8*, *GPC6*, *MECOM*) (Supplementary Tables 5 and 6). To further test the hypothesis that some of these genes were indeed within the GWAS loci, we proceeded to evaluate two candidates: *Cpne8* and *Sorbs2*.

*Cpne8* encodes copine 8, a member of a family of calcium-dependent membrane proteins found in a variety of eukaryotes. *Cpne8* is associated with numerous CV traits, including arterial stiffness and heart rate[23,24]. *Cpne8* expression in vSMCs is highly enriched for aortic arch vSMCs (Fig. 5d). Query of aortic eQTLs in GTEx identified 174 eQTLs (peak eQTL *P* value = $1.47 \times 10^{-15}$) residing within a 1-Mb window of the gene, including rs1486346, which was located 2,274 bp away from the peak arterial stiffness association at rs7979541 (Fig. 5e). Another gene, *Sorbs2*, also known as ArgBP2, encoding a protein located in the sarcomeric Z-disk, was associated with multiple blood pressure traits across multiple ethnic groups[25]. Eighty-nine significant aortic eQTLs (peak eQTL *P* value = $1.47 \times 10^{-17}$) resided within a 500-kb window around the *Sorbs2* gene and within proximity of the peak blood pressure association rs75305034 (Fig. 5f,g). Furthermore, 48.6% of GTEx-identified aortic arch eQTLs (53 of 109, peak SNP *P* value ≤ $1 \times 10^{-5}$) showed replication in the Stockholm–Tartu Atherosclerosis Reverse Networks Engineering Task (STARNET), a secondary database of human aortic root-derived eQTLs, further supporting human relevance of the arch-enriched transcripts identified by this work across multiple eQTL datasets (Fig. 5g).



https://doi.org/10.1038/s44161-025-00692-4

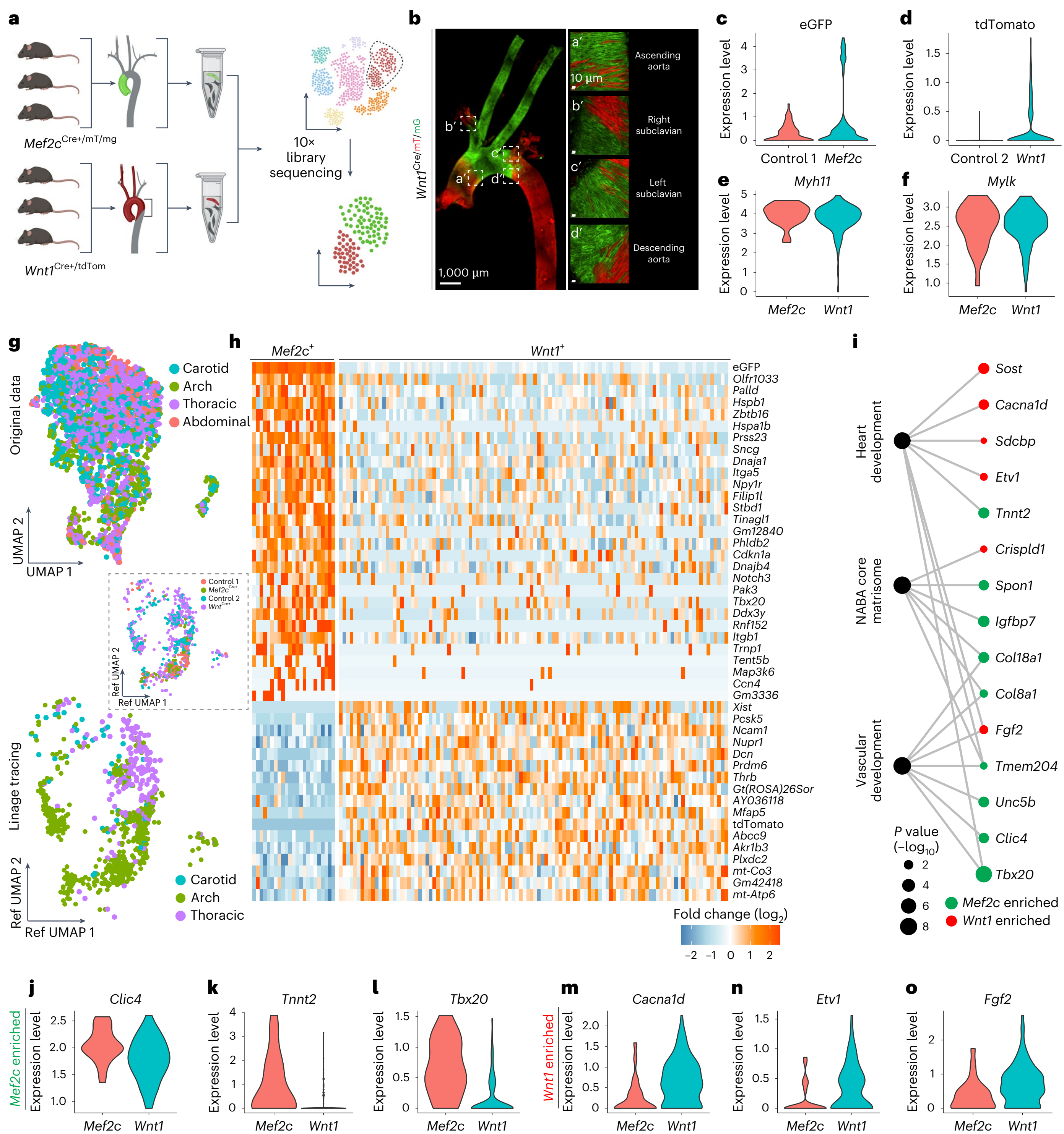



**Fig. 3 | Lineage tracing of the SHF and NC aortic arch highlights respective contributions to the arch regional signature.** **a**, Experimental design. Single-cell suspensions were generated from isolated aortic arches of mice expressing eGFP after *Mef2c*-Cre recombination (SHF-derived cells) or tdTomato after recombination with *Wnt1*-Cre (NC derived). Lineage-traced cells were bioinformatically segregated from the data using transcript expression of the indicated fluorescent proteins before downstream analysis. **b**, Representative IF images of NC-vSMCs (green) after *Wnt1*-Cre-mediated recombination visualized in the adult mouse aorta, $n = 6$. **c**–**f**, Violin plots for expression of eGFP (**c**), tdTomato (**d**) and vSMC contractile markers (*Myh11* (**e**), Mylk (**f**)) in the *Mef2c* versus *Wnt1* datasets. **g**, UMAP plot visualizing the distribution of $Mef2c^{+}$ and $Wnt1^{+}$ lineage-tracing experiments and their respective controls. Ref, reference. **h**, Heatmap of the top 45 genes differentially expressed between $eGFP^{+}$ *Mef2c*-derived vSMCs and $tdTomato^{+}$ *Wnt1*-derived vSMCs. *Gm42418*, *Rn18s-rs5* (gene symbol); *Olfr1033*, *Or5m3b* (gene symbol). **i**, Gene ontology enrichment for arch-derived vSMCs highlighting signature genes from each of the two developmental origins calculated via hypergeometric distribution. *P* values for enrichment categories are as follows: vascular development, $3.88 \times 10^{-9}$; NABA core matrisome, $5.53 \times 10^{-9}$; heart development, $2.63 \times 10^{-8}$. **j**–**o**, Violin plots visualizing the expression of lineage-derived vSMC markers (*Clic4* (**j**), *Tnnt2* (**k**), *Tbx20* (**l**), *Cacna1d* (**m**), *Etv1* (**n**), *Fgf2* (**o**)) in aortic arch $Mef2c^{+}$ and $Wnt1^{+}$ vSMCs. Panel **a** created using BioRender.com.

Given the complex developmental origin of aortic arch vSMCs, GWAS candidates identified in this region were also overlaid with NC and SHF lineage-tracing data. Twenty-eight percent of the lineage-associated arch vSMC regional candidates were also found in GWASs for CV-related traits, including ascending aortic diameter, a well-known risk factor for aneurysm development (Supplementary Tables 5–7).

We also found that the arch vSMC genes *Tbx20* and *Crispld1* were candidates for ascending aortic diameter across several independent GWASs[26,27]. *Tbx20* is a known SHF vSMC gene encoding a protein acting as a regulator of outflow tract development; by contrast, very little is known about NC-vSMC-associated *Crispld1*. Neither gene showed significant evidence of eQTL in aortic-derived data from GTEx, which takes samples from the descending aortic arch, a region that highlights the site where NC-vSMCs and mesoderm-derived thoracic vSMCs merge. By contrast, STARNET samples are mostly from the aortic root. Tapping into STARNET eQTL data, we found that both *Tbx20* and *Crispld1* contained multiple significant eQTLs ($P_{adj} < 3.0 \times 10^{-8}$) for CV pathology. This apparent discrepancy between the two eQTL datasets emphasizes that, even within a small region (aortic arch), heterogeneity in vSMCs can have a profound impact on the identification of disease-associated genes. Collectively, our findings strongly support the notion that site-specific expression could be used to mine GWAS and eQTL data to uncover unknown disease-associated genes.

## Effect of sex on the vSMC transcriptome

Many CV diseases including AAA show significant sex-related differences in prevalence and progression. In AAA, disease burden is higher in males and manifests at a younger age; whereas onset of AAA is delayed in females, but the outcome of disease is poorer with a higher incidence of lethal events[28]. To better understand the role of biological sex in gene expression profiles of vSMCs, we performed a sex-matched analysis of four arterial locations (Fig. 6a,b). Hierarchical clustering of cells showed that the arch and carotids segregated together according to site; however, thoracic and abdominal vSMCs segregated by sex. Curiously, the thoracic and the abdominal regions share the same embryonic origin (somites), unlike the arch and carotids, further supporting the role of ancestry in expression patterns. Thoracic and abdominal vSMCs showed a 2.95-to-10.4-fold increase in the number of sex-associated genes relative to carotid and aortic arch vSMCs (Fig. 6c,d, Extended Data Fig. 3 and Supplementary Table 8).

In the thoracic region, female-enriched genes were found to be associated with cell polarity, axon guidance and the electron transport chain (Fig. 6e). Male taSMCs showed enrichment for proteins associated with extracellular matrix organization and peptide metabolic processes (Fig. 6e). IF was performed for a subset of sex-enriched regional genes to confirm that the observed transcriptional differences represented functional changes at the protein level. Biglycan (*Bgn*), a known contributor to TAA, resides on the X chromosome but has been reported to be positively regulated by the presence of the Y chromosome in male cells[29]. We found that *Bgn* is one of the top differentially expressed genes in male thoracic vSMCs (Fig. 6f). IF staining for BGN showed predominant expression in all vSMC layers of the thoracic aorta in males relative to females, suggesting that males produced more BGN protein as a functional consequence of the transcriptional upregulation (Fig. 6g,h). Expression of periostin (*Postn*), encoding an extracellular matrix protein associated with tissue reorganization and wound healing, was significantly increased in female taSMCs relative to males (Fig. 6i). Staining of thoracic aortic sections showed significant increases in protein expression in all vSMC layers of the thoracic aorta in females relative to males, consistent with transcriptional differences observed in the scRNA-seq data (Fig. 6j,k).

In the abdominal aorta, female abdominal SMCs were found to be enriched for transcripts associated with modulation of the host immune response, ATP biosynthetic processes and diabetic cardiomyopathy (Fig. 6l). By contrast, male abdominal vSMCs were found to express transcriptional profiles associated with matrix proteins, muscle contraction and response to growth factors (Fig. 6l). The transcript for the gene *Tgfbi* was significantly enriched in male abdominal SMCs (Fig. 6m). which was confirmed by TGFBI protein expression (Fig. 6n,o). By contrast, *Mcam* (also known as CD146 and MUC18) transcript expression was significantly enriched in female abdominal SMCs relative to those of males, residing within the top 25 female abdominal-enriched genes (Fig. 6p and Extended Data Fig. 4a), a finding also confirmed by immunohistochemistry (Fig. 6q,r).

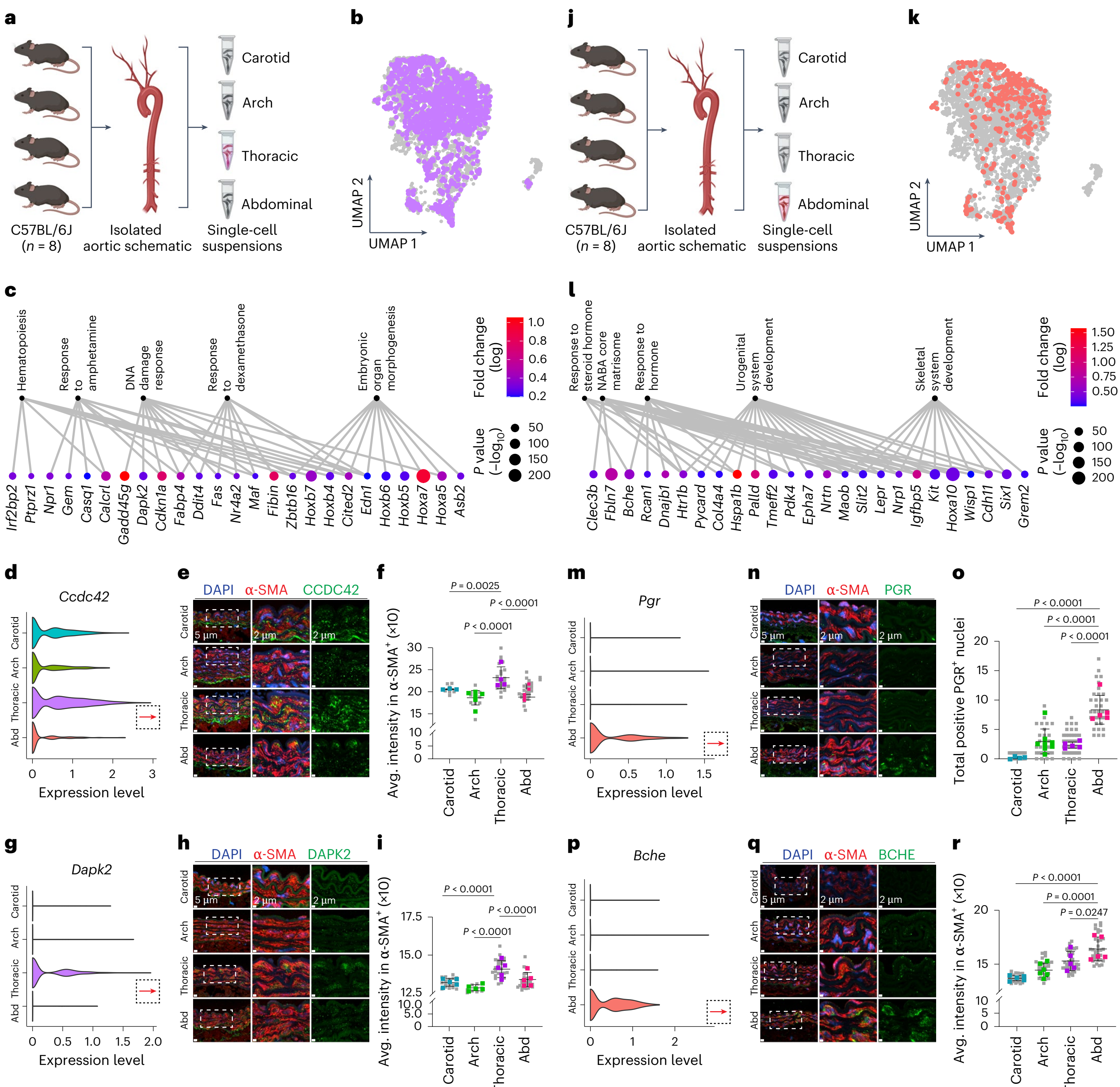
## Elucidating sex-specific molecular processes altered by MCAM expression

In addition to sex enrichment, further examination of MCAM protein expression in the mouse aorta showed a peculiar distribution. Specifically, the ventral region of the aorta displayed higher levels of MCAM than the dorsal region (Fig. 7a). Given these findings, we hypothesized that MCAM may serve as a protective factor in aortic vSMCs with sex-specific molecular interactions. We first confirmed that sex enrichment of MCAM was also present in human vSMCs. Western

