

Integrative functional genomics analysis identifies pleiotropic genes for vascular diseases

Received: 2 March 2025

Accepted: 29 January 2026

Published online: 05 February 2026

 Check for updates

Charles U. Solomon^{1,2,3,12}, David G. McVey^{1,2,3,12}, Catherine Andreadi^{1,2,3}, Peng Gong^{1,2,3}, Lenka Turner^{1,2,3}, Dedrick S. S. Song⁴, Heming Zhang⁴, Dominic P. Lee⁴, Elisavet Karamanavi^{1,2,3}, Wei Yang^{5,6}, Jiapeng Chu⁴, Runji Chen^{5,7}, Kim E. Haworth^{1,2,3}, Chukwuemeka George Anene-Nzelu^{4,8}, Hui Li⁴, Matthew J. Denniff^{1,2,3}, Peter Y. Li⁴, Yanhong Zhang⁵, Xiaoxin Huang⁵, Gavin E. Morris^{1,2,3}, Peter A. Greer⁹, Emma J. Stringer^{1,2,3}, Haojie Yu⁴, Roger S. Y. Foo⁴, Gillian Douglas^{10,11}, Nilesh J. Samani^{1,2,3,13}, Tom R. Webb^{1,2,3,13} & Shu Ye^{1,2,3,4,5,7,13} ✉

Several vascular diseases including coronary artery disease, hypertension, stroke, and abdominal aortic aneurysm, have significant genetic underpinnings. Genome-wide association studies have unveiled many genetic loci associated with one or more of these diseases. However, the causative genes at most of these loci are yet to be determined, which hampers the translation of the genetic findings into a better understanding of the disease mechanisms and the identification of new therapeutic targets. Here, in an integrative functional genomics analysis of these loci, we identify a panel of likely causal genes, some of which are pleiotropic for more than one of these vascular diseases. Pooled CRISPR knockout screen analyses of these likely causal genes indicate that many of them influence vascular smooth muscle cell behaviour, and validation experiments of selected genes confirm that *FES*, *BCAR1*, *CARF* and *SMARCA4* exert such effects. Further functional experiments focusing on *FES*, a pleiotropic gene for both coronary artery disease and hypertension, show that it modulates the expression of genes involved in vascular remodeling and that *Fes* knockout in mice promotes atherosclerosis as well as raises blood pressure. These findings provide an insight into the genetic basis of vascular diseases and inform targets for therapeutic development.

In recent years, genome-wide association studies (GWAS) have unveiled genomic loci where genetic variants are associated with vascular diseases such as coronary artery disease (CAD)^{1–4}, hypertension⁵, stroke^{6,7}, and abdominal aortic aneurysm (AAA)^{5,8}. To facilitate the medical translation of these genetic discoveries, it is important to identify and investigate the causal genes at these loci.

Vascular smooth muscle cells (VSMC) are a major player in vascular diseases. VSMC phenotypic switching, migration, proliferation and apoptosis play important roles in the pathogenesis of atherosclerosis, which promotes CAD, hypertension, stroke, and AAA⁹. Further, VSMC hypercontractility and vascular remodelling involving VSMC migration, proliferation and apoptosis contribute to elevated blood pressure¹⁰, VSMC loss and impaired contractility are hallmarks

of AAA¹¹, while VSMC degeneration causes cerebral small-vessel disease that can lead to ischaemic and haemorrhagic strokes¹².

As most disease-associated genetic variants revealed by GWASs are located in non-coding regions of the genome^{13,14}, a prevailing hypothesis is that these variants can alter the individual's disease risk through their effect on gene expression^{13,14}. Studies have shown that there is considerable inter-individual variation in VSMC behaviour, largely due to genetic variation among individuals^{15,16}. Interestingly, some vascular disease-associated genetic variants have been reported to modulate gene expression in VSMCs and affect VSMC behaviour¹⁵⁻¹⁸. Therefore, it is likely that the genetic influence on vascular disease acts partly through VSMCs.

Here, we report results from an integrative functional genomics analysis on VSMCs (outlined in Fig. 1), with the aim to identify causal genes at the genomic loci that have been found to be associated with CAD, hypertension, stroke and/or AAA in GWAS.

Results

Influences of vascular disease-associated genetic variants on gene expression

We recently established a large collection of VSMCs from different individuals ($n = 1486$), and on this cell bank, we performed genome-

wide genotyping and whole-genome RNA-sequencing (RNA-seq) to generate a comprehensive catalogue of expression quantitative trait loci (eQTL) with stringent statistical significance¹⁶. We compiled a list of tagging single nucleotide polymorphisms (SNPs) that marked the genetic variants/haplotypes reported to be associated with CAD in GWAS¹⁻⁴ and then interrogated these tagging SNPs in the VSMC eQTL catalogue. As reported previously¹⁶, we observed that ~60% of the CAD-associated SNPs had an eQTL effect on gene expression, with two-thirds of them affecting more than one gene.

Here we compiled lists of tagging SNPs that marked the genetic variants/haplotypes reported to be associated with blood pressure/hypertension^{5,19}, stroke⁵⁻⁷ and/or AAA^{5,8} in GWAS, and interrogated these tagging SNPs in the abovementioned VSMC eQTL catalogue. We found that ~59% of the blood pressure/hypertension-associated SNPs, ~55% of the stroke-associated SNPs, and ~59% of the AAA-associated SNPs, had eQTL effects in VSMCs (Supplementary Data 1-3 and Supplementary Fig. 1). Of these SNPs, ~66% showed eQTL effects on more than one gene. Of the associated genes, ~77% were protein-coding, and ~15% were long non-coding RNAs (lncRNAs) (highlighted in green and yellow, respectively, in Supplementary Data 1-3). When male samples ($n = 780$) and female samples ($n = 706$) were analysed separately, the eQTL effects of approximately 50% of CAD-, blood pressure/