**Fig. 4 | Transcriptional specificity of thoracic and abdominal vSMCs.** **a**, Schematic of experimental design. **b**, UMAP of vSMCs highlighting the distribution of thoracic-derived cells. **c**, Visualization of the top five gene ontology categories from the top 50 unique thoracic vSMC-expressed genes calculated via hypergeometric distribution. *P* values for enrichment categories are as follows: embryonic organ morphogenesis, $3.64 \times 10^{-11}$; response to dexamethasone, $7.23 \times 10^{-7}$; DNA damage response, $4.63 \times 10^{-6}$; response to amphetamine, $1.58 \times 10^{-5}$; hematopoiesis, $8.09 \times 10^{-5}$. **d**, Violin plot visualizing the expression pattern of *Ccdc42*. **e**, Representative IF images of CCDC42 expression (green) per site. **f**, Quantification of CCDC42 expression. Carotid, $n = 7$ images, 3 animals; arch, $n = 12$ images, 6 animals; thoracic, $n = 24$ images, 4 animals; abdominal, $n = 27$ images, 5 animals; 73 images in total. Brown–Forsythe and Welch's ANOVA, *P* values: $P = 0.0025$, thoracic versus carotid; $P < 0.0001$, thoracic versus arch; $P < 0.0001$, thoracic versus abdominal. **g**, Violin plot of *Dapk2* transcript expression. **h**, Representative IF images of DAPK2 expression (green) in the four anatomic sites. **i**, Quantification of DAPK2 expression in the four anatomic sites. Carotid, $n = 18$ images, 5 animals; arch, $n = 12$ images, 6 animals; thoracic, $n = 26$ images, 5 animals; abdominal, $n = 28$ images, 5 animals; 82 images in total. Kruskal–Wallis test, *P* values: $P < 0.0001$, thoracic versus carotid; $P < 0.0001$, thoracic versus arch; $P < 0.0001$, thoracic versus abdominal. **j**, Schematic of experimental design highlighting the abdominal-derived vSMC population. **k**, UMAP of abdominal-derived vSMCs within the total cluster of vSMCs. **l**, Visualization of the top five gene ontology categories from the top 50 unique abdominal vSMC-expressed genes calculated via hypergeometric distribution. *P* values for enrichment categories are as follows: skeletal system development, $6.09 \times 10^{-10}$; urogenital system development, $9.53 \times 10^{-8}$; response to hormone, $4.21 \times 10^{-7}$; NABA core matrisome, $6.17 \times 10^{-6}$; response to steroid hormone, $1.05 \times 10^{-4}$. **m**, Violin plot of *Pgr* transcript expression. **n**, Representative IF images of PGR expression (green) in the four anatomic sites. **o**, Quantification of PGR expression in the four anatomic sites. Carotid, $n = 22$ images, 4 animals; arch, $n = 30$ images, 8 animals; thoracic, $n = 46$ images, 5 animals; abdominal, $n = 36$ images, 5 animals; 134 images in total. Kruskal–Wallis test, *P* values: $P < 0.0001$, abdominal versus carotid; $P < 0.0001$, abdominal versus arch; $P < 0.0001$, abdominal versus thoracic. **p**, Violin plot of *Bche* transcript expression. **q**, Representative IF images of BCHE expression (green) in the four anatomic sites. **r**, Quantification of BCHE expression in the four anatomic sites. Carotid, $n = 25$ images, 5 animals; arch, $n = 26$ images, 6 animals; thoracic, $n = 42$ images, 5 animals; abdominal, $n = 34$ images, 5 animals; 127 images in total. Kruskal–Wallis test, *P* values: $P < 0.0001$, abdominal versus carotid; $P = 0.0001$, abdominal versus arch; $P = 0.0247$, abdominal versus thoracic. For **c**,**l**, color indicates the degree of enrichment (log) of expression, while circle size indicates the *P* value ($-\log_{10}$). For **e**,**h**,**n**,**q**, the boxed red arrow indicates the anatomic site of interest. For **f**,**i**,**o**,**r**, error bars represent mean ± s.d. Two-sided Brown–Forsythe and Welch's ANOVA (*Ccdc42*) and two-sided Kruskal–Wallis test (*Dapk2*, *Pgr* and *Bche*) with multiple-testing comparison. Panels **a** and **j** created using BioRender.com.

blots of MCAM protein expression showed higher expression levels in female human aortic vSMCs than in males (Fig. 7b,c). Reduction of MCAM expression by small interfering RNA (siRNA) in a cohort of six male and female human aortic vSMC lines identified key molecular pathways regulated by MCAM expression (Fig. 7d–f and Supplementary Table 9). Decreased MCAM expression, representing a 90–95% knockdown (Fig. 7d,e) contributed to significant alterations to the vSMC transcriptome with 8,993 (male) to 9,229 (female) genes with expression altered by reduced MCAM expression ($P_{adj} \leq 0.05$; Fig. 7f and Supplementary Table 9). Due to the substantial number of genes in these datasets, we chose to focus on the top 3,000 differentially expressed genes as ranked by $P_{adj}$ ($P_{adj} \leq 7.92 \times 10^{-18}$ (sex invariant, combined), $P_{adj} \leq 3.03 \times 10^{-10}$ (female donors) and $P_{adj} \leq 1.01 \times 10^{-9}$ (male donors)). Comparison of the combined data (representing the six independent vSMC donors) with each sex-segregated cohort showed 70% overlap between both sexes (Fig. 7f). Gene ontology analysis of the differentially expressed transcripts upon reduced MCAM expression that were shared by both sexes identified the cell cycle, DNA repair and DNA metabolism (Fig. 7g). These transcripts included ones encoding the DNA suppressor *BCRA2* and kinetochore proteins and *CENPA* and *CENPE* as well as one encoding the enzyme thymidine kinase, the protein responsible for the generation of deoxythymidine monophosphate, a critical step in the DNA replication process. The strength of enrichment for MCAM-modulated genes in these processes strongly implies that maintenance of MCAM expression across both sexes is important for maintaining replicative capacity and preventing senescence in vSMCs.

In males, we identified 498 genes with expression only altered in males. Gene ontology enrichment of these genes highlighted proteins associated with cytokine signaling, cell migration and cholesterol synthesis (Fig. 7h). These male-specific MCAM-associated transcripts



included *CX3CL1* (encoding a cytokine), *SREBF2* and *INSIG1* and *PTGS2*. The products of *SREBF2* and *INSIG1* serve as key regulators of cholesterol synthesis. *PTGS2*, also known as COX-2, is involved in the production of prostaglandin H2, a bioactive molecule that can stimulate platelet aggregation and modulate vessel contraction. Finally, *CX3CL1* encodes a chemokine that promotes the recruitment of immune cells. In females, we identified 418 genes with significantly altered expression. Gene ontology enrichment identified vascular development, intracellular signal transduction and vascular endothelial growth factor (VEGF) signaling (Fig. 7i). These female-enriched MCAM-associated transcripts included *F3*, *GDF15 and MYLK*. F3, also known as tissue factor, is essential for the formation of fibrin clots upon injury, including during aortic rupture. GDF15, a stress response cytokine, was previously shown to alter aortic contractility and relaxation[30]. Finally, mutations in *MYLK*, which encodes the protein myosin light-chain kinase, were identified as causational in familial thoracic aortic aneurysm disorders[31].

### Sex-enriched MCAM (CD146) in the context of abdominal aortic aneurysm

The vast transcriptional alterations associated with *MCAM* knockdown in vSMCs suggested that this gene may be protective and prevent the development of AAA. To further test this, we first performed scRNA-seq profiling of vascular segments in an AAA model that includes AngII (AngII$^{+}$) delivery in hyperlipidemic mice in comparison to control mice (Fig. 8a–d). In total, 813–2,440 cells per group were sequenced with six clusters representing distinct cell populations[32] (Extended Data Fig. 4 and Fig. 8d–f). vSMCs positive for *Mcam* were identified in both control and AAA vessels. Comparison of expression patterns from these cells showed a substantial decrease in the number of *Mcam* transcripts expressed per cell in AAA-derived SMCs (Fig. 8f,g). We performed staining on aortic samples from control and AAA mice, finding decreased expression of MCAM protein in AAA samples despite both SMC populations expressing substantial levels of α-smooth muscle actin (α-SMA), consistent with the scRNA-seq findings (Fig. 8h). Critical discrepancies have been noted in AngII$^{+}$ model AAAs versus typical human lesions; therefore, we also assessed *Mcam* expression in a second model of AAA, elastase treatment[33] (Extended Data Fig. 4k–r). Elastase treatment of the aorta significantly decreases *Mcam* expression in vSMCs, suggesting *Mcam* loss as a general mechanism of AAA progression.

To confirm causation, *Mcam*-knockout (KO) mice were crossed with hypercholesterolemic (*Apoe*$^{-/-}$) animals, and the resulting progeny were treated with AngII infusion for 28 d (Fig. 8i–l). Mice lacking MCAM showed decreased survival during the treatment period, with only 50% of *Mcam*-null animals surviving the full 28-d period, relative to 60% of wild-type littermates (Fig. 8). However, it should be stressed that the difference between 50% and 60% did not reach significance. Measurement of the external diameter of the thoracic aorta at killing showed no significant differences between *Apoe*$^{-/-}$;*Mcam*$^{+/+}$ and *Apoe*$^{-/-}$;*Mcam*$^{-/-}$ groups (Fig. 8k). By contrast, abdominal aorta measurements showed greater dilation in *Apoe*$^{-/-}$;*Mcam*$^{-/-}$ animals than in wild-type controls (Fig. 8l). MCAM expression was also evaluated at the protein level in healthy (control) and AAA sections from human patients (Fig. 8m–p). In humans, AAA specimens also showed a significant decrease in MCAM expression by both IF and immunoblot, indicating that loss of MCAM is likely involved in the regulation of AAA pathogenesis, albeit through unclear mechanisms (Fig. 8m–p). In sum, these findings support the conclusion that expression of MCAM (CD146) in the descending abdominal aorta is sex regulated with preferential expression in female abdominal SMCs with mechanistic experiments demonstrating a protective role against the development of AAA.

## Discussion

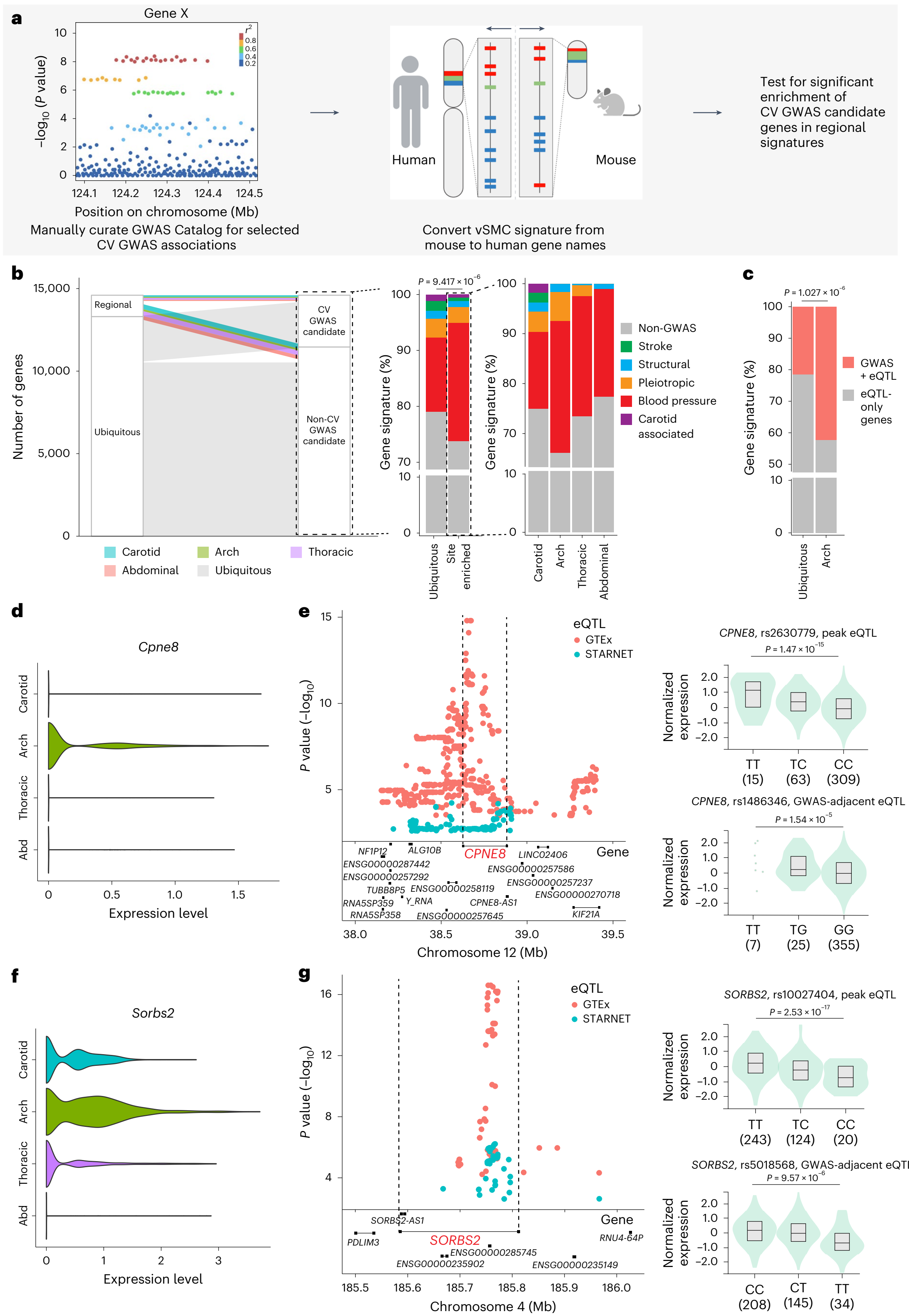
Substantial heterogeneity exists within vSMC populations across different vessel types, leading to important physiological and pathological outcomes[34–36]. Furthermore, vSMCs are endowed with impressive transcriptional plasticity, particularly when exposed to stressors[37,38]. Identifying vSMC heterogeneity within distinct regions of the same vessel can provide insight into subpopulations and molecular drivers that underpin vascular physiology[21,39] and pathology[39–41]. Transcriptomic studies have recognized subpopulations of vSMCs associated with the progression of atherosclerosis (*Ly6a*) and AAA (*Malat1*)[21,39]. These findings raise questions about what drives vSMC diversity, such as mosaicism, epigenetic modifications, local physical forces and signaling gradients. The enduring influence of embryological origin versus topological identity remains unresolved and is a focus of this study.

Despite their structural and hemodynamic differences, vSMCs of the carotids and much of the aortic arch share NC ancestry, justifying the inclusion of carotids in the analysis. We paired single-cell transcriptomics with lineage tracing and assessed embryological origin and positional identity, profiling 18,267 transcripts across 3,124 vSMCs from four anatomic regions. Although transcriptomes were broadly similar, 1,315 transcripts (7.2%) showed regionalized expression.

scRNA-seq data showed that vSMCs from the aortic arch and carotids were more similar to each other than to those from the thoracic and abdominal aorta, supporting retention of embryological memory. Lineage tracing combined with scRNA-seq also showed unique gene signatures in aortic arch vSMC populations identified as cells of NC or SHF origin. These results reinforce the notion that embryonic origin shapes the adult vSMC transcriptome and align with other studies indicating that origin influences function and stress responses[42].