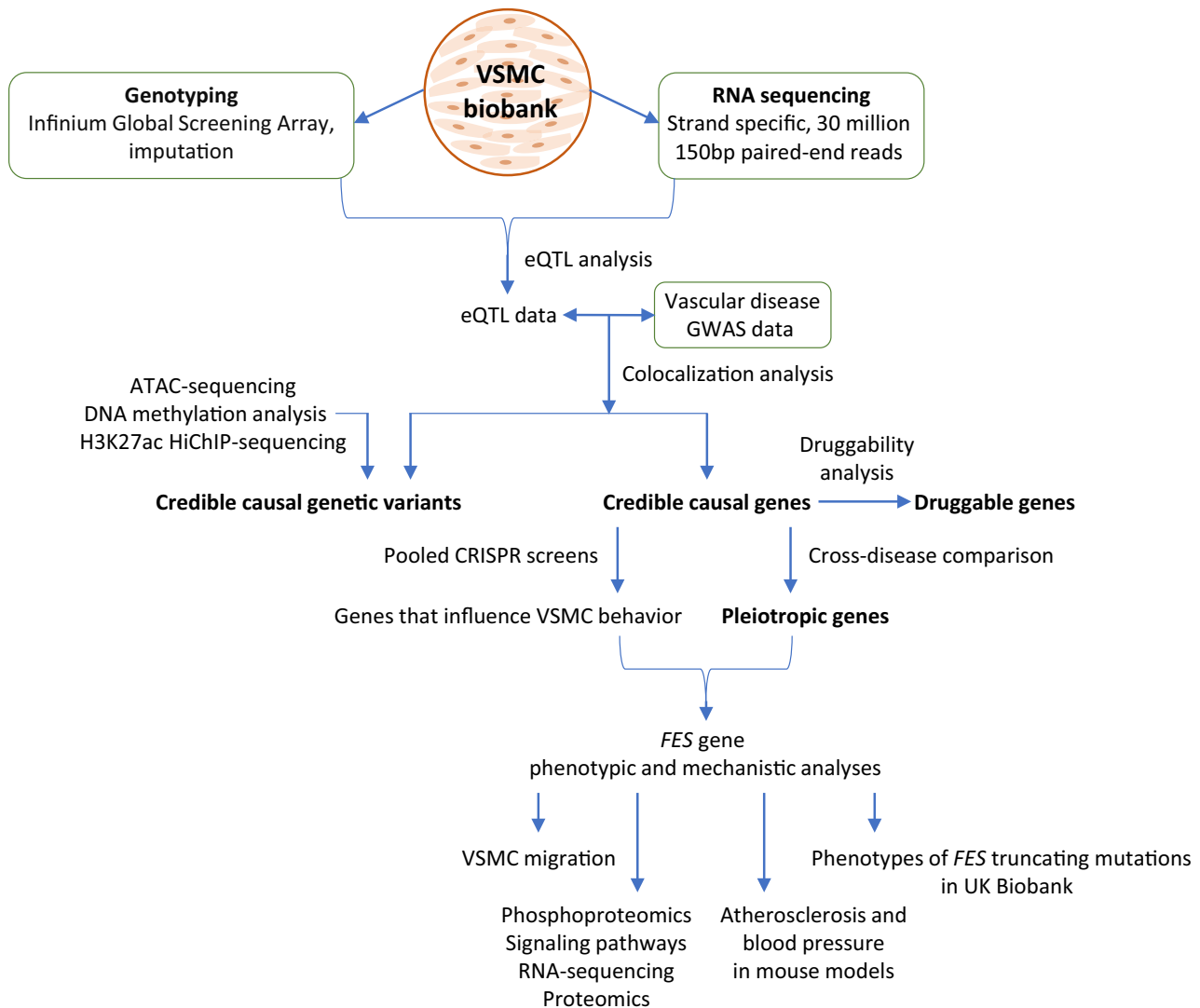


Fig. 1 | Flow diagram of the study design. ATAC-seq: assay for transposase-accessible chromatin with sequencing; eQTL: expression quantitative trait locus; GWAS: genome-wide association study; H3K27ac: histone 3 lysine 27

acetylation; HiChIP: HiC (High-throughput chromosome conformation capture) with ChIP-seq (Chromatin immunoprecipitation with sequencing); VSMC: vascular smooth muscle cells.

hypertension-, stroke-, or AAA-associated SNPs were still detected (Supplementary Data 1–4), despite the reduction in the sample sizes per group, indicating that they are common to both sexes.

Likely causative variants/genes for CAD, hypertension, stroke, and AAA

To explore causal variants and related causal genes for each of the vascular diseases mentioned above, we performed colocalization tests, including eCAVIAR (eQTL and GWAS CAusal Variant Identification in Associated Regions)²⁰ and SMR/HEIDI (summary-data-based Mendelian Randomisation/Heterogeneity in Dependent Instruments)²¹, utilising reported disease GWAS summary statistics in conjunction with the above-described VSMC eQTL dataset. These analyses detected panels of likely causal variants and likely causal genes for each of these vascular diseases. Supplementary Data 5–8 show the results of these colocalization analyses, including the identified likely causal variants (columns A–F), the identified likely causal genes together with eQTL data in relation to the likely causal variants (columns G–M), summary statistics of reported associations of the likely causal variants with the relevant vascular diseases (columns N–R), and the statistics of the eCAVIAR and SMR/HEIDI analyses (columns S–Y).

To investigate if any of the likely causal variants identified in the above colocalization analyses were located in transcriptionally active regions of the genome, we performed VSMC ATAC-seq (assay for transposase-accessible chromatin with sequencing) to delineate open chromatin regions, as well as DNA methylation assay and H3K27ac HiChIP-seq (histone H3 acetylation at lysine 27 high throughput chromosome conformation capture with chromatin immunoprecipitation with sequencing) since DNA demethylation and H3K27 acetylation upregulate gene expression. We interrogated the likely causal variants in regions of ATAC-seq peaks, low DNA methylation and H3K27ac, respectively. The analyses showed that a large proportion of the likely causal variants were located in regions of ATAC-seq open chromatin (Supplementary Data 9), low DNA-methylation (variants highlighted in yellow in Supplementary Data 10), and/or H3K27ac (Supplementary Data 11). The likely causal variants in relation to peaks of ATAC-seq, low DNA methylation and H3K27ac, as well as their associated eQTL genes and vascular diseases, are summarised in Supplementary Data 12. As an example, we found that the CAD-associated variant rs1894401 in the *FES* gene resided in an ATAC-seq open chromatin region with lessened DNA-methylation (Supplementary Figs. 2, 3) and with H3K27ac (Supplementary Fig. 4).

Because a genetic variant could possibly exert a long-range effect on the transcription of a gene located at a distance due to DNA looping, we utilised the VSMC H3K27ac HiChIP-seq data to investigate if any of the likely causal variants resided at a DNA interaction anchor, with a likely causal gene located at a linked DNA interaction anchor. This analysis detected multiple variant-gene pairs where a likely causal variant was detected within an anchor and a likely causal gene was found within a linked anchor (Supplementary Data 11 and Supplementary Fig. 4), suggesting that the likely causal gene was brought physically closer to the likely causal variant through DNA looping. For instance, the CAD-associated variant rs1894401 was detected in an anchor, and the *FES* gene 5'-region was observed in a linked anchor (Supplementary Fig. 4A).

The aforementioned colocalization analyses identified 134 likely causal genes for CAD, whose eQTL signals in VSMCs showed significant colocalization with CAD GWAS signals, including 84 genes that we reported previously¹⁶ and a further 50 genes now identified using more recently available CAD GWAS data¹ (Fig. 2, Supplementary Fig. 5A and Supplementary Data 5). Further, the analyses identified 74 likely causal genes for hypertension (Fig. 2, Supplementary Fig. 5B and Supplementary Data 6), 13 likely causal genes for stroke (Fig. 2, Supplementary Fig. 5C, and Supplementary Data 7), and 43 likely causal genes for AAA (Fig. 2, Supplementary Fig. 5D and Supplementary Data 8).

In addition to the abovementioned colocalization analyses using eQTL data from the VSMC bank, we performed additional colocalization analyses using eQTL data from coronary artery and aortic artery tissues (from the GTEx Portal) together with the disease GWAS statistics. Noteworthy, most of the likely causal genes identified in the colocalization analyses using eQTL data from the VSMC bank were also significant in the colocalization analyses using eQTL data from coronary artery and aortic artery tissues (Supplementary Data 13), further indicating that they are likely to be causal genes.

Gene ontology enrichment analyses of the likely causal genes identified showed enrichment in several major biological processes and pathways, including transforming growth factor-beta (TGF β) signalling in CAD (Supplementary Fig. 6), vascular endothelial growth factor (VEGF) signalling in hypertension (Supplementary Fig. 7), and signalling pathways of vascular endothelial growth factor receptor (VEGFR), platelet-derived growth factor receptor (PDGFR), integrin and cadherin in AAA (Supplementary Fig. 8), suggesting that the genetic influence on these vascular diseases is, at least partly, mediated by these genes acting through these biological pathways.

Pleiotropic genes implicated in more than one vascular disease

A cross-trait linkage disequilibrium score regression analysis of GWAS summary statistics revealed significant genetic correlations between the various vascular diseases investigated in this study, indicating shared genetic influences (Fig. 3A and Supplementary Data 14).

Further, comparing the likely causal genes identified for CAD, hypertension, stroke and AAA, respectively, we noticed that 18 genes were common for two or more of these diseases, suggesting that they have pleiotropic effects. Specifically, *ARNTL*, *CSNK2B*, *FES*, *SNF8*, *TARID*, and *TCF21* were detected for both CAD and hypertension; *CARF*, *NBEAL1*, and *PROCR* were common in CAD and stroke; *BCAR1*, *ILF3-DT*, *LPAL2*, and *SMARCA4* were shared by CAD and AAA; while *PLCE1-AS1* was found for CAD and hypertension as well as AAA (Fig. 3B). In addition, *ADO* was identified for both hypertension and AAA, whilst *SURF1* and 2 lncRNA genes were detected for both hypertension and stroke (Fig. 3B).