In the era of precision medicine, the development of CV diseases is inherently linked to unique interactions between an individual's genetic profile, environmental risk factors and lifestyle choices[43]. Since the first GWAS in 2005 (ref. 44), thousands of genetic loci have been linked to coronary artery disease, hypertension and aortic aneurysm. However, the transition from single-gene to multi-gene models of disease has been substantially confounded by many disease-causing variants residing in noncoding regions. Attempts to elucidate the functional effects of identified genetic variants have led to quantitative trait locus mapping, which links disease-associated genetic variants with molecular traits (gene expression, chromatin accessibility)[45]. The identification of regional vSMC signatures allowed us to explore whether site-enriched transcripts contributed to localized disease. Overlaying available datasets showed that 339 site-enriched genes were

**Fig. 5 | Site-enriched genes show evidence of disease causality in humans.** **a**, Pipeline for identifying enrichment of regional vSMC gene signatures with CV disease incidence. **b**, Quantification of regional signature genes and known CV GWAS candidate genes assessed by one-sided Fisher's exact test. **c**, Visualization of regional vSMC gene signatures with known eQTLs in the GTEx database assessed by one-sided Fisher's exact test. **d**, Expression of *Cpne8* in the four mouse vascular regions. **e**, eQTL plot visualizing SNPs residing within 1 Mb of *CPNE8* in GTEx and STARNET as well as individual plots for the top *CPNE8* GTEx eQTL (rs2630779) and the GTEx eQTL closest to the peak arterial stiffness GWAS SNP (rs1486346). **f**, Expression of *Sorbs2* in the four mouse vascular regions. **g**, eQTL plot visualizing SNPs residing within 500 kb of *SORBS2* in GTEx and STARNET as well as individual plots for the top *SORBS2* GTEx eQTL (rs10027404) and the GTEx eQTL closest to the peak diastolic and systolic blood pressure GWAS SNP (rs5018568). For **e**,**g**, *n* is the number of individuals with indicated genotype per SNP as follows: rs2630779: TT, 15; TC, 63; CC, 309; $P = 1.47 \times 10^{-15}$. rs1486346: TT, 7; TG, 25; GG, 355; $P = 1.54 \times 10^{-5}$. rs10027404: TT, 243; TC, 124; CC, 20; $P = 2.53 \times 10^{-17}$. rs5018568: CC, 208; CT, 145; TT, 34; $P = 9.57 \times 10^{-6}$. Box plot lower and upper boundaries represent 25th and 75th quartiles, while the midline represents the mean normalized transcript expression per genotype group. Nominal *P* values were generated for each variant–gene pair (eQTL) by testing against a linear regression model between genotype and expression using FastQTL as described by GTEx. Panel **a** created using BioRender.com.



a
Gene X
−log10 (P value)
r2
0.8
0.6
0.4
0.2
Position on chromosome (Mb)
Manually curate GWAS Catalog for selected CV GWAS associations
Human
Mouse
Convert vSMC signature from mouse to human gene names
Test for significant enrichment of CV GWAS candidate genes in regional signatures
b
Number of genes
Regional
Ubiquitous
CV GWAS candidate
Non-CV GWAS candidate
Carotid
Arch
Thoracic
Abdominal
Ubiquitous
P = 9.417 × 10−6
Gene signature (%)
Ubiquitous
Site enriched
Carotid
Arch
Thoracic
Abdominal
Non-GWAS
Stroke
Structural
Pleiotropic
Blood pressure
Carotid associated
c
P = 1.027 × 10−6
Gene signature (%)
GWAS + eQTL
eQTL-only genes
Ubiquitous
Arch
d
Cpne8
Carotid
Arch
Thoracic
Abd
Expression level
e
P value (−log10)
eQTL
GTEx
STARNET
Gene
CPNE8
Chromosome 12 (Mb)
CPNE8, rs2630779, peak eQTL
P = 1.47 × 10−15
Normalized expression
TT (15)
TC (63)
CC (309)
CPNE8, rs1486346, GWAS-adjacent eQTL
P = 1.54 × 10−5
TT (7)
TG (25)
GG (355)
f
Sorbs2
Carotid
Arch
Thoracic
Abd
Expression level
g
P value (−log10)
eQTL
GTEx
STARNET
Gene
SORBS2
Chromosome 4 (Mb)
SORBS2, rs10027404, peak eQTL
P = 2.53 × 10−17
Normalized expression
TT (243)
TC (124)
CC (20)
SORBS2, rs5018568, GWAS-adjacent eQTL
P = 9.57 × 10−6
CC (208)
CT (145)
TT (34)

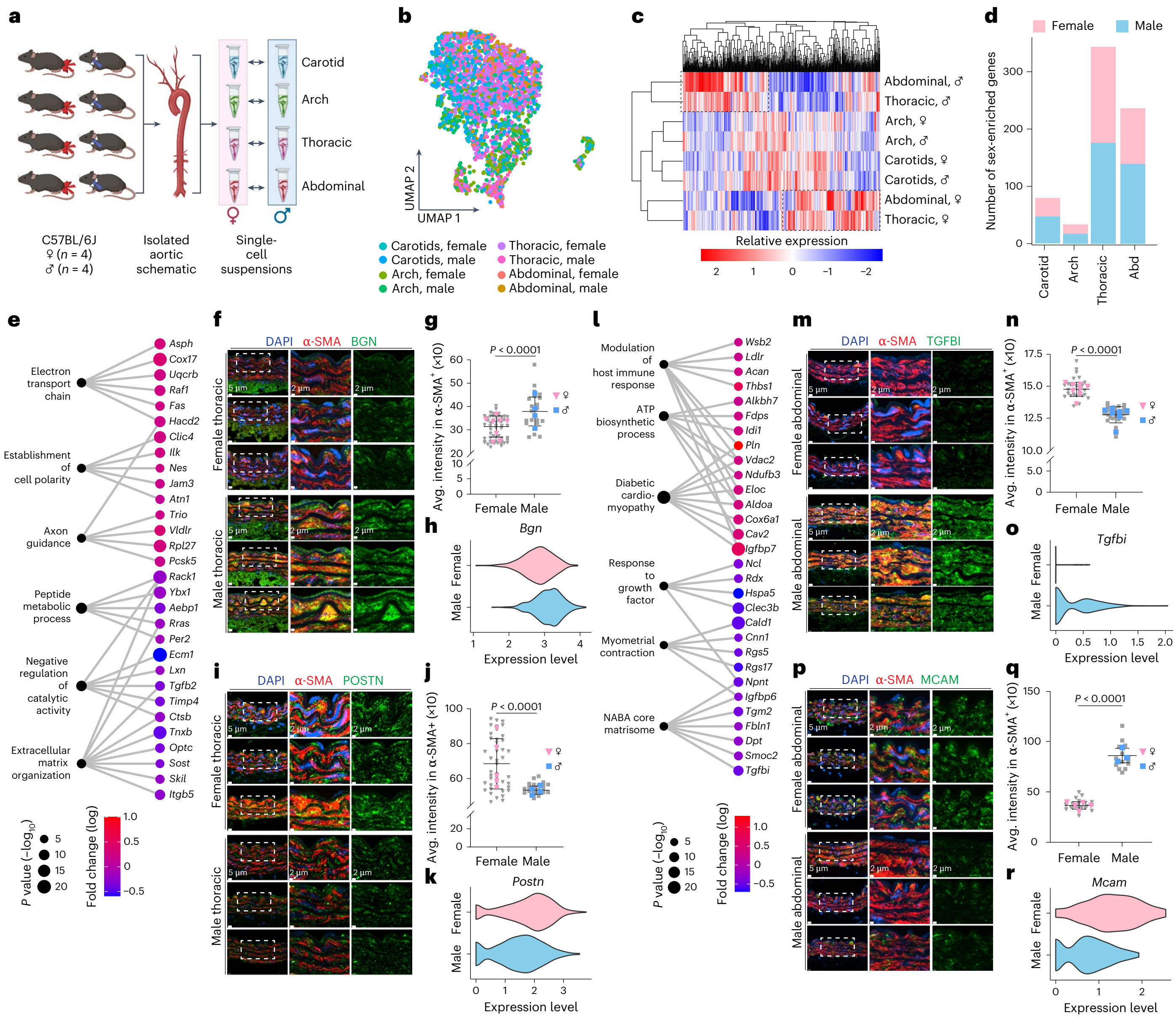
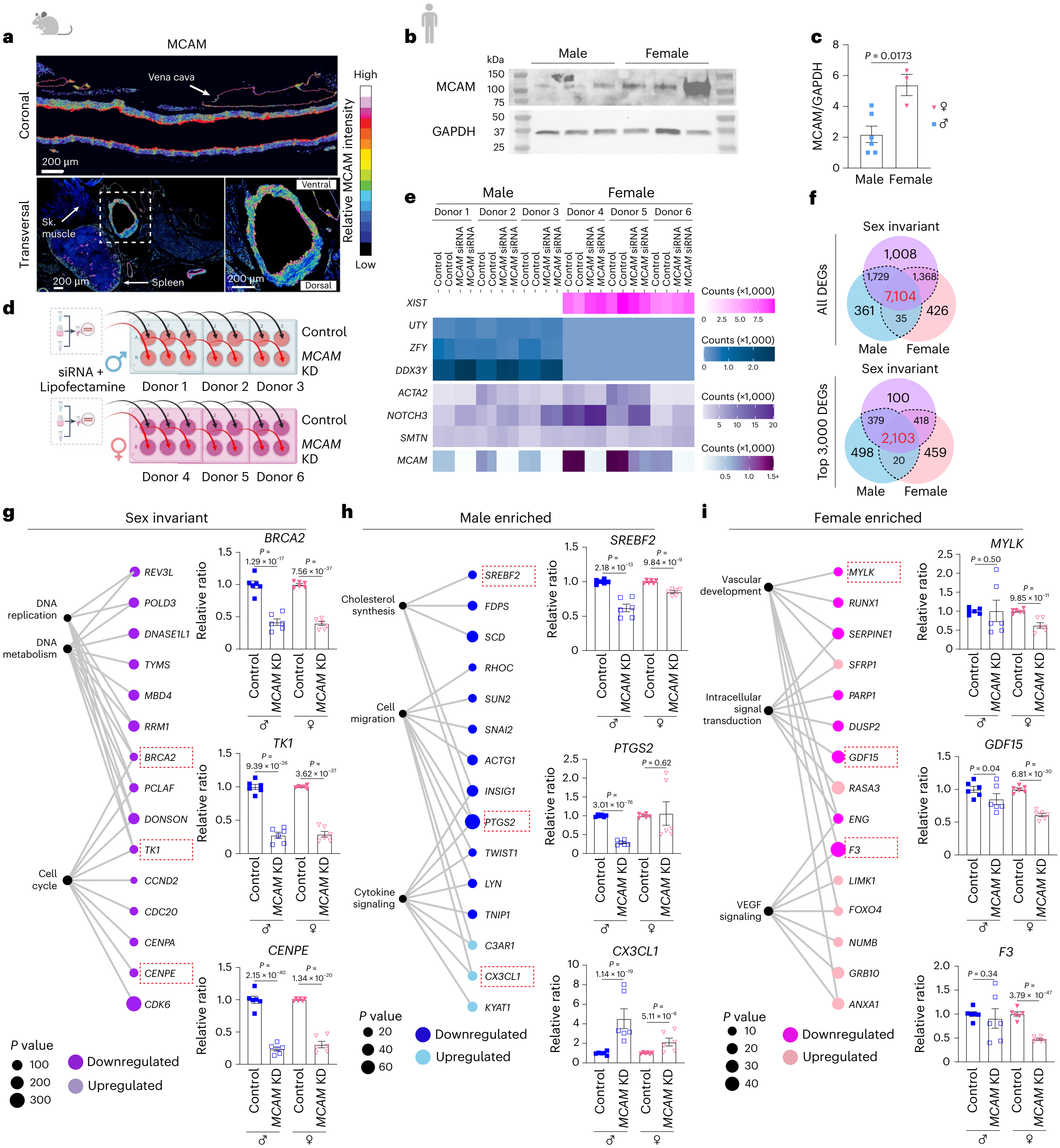



**Fig. 6 | Regionally restricted sex differences in vascular smooth muscle.** **a**, Schematic of experimental design. Twelve-week-old C57BL/6J female and male mice (four per group or sex) were dissected and digested into eight single cells for generation of libraries based on location and sex for scRNA-seq. **b**, UMAP of vSMCs from the eight libraries. **c**, Heatmap of all 527 sex-enriched transcripts in vSMCs. Dashed boxes indicate concurrent gene expression patterns across sex rather than anatomic site. **d**, Number of sex-enriched genes by anatomic location and sex (see Extended Data Fig. 2 for additional specifics). **e**, Gene ontology plot of uniquely enriched, autosomal thoracic-derived vSMC genes in both sexes calculated via hypergeometric distribution. *P* values for enrichment categories are as follows: female: electron transport chain, $9.82 \times 10^{-8}$; establishment of cell polarity, $7.19 \times 10^{-7}$; axon guidance, $7.94 \times 10^{-7}$; male: peptide metabolic process, $6.15 \times 10^{-12}$; negative regulation of catalytic activity, $1.80 \times 10^{-10}$; extracellular matrix organization, $1.11 \times 10^{-7}$. **f**, Representative IF images of BGN expression (green) in female and male thoracic vSMCs. **g**, Quantification of BGN intensity in female and male vSMCs. Females, $n = 54$ images, 8 animals; males, $n = 21$ images, 4 males; 74 images in total. $P < 0.0001$. **h**, Violin plot visualizing the expression pattern of *Bgn* in thoracic vSMCs by sex. **i**, Representative IF images of POSTN expression (green) in female and male thoracic vSMCs. **j**, Quantification of POSTN intensity in female and male vSMCs. Females, $n = 31$ images, 5 animals; males, $n = 18$ images, 5 males; 49 images in total. $P < 0.0001$. **k**, Violin plot visualizing the expression pattern of *Postn* in thoracic vSMCs by sex. **l**, Gene ontology plot of uniquely enriched, autosomal abdominal vSMC genes in both sexes calculated via hypergeometric distribution. *P* values for enrichment categories are as follows: females: diabetic cardiomyopathy, $2.27 \times 10^{-22}$; ATP biosynthetic process, $1.99 \times 10^{-10}$; modulation of host immune response, $1.99 \times 10^{-7}$; males: NABA core matrisome, $4.27 \times 10^{-7}$; myometrial contraction, $2.60 \times 10^{-6}$; response to growth factor, $3.42 \times 10^{-6}$. **m**, Representative IF images of TGFBI expression (green) in female and male abdominal vSMCs. **n**, Quantification of TGFBI intensity in female and male vSMCs. Females, $n = 21$ images, 5 animals; males, $n = 14$ images, 3 males; 35 images in total; $P < 0.0001$. **o**, Violin plot visualizing the expression pattern of *Tgfbi* in abdominal vSMCs by sex. **p**, Representative IF images of MCAM expression (green) in female and male abdominal vSMCs. **q**, Quantification of MCAM intensity in female and male vSMCs. Females, $n = 30$ images, 8 animals; males, $n = 26$ images, 7 males; 56 images in total; $P < 0.0001$. **r**, Violin plot visualizing the expression pattern of *Mcam* in abdominal vSMCs by sex. For **g**,**j**,**n**,**q**, error bars represent mean ± s.d. $n = 14$–25 images per animal (gray), 3–8 animals per sex (blue squares, male; pink inverted triangles, female). Two-sided Mann–Whitney test (*Postn*, *Bgn*, *Mcam*) and two-sided Welch's *t*-test (*Tgfbi*). Panel **a** created using BioRender.com.




residing within CV GWAS loci (Supplementary Tables 3, 5 and 6). While exciting, additional mechanistic experiments are needed to reveal the contributions of each of the hundreds of markers with regional specificity, and this constitutes a limitation of the present study. Nonetheless, the information provided here could aid in narrowing the search to identify causative genes.

Since the Framingham Heart Study identified male sex as a coronary disease risk factor, sex has emerged as a key variable in multiple CV pathologies, including aortic aneurysm[46,47]. Pathogenesis of aortic aneurysm shows exquisite regional differences, with 60% of TAA associated with the aortic root and AAA associated with the infra-renal aorta[47]. TAA and AAA also show substantial sex bias, with disease prevalence in males ranging from 2:1 (TAA) to 5:1 (AAA)[47]. Despite increased male prevalence, females are predicted to have worse outcomes upon diagnosis, with up to threefold increased risk of dissection and rupture[47]. Here, we characterize sex differences in region-specific vSMC populations, finding that sex-biased transcriptional differences were strongest in the thoracic and abdominal aorta but less pronounced in the arch and carotids. Among differentially expressed genes, *MCAM* was selected for further analysis.