Deleterious protein-coding variants in likely causal genes

The above colocalization analyses focused on regulatory SNPs that influenced gene expression. Next, we further explored if there also existed deleterious variants (such as frameshift mutations and deleterious missense substitutions) in the protein-coding regions of the likely causal genes that we had identified and if these variants were associated with the diseases under this study, because such evidence would further indicate that these genes were causal. Thus, we interrogated the likely causal genes in databases from the UK Biobank^{22,23} and Finnish Biobank²⁴ (through the AstraZeneca PheWAS Portal²⁵ and the PheWeb browser, respectively) for deleterious coding variants identified by exome sequencing. This interrogation identified such variants in several of the likely causal genes, including the *ADAMTS7*, *FES*, *ICAIL*, *MIA3*, and *MTAP* genes in relation to CAD and the *ADO*, *FES*, and *MAP4* genes in association with hypertension (Supplementary Data 15). However, no coding variants in any gene were associated with stroke or AAA at the genome-wide significance level ($P < 1 \times 10^{-8.7}$ in PheWAS and $P < 5 \times 10^{-8}$ in PheWeb) in the UK Biobank and Finnish Biobank databases.

Druggability of likely causal genes

An interrogation of the likely causal genes in the Drug Gene Interaction Database (DGIdb)²⁶ identified 63 potentially druggable targets for CAD, 39 for hypertension, 5 for stroke, and 25 for AAA (Fig. 4 and Supplementary Data 16–23). Notably, 5 of these targets were applicable for both CAD and hypertension, 2 for both CAD and stroke, 2 for both CAD and AAA, and 1 for both hypertension and AAA (Fig. 4).

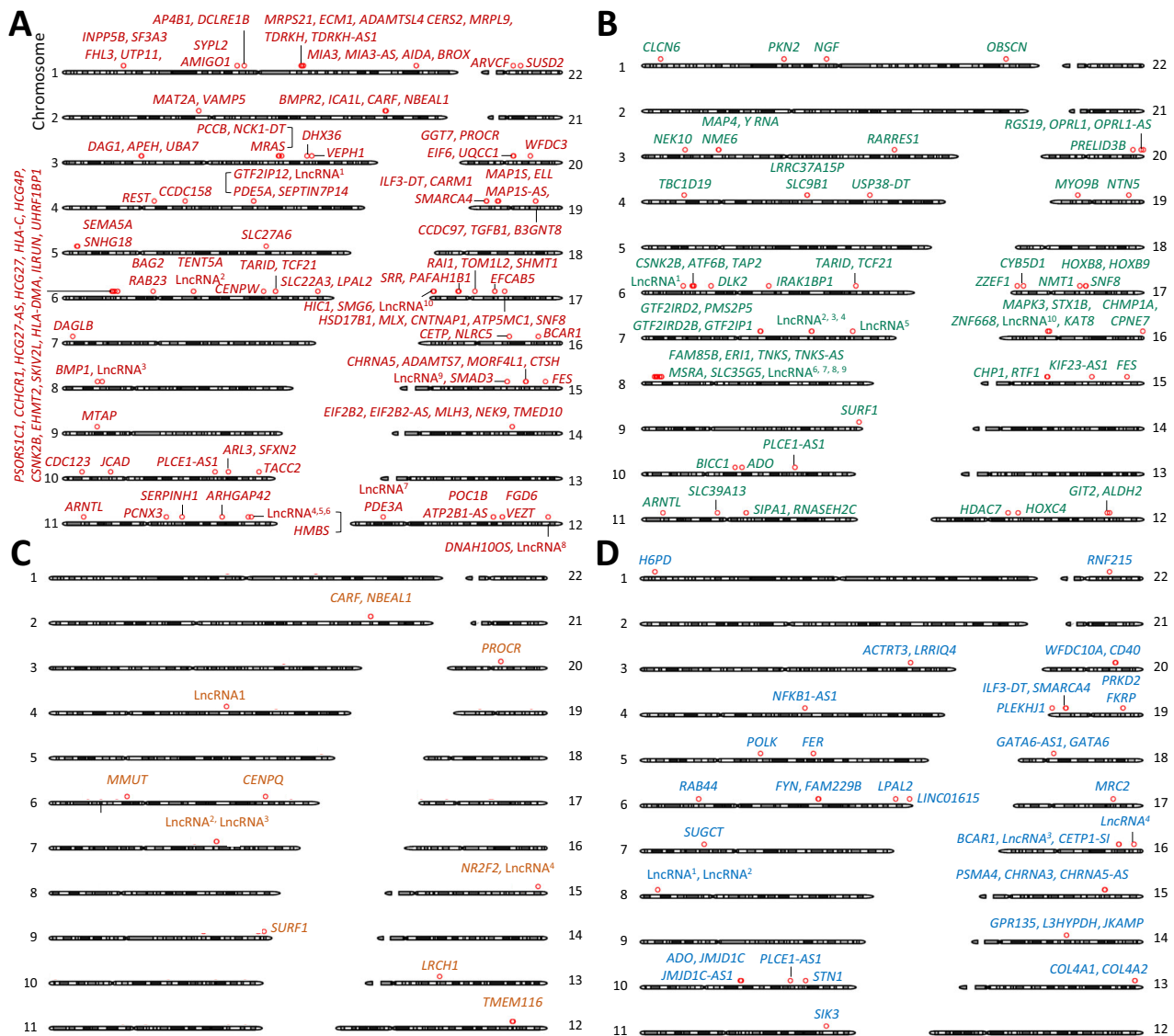


Fig. 2 | Credible causal genes. Chromosomal plots of credible causal genes for CAD (A), hypertension (B), stroke (C) and AAA (D), respectively. The identities of the long non-coding RNA (lncRNA) genes are described in Supplementary Data 5–8.

Among these potential targets, 38 encode proteins for which there are existing drugs²⁶ (Fig. 4 and Supplementary Data 17,19,21 & 23).

Influence of likely causal genes on VSMC behaviour

To assess possible effects of the likely causal genes on VSMC behaviour, we performed high-throughput pooled CRISPR-Cas9-nuclease knockout screens targeting the protein-coding genes among these likely causal genes (this technique is unable to test non-coding RNA genes). The results of these screens suggested that some of these genes might influence VSMC proliferation or migratory ability (Supplementary Data 24, 25, and Fig. 5A, B). A selection of these genes was subjected to validation experiments, with the following selection criteria: (1) among the top 10 genes ranked by score in the pooled CRISPR-Cas9-nuclease knockout screens; (2) considered to be a pleiotropic gene for more than one vascular disease in the aforementioned cross-trait analysis; and (3) likely to be druggable in the abovementioned druggability analysis. In the validation experiments, VSMCs were transfected with either siRNA for the selected gene or negative control siRNA, followed by a VSMC proliferation or migration

assay. The following genes met the selection criteria and were tested in the validation experiments: *BCARI*, *CARF*, and *SMARCA4* for VSMC proliferation, and *CSNK2B* and *FES* for VSMC migration. The validation experiments confirmed the CRISPR-Cas9-nuclease knockout screen findings for *BCARI*, *CARF*, *FES*, and *SMARCA4*, with the same direction of effect in both the screen and validation experiments (Fig. 5C, D and Supplementary Fig. 9A, B), but not for *CSNK2B* (Supplementary Fig. 9C). For instance, the CRISPR-Cas9-nuclease screen showed that *FES* knockout increased VSMC migration, and this finding was supported in the validation experiment where we observed that siRNA-induced knockdown of *FES* increased the migratory ability of VSMCs (Fig. 5D).

Effect of FES on VSMC migration and production of matrix metalloproteinases

Following the validation experiments, we carried out further functional experiments with a particular focus on *FES*. As the *FES* protein has been reported to participate in intracellular signalling²⁷, we investigated if *FES* knockdown might influence certain signalling

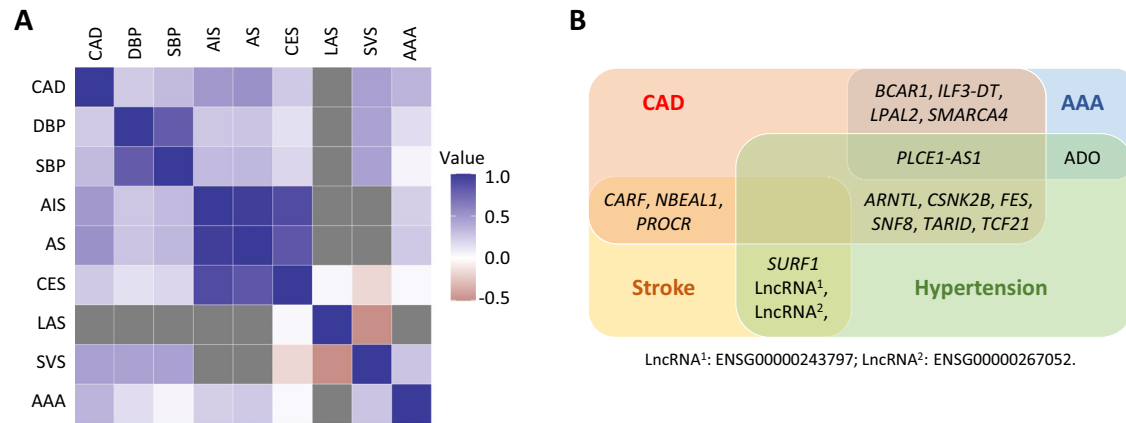


Fig. 3 | Genetic correlations and pleiotropic genes. A Genetic correlations between CAD, blood pressure, hypertension, various types of stroke, and AAA. Abbreviations: AAA abdominal aortic aneurysm; AIS arterial ischaemic stroke; AS all stroke; CAD coronary artery disease; CES cardioembolic stroke; DBP diastolic blood

pressure; LAS large artery stroke; SBP systolic blood pressure; SVS small vessel stroke. Grey box indicates data unavailable. **B** Names of candidate causal genes shared by more than one vascular disease.