**Fig. 7 | Sex as a modifier of the MCAM-associated transcriptional signature in aortic vSMCs. a**, Localization of MCAM in the thoracic–abdominal region in the mouse aorta. Sk. muscle, skeletal muscle. **b**, Protein lysates from a cohort of female and male vSMC cell lines were probed for expression of MCAM and GAPDH (loading control). **c**, Quantification of the relative ratio of MCAM to GAPDH in each donor. Data are shown as mean ± s.e.m; $n = 3$ (female), $n = 6$ (male); two-sided Welch's $t$-test, $P = 0.0173$. **d**, Experimental design. Aortic vSMCs isolated from three female and three male donors were cultured and transfected with siRNA against *MCAM*. Forty hours after transfection, RNA was isolated for transcriptomic profiling. KD, knockdown. **e**, Heatmap visualizing transcriptional expression of selected markers of biological sex and vSMC identity across experimental conditions and donors. **f**, Venn diagram visualizing the overlap of differentially expressed genes (DEGs) in male versus female vSMCs following knockdown of MCAM. **g**–**i**, Gene ontology enrichment for sex-invariant (**g**), male-enriched (**h**) and female-enriched (**i**) differentially expressed genes in *MCAM*-knockdown vSMCs, 40 h after transfection. Dot color indicates the direction of expression change upon loss of *MCAM* expression, while size indicates significance of enrichment; $P$ values are as follows: sex invariant (**g**): cell cycle, $P = 5.18 \times 10^{-78}$; DNA metabolism, $P = 6.65 \times 10^{-55}$; DNA replication, $P = 1.84 \times 10^{-33}$; male enriched (**h**): cytokine signaling, $P = 4.11 \times 10^{-10}$; cell migration, $P = 7.47 \times 10^{-10}$; cholesterol synthesis, $P = 2.19 \times 10^{-9}$; female enriched (**i**): VEGF signaling, $P = 4.30 \times 10^{-12}$; intracellular signal transduction, $P = 4.49 \times 10^{-11}$; vascular development, $P = 1.09 \times 10^{-10}$. For **g**–**i**, selected graphs of transcript levels of member genes in each ontology category data are visualized using mean ± s.e.m. $n = 6$. Displayed $P$ values are Benjamini–Hochberg corrected Wald test values from DESeq2 for the following genes: *BRCA2* (male, $P = 1.29 \times 10^{-17}$; female, $P = 7.56 \times 10^{-37}$), *TK1* (male, $P = 9.39 \times 10^{-26}$; female, $P = 3.62 \times 10^{-37}$), *CENPE* (male, $P = 2.15 \times 10^{-40}$; female, $P = 1.34 \times 10^{-20}$), *SREBF2* (male, $P = 2.18 \times 10^{-13}$; female, $P = 9.84 \times 10^{-9}$), *PTGS2* (male, $P = 3.01 \times 10^{-76}$; female, $P = 0.62$), *CX3CL1* (male, $P = 1.14 \times 10^{-19}$; female, $P = 5.11 \times 10^{-6}$), *MYLK* (male, $P = 0.50$; female, $P = 9.85 \times 10^{-11}$), *GDF15* (male, $P = 0.04$; female, $P = 6.81 \times 10^{-30}$), *F3* (male, $P = 0.34$; female, $P = 3.79 \times 10^{-47}$). Mouse pictograph (**a**), human pictograph (**b**) and panel **d** created using BioRender.com.

*Mcam* was preferentially expressed in the aorta of females. MCAM is a transmembrane glycoprotein expressed by vSMCs and several other vascular cell types (endothelial, pericytes) with roles in adhesion and platelet-derived growth factor receptor, β polypeptide (PDGFRβ) signaling[48,49]. During embryonic development, MCAM is highly expressed in early SMC progenitors and it regulates the proliferation–differentiation balance[50]. In adults, expression is higher in disturbed flow regions such as branches[50] and in the ventral aspect of the dorsal aorta. Coincidentally, the ventral side of the aorta is a preferred site for AAA development and dissection, likely due to mechanical factors and intrinsic biological differences. Dorsally, the aorta is in close contact with the vertebral bodies, and aneurysm geometry flattens on the posterior aspects of the aorta[51]. Furthermore, the aortic wall is circumferentially heterogeneous in gene and protein expression[52]. Our study builds on this previous evidence and adds the contribution of *Mcam* in vSMCs to the spatial resolution of aortic aneurysm development.

MCAM expression reduction by siRNA altered vSMC transcriptional programs, including genes linked to cell division. Sex-specific differences were evident: in male vSMCs, *MCAM* knockdown affected cholesterol synthesis and inflammation, consistent with its proposed immunomodulatory roles[50,51]; in females, genes related to vascular development and VEGF signaling were impacted. To mechanistically test the role of MCAM in AAA development, we turned to MCAM-null mice. Using MCAM-null mice, we observed decreased survival following AAA induction and increased abdominal aortic dilation. We acknowledge that an important limitation of these studies is that the inactivation of MCAM was global rather than vSMC specific. Importantly, we also found that MCAM was reduced in SMCs at sites of human aneurysm, further supporting a possible contribution of MCAM to the structural integrity of the tunica media.

Overall, this work builds on scRNA-seq efforts to uncover transcriptional heterogeneity in the vascular tree. Our segmented approach revealed regional and sex-specific transcriptional differences relevant to aneurysms and dissection-prone sites. Finally, we show that high-resolution scRNA-seq can enhance GWAS and eQTL interpretation in CV research.

### Limitations of our study

We note that, although several animals were used to generate each of the eight scRNA-seq libraries, only two libraries (one from males and one from females) were part of the direct regional comparisons. Naturally, inclusion of multiple libraries from the same region would have strengthened the data. Furthermore, while the top differentially expressed genes were validated in at least three mice per sex per region (24 different mice with over 1,000 images for quantification), additional studies will be needed to expand on the relevance of the identified differences.

The selection of *Mcam* for in-depth analysis was based on its peculiar ventral–dorsal expression profile and the availability of a complete KO model for mechanistic interrogation. Given that AAA induction is more effective in male mice, we were only able to assess the consequences of MCAM loss in males. This sex-dependent variability is attributed to intrinsic differences in arterial wall composition and reduced inflammatory response in females[53].

Overall, our data suggest that *Mcam* contributes to but is not the primary driver of AAA. Notably, homozygous loss-of-function mutations in *MCAM* have not been reported in humans. Mechanistic studies of other candidate genes identified in this work are ongoing, using cell-specific approaches and multiple AAA induction models. These efforts should continue to clarify the molecular basis for innate protection against AAA observed in female mice (and humans).

## Methods

### Mice

C57BL6/J mice (000664) were purchased from Jackson Laboratory. *Mef2c*$^{Cre}$ mice (030262) crossed with ROSA26R$^{mT/mG}$ (007676) mice were a kind gift from A. Daugherty (University of Kentucky, Jackson Laboratory), *Wnt1*$^{Cre}$ mice, a kind gift from B. Thomson (Northwestern University, Jackson Laboratories), were crossed with B6.Cg-Gt(ROSA)26Sor$^{tm14(CAG\text{-}tdTomato)Hze}$/J mice (07914) before scRNA-seq experiments. For studies using CD146 (*Mcam*)-KO mice[54], animals were backcrossed for more than ten generations on the C57BL/6J background. Primers for genotyping CD146 (*Mcam*) alleles were as follows: 5′-TCACTTGACAGTGTGATGGT-3′ (forward primer used to detect CD146 (*Mcam*) WT, floxed and KO alleles), 5′-CCTTAGAAAGCAGGGATTCA-3′ (reverse primer used to detect CD146 (*Mcam*) WT and floxed alleles) and 5′-CCCAAATCCTCTGGAAGACA-3′ (reverse primer used to detect CD146 (*Mcam*) KO allele). Genotyping primers for the *Apoe* alleles were as follows: 5′-GCCTAGCCGAGGGAGAGCCG-3′, 5′-TGTGACTTGGGAGCTCTGCAGC-3′ and 5′-GCCGCCCCGACTGCATCT-3′. Mice were housed at the University of California, Los Angeles, Northwestern University, New York University Langone Health and at the Cardiovascular Research Center, Université de Paris. Animals were maintained in ventilated racks, with no more than five mice per cage and a 14–10-hour light–dark cycle. Room temperature was between 68 °F and 77 °F, and humidity was between 30% and 70%. Mice received food and water ad libitum. Experimental procedures were reviewed and approved by the institutional animal care and use committee at the respective institutions. Protocols were conducted in accordance with federal regulations as outlined in the 'Guide for the Care and Use of Laboratory Animals' for US laboratories and European Community Guidelines for European laboratories.

https://doi.org/10.1038/s44161-025-00692-4

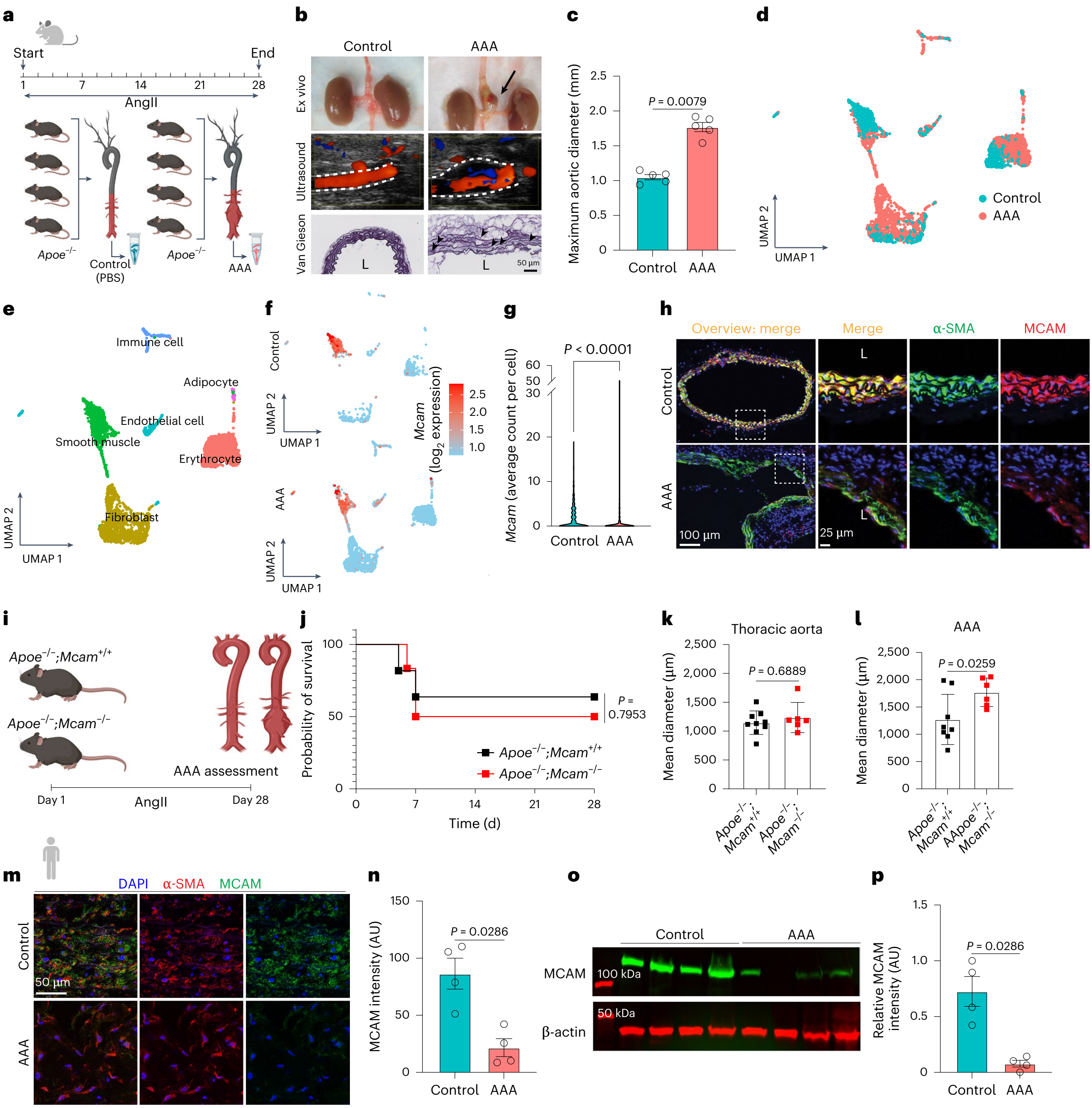



**Fig. 8 | Characterization of *Mcam* (CD146) in AAA. a**, Experimental design. Male *Apoe*$^{-/-}$ mice were treated with AngII ($n = 3$) or phosphate-buffered saline (PBS) ($n = 3$) for 28 d to promote the development of AAAs. Aortas were collected, digested and analyzed by scRNA-seq. **b**, Morphological and histological assessment of control and diseased (AAA) aortas. **c**, Diameter at sacrifice. $n = 5$ per group, male animals, two-tailed Mann–Whitney test. Error bars represent mean ± s.e.m. **d**, UMAP plot of scRNA-seq of diseased aortas and healthy aortas ($n = 3$, pooled). **e**, UMAP plot of cell type distribution in AAA and control aortas. **f**, Feature plot of *Mcam* expression across control (top) and AAA (bottom) cells. **g**, Average *Mcam* expression in healthy and diseased (AAA) vSMCs. Unpaired two-sided *t*-test, $n = 813$ (control) to 2,440 (AAA). **h**, IF MCAM staining in control versus AAA mouse aortas at low and high magnification, $n = 3$. **i**, Experimental design. *Apoe*$^{-/-}$ mice lacking *Mcam* (*Apoe*$^{-/-}$;*Mcam*$^{-/-}$) and wild-type littermates (*Apoe*$^{-/-}$;*Mcam*$^{+/+}$) were treated with AngII for 28 d. **j**, Kaplan–Meier curves of control and experimental groups over 28 d. $n = 6$ (*Apoe*$^{-/-}$;*Mcam*$^{-/-}$) to 11 (*Apoe*$^{-/-}$;*Mcam*$^{+/+}$). **k**, Quantification of thoracic aorta (non-aneurysmal) diameter following treatment. $n = 6$ (*Apoe*$^{-/-}$;*Mcam*$^{-/-}$) to 9 (*Apoe*$^{-/-}$;*Mcam*$^{+/+}$). **l**, Quantification of abdominal aorta (aneurysm) diameter after treatment. $n = 6$ (*Apoe*$^{-/-}$;*Mcam*$^{-/-}$) to 8 (*Apoe*$^{-/-}$;*Mcam*$^{+/+}$). **m**, IF staining of human healthy aortas (control) or AAA with α-SMA (red) and MCAM (green). **n**, Quantification of human IF samples. AU, arbitrary units. **o**, Western blot of control or AAA human aortas probed for MCAM (green) and loading control (β-actin; red). **p**, Quantification of western blot samples. For **b**, the arrow indicates the region of aneurysm formation. Arrowheads indicate elastin breaks. L, lumen. For **j**–**l**, two-tailed Mann–Whitney test (**k**) and two-tailed Welch's *t*-test (**l**). Error bars represent mean ± s.d. For **n**–**p**, $n = 4$ per group, two-tailed Mann–Whitney test. Error bars represent mean ± s.d. Controls are all male, while AAA samples are male and female. Note that **d**–**h** represent a reanalysis of Hadi et al.[32]. Panels **a** and **i**, mouse pictograph (**a**) and human pictograph (**i**) created using BioRender.com.

### Human samples

All protocols involving human participants are compliant with the Declaration of Helsinki principles. All studies were conducted in accordance with policies set forth by the New York University Langone Medical Center Institutional Review Board. The study was authorized with institutional review board approval number i16-01807. Patients were not discriminated against according to age, gender, genotype information or past or current diagnosis. Cadaver tissues from multi-organ donors who had been confirmed as brain dead were provided by the LiveOnNY organization and used as controls. Samples of unruptured human aneurysmal (>50 mm) aortic walls were collected during open aortic repair procedures. Informed consent was obtained for each participant before surgery. All tissues were macroscopically inspected, oriented, formalin fixed and paraffin embedded before sectioning. Aortic tissue for immunoblots was snap frozen in liquid nitrogen and pulverized with a mortar and a pestle before resuspension in radioimmunoprecipitation assay (RIPA) buffer.

### scRNA-seq sample preparation for site-enriched transcriptomics

Four male and four female 12-week-old C57BL/6J mice were killed and perfused with 10 ml Versene. Following perfusion, the carotid arteries and the aorta (from the aortic root to the iliac bifurcation) were harvested and further dissected under a microscope to remove the adventitial layer as efficiently as possible, ensuring that the process did not exceed 30 min. Subsequently, the four vascular fragments (carotid arteries (right and left common but not their branches), the aortic arch (from the root including the bulb to the subclavian branch), the thoracic aorta (below the subclavian up to the diaphragm but excluding all branches) and the abdominal aorta (below the diaphragm to the iliac bifurcation, excluding all branches)) were obtained from each animal, and samples were pooled into eight groups for digestion and library preparation segregated by sex and site (Supplementary Table 2). The eight pooled vessel samples were then individually minced into small fragments and incubated under agitation in 500 µl digestion buffer containing freshly prepared Liberase solution (2.5% Liberase TH from frozen stock (Sigma Aldrich), 5 Kunitz units per ml DNase I, 1 M HEPES in 1× HBSS). Tissue preparation was carried out at room temperature. Digestions to isolate single-cell suspensions were conducted at 37 °C for approximately 20 min. The resulting suspensions were neutralized with DMEM + 10% FBS, run through a 40-µM filter and then centrifuged for 5 min at 10,000*g*. Following centrifugation, cell pellets were resuspended in 0.4% BSA–PBS, RBC lysis was performed, and cells were pelleted by centrifugation and assessed for viability before library preparation using trypan blue staining with the TC20 cell counter (Bio-Rad). Cell viability ranged from 92% to 95%. All libraries were processed concurrently. Bioinformatic analysis showed that total cell numbers per library were as follows: carotids, 7,249 cells; aortic arch, 7,929 cells; thoracic aorta, 8,401 cells; abdominal aorta, 4,930 cells. Importantly, we recognize that blood might be a frequent contaminant of any vascular tissue[55]. In fact, we can detect hemoglobin genes in almost all our datasets. To mitigate the potential role of ambient RNA in our libraries, we applied two algorithms: EmptyDrops[56] and SoupX[57]. These tools identify empty droplets in the library and assess transcript abundance in these droplets. We noted that the only substantial contaminant was hemoglobin (in the carotid libraries) and thus decided against further modifying the original data, as the levels of ambient RNA did not alter the findings.