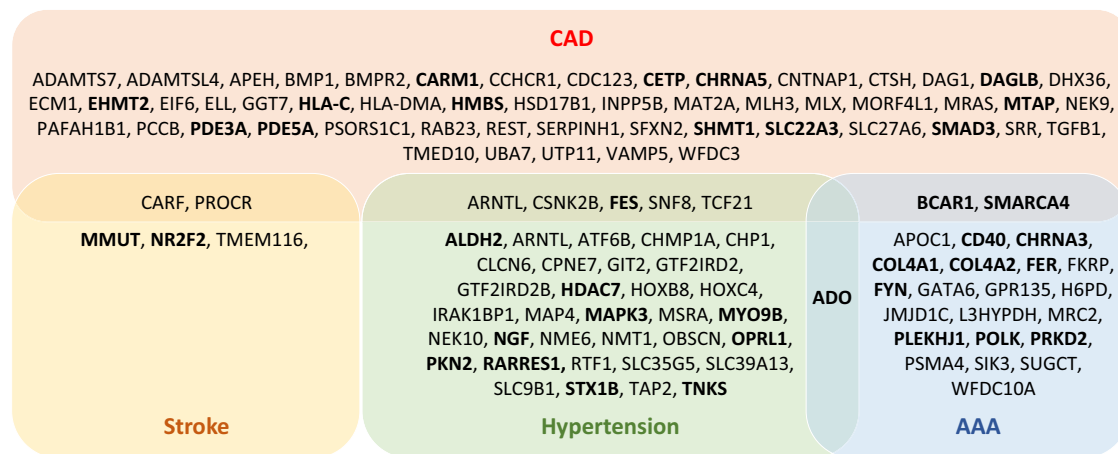


Fig. 4 | Druggability of credible causal genes. The boxes show the names of genes that encode potentially druggable proteins. Bold font indicates genes that encode proteins for which there are existing drugs.

pathways and result in changes in the production of certain proteins. Thus, we performed whole-genome RNA-seq, quantitative proteomics and phosphoproteomics assays on VSMCs with and without siRNA-induced *FES* knockdown. The phosphoproteomics analysis identified a number of proteins whose phosphorylation status might be influenced by *FES* knockdown (Supplementary Data 26) and a pathway analysis of these proteins suggested that they were enriched in several pathways including the epidermal growth factor (EGF)-, VEGF-, and CCKR (cholecystokinin A receptor)-signalling pathways (Fig. 6A). The RNA-seq and quantitative proteomics analyses indicated that *FES* knockdown could potentially influence the production of a number of genes/proteins, including the extracellular matrix protein-degrading enzymes matrix metalloproteinase-1 (MMP1) and MMP3 (Fig. 6B, C and Supplementary Data 27–29). Reverse transcription-polymerase chain reaction (RT-PCR) and Western blot analyses confirmed increased MMP1 and MMP3 production in VSMCs with *FES* knockdown (Fig. 7A–C) and additionally showed that *FES* knockdown decreased the levels of contractile VSMC phenotype markers (Fig. 7D, E).

Effect of *FES* on atherosclerosis development and blood pressure in mouse models

In the colocalization analysis described early, *FES* was identified as a pleiotropic gene for both CAD and hypertension.

We previously showed that *Fes*^{-/-}/*Apoe*^{-/-} mice had larger aortic-root atherosclerotic lesions with a higher content of VSMCs and macrophages, compared with *Fes*^{+/-}/*Apoe*^{-/-} littermates, after a high-fat diet for 12 weeks²⁸. To further assess the effect of *Fes* on atherosclerosis, we performed *en face* Oil Red O staining on aortae taken from the *Fes*^{-/-}/*Apoe*^{-/-} mice and their *Fes*^{+/-}/*Apoe*^{-/-} littermates, and found that *Fes*^{-/-}/*Apoe*^{-/-} mice had significantly larger areas of atherosclerotic lesions than *Fes*^{+/-}/*Apoe*^{-/-} littermates (mean ± standard deviation: 8.34 ± 2.54 versus 6.06 ± 2.35; *P* = 0.013) (Fig. 8A), which was observed in male mice but not in female mice when data from male and female mice were analysed separately (Supplementary Fig. 10).

To assess the impact of *FES* on blood pressure, we performed hemodynamic measurements in unconscious *Fes*^{-/-}/*Apoe*^{-/-} mice and their *Fes*^{+/-}/*Apoe*^{-/-} littermates, both of which had been on a chow diet. At baseline, *Fes*^{-/-}/*Apoe*^{-/-} mice had a significant elevation in both systolic (88.0 ± 10.1 mmHg versus 104.4 ± 6.7 mmHg, *Fes*^{+/-}/*Apoe*^{-/-} versus *Fes*^{-/-}/*Apoe*^{-/-}, *P* = 0.042) and diastolic (58.8 ± 8.9 mmHg versus 74.9 ± 9.0 mmHg, *Fes*^{+/-}/*Apoe*^{-/-} versus *Fes*^{-/-}/*Apoe*^{-/-}, *P* = 0.042) blood pressure (Fig. 8B). The increase in blood pressure in *Fes*^{-/-}/*Apoe*^{-/-} mice did not lead to a compensatory change in heart rate, with no difference observed in heart rate at baseline between the two groups (Supplementary Fig. 11A). As expected, acetylcholine caused a reduction in both systolic and diastolic blood pressure in both groups. However, in