### Single-cell RNA sequencing

Libraries were prepared using 10x Genomics Chromium Single Cell 3′ Library & Gel Bead Kit version 3 according to the manufacturer's protocol. For the generation of single-cell gel beads in emulsion, cells were loaded on a Chromium Single Cell instrument (10x Genomics) with an estimated targeted cell recovery of between 5,000 and 10,000 cells per library.

### Bioinformatic analysis for single-cell transcriptomics

Sequencing of scRNA-seq libraries was performed on the same run of an Illumina NovaSeq 6000 sequencer. The data pipeline was as follows: the digital expression matrix was generated by demultiplexing, barcode processing and gene unique molecular index counting using the Cell Ranger pipeline. Expression matrices for different samples were concatenated, and the R package Seurat (version 4.3.0) was used to analyze the concatenated expression matrix. Cells with less than 100 genes or more than 50,000 unique molecular identifiers or greater than 10% mitochondrial expression were removed from further analysis (Extended Data Figs. 1 and 2). The Seurat NormalizeData function was used to normalize raw counts. Variable genes were identified using the FindVariableGenes function; the top 2,000 variable genes were selected for further analysis. The Seurat ScaleData function was used to scale and center expression values in the dataset for dimensional reduction. Principal-component analysis was performed with the RunPCA function. The first 30 principal components were input for the RunHarmony function to remove potential batch effects among different samples. The first 30 Harmony dimensions were then used for UMAP, and the first two UMAP dimensions were used in plots. A graph-based clustering approach was later used to cluster the cells on the first 30 Harmony dimensions; next, signature genes were identified using the FindAllMarkers function and used to define cell types for each cluster. The Seurat FindMarkers function was used for differential expression analysis between any two groups of cells, in which the Wilcoxon rank-sum test was implemented. The Benjamini–Hochberg procedure was used to calculate the false discovery rate (FDR), and FDR < 0.05 was used to select significantly differentially expressed genes.

### Site-specific gene signature determination

The Seurat FindAllMarkers function was used to identify positive markers of each of the four anatomic locations associated with the vSMC population (Supplementary Table 1). Following gene list generation, the datasets were filtered for significance (FDR ≤ 0.05). Genes that were associated with more than one anatomic location were removed. Regional signature genes were then ranked by the ratio of transcript-producing cells in the region (percentage of cells where the feature is detected in the first group (pct.1)/percentage of cells where the feature is detected in the second group (pct.2)) and the significance of fold enrichment for that transcript to generate a sum rank score. Genes with the smallest sum rank score showed drastic increases in transcript fold change relative to other locations and an increased number of vSMCs in this region expressing these transcripts. Following rank assignment, genes residing in the top 100 genes for the anatomic region were plotted using a violin plot to assess the distribution of transcript levels. Based on the availability of antibodies (commercially), subsets of each site were selected for further testing by antibody-based methods to study protein expression levels and distribution patterns.

### Lineage-tracing overlap with the arch anatomic signature

The Seurat FindAllMarkers function was used to find positive markers of each lineage-specific vSMC population (eGFP$^+$ for *Mef2c*$^{Cre}$ and tdTomato$^+$ for *Wnt1*$^{Cre}$), and a pairwise comparison between cells from the two datasets was analyzed. The resultant gene list was then selected for those subsets of genes that were identified in the aortic arch location-enriched vSMC signature and then filtered for significance ($P$ value ≤ 0.05). To map the lineage-specific vSMC datasets to vSMC datasets from different anatomic locations, the Seurat FindTransferAnchors and MapQuery functions were used to transfer the anomic labels and UMAP coordinates.

### Sex-enriched gene signature determination

Pairwise comparisons between male and female vSMCs for each anatomic site were performed using Seurat, and datasets were filtered for significance (FDR ≤ 0.05). This subset of genes was then ranked in descending order in relation to fold enrichment for the anatomic location and assigned a number in the rank list. Genes were then ranked in descending order by calculating the ratio of pct.1/pct.2. Scores for each rank list (fold change enrichment and pct1/pct2 ratio) were summed, and genes were ranked by their final score in ascending order.

### Gene ontology analysis

Pathway enrichment analysis was performed with Metascape express analysis using default options. Metascape uses the hypergeometric distribution to calculate significance for each gene enrichment category. Scores (log (*P* values)) are reflective of the number of genes in the gene list defined by the category in question relative to the total number of genes in the category relative to background. For gene ontology enrichment of the anatomic gene signatures, the top 50 ranked genes from each anatomic signature were used to form the gene list, while, for the regionally sex-enriched signatures, the entirety of the unique gene signature was used. From the output, the top summary gene ontology categories (as ranked by *P* value of enrichment) for each of the anatomic locations (five categories) and sex enrichment (three categories) were used for visualization of gene ontology following curations and removal of duplicate or highly similar terms. From these data, two tabular matrices were constructed. The first matrix (the vertex file) contained the fold enrichment for the individual gene ontology categories as well as top gene ontology category membership, fold enrichment for anatomic location and *P* value for anatomic location. The second matrix (Edge file) contained all gene ontology memberships for visualized genes. Both matrices were then visualized into a custom ontology network using a combination of the ggplot2, igraph and ggraph packages from R.

### Immunostaining

**Selection process of candidate genes for IHC evaluation.** Following gene list generation, the datasets were filtered for significance (FDR ≤ 0.05). Transcripts associated with more than one anatomic location were removed. Regional signature genes were then ranked by the ratio of transcript-producing cells in the region (pct.1/pct.2) and the significance of fold enrichment for that transcript to generate a sum rank score. Genes with the smallest sum rank score showed the highest transcript fold change relative to other locations and an increased number of vSMCs in this region expressing these transcripts. Following rank assignment, genes in the top 100 for each anatomic region were plotted using a violin plot to assess the distribution of transcript levels. Subsets of each site were selected for further testing based on ranking and the availability of antibodies. Importantly, antibodies were tested for specificity by WB, protein competition and/or CRISPR KO.

**Mouse tissue.** Paraformaldehyde (2%)-fixed, paraffin-embedded specimens from mouse aortic and carotid vessels were sectioned at 5 µm, deparaffinized, processed for rehydration and antigen retrieval (citrate buffer, pH 6) and incubated with primary antibodies overnight followed by species-specific secondary antibodies for 1 h before mounting in ProLong Gold (Thermo Scientific).

**Human tissue.** Paraffin-embedded aortic samples were transversally sectioned into 7-µm-thick sections, subsequently deparaffinized and processed for rehydration and antigen retrieval as previously described[32]. Briefly paraffin-embedded aorta sections were deparaffinized and rehydrated by successive washes with xylene, xylene–ethanol (vol/vol), 100% ethanol, 95% ethanol, 70% ethanol and water. Antigen retrieval was performed by boiling in 10 mM Tris, 1 mM EDTA, 0.05% Tween-20, pH 9 retrieval buffer. Sections were then blocked before overnight incubation with primary antibodies followed by species-specific secondary antibodies for 1 h before mounting. Specific antibody information including antibody concentration, vendor information and RRID are included in Supplementary Table 10.

**Confocal microscopy for mouse samples.** Samples were evaluated and photographed using a spinning-disk confocal microscope (Nikon, Yokogawa, CSU-W1) equipped with ×20 air and ×40 and ×100 oil objectives. For a subset of candidates (*Mcam*), samples were evaluated and photographed using an LSM 880 confocal microscope (Carl Zeiss) equipped with Zeiss Plan-Apochromat ×20/0.8 M27 for acquisition.

**Confocal microscopy for human samples.** Images were acquired on a Zeiss LSM 710 confocal microscope (Carl Zeiss) using Zeiss Efficient Navigation software (Carl Zeiss). Identical acquisition parameters were set for control and AAA samples.

**Image quantification.** Images were first processed with the Imaris file converter and then transferred to Imaris 10.0.0 software. Visualization of all channels was normalized across images for the candidate gene being analyzed. Imaris surfaces were generated using the α-SMA$^+$ channel (for cytoplasmic staining for α-SMA) using background subtraction for thresholding. This approach allowed us to segregate and restrict our analysis of channel intensity to only those regions in the image that were α-SMA$^+$. Following surface creation, all α-SMA$^+$ surfaces in a given image were selected, and the mean intensity of the site-specific candidate (568 channel mean intensity) per surface was extracted from the statistics table and exported as a new tab-delimited file. This process of surface generation, selection and channel intensity statistic export was repeated for all four anatomic sites across two to three images per location, across three to six animals per candidate gene investigated for anatomic location signatures. For sex-enriched genes, surface generation, selection and channel intensity, statistic export was repeated across two to three images per location, across three to eight animals per site and sex.

From each export file, the mean channel intensity statistic export value was averaged across all surfaces to produce a single value per image and then compiled into a master table with all intensity values across all images for all regions across all animals. The final table was then imported into GraphPad Prism for statistical analysis with GraphPad Prism 10.0. Based on the distribution of the data, significance of expression was assessed using either Brown–Forsythe and Welch's ANOVA or the Kruskal–Wallis test with correction for multiple-testing comparison.

For those candidate genes with predominant nuclear staining (that is, PGR), the total number of positive nuclei in an α-SMA$^+$ area per image was quantified by hand for all four anatomic sites across two to three images per site per three to six animals. The quantification across all images was then imported into GraphPad Prism for statistical analysis with GraphPad Prism 10.0. Based on the distribution of the data, significance of expression was assessed using either Brown–Forsythe and Welch's ANOVA or the Kruskal–Wallis test with correction for multiple-testing comparison.

**Enrichment analysis of cardiovascular GWAS genes in site-enriched vSMC signatures.** Entries from the GWAS Catalog were downloaded on 30 September 2023 and manually curated. The catalog entries identified with CV-associated Experimental Factor Ontology (EFO)-reported traits with known vSMC-driven etiology were selected for initial analysis. From this initial curation, 69 EFO categories were categorized into four broad CV representative groups: vascular aneurysm and/or structural alterations, blood pressure, stroke and carotid artery phenotypes. Entries under the Mapped_Genes category from each EFO entry in the four CV groups were then filtered to remove duplicate gene entries. If a gene was identified in at least two CV GWAS categories, it was moved to a new fifth category, the pleiotropic CV category.

The list of all vSMC-expressed genes was extracted from the Seurat expression matrix using the function AverageExpression. Gene symbols from the expression matrix were then converted from mouse to human using Metascape. To correct for cell-specific expression, the total number of possible CV-enriched GWAS genes was overlaid with the list of vSMC-expressed genes. Fisher's exact test was performed in R using the stats package with numbers generated from four key gene lists in a two-by-two contingency table: (1) regional vSMC CV GWAS candidates, (2) regional-enriched vSMC genes, (3) CV GWAS genes, (4) vSMC-expressed genes. Data visualization graphs were generated using the ggplot2, ggalluvial and ggbreak packages for R.

**Expression quantitative trait locus overlap with the aortic transcriptome from GTEx and STARNET**

Tissue eQTLs for the aortic transcriptome were downloaded from GTEx (https://www.gtexportal.org/home/) on 9 January 2024 from the GTEx_Analysis_v8_eQTL directory. Identification of significant aortic eQTLs was performed using the file Artery_Aorta.v8.signif_variant_gene_pairs.txt. Overlay between the aortic gene set and the vSMC gene list was performed using the table merge function in R and ENSMBL_IDs to overlay the gene list. Fisher's exact test for enrichment for GWAS genes with aortic eQTLs within the carotid and aortic arch gene lists was performed in R using the stats package with numbers generated from four key gene lists in a two-by-two contingency table: (1) regional vSMC CV GWAS with aortic eQTL candidates, (2) regional-enriched vSMC genes with eQTLs, (3) CV GWAS genes with aortic eQTLs, (4) vSMC-expressed genes with eQTLs. Tissue eQTLs for the aortic root transcriptome were manually curated from the STARNET database (http://starnet.mssm.edu/) on 24 February 2024. Genomic locations of the rsIDs for eQTL-associated variants from the STARNET database were obtained using the UCSC Genome Browser's table browser for human assembly GRCh38, common dbSNP (155). Data visualization graphs were generated using the ggplot2 and ggbreak packages for R.

**Cell culture.** Aortic SMCs were purchased from Lonza (CC-2571), ATCC (PCS-100-012) and PromoCell (C-12533). An additional subset of six aortic SMC donor samples was a kind gift from M. Civelek at the University of Virginia. These additional aortic SMCs were isolated from aorta explants collected in the same region of the vessel at the University of California at Los Angeles transplant program as described previously[58]. Cells were maintained in Smooth Muscle Cell Basal Medium (SmBM, CC-3181, Lonza) supplemented with Smooth Muscle Medium-2 SingleQuots Kit (SmGM-2, CC-4149, Lonza) containing 20% FBS. For *MCAM*-knockdown experiments, vSMCs were cultured to 70% confluency and transfected with Silencer Select Pre-Designed *MCAM* siRNA (4392420, siRNA ID s8571; sense, 5′-CCAACGACCUGGGCAAAAtt-3′; antisense, 5′-UUUUUGCCCAGGUCGUUGGag-3′) or Silencer Select Negative Control No. 1 siRNA (4390843) using Lipofectamine RNAiMAX (13778150, Thermo Fisher Scientific). RNA isolation was performed on all donor samples 40 h after transfection using the RNeasy Mini Kit (74104, Qiagen).

**Immunoblotting preparation for vascular smooth muscle cell culture samples.** Human aortic vSMCs were cultured in six-well plates to confluency in the conditions mentioned above. Once confluent, cells were washed with 1× PBS before lysis in RIPA buffer (89901, Thermo Fisher Scientific) supplemented with cOmplete Protease Inhibitor (11697498001, Roche) and sodium orthovanadate (200-mM stock). Protein concentrations were determined using a colorimetric protein assay kit (22660, Thermo Fisher). Twenty micrograms of total protein was loaded onto a 4–20% Mini-PROTEAN TGX gel (456-1094, Bio-Rad) for SDS–PAGE and transferred to nitrocellulose membranes (170-4271, Bio-Rad) using the Trans-Blot Turbo Transfer System (Bio-Rad). Membranes were incubated with primary antibodies overnight followed by species-specific secondary antibodies for 1 h. Membranes were imaged on a Bio-Rad ChemiDoc Imaging System (Bio-Rad). Relative band intensities were quantified using Image Lab software.

**Immunoblotting preparation for aortic aneurysm samples.** Human aortic tissues were lysed in RIPA buffer (98065, Cell Signaling Technology) supplemented with cOmplete Protease Inhibitor Cocktail (11697498001, Roche). Samples were sonicated twice on ice at 50% power for 10 seconds (QSonica, Q125). Protein concentration was determined using a colorimetric protein assay kit (22660, Thermo Fisher). Thirty micrograms of total protein was loaded onto a 10% Mini-PROTEAN TGX gel (456-8034, Bio-Rad) for SDS–PAGE and transferred to polyvinylidene fluoride membranes (BR20180416, Bio-Rad) using a Trans-Blot Turbo Transfer System (Bio-Rad). Membranes were incubated with primary antibodies overnight followed by species-specific secondary antibodies for 1 h. Membranes were imaged on a Bio-Rad ChemiDoc Imaging System (Bio-Rad). Mean band intensities were normalized to those of the loading control and quantified using ImageJ software.

**RNA extraction, sequencing, mapping and quantification.** Total RNA was extracted using the RNeasy Mini Kit (Qiagen) and the RNase-free DNase Set. RNA integrity was assessed using the Agilent TapeStation, and samples with RIN > 9 were used for sequencing. Libraries were prepared with the Illumina TruSeq Stranded mRNA Library Prep Kit and sequenced on the Illumina NovaSeq X with 150-bp paired-end reads. Reads were aligned to the hg38 version of the human genome using STAR.