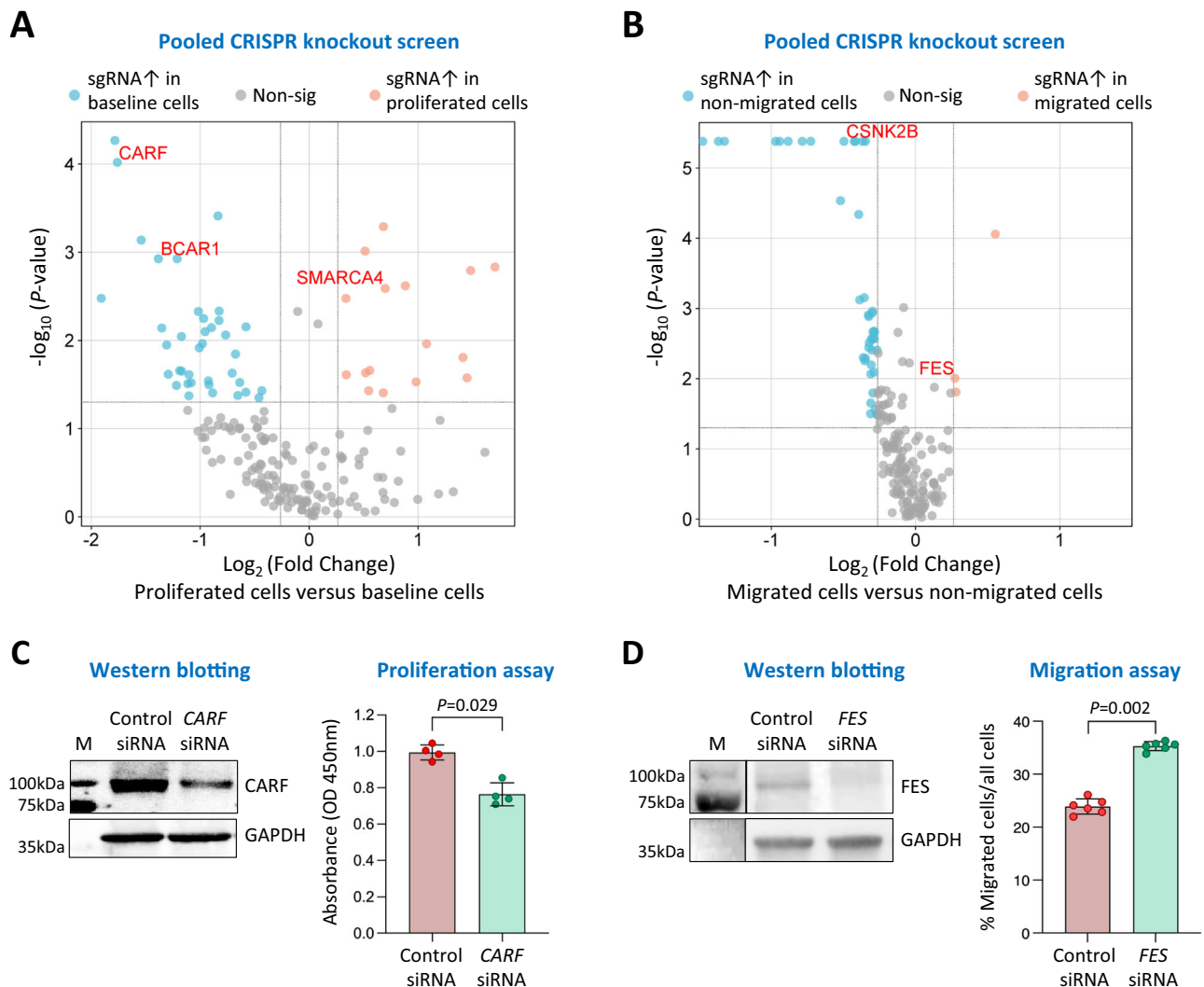


Fig. 5 | CAD/hypertension/stroke/AAA credible causal genes that affect VSMC proliferation or migration. A, B Pooled CRISPR screens of CAD/hypertension/stroke/AAA credible causal genes that affect VSMC proliferation (A) or migration (B). Pink symbols indicate genes whose knockout promotes proliferation (A) or migration (B); blue symbols indicate genes whose knockout reduces proliferation (A) or migration (B). Shown in the y-axis are \log_{10} transformed nominal P -values from a two-sided permutation test. **C** Results of Western blot analysis of CARF (left)

and proliferation assay (right) of VSMCs transfected with either CARF siRNA or control siRNA. OD: Optical density. **D** Results of Western blot analysis (left) and migration assay (right) of VSMCs transfected with either FES siRNA or control siRNA. Charts in (C, D) show mean \pm standard deviation (error bars); $n=4$ (C) or 6 (D) technical replicates; P -values from a two-tailed Mann-Whitney test; source data are provided as a Source Data file. M, molecular weight marker.

Fes^{-/-}/*Apoe*^{-/-} mice this response was significantly blunted with systolic blood pressure reduced by only 14.6 ± 3.8 mmHg in *Fes*^{-/-}/*Apoe*^{-/-} mice versus 22.2 ± 4.3 mmHg in *Fes*^{+/-}/*Apoe*^{-/-} mice ($P=0.027$) (Fig. 8C). A similar result was observed in diastolic responses with a greater reduction in diastolic blood pressure observed in *Fes*^{+/-}/*Apoe*^{-/-} mice compared with *Fes*^{-/-}/*Apoe*^{-/-} ($P=0.014$) (Fig. 8C). In contrast, no difference was observed in either systolic or diastolic blood pressure between the two groups in response to phenylephrine (Supplementary Fig. 11B).

Association of FES gene mutations with susceptibility to CAD and hypertension in humans

In addition to the above mouse model experiments, through the AstraZeneca PheWAS Portal²⁵, we collated human data of deleterious genetic variants in the protein-coding regions of the human *FES* gene in relation to blood pressure, hypertension, and the CAD traits angina and myocardial infarction, based on whole genome sequencing and phenotypic measurements of European participants in the UK

Biobank. Collapsing analyses of *FES* protein truncating variants showed that, collectively, they were significantly associated with elevated systolic and diastolic blood pressure ($P=4.70 \times 10^{-5}$ for systolic blood pressure and 2.70×10^{-4} for diastolic blood pressure)(Table 1), and a higher likelihood of hypertension [odds ratio (95% confidence interval) = 1.93 (1.30–2.86)](Table 2), as well as a >2-fold increased risk of myocardial infarction [2.30 (1.19–4.44)] and angina [2.05 (1.16–3.61)]²⁵ (Table 2).

Discussion

To facilitate medical translation from the recent discovery of the many genomic loci associated with CAD, hypertension, stroke and/or AAA, it is important to identify the causal genes at these loci to better understand the disease pathogenesis and utilise such information to identify new targets for therapeutic development. Our study, focused on VSMCs, a key cell type involved in the pathogenesis of these diseases, was aimed at this purpose and provides important findings in several aspects.

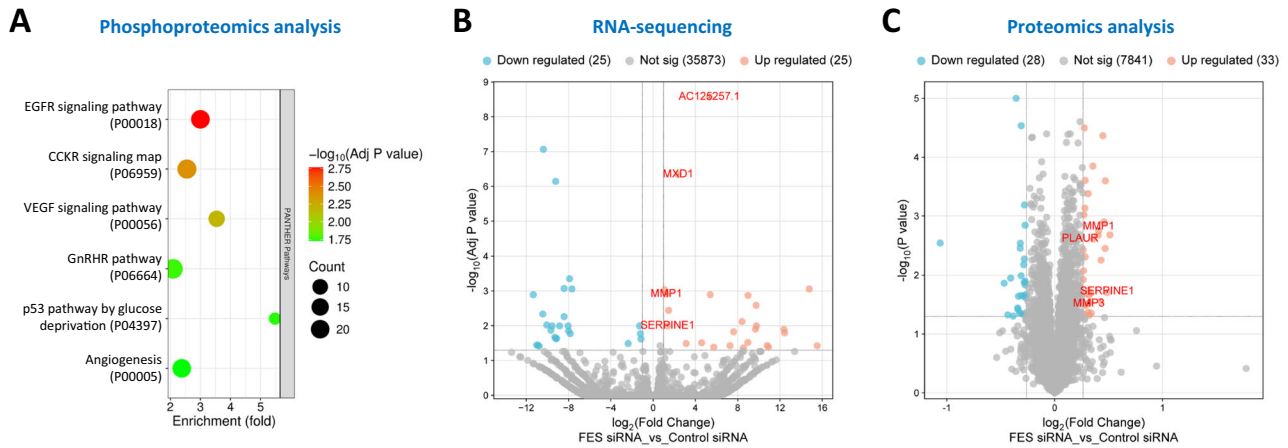


Fig. 6 | Results of phosphoproteomics, RNA-sequencing, and quantitative proteomics analyses of vascular smooth muscle cells transfected with either FES siRNA or control siRNA. A Pathways enriched with phosphorylated proteins in vascular smooth muscle cells (VSMCs), transfected with FES siRNA versus VSMCs transfected with control siRNA. CCKR: cholecystokinin A receptor; EGFR: epidermal growth factor receptor; GnRHR: gonadotropin-releasing hormone receptor; VEGF: vascular endothelial growth factor. **B** Up- or down-regulated genes in VSMCs transfected with FES siRNA versus VSMCs transfected with control siRNA. **C** Up- or down-regulated proteins in VSMCs transfected with FES siRNA versus VSMCs transfected with control siRNA. *P*-values from two-sided Fisher's Exact (**A**) or two-sided Student's *t* test (**B, C**).

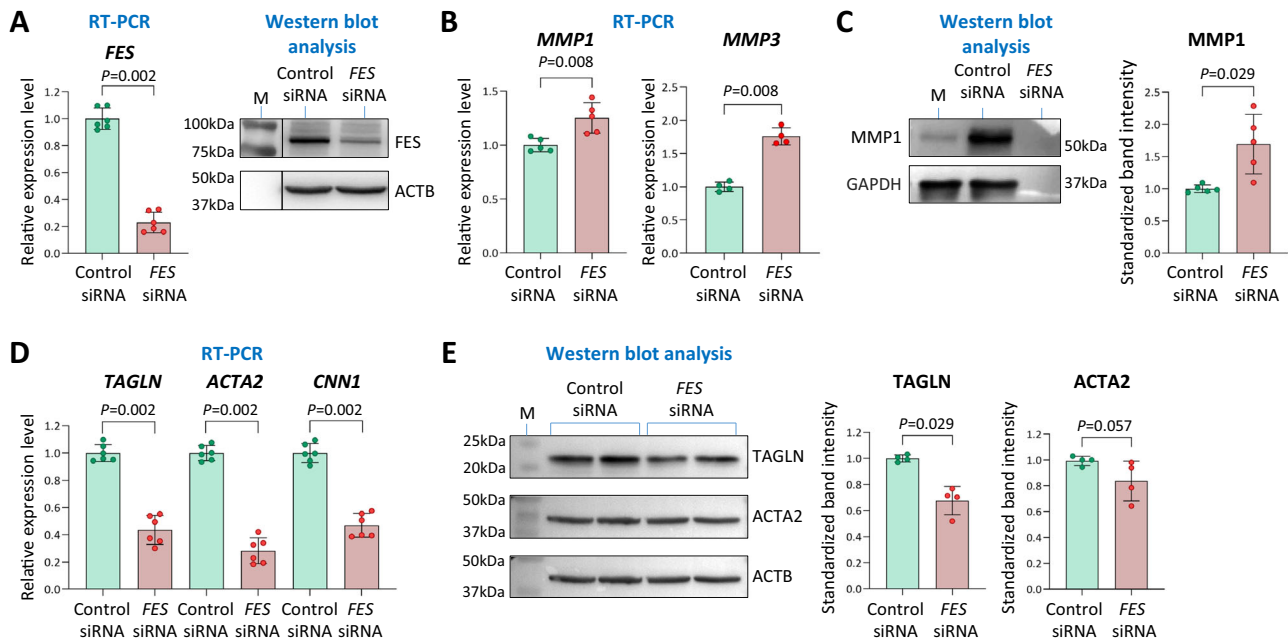


Fig. 7 | Effect of FES on the levels of MMPs and contractile VSMC phenotype markers. A Reverse transcriptase polymerase chain reaction (RT-PCR) and Western blot analyses confirming FES knockdown in VSMCs transfected with FES siRNA compared with VSMCs transfected with control siRNA. **B** RT-PCR analysis of matrix metalloproteinase 1 (*MMP1*), *MMP2*, and *MMP3*. **C** Western blot analysis of *MMP1*. **D** RT-PCR analysis of the contractile VSMC phenotype markers *ACTA2*, *TAGLN*, and *CNN1*. **E** Western blot analysis of *ACTA2* and *TAGLN*. **A, C, E** Average representative Western blot images are shown. M: Protein size markers; *ACTB*: β -actin. **A–E** Column charts show mean \pm standard deviation (error bars) values; $n = 4$ (**C, E**), 5 (**B**) or 6 (**A, D**) technical replicates; *P*-values shown are from two-tailed Mann-Whitney test; source data are provided as a Source Data file.

First, using an integrative functional genomics approach to systematically analyse regulatory genetic variants and their associated genes in conjunction with disease GWAS data, our study was able to identify panels of genes that are likely to be causal at the loci associated with CAD, hypertension, stroke, and/or AAA. Reassuringly, some of these genes have been demonstrated in previous studies using other methods to be important players in relevant disease processes, in particular, *ADAMTS29*, *JCAD*³⁰, *SMAD3*³¹ and *TCF21*¹⁷ in atherogenesis; *ARNTL*³², *CLCN6*³³ and *PKN2*³⁴ in blood pressure regulation, and *COL4A1/COL4A2*²⁵ and *SMARCA4*³⁶ in aortic aneurysm development.

This indicates the robustness of our analyses and provides further evidence supporting the hypothesis that these genes have a pathological role. Importantly, our analyses additionally identified multiple other genes that are likely to be causal but have hitherto not been implicated in the vascular diseases under study. These findings provide insight into the genetically mediated pathogenesis of these diseases.

The functional pathway analysis of the likely causal genes identified highlights TGF β signalling as a key pathway mediating the genetic influence on CAD susceptibility, substantiating findings from previous studies^{15,16,37}. TGF β is known to be a key regulator of

Table 2 | Association of FES protein truncating variants with hypertension, angina, and myocardial infarction in UK Biobank subjects

Phenotype	Cases	Cases with variants	Controls	Controls with variants	Odds ratio (95% CI)	P-value
Hypertension	n = 188129	n = 86 (0.046%)	n = 147753	n = 35 (0.024%)	1.93 (1.30–2.86)	0.001
Angina	n = 34976	n = 14 (0.054%)	n = 120611	n = 32 (0.026%)	2.05 (1.16–3.61)	0.018
Myocardial infarction	n = 26883	n = 14 (0.052%)	n = 105893	n = 24 (0.023%)	2.30 (1.19–4.44)	0.015

P-values were from a two-sided logistic regression analysis.

In support, we observed that mice with knockout of *FES*, one of the pleiotropic genes, had increased atherosclerosis as well as raised blood pressure, and additionally gathered evidence of *FES* gene deleterious variants significantly increasing the risk of both hypertension and CAD in humans among UK Biobank participants. Our experiments suggested that *FES* knockdown in VSMCs could influence the EGF-, VEGF-, and CCKR-signalling pathways, reduce contractile VSMC phenotype markers, promote VSMC migration, and increase production of the proteases MMP1 and MMP3. These findings are in agreement with previous reports that *FES* can interact with the EGF receptor^{44,45}, that EGF induces transition of VSMCs from a differentiated/contractile state to a dedifferentiated/synthetic phenotype⁴⁶, and that EGF-, VEGF-, and CCKR-signalling upregulates MMP expression^{41,47–50}.

Taken together, these findings suggest that *FES* can interact with these signalling pathways, whilst *FES* reduction can induce VSMC dedifferentiation to a synthetic phenotype with increased mobility and MMP expression. It has long been documented and is recently re-emphasised^{51,52} that VSMC dedifferentiation and their accumulation significantly contribute to atherosclerotic plaque formation and enlargement, although VSMCs in the fibrotic cap are beneficial for plaque stability. Furthermore, both MMP1 and MMP3 are capable of degrading a variety of extracellular proteins in the vascular wall and have been shown to facilitate VSMC migration, matrix degradation and vascular remodelling⁵³. These processes are instrumental in the pathogenesis of both atherosclerosis and hypertension^{53–55}.

Recent studies have indicated that using genetic data to nominate therapeutic targets can significantly increase the likelihood of success in the development of new drugs^{56–58}. In this context, another notable finding of our study is the identification of a panel of potential targets that are known or expected to be druggable. Remarkably, several of the pleiotropic genes identified are potentially druggable, raising the possibility of developing drugs to target these genes for the treatment of multiple co-existing vascular diseases.

We note some limitations of our work. First, our study used human umbilical cord-derived VSMCs for eQTL mapping in the context of adult cardiovascular disease. At present, a large collection of adult VSMCs is unavailable. Therefore, in this study, we utilised a large collection of human umbilical cord-derived VSMCs (VSMC eQTL data from 1486 individuals) to provide high statistical power for the comprehensive genome-wide eQTL analysis. As the cells were isolated from umbilical cords from individuals of the same biological age, the analysis also benefited from circumventing age-related confounding. However, there is a caveat that neonatal cells can differ from adult cells in gene regulation and expression. Notwithstanding, in a recent work¹⁶, we compared the transcriptomes of the umbilical cord-derived VSMCs utilised in this study with reported transcriptomic data from other types of cell/tissue^{59,60}, and found that the gene expression profile of the human umbilical cord-derived VSMCs mapped closer to adult human coronary artery smooth muscle cells than any other cell/tissue types¹⁶. Second, as *FES* is expressed in not only VSMCs but also some other cell types such as monocytes and endothelial cells, the total atherogenic contribution of *FES* is expected to derive from its effect on the various cell types, including VSMCs, as suggested here and monocytes as indicated in our recent study²⁸. The relative contributions from the different cell types are yet to be established.