**Differential gene expression and functional enrichment analysis.** Differential expression analysis was performed using DESeq2 with a significance cutoff of $P_{adj} \leq 0.05$. To characterize the functional consequences of gene expression changes associated with loss of MCAM in vSMCs as well as the effects of sex, we performed gene ontology enrichment analysis using Metascape. Given the number of genes identified in the three major comparisons, we further filtered our gene list to focus on the top 3,000 differentially expressed genes (ranked by $P_{adj}$) for downstream analysis. $P_{adj}$ cutoffs were $7.92 \times 10^{-18}$ (sex invariant, combined), $3.03 \times 10^{-10}$ (female donors) and $1.01 \times 10^{-9}$ (male donors). The top three gene ontology categories identified per gene list for comparison were then used for visualization.

**Abdominal aortic aneurysm induction.** For AAA induction, 8-week-old male apolipoprotein-deficient (*Apoe*$^{-/-}$) mice and control littermates were implanted with Alzet osmotic pumps (model 2004, 0000298, Durect) subcutaneously to deliver a dosage of 1 µg per kg per min of PBS or AngII (H-1705, Bachem) for 28 d[32]. Aneurysm progression was monitored by Doppler ultrasound, and aortic measurements were quantified on images of the abdominal aorta captured before pump implantation and at the end of the experiment. For *Mcam*-KO experiments, AAA induction was performed as described above using *Apoe*$^{-/-}$;*Mcam*$^{+/+}$ and *Apoe*$^{-/-}$;*Mcam*$^{-/-}$ animals.

**Doppler ultrasound imaging.** Aortic diameter measurements were performed using a Vevo 2100 ultrasound imaging platform (FUJIFILM VisualSonics). Mice were anesthetized with 2% isoflurane inhalation and placed supine on a heated platform. The abdominal region was shaved, and images were acquired with a linear probe after application of an aquasonic ultrasound transmission gel (NC9861677, Parker Laboratories). Heart rate and basal temperature were monitored throughout the ultrasound imaging procedure. Pulse wave and color mode were used to identify the aortic pararenal region. Aortic diameter was acquired on the suprarenal region for the AngII model in the region of maximum aortic diameter. Measurements were performed by a blinded investigator.

**Elastin staining.** Aortas were embedded in optimal cutting temperature compound (4585, Fisher Scientific) and cryosectioned transversally at a thickness of 7 µm. Frozen cryosections were fixed in 10%

formalin before staining. Elastin was stained with Weigert's Resorcin Fuchsin (Electron Microscopy Sciences, 26370-01), and nuclear counterstaining was performed with 1% methyl green (R&D Systems, 4800-30-18) according to the manufacturer's instructions.

**Single-cell abdominal aortic aneurysm model analysis.** Datasets of scRNA-seq from the AAA *Apoe* + AngII mouse model and elastase-induced AAA and control sham aortas were analyzed as previously published[32]. For the AngII + *Apoe* model of AAA, alignment, barcode assignment and unique molecular identifier counting was performed with the Cell Ranger Single Cell Software Suite followed by downstream analysis in the R package Seurat. For elastase AAA model data, raw scRNA-seq data were processed using 10x Genomics Cell Ranger software followed by aggregation of samples. Seurat version 3.0 was used for cell filtration, normalization, principal-component analysis, variable gene finding, clustering analysis and UMAP dimensional reduction.

### Statistics and reproducibility

Sampling size for animal experiments was determined by power analysis with a type 1 error rate of 5% and a minimum detectable effect of 20%. The region-specific scRNA-seq experiment was performed twice with distinct cohorts of animals (eight mice per region per experiment). The final experiment presented in this study, included four anatomic locations per sex, resulting in eight independent libraries (four male and four female mice; GEO GSE255696). For sex-invariant gene signatures, both males and females at the indicated ages were used to control for both chromosome and hormone effects. However, for experiments assessing sex differences, mice were allocated by biological sex. For validation studies, site-enriched candidate reagents were tested using a separate subset of animals. For experiments for which the outcome was IF, the figures show representative images; however, the numbers of independent images and animals assessed are provided in the legend. For *MCAM* RNA-seq studies, cells from multiple donors from distinct vendor and collaborative sources were tested against three siRNAs. Silencer Select Pre-Designed *MCAM* siRNA (4392420, siRNA ID s8571) represents the siRNA with the most-consistent downregulation of *MCAM* across multiple donors. Investigators were blinded to allocation for outcome assessment.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

The scRNA-seq and RNA-seq datasets in this article are deposited in the international public repository Gene Expression Omnibus database under accession codes GSE255696 and GSE292995. Image data are available at BioImage Archive (https://www.ebi.ac.uk/biostudies/bioimages) under the accession code S-BIAD2075. For the data used in the AAA studies, these findings were initially characterized by Hadi et al.[32] and Zhao et al.[33]. Source data are provided with this paper.

## References


1. Pinard, A., Jones, G. T. & Milewicz, D. M. Genetics of thoracic and abdominal aortic diseases. *Circ. Res.* **124**, 588–606 (2019).
2. Davis, F. M., Daugherty, A. & Lu, H. S. Updates of recent aortic aneurysm research. *Arterioscler. Thromb. Vasc. Biol.* **39**, e83–e90 (2019).
3. Quintana, R. A. & Taylor, W. R. Introduction to the compendium on aortic aneurysms. *Circ. Res.* **124**, 470–471 (2019).
4. Evans, G. H., Stansby, G. & Hamilton, G. Suggested standards for reporting on arterial aneurysms. *J. Vasc. Surg.* **15**, 456 (1992).
5. Isselbacher, E. M. Thoracic and abdominal aortic aneurysms. *Circulation* **111**, 816–828 (2005).
6. Pham, M. H. C. et al. Aortic aneurysms in a general population cohort: prevalence and risk factors in men and women. *Eur. Heart J. Cardiovasc. Imaging* **25**, 1235–1243 (2024).
7. AlSiraj, Y. et al. Monosomy X in female mice influences the regional formation and augments the severity of angiotensin II-induced aortopathies. *Arterioscler. Thromb. Vasc. Biol.* **41**, 269–283 (2021).
8. Cheung, C., Bernardo, A. S., Trotter, M. W., Pedersen, R. A. & Sinha, S. Generation of human vascular smooth muscle subtypes provides insight into embryological origin-dependent disease susceptibility. *Nat. Biotechnol.* **30**, 165–173 (2012).
9. Owens, A. P. 3rd et al. Angiotensin II induces a region-specific hyperplasia of the ascending aorta through regulation of inhibitor of differentiation 3. *Circ. Res.* **106**, 611–619 (2010).
10. Topouzis, S. & Majesky, M. W. 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-β. *Dev. Biol.* **178**, 430–445 (1996).
11. Gadson, P. F. Jr. et al. Differential response of mesoderm- and neural crest-derived smooth muscle to TGF-β1: regulation of c-myb and α1 (I) procollagen genes. *Exp. Cell Res.* **230**, 169–180 (1997).
12. Majesky, M. W. Developmental basis of vascular smooth muscle diversity. *Arterioscler. Thromb. Vasc. Biol.* **27**, 1248–1258 (2007).
13. Wasteson, P. et al. Developmental origin of smooth muscle cells in the descending aorta in mice. *Development* **135**, 1823–1832 (2008).
14. Harmon, A. W. & Nakano, A. Nkx2-5 lineage tracing visualizes the distribution of second heart field-derived aortic smooth muscle. *Genesis* **51**, 862–869 (2013).
15. Sawada, H., Rateri, D. L., Moorleghen, J. J., Majesky, M. W. & 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.* **37**, 1722–1726 (2017).
16. MacFarlane, E. G. et al. Lineage-specific events underlie aortic root aneurysm pathogenesis in Loeys–Dietz syndrome. *J. Clin. Invest.* **129**, 659–675 (2019).
17. Chang, H. Y. Anatomic demarcation of cells: genes to patterns. *Science* **326**, 1206–1207 (2009).
18. Trigueros-Motos, L. et al. Embryological-origin-dependent differences in homeobox expression in adult aorta: role in regional phenotypic variability and regulation of NF-κB activity. *Arterioscler. Thromb. Vasc. Biol.* **33**, 1248–1256 (2013).
19. Visconti, R. P. & Awgulewitsch, A. Topographic patterns of vascular disease: HOX proteins as determining factors? *World J. Biol. Chem.* **6**, 65–70 (2015).
20. Pruett, N. D. et al. Changing topographic Hox expression in blood vessels results in regionally distinct vessel wall remodeling. *Biol. Open* **1**, 430–435 (2012).
21. Kalluri, A. S. et al. Single-cell analysis of the normal mouse aorta reveals functionally distinct endothelial cell populations. *Circulation* **140**, 147–163 (2019).
22. Pasek, R. C. et al. Coiled-coil domain containing 42 (Ccdc42) is necessary for proper sperm development and male fertility in the mouse. *Dev. Biol.* **412**, 208–218 (2016).
23. Fung, K. et al. Genome-wide association study identifies loci for arterial stiffness index in 127,121 UK Biobank participants. *Sci. Rep.* **9**, 9143 (2019).
24. den Hoed, M. et al. Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. *Nat. Genet.* **45**, 621–631 (2013).
25. Hoffmann, T. J. et al. Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat. Genet.* **49**, 54–64 (2017).
26. Pirruccello, J. P. et al. Deep learning enables genetic analysis of the human thoracic aorta. *Nat. Genet.* **54**, 40–51 (2022).

27. Pirruccello, J. P. et al. The genetic determinants of aortic distention. *J. Am. Coll. Cardiol.* **81**, 1320–1335 (2023).
28. Martin, S. S. et al. 2024 Heart Disease and Stroke Statistics: a report of US and global data from the American Heart Association. *Circulation* **149**, e347–e913 (2024).
29. Geerkens, C. et al. The X-chromosomal human biglycan gene *BGN* is subject to X inactivation but is transcribed like an X–Y homologous gene. *Hum. Genet.* **96**, 44–52 (1995).
30. Mazagova, M. et al. Growth differentiation factor 15 impairs aortic contractile and relaxing function through altered caveolar signaling of the endothelium. *Am. J. Physiol. Heart Circ. Physiol.* **304**, H709–H718 (2013).
31. Wang, L. et al. Mutations in myosin light chain kinase cause familial aortic dissections. *Am. J. Hum. Genet.* **87**, 701–707 (2010).
32. Hadi, T. et al. Macrophage-derived netrin-1 promotes abdominal aortic aneurysm formation by activating MMP3 in vascular smooth muscle cells. *Nat. Commun.* **9**, 5022 (2018).
33. Zhao, G. et al. Single-cell RNA sequencing reveals the cellular heterogeneity of aneurysmal infrarenal abdominal aorta. *Cardiovasc. Res.* **117**, 1402–1416 (2021).
34. Yoshida, T. & Owens, G. K. Molecular determinants of vascular smooth muscle cell diversity. *Circ. Res.* **96**, 280–291 (2005).
35. Liu, M. & Gomez, D. Smooth muscle cell phenotypic diversity. *Arterioscler. Thromb. Vasc. Biol.* **39**, 1715–1723 (2019).
36. Allahverdian, S., Chaabane, C., Boukais, K., Francis, G. A. & Bochaton-Piallat, M. L. Smooth muscle cell fate and plasticity in atherosclerosis. *Cardiovasc. Res.* **114**, 540–550 (2018).
37. Feil, S. et al. Transdifferentiation of vascular smooth muscle cells to macrophage-like cells during atherogenesis. *Circ. Res.* **115**, 662–667 (2014).
38. Speer, M. Y. et al. Smooth muscle cells give rise to osteochondrogenic precursors and chondrocytes in calcifying arteries. *Circ. Res.* **104**, 733–741 (2009).
39. Yu, L. 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.* **7**, 125 (2022).
40. Xu, C. et al. Single-cell RNA sequencing reveals smooth muscle cells heterogeneity in experimental aortic dissection. *Front. Genet.* **13**, 836593 (2022).
41. Li, Y., LeMaire, S. A. & Shen, Y. H. Molecular and cellular dynamics of aortic aneurysms revealed by single-cell transcriptomics. *Arterioscler. Thromb. Vasc. Biol.* **41**, 2671–2680 (2021).
42. Sambasivan, R. & Tajbakhsh, S. Skeletal muscle stem cell birth and properties. *Semin. Cell Dev. Biol.* **18**, 870–882 (2007).
43. Leopold, J. A. & Loscalzo, J. Emerging role of precision medicine in cardiovascular disease. *Circ. Res.* **122**, 1302–1315 (2018).
44. Loos, R. J. F. 15 years of genome-wide association studies and no signs of slowing down. *Nat. Commun.* **11**, 5900 (2020).
45. Lappalainen, T. & MacArthur, D. G. From variant to function in human disease genetics. *Science* **373**, 1464–1468 (2021).
46. Andersson, C., Johnson, A. D., Benjamin, E. J., Levy, D. & Vasan, R. S. 70-year legacy of the Framingham Heart Study. *Nat. Rev. Cardiol.* **16**, 687–698 (2019).
47. Bossone, E. & Eagle, K. A. Epidemiology and management of aortic disease: aortic aneurysms and acute aortic syndromes. *Nat. Rev. Cardiol.* **18**, 331–348 (2021).
48. Chen, J. et al. CD146 coordinates brain endothelial cell–pericyte communication for blood–brain barrier development. *Proc. Natl. Acad. Sci. USA* **114**, E7622–E7631 (2017).
49. Leroyer, A. S. et al. CD146 (cluster of differentiation 146). *Arterioscler. Thromb. Vasc. Biol.* **39**, 1026–1033 (2019).
50. Roostalu, U. et al. Distinct cellular mechanisms underlie smooth muscle turnover in vascular development and repair. *Circ. Res.* **122**, 267–281 (2018).
51. Kwon, S. T. et al. Interaction of expanding abdominal aortic aneurysm with surrounding tissue: retrospective CT image studies. *J. Nat. Sci.* **1**, e150 (2015).
52. Erhart, P. et al. Gene expression profiling in abdominal aortic aneurysms after finite element rupture risk assessment. *J. Endovasc. Ther.* **24**, 861–869 (2017).
53. Fashandi, A. Z. et al. Female mice exhibit abdominal aortic aneurysm protection in an established rupture model. *J. Surg. Res.* **247**, 387–396 (2020).
54. Jouve, N. et al. CD146 mediates VEGF-induced melanoma cell extravasation through FAK activation. *Int. J. Cancer* **137**, 50–60 (2015).
55. Ma, F., Hernandez, G. E., Romay, M. & Iruela-Arispe, M. L. Single-cell RNA sequencing to study vascular diversity and function. *Curr. Opin. Hematol.* **28**, 221–229 (2021).
56. Lun, A. T. L. et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* **20**, 63 (2019).
57. Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* **9**, giaa151 (2020).
58. Aherrahrou, R. et al. Genetic regulation of atherosclerosis-relevant phenotypes in human vascular smooth muscle cells. *Circ. Res.* **127**, 1552–1565 (2020).



## Acknowledgements

We thank the Northwestern Sequencing Core Facility and the Mouse Histology and Phenotyping Laboratory. We acknowledge the Lurie Comprehensive Cancer Center's Quantitative Data Sciences Core, especially M. Kocherginsky and E. T. Bartom, for statistical analysis assistance. We also thank B. E. Stranger (Northwestern University) for her insight on GTEx data and sex differences. We acknowledge the kind gifts of *Mef2c*$^{Cre}$/ROSA26R$^{mT/mG}$ mice from A. Daugherty (University of Kentucky) and *Wnt1*$^{Cre}$ mice from B. Thomson (Northwestern University) for lineage-tracing experiments. We also acknowledge the kind gift of vSMC lines from M. Civelek (University of Virginia). We thank members of the Arispe laboratory for valuable discussions. This work was supported by funds from the Ruth L. Kirschstein National Research Service Award (T32 HL069766 and T32 HL1346334) to M.C.R. and grants from the National Institutes of Health to M.L.I.-A. (R35HL140014) and B.R. (R01HL146627). M.S. is supported by an American Heart Association postdoctoral fellowship (907602).


## Author contributions

M.C.R. designed and performed experiments and wrote and edited the paper. F.M. performed bioinformatic analysis. A.M., M.S., G.E.H., J.S., A.L.W., M.V., N.B. and M.B.-C. performed experiments. A.S.L., H.A.-O. and B.R. provided intellectual discussion. M.L.I.-A. conceived the study, designed experiments and wrote and edited the paper. All authors had the opportunity to comment on the final paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Extended data** is available for this paper at https://doi.org/10.1038/s44161-025-00692-4.

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s44161-025-00692-4.