Notwithstanding, the primary aim of our present study is not to test a cell type-specific effect on vascular diseases, but to identify potential causal genes at the genomic loci that have been found to be associated with CAD, hypertension, stroke and/or AAA in GWAS. Third, the in vivo experiments of *Fes* knockout were performed on small numbers of mice and were not powered to test possible sex-specific effects of *Fes* on atherosclerosis and blood pressure. Fourth, the CRISPR knockout screens and follow-up experiments were carried out on VSMCs from a single donor and so could not account for potential inter-individual biological variations. Therefore, the findings from these experiments are exploratory and for hypothesis generation, rather than definitive.

In summary, the findings from this systematic, integrative functional genomics study provide a better understanding of the genetic underpinnings and treatment targets of several major vascular diseases.

Methods

Generation of a VSMC eQTL catalogue

The study was approved by the East Midlands - Derby Research Ethics Committee (20/EM/0028), and the parents of the umbilical cord donors gave written informed consent.

The collection of VSMCs from different individuals and the methods of genotyping, imputation, RNA-Seq, eQTL analysis, and VSMC behaviour assays have been described previously¹⁶. In brief, VSMCs were isolated from the artery of umbilical cords from 2114 different donors, with the use of a well-established method⁶¹. DNA from 1992 samples was genotyped for 760,000 variants using the Infinium Global Screening Array-24 2.0 BeadChips (Illumina), and imputation was subsequently performed to obtain genotypic information of 7,334,165 variants. RNA-seq was conducted on total RNA extracted from passage 3 VSMCs from 1499 donors. For RNA-seq, a strand-specific library with rRNA removal was prepared from total RNA of each of the 1499 VSMC samples, and 150 bp paired-end sequencing, at a read depth of 30 million, was performed using the Illumina platform. An eQTL analysis was carried out on 1486 VSMC samples that had both genotype and RNA-Seq data, applying the eigenMT-BH method⁶² for multiple testing correction.

Compilation of tagging SNPs that marked the genetic variants/haplotypes reported to be associated with CAD, blood pressure/hypertension, stroke, and AAA, respectively

We previously identified 424 tagging SNPs which marked the genetic variants/haplotypes that had been reported to be associated with CAD at the genome-wide significance level ($P < 5 \times 10^{-8}$) in GWAS^{3,4,63,64}, did not have significant linkage disequilibrium (LD) between them, and were typed in the abovementioned VSMC collection either directly by using the Global Screening Array or indirectly by imputation¹⁶. In the present study, we further compiled lists of tagging SNPs which marked the genetic variants/haplotypes that had been reported to be associated with blood pressure/hypertension^{5,19}, stroke^{5–7}, and AAA^{5,8}, respectively, at the genome-wide significance level ($P < 5 \times 10^{-8}$) in GWAS, did not have significant LD between them, and were typed in the abovementioned VSMC collection either directly by using the Global Screening Array or indirectly by imputation¹⁶. LD between the SNPs was assessed using LDlink's SNPclip programme, which removed

SNPs based on LD in a given population and kept the first SNP (based on genomic position). We used an R^2 threshold of 0.8 and an MAF threshold of 0 with the “European” option. If there was available LD data, SNPs were included/excluded as required, but if not, the SNPs were kept in the list and assumed to be independent signals. In total, 1402 independent SNPs were identified for blood pressure/hypertension (Supplementary Data 30), 139 for stroke (Supplementary Data 31), and 157 for AAA (Supplementary Data 32). Of these, 1203 for blood pressure/hypertension, 119 for stroke, and 136 for AAA, were available in this study (either directly or via proxy SNPs with $R^2 > 0.8$), which were typed in the abovementioned VSMC collection¹⁶ by the Global Screening Array or imputed, and were not filtered out (missingness < 0.05 , MAF > 0.01 ; HWE P -value $> 1 \times 10^{-6}$).

Analyses of colocalization between disease GWAS and VSMC eQTL signals

Source and quality control of CAD GWAS datasets. CAD GWAS summary statistics from Nelson et al.², van der Harst and Verweij³, Howson et al.⁴, and Aragam et al.¹ were downloaded from <http://www.cardiogramplusc4d.org> (accessed on February 18, 2021), ftp://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST005001-GCST006000/GCST005194 (accessed on February 18, 2021), <https://grasphlbinihgov.ezproxy3.lib.le.ac.uk/FullResults.aspx> (accessed on March 29, 2021), and http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST90132001-GCST90133000/GCST90132314 (accessed December 13, 2022), respectively. Blood pressure GWAS summary statistics from Evangelou et al.¹⁹ (GCST006624, GCST006630) were downloaded from http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST006001-GCST007000 (accessed on February 02, 2022), and hypertension GWAS summary statistics from Neale⁵ were downloaded from <http://www.nealelab.is/uk-biobank>. Stroke GWAS summary statistics from Malik et al.⁶ and Mishra et al.⁷ (GCST006909, GCST006907, GCST006910, GCST006906, GCST90104543, GCST90104542, GCST90104541, GCST90104539) were downloaded from GWAS catalogue <https://www.ebi.ac.uk/gwas/> (accessed April 02, 2022 and December 13, 2022), and AAA GWAS summary statistics from Roychowdhury et al.⁸ were obtained from the study authors, respectively.

Cleaning and quality control steps were applied to the raw summary statistics data before they were used. Genome coordinates were converted to the hg38 build with CrossMap⁶⁵ and SNP marker names were updated with the new coordinates. If absent, variance, z-scores and minor allele frequencies were calculated in R v3.6.1 (<https://www.R-project.org>).

eCAVIAR analyses. We used eCAVIAR²⁰ to test for colocalization of disease GWAS with VSMC *cis*-eQTL. The eCAVIAR analysis was performed on SNPs that had been reported to associate with CAD¹⁻⁴, hypertension⁵, stroke^{6,7} or AAA⁸, and QTL data from the abovementioned VSMC eQTL catalogue. The top eQTL SNP (i.e., SNPs with the lowest P -value) of genes were selected and SNPs within 500 kb of the top SNP were overlapped with GWAS summary data. We analysed 500 kb windows as in several other studies^{16,59,66}, to include most, if not all, SNPs in LD with the top SNPs, since significant LD between SNPs typically decays within 100 kb in outbred human populations⁶⁷⁻⁶⁹. Supplementary Fig. 12 shows a locus zoom plot of a CAD-associated locus on chromosome 15, at which all CAD-associated SNPs are located within 100 kb. 500 kb windows were more extensive than 200 kb windows used in the originally reported eCAVIAR method²⁰; nevertheless, fixed windows of any size have the limitation of not accounting for between-region differences in LD structure. The eCAVIAR colocalization test was performed for loci that had 5 or > 5 overlapping SNPs present in both the eQTL and GWAS datasets, to reduce potential false positives due to overlaps by chance. We estimated LD for overlapping GWAS SNPs from individuals of European ancestry in the 1000

Genomes Phase 3 using PLINK⁷⁰. LD of *cis*-eQTL overlapping SNPs was estimated from VSMC genotype data. The number of causal SNPs was set to be up to 2 per locus, and colocalization events with > 0.05 colocalization posterior probability (CLPP) were considered significant.

The above analyses were performed on SNPs reported to be associated with CAD¹⁻⁴, hypertension⁵, or AAA⁸, at the genome-wide significant level ($P < 5 \times 10^{-8}$), and on SNPs reported to be associated with stroke^{6,7} at $P < 5 \times 10^{-5}$ as the number of SNPs reported to be associated with stroke at $P < 5 \times 10^{-8}$ was small. However, only the SNPs with both CLPP > 0.05 in eCAVIAR and $P < 5 \times 10^{-8}$ in stroke GWAS^{6,7} were considered to be likely causal SNPs for stroke.

SMR/HEIDI analyses. We also used SMR/HEIDI²¹ to test for colocalization between disease GWAS and VSMC *cis*-eQTL. We performed SMR/HEIDI analyses with default settings. LD was estimated with individuals of European ancestry in the 1000 Genomes Phase 3 as reference. Although we ran transcriptome-wide SMR/HEIDI analysis, we focused only on SNPs that had been reported to associate with CAD¹⁻⁴, hypertension⁵, stroke^{6,7}, or AAA⁸, at the genome-wide significance level (P -value $< 5 \times 10^{-8}$) in GWAS. We applied a threshold of $p_{SMR} < 0.05$ & $p_{HEIDI} > 0.05$ to identify significant colocalizations of CAD GWAS signals with VSMC eQTL signals.