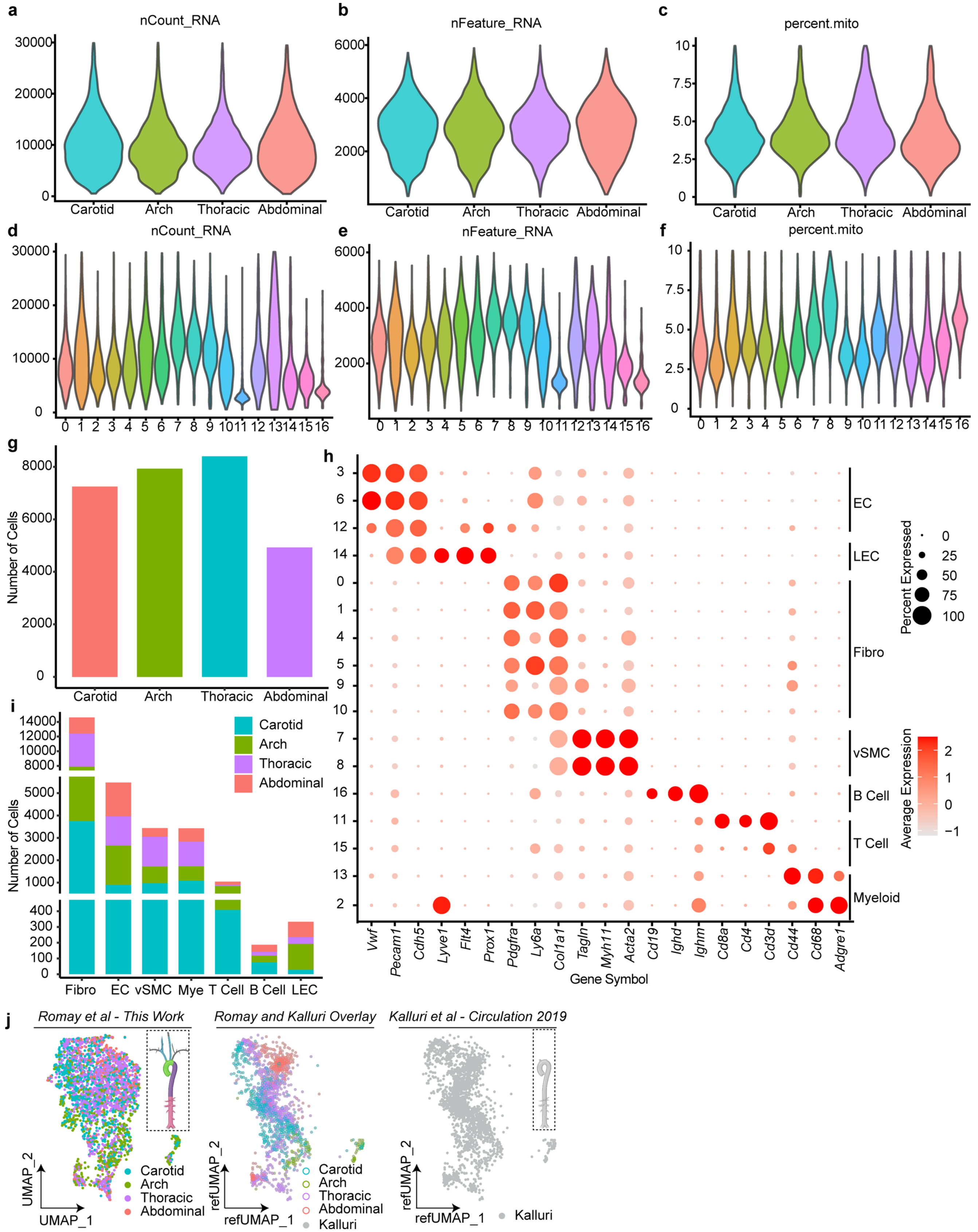
**Correspondence and requests for materials** should be addressed to M. Luisa Iruela-Arispe.

**Peer review information** *Nature Cardiovascular Research* thanks Sanjay Sinha and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at www.nature.com/reprints.



[1]Department of Cell and Developmental Biology, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [2]Division of Vascular and Endovascular Surgery, Department of Surgery, New York University Langone Health, New York, NY, USA. [3]Department of Cell Biology, New York University Langone Health, New York, NY, USA. [4]Molecular Biology Institute, University of California, Los Angeles, Los Angeles, CA, USA. [5]Department of Biological Sciences, Northwestern University, Evanston, IL, USA. [6]INSERM U970, Cardiovascular Research Center, Université de Paris, Paris, France. [7]INSERM 1263, INRA 1260, Aix-Marseille Université, Marseille, France. [8]Medical Intensive Care Unit, Hopital Saint-Antoine, Assistance Publique—Hopitaux de Paris, Sorbonne Universite, Paris, France. ✉e-mail: arispe@northwestern.edu

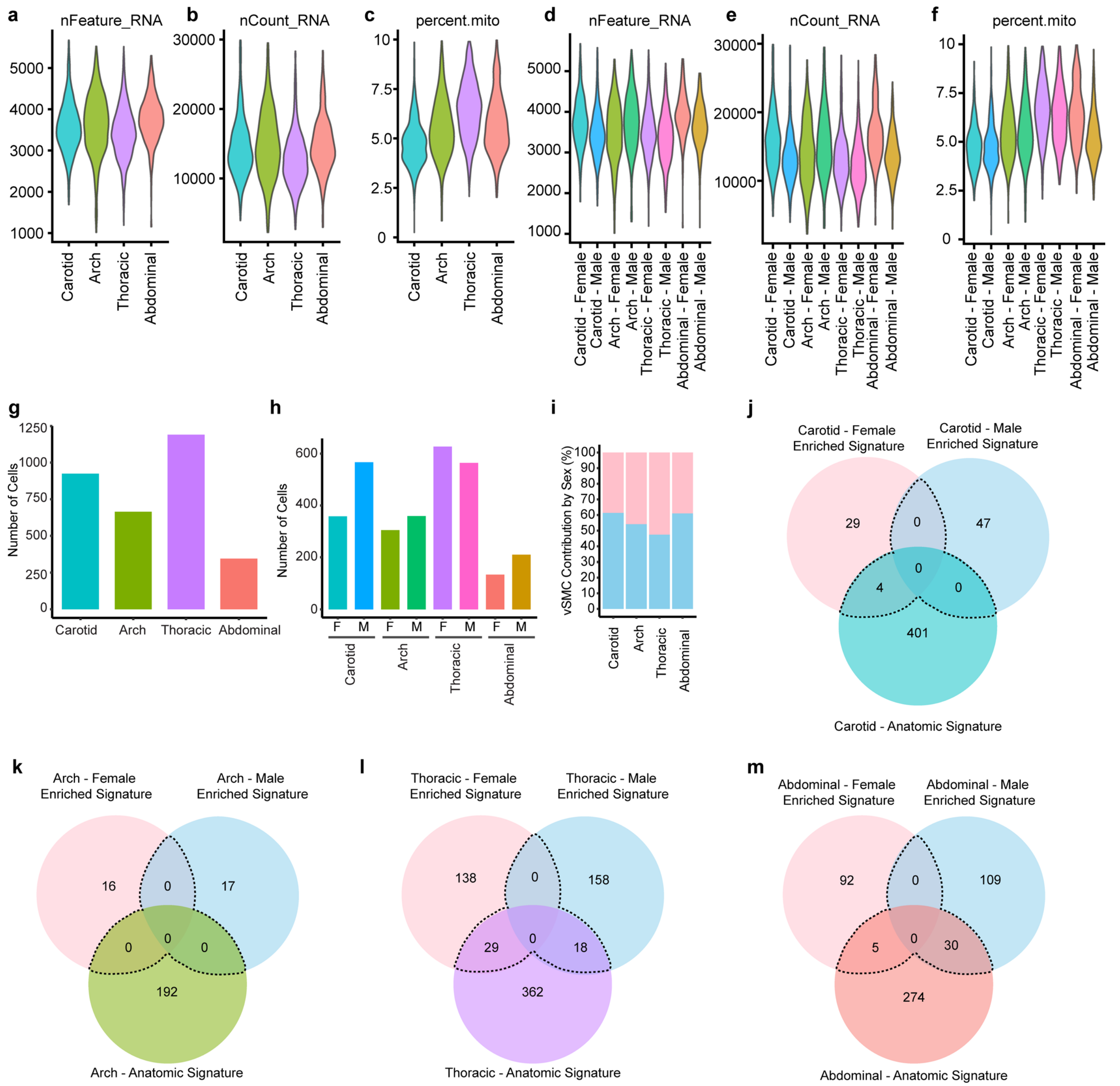



**Extended Data Fig. 1 | See next page for caption.**

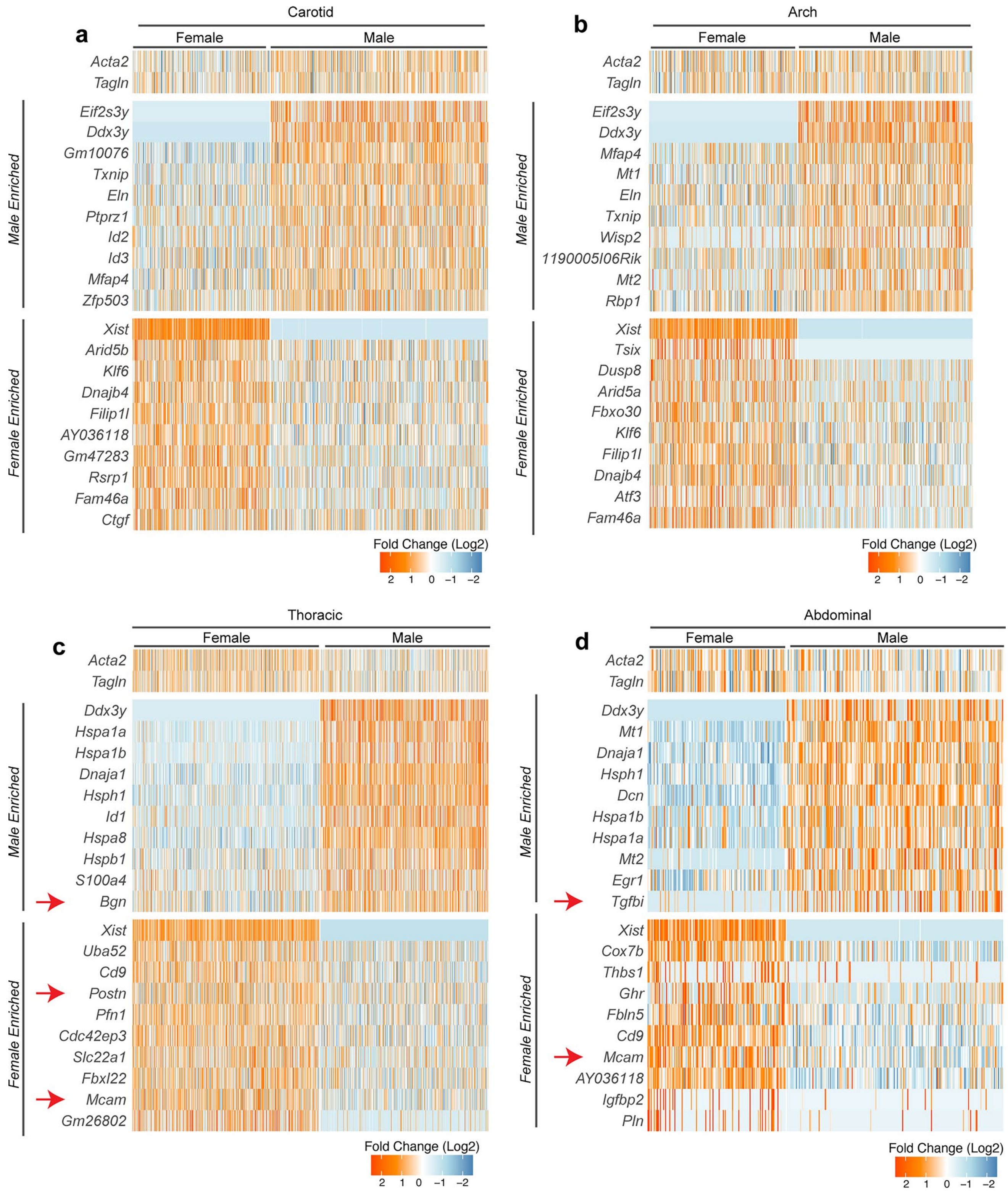
 https://doi.org/10.1038/s44161-025-00692-4

**Extended Data Fig. 1 | QC Metrics for scRNA-seq for all cells. a-c**, Violin plots representing number of genes detected in each cell (nFeature_RNA) **(a)**, total number of RNA molecules detected within a cell (nCount_RNA) **(b)**, and percentage of transcript counts arising from the mitochondrial genome (percent.mito) **(c)** per each anatomic location. **d-e**, Violin plots representing number of genes detected in each cell (nFeature_RNA) **(d)**, total number of RNA molecules detected within a cell (nCount_RNA) **(e)**, and percentage of transcript counts arising from the mitochondrial genome (percent.mito) **(f)** for each individual cell cluster identified as determined using Seurat. **g**, Bar graph representing the number of cells analyzed per anatomic location. **h**, Dot plot visualization of classical markers (three per cell type) in all cells in the 4 anatomic locations. The cells are subdivided into clusters as determined by Seurat**. i**, bar plot visualizing the number of cells per cell type and anatomic location. **j**, Overlay of scRNA-seq data generated in this study with Kalluri et al.[21]. For **j** vessel schematics in corner indicate the segments of the aorta and large arteries taken for sequencing Abbreviations, EC – endothelial cell, LEC – lymphatic enodthelial cell, Fibro – fibroblast, vSMC – vascular smooth muscle cell, Mye – myeloid. Aortic schematic in panel **j** created using BioRender.com.

https://doi.org/10.1038/s44161-025-00692-4

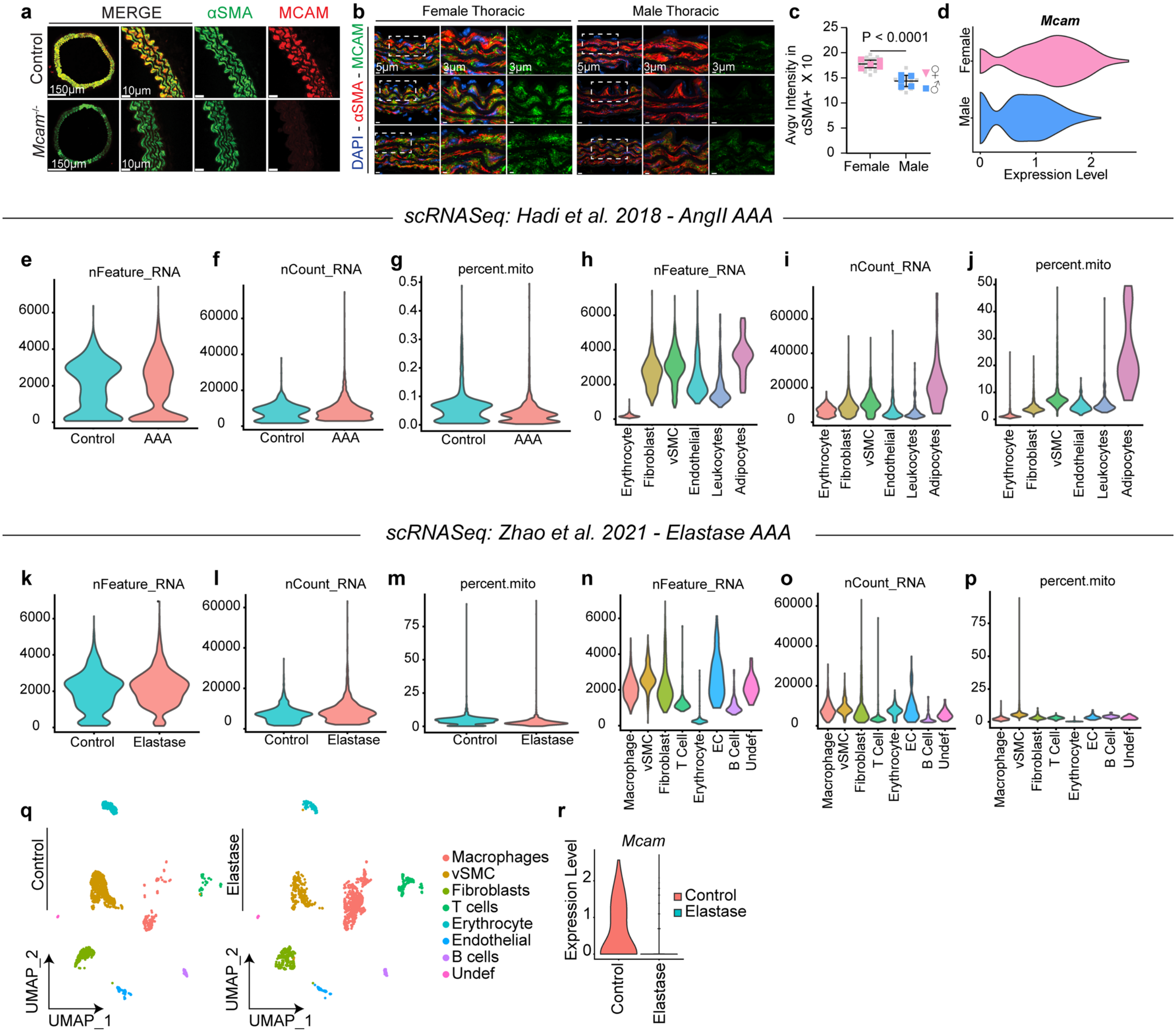



**Extended Data Fig. 2 | QC metrics for scRNA-seq data – vSMC only analysis. a-c**, Violin plots representing number of genes detected in each cell (nFeature_RNA) **(a)**, total number of RNA molecules detected within a cell (nCount_RNA) **(b)**, and percentage of transcript counts arising from the mitochondrial genome (percent.mito) **(c)** per each anatomic location. **d-f**, Violin plots representing number of genes detected in each cell (nFeature_RNA) **(d)**, total number of RNA molecules detected within a cell (nCount_RNA) **(e)**, and percentage of transcript counts arising from the mitochondrial genome (percent.mito) **(f)**, for each individual cell cluster identified by sex and anatomic site. **g**, Bar plot visualizing the number of vSMCs analyzed by anatomic location. **h**, Bar plot visualizing the number of vSMCs analyzed by anatomic location and biological sex. **i**, Stacked barplot visualizing the contribution of vSMCs by sex. **j**, Venn diagram visualizing the overlap between carotid vSMC anatomic signature, female-enriched carotid vSMC genes and male- enriched carotid vSMC signature. **k**, Venn diagram visualizing the overlap between the aortic arch vSMC anatomic gene signature, female-enriched aortic arch vSMC genes and male- enriched aortic arch vSMC genes. **l**, Venn diagram visualizing the overlap between the thoracic aorta vSMC anatomic signature, female-enriched thoracic vSMC genes and male - enriched thoracic vSMC signature. **m**, Venn diagram visualizing the overlap between the abdominal aorta vSMC anatomic signature, female-enriched abdominal aorta vSMC genes and male- enriched abdominal vSMC signature




**Extended Data Fig. 3 | Top sex-enriched vSMC genes per anatomic location.** **a**, Heatmap visualizing the top 10 sex-enriched genes in carotid vSMCs, per sex. **b**, Heatmap visualizing the top 10 sex-enriched genes in aortic arch vSMCs, per sex. **c**, Heatmap visualizing the top 10 sex-enriched genes in thoracic vSMC, per sex. **d**, Heatmap visualizing the top 10 sex-enriched abdominal vSMCs. Red arrows indicate genes validated using IF.

 https://doi.org/10.1038/s44161-025-00692-4




**Extended Data Fig. 4 | *Mcam* in health and disease. a**, Immunofluorescent staining of aortic rings taken from control and $Mcam^{-/-}$ animals stained with MCAM (red) and alpha-smooth muscle actin (aSMA, green). **b**, Representative IF images of MCAM expression (green) in female and male thoracic vSMCs. **c**, Quantification of MCAM intensity in female and male vSMCs. Females n = 30 images, 5 animals and males n = 17 images, 5 animals, 47 images total. Error bars represent mean ± SD. Two-tailed unpaired t-test with Welch's correction. p < 0.0001 **d**, Violin plot visualizing the expression pattern of *Mcam* in thoracic vSMCs by sex. **e-g**, Violin plots representing: **(e)** number of genes detected in each cell (nFeature_RNA), **(f)** total number of RNA molecules detected within a cell (nCount_RNA), **(g)** and percentage of transcript counts arising from the mitochondrial genome (percent.mito) in control and AngII-driven abdominal aortic aneurysm (AAA) model. **h-j**, Violin plots representing: **(h)** number of genes detected in each cell (nFeature_RNA), **(i)** total number of RNA molecules detected within a cell (nCount_RNA), and **(j)** percentage of transcript counts arising from the mitochondrial genome (percent.mito) for each individual cell type in the AAA dataset using Seurat. **k-m**, Violin plots representing: **(k)** number of genes detected in each cell (nFeature_RNA), **(l)** total number of RNA molecules detected within a cell (nCount_RNA) **(m)**, and percentage of transcript counts arising from the mitochondrial genome (percent.mito) in control and elastase-treatment induced AAA model. **n-p**, Violin plots representing: **(n)** number of genes detected in each cell (nFeature_RNA), **(o)** total number of RNA molecules detected within a cell (nCount_RNA), **(p)** and percentage of transcript counts arising from the mitochondrial genome (percent.mito) for each individual cell type in the elastase dataset using Seurat. **q**, UMAP plot visualizing cell distribution by celltype and dataset in control or elastase-treated aortas. **r**, Violin plot visualizing the expression of *Mcam* in control or elastase treated aortas. Data **e-g** represents a reanalysis of Hadi 2019, whereas Data **k-r** represents a reanalysis of Zhao[33]. Abbreviations, EC - endothelial cell, vSMC - vascular smooth muscle cell.

nature portfolio

Corresponding author(s): M. Luisa Iruela-Arispe

Last updated by author(s): 07/01/2025

# Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | scRNAseq libraries were generated using the Chromium Single Cell 3' Library and Gel Bead Kit v3 (10X Genomics) and sequences on an Illumina NovaSeq 6000 sequencer. Mouse blood vessel section images were acquired using spinning disk confocal microscope (Nikon Yokogawa CSU-W1) equipped with 20X air and 40X, 100X oil objectives. For a subset of candidates (Mcam), samples were evaluated and photographed using an LSM880 confocal microscope (Carl Zeiss) equipped with Zeiss Plan-Apochromat 20x/0.8 M27 for acquisition.<br>For human samples, images were acquired on a Zeiss LSM 710 confocal microscope (Carl Zeiss) using the Zeiss Efficient Navigation software (Carl Zeiss).Western Blot images were acquired on a Bio-Rad ChemiDoc imaging system. |
| Data analysis | scRNA-seq libraries were processed using the Cell Ranger pipeline (10X Genomics) followed by R package Seurat (version 4.3.0). Gene ontology enrichment was performed using Metascape. Data visualization for all scRNA-seq and gene ontology was performed with ggplot2, igraph, ggraph, ggbreak and ggalluvial. For immunofluorescence images, images were quantified using IMARIS 10.0.0 and the values from quantification were moved into Graphpad Prism 10.0 for statistical analysis and visualization. For GWAS and eQTL, data was acquired from the GWAS catalog on 2023_09_30, from GTEx database v8 on 2024_01_30, from STARNET 2024_02_24. Metascape was used to convert mouse genomic identifiers to human and statistical analysis (fisher's exact test) was perfomed using the package stat in R version 4.2.2. and visualizations were generated with ggplot2, ggbreak and ggalluvial. Western blot images were quantified with ImageJ and ImageLab. For RNASeq reads were mapped with STAR 2.7.11a to the human genome (hg38). The counts for each gene were obtained using quantMode GeneCounts in STAR commands, and the other parameters during alignment were set to default. Differential Gene Expression analysis was performed with DESeq2. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The authors declare that all supporting data will be made available to researchers within the article [and its online supplemental files]. The scRNA-seq and RNASeq data sets in this article are deposited in the international public repository Gene Expression Omnibus database under accession code GSE255696 and GSE292995. The image data is available at Bioimages Archive (https://www.ebi.ac.uk/biostudies/bioimages) under the accession code: S-BIAD2075. All supporting data are available within the article and the Data Supplement. For the data used in the AAA studies, these findings were initially characterized in Hadi 2018(32) and Zhao 2021 (33).

## Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

| | |
|---|---|
| Reporting on sex and gender | For human derived samples, sex-based analysis was only performed in specific experiments and in specific samples (human derived primary cell lines) where within the given experiment, confirmation of biological sex could be determined, such as expression of Xist or Uty via RNASeq. |
| Reporting on race, ethnicity, or other socially relevant groupings | N/A |
| Population characteristics | Patients were not discriminated according to age, gender genotype information or past/current diagnosis |
| Recruitment | Aneurysmal tissue was collected from individuals undergoing open aortic aneurysm repair. Informed consent was obtained for each subject. The study was authorized with IRB approval number i16-01807. Healthy cadaver tissues from multi- organ donors who had been confirmed as brain-dead were provided by LiveOnNY organization (New York). |
| Ethics oversight | New York University Langone Medical Center Institutional Review Board (IRB): i16-01807. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | Sampling size for animal experiments was determined by power analysis with type 1 error rate of 5%, and a minimum detectable effect of 20%. For experiments where the outcome was immunofluorescence, the figures show representative images; however the number of independent images and animals assessed were provided in the legend. |
| Data exclusions | No samples or animals were excluded. |
| Replication | The region-specific scRNA – seq experiment was performed twice with distinct cohorts of animals (8mice/region/experiment). The final experiment presented in this manuscript, included 4 anatomic locations per sex, resulting in 8 independent libraries (4 male and 4 female mice) GEO GSE255696. For validation studies, site-enriched candidate reagents were tested using a separate subset of animals. For the MCAM RNASeq studies, cells from multiple donors from distinct vendor and collaborative sources were tested against three siRNAs. Silencer™ Select Pre-Designed MCAM siRNA (4392420, siRNA ID - s8571) represents the siRNA with the most consistent down regulation of MCAM across multiple donors. |
| Randomization | For sex-invariant gene signatures, both males and females at the indicated ages were used to control for both chromosome and hormone effects. However, for experiments assessing sex differences mice were allocated by biological sex. |
| Blinding | Investigators were blinded to allocation for outcome assessment. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
| --- | --- |
| ☐ | ☒ Antibodies |
| ☐ | ☒ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology and archaeology |
| ☐ | ☒ Animals and other organisms |
| ☒ | ☐ Clinical data |
| ☒ | ☐ Dual use research of concern |
| ☒ | ☐ Plants |

## Methods

| n/a | Involved in the study |
| --- | --- |
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Antibodies

| | |
| --- | --- |
| Antibodies used | The following antibodies were used in immunofluorescence experiments. alpha-smooth muscle actin (1:500), Sigma-Aldrich, Cat#F3777, AB_476977. Goat anti-mouse ACE polyclonal antibody, (1:200), R&D Systems Cat# AF1513, RRID:AB_354832. Rabbit Anti-Aggrecan, amino acids 1177-1326 Polyclonal antibody, Unconjugated (1:40), EDM Millipore Cat# AB1031, RRID:AB_90460. Mouse/Rat Butyrylcholinesterase/BCHE Antibody (1:100), R&D Systems Cat# AF9024, RRID:AB_3076728. Anti-Biglycan (Mouse/Rat) [LF-159] Antibody (1:100), Kerafast Cat# ENH020-FP, RRID:AB_2920701. Ccdc42 (1:200), Aviva Systems Biology Cat# ARP52735_P050, RRID:AB_3086800. CD146 antibody [EPR3208] (1/200), Abcam Cat# ab75769, RRID:AB_2143375. Anti-DAPK2 (1:200) EDM Millipore Cat# 07-1229, RRID:AB_11215120. Rabbit Anti-Human Progesterone Receptor Monoclonal Antibody, Unconjugated, Clone SP2 (1:100), Fisher Scientific Cat# RM-9102-S1, RRID:AB_149905. Rabbit Anti-Periostin Polyclonal Antibody, Unconjugated (1:100), Abcam Cat# ab14041, RRID:AB_2299859. Rspondin-3 antibody. (1:200), Epigentek Cat# A60722, RRID:AB_3086801. TGFBI / BIGH3 antibody(1:200); Proteintech Cat# 10188-1-AP, RRID:AB_2202311. Anti-Wif1 (1:200), Abcam Cat# ab155101, RRID: AB_3105780. Goat Anti-Rabbit IgG Antibody (H+L), Biotinylated (1:500) Vector Laboratories Cat# BA-1000, RRID:AB_2313606. Anti-Goat IgG (H+L), made in horse, Biotinylated (1:500), Vector Laboratories Cat# BA-9500, RRID:AB_2336123. Alexa Fluor 568 – donkey anti-rabbit IgG (1:500), Thermo-Fisher Cat#A10042, AB_2534017. Alexa Fluor 568 – donkey anti goat IgG (1:500), A11057, AB_142581. The following antibodies were used for western blot. CD146 antibody [EPR3208] (1/1000), Abcam Cat# ab75769, RRID:AB_2143375. Secondary - StarBright Blue 700 Goat Anti-Rabbit IgG (1:2000), Bio-Rad Cat# 12004161, RRID:AB_2721073). hFAB™ Rhodamine Anti-Actin Primary Antibody(1:4000), Bio-Rad Cat# 12004163, RRID:AB_2861334. |
| Validation | Antibodies were validated by the respective vendors for the applications described in this manuscript. Validation of CD146 were done by immunocytochemistry using a KO mouse side by side with a wild-type (control) mouse. |

## Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

| | |
| --- | --- |
| Cell line source(s) | Aortic smooth muscle cells were purchased from Lonza (CC-2571), ATCC (PCS-100-012) and Promo Cell (C-12533). An additional subset of 6 aortic smooth muscle cell donors from the University of Virginia. These additional aortic smooth muscle cells were isolated from the explants of ascending aorta at the University of California at Los Angeles (UCLA) transplant program as described previously (PMID: 1752961, 33040646). |
| Authentication | Expression of classical vascular smooth muscle markers using both RNASeq and WB were confirmed in all cell lines |
| Mycoplasma contamination | All Cell Lines were tested for Mycoplasma contamination and tested negative. |
| Commonly misidentified lines (See ICLAC register) | *Name any commonly misidentified cell lines used in the study and provide a rationale for their use.* |

## Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

| | |
| --- | --- |
| Laboratory animals | Female and Male C57BL6/J mice (IMSR_JAX:000664) at 12 weeks of age were used for the initial single cell experiments with an additional cohort of mice for sex at 12 weeks of age used for the immunofluorescence validation. For the lineage tracing experiments). Male Mef2c-Cre mice (MMRRC_030262-UNC) crossed with ROSA26RmT/mG (IMSR_JAX:007676) and littermate controls at 16 weeks of age. For neural crest linage tracing male and female 16 week old B6.Cg-H2az2Tg(Wnt1-cre)11Rth Tg(Wnt1-GAL4)11Rth/J crossed B6.Cg-Gt(ROSA)26Sortm14(CAG-tdTomato)Hze/J (IMSR_JAX:007914) were used. For studies using CD146/Mcam knock out mouse(pmid: 254497730), animals were backcrossed for more than 10 generations on the C57BL/6J background and used for AAA induction at 8 weeks of age. Primers for genotyping CD146/Mcam alleles were as follows:5’- |

TCACTTGACAGTGTGATGGT-3' (forward primer used to detect CD146/Mcam WT, floxed and KO alleles), 5'-CCTTAGAAAGCAGGGATTCA-3' (reverse primer used to detect CD146/Mcam WT and floxed alleles) and 5'-CCCAAATCCTCTGGAAGACA-3' (reverse primer used to detect CD146/Mcam KO allele).Genotyping primers for the ApoE alleles are as followed: 5'-GCCTAGCCGAGGGAGAGCCG-3', 5'-TGTGACTTGGGAGCTCTGCAGC -3' and 5'-GCCGCCCCGACTGCATCT-3'. Mice were housed at the University of California, Los Angeles (UCLA), Northwestern University, New York University Langone Health and at the Cardiovascular Research Center, Université de Paris, France.

Animals were maintained in ventilated racks, no more than 5 mice per cage with a 14/10 hour light/dark cycle. Diet and water were provided ab libitum. Room temperature was between 68 - 77 F and humidity between 30-70%.

| | |
|---|---|
| Wild animals | N/A |
| Reporting on sex | Male and Female mice were used in the initial scRNA-seq experiments and when possible in all validations experiments. |
| Field-collected samples | N/A |
| Ethics oversight | Northwestern University Institutional Animal Care and Use Committee - IS00013945 |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Plants

| | |
|---|---|
| Seed stocks | *Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.* |
| Novel plant genotypes | *Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.* |
| Authentication | *Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.* |