Analyses of colocalization between disease GWAS and arterial tissue eQTL signals

With the use of LocusCompare V2 (<https://www.locuscompare2.com/>) and its default settings, additional colocalization analyses were performed on the abovementioned summary statistics of GWASs of CAD¹⁻⁴, hypertension⁵, stroke^{6,7}, and AAA⁸, respectively, in conjunction with eQTL data from coronary artery tissues and aortic artery tissues, respectively, from the GTEx Portal. eCAVIAR CLPP > 0.05 , SMR P -value $< 10^{-6}$, Coloc (H4) > 0.8 , PrediXcan P -value $< 10^{-6}$, fastENLOC GLCP > 0.05 , and/or FUSION P -value $< 10^{-6}$, were considered significant.

ATAC-seq

ATAC-seq was performed on 3 biological replicates of VSMCs derived from umbilical arteries from different individuals, as described previously⁷¹. In brief, 5×10^4 VSMCs were lysed with cold lysis buffer, and then the nuclei were incubated with a transposition reaction buffer containing Tn5 transposase (Illumina, FC-121-1030) for 30 minutes at 37 °C. The resulting libraries were subjected to polymerase chain reaction (PCR) with high-fidelity polymerase (New England Labs, M0541) and Nextera barcodes. The libraries were then multiplexed and sequenced using the Novaseq 6000 SP Flow Cell (Illumina) with > 100 million reads (50 bp pair end) per individual indexed library.

ATAC-seq data were processed using Galaxy (<https://usegalaxy.eu> and <https://usegalaxy.org>)^{72,73}. In brief, the quality of the raw data was evaluated with FastQC (v0.11.9). Adaptors were removed using trimmomatic (v0.38), and paired-end reads aligned to the human genome (hg38) using Bowtie2 (v2.4.2) with the `--very-sensitive`, `--dovetail` and `--X 1000` maximum insert size parameters. Mitochondrial and duplicate reads were removed using bamtools (v2.4.0) and Picard MarkDuplicates (v2.18.2), resulting in an average of 45 million filtered reads per sample for downstream analyses. Peaks were called using MACS2 (v2.1.1.20160309) with the start sites of the reads extending by 200 bp (100 bp in each direction) to assess coverage and a P -value 0.05 cutoff to reveal open chromatin peaks. The intersect function of BedTools v2.30.0 was used to identify overlaps of MACS2 peaks with the chromosome 7p15.3 and 7q23.3 loci at which variants were found to be associated with VSMC apoptosis/ACD in this study. PyGenomTracks (v3.6) was used for visualisation of genomic data tracks with publicly available ENCODE ChIP-seq SMC embryo-origin datasets (ENCSR116JEF: H3K27ac experiment ENCSR210ZPC, H3K4me3

experiment ENCSR515PKY and H3K27me3 experiment ENCSR143RMH) and the generation of MACS2 peak and coverage files.

H3K27ac HiChIP-seq

H3K27ac HiChIP-seq was performed as described previously^{74,75}, with some modifications. Two independent H3K27ac ChIP-seq on human coronary artery smooth muscle cells ($>1 \times 10^7$) (ThermoFisher Scientific, C-017-5C) were performed, prior to enriching ligated junctions. Briefly, cells were crosslinked with 1% formaldehyde. Hi-C was performed with the use of the Arima High Coverage HiC Kit (Arima Genomics, A410110) according to the manufacturer's protocol. Thereafter, sheared chromatin samples were subjected to immunoprecipitation by an incubation with 5 μ g ChIP-grade anti-histone H3 (acetyl K27) antibody (Abcam, ab4729) and 50 μ l Dynabeads Protein G for Immunoprecipitation (ThermoFisher Scientific, 10004D) overnight at 4 °C. After washing, the immunoprecipitated DNA was de-crosslinked at 67 °C for 2 h, eluted in a solution containing 50 mM sodium bicarbonate and 1% sodium dodecyl sulphate, and enriched with the use of Dynabeads MyOne Streptavidin C1 (ThermoFisher Scientific, 65002). Thereafter, a library was prepared with the use of NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, E7645L) according to the manufacturer's protocol. The library was subjected to PCR amplification (12 cycles) with indexed primers, and subsequently amplicons of 300–500 bp in length were selectively collected with the use of Omega Mag-Bind TotalPure NGS (Omega BIO-TEK, M1378). The library was subjected to next-generation sequencing at >300 million paired-end, 2×151 bp reads, on an illumina HiSeq 4000 platform.

Raw Fastq format paired-end reads were processed with the HiChIP pipeline by Dovetail Genomics (<https://hichip.readthedocs.io/en/latest/index.html>). The hg38 human genome index generated with Burrows-Wheeler Aligner (<http://arxiv.org/abs/1303.3997>) was used as reference for alignments and valid pair selection. Duplicate pairs were removed using pairtools⁷⁶. FitHiChIP4 (version 11.0) was used to call significant HiChIP loops at 2.5, 5, 10, 15, and 20Kb resolution. Using ChIP-Seq peak files as input, and initial paired interaction files, high-confidence peak-to-all interactions were processed with the following parameters: loose background (UseP2PBackgrnd = 0), coverage bias regression (BiasType = 1), merge filtering (MergeInt = 1), and 1% FDR (QVALUE = 0.01). High-confidence chromatin interactions were sorted and indexed in bed format, and plotted in R with Givz⁷⁷ (version 3.20).

DNA methylation analysis

DNA methylation assay was performed on 3 biological replicates of human umbilical artery smooth muscle cells (HUASMC, from different individuals) and on human aortic smooth muscle cells (HAoSMC, from 1 individual). Genomic DNA was extracted from the studied cells using the Sigma GenElute Mammalian DNA kit (Sigma Aldrich, G1N70). The DNA samples were subjected to analysis of methylation using the Illumina Infinium MethylationEPIC array. The resulting IDAT files were processed to derive CpG methylation beta values from the array data using the EPIC methylation array-specific ChAMP pipeline^{78,79}.

The hg38 genomic positions of the CpGs in the Infinium MethylationEPIC array were identified, and the mean methylation beta value (taken from 3 HUASMC and 1 HAoSMC cell lines) was calculated for each CpG. A CpG was defined as having low methylation levels if the mean beta value was ≤ 0.25 . The distance between each CpG and the likely causal variant identified from our colocalization analyses was calculated, and CpG-variant pairs with an absolute distance of ≤ 200 bp were considered to be within close proximity of each other (for variants with more than one CpG within close proximity, the nearest CpG was selected). Of the close-proximity CpG-variant pairs identified, the proportion that had low methylation (defined above) was calculated (both separately for each vascular disease and combined).

Pathway analyses

The likely causal genes identified in the colocalization analyses were subjected to gene ontology (GO) enrichment analysis⁸⁰. For each disease, the names of all of the likely causal genes were entered in the Gene Ontology browser (<http://geneontology.org/>) and searched for various GO aspects, including molecular function, biological process, cellular component, and PANTHER pathways, with the default setting for overrepresentation analysis by Fisher's exact test and with false discovery rate correction. For each GO aspect, the 3 most significant (with the smallest P-values and smallest false discovery rates) GO terms are presented as enrichment bubble plots, which were drawn using the SRplot tool⁸¹.

Cross-trait LD score regression analysis

We assessed the genetic correlation between each of the cardiovascular diseases and traits using default settings of the software package LDSC (v1.0.1)⁸². The pairwise product of the SNP z statistic was regressed against LD scores of Europeans available in the LDSC software.

Investigation of phenotypic effects of deleterious protein-coding variants in candidate causal genes

To investigate if deleterious variants in the protein-coding regions of the likely causal genes identified in the colocalization analyses were associated with the disease of interest, we interrogated the candidate genes in the AstraZeneca PheWAS Portal²⁵ (<https://azpewas.com/>) (UK Biobank 500k WGS (v2) Public) and PheWeb (<https://r8.finnngen.fi/>). Each of these genes was individually interrogated in both the AstraZeneca PheWAS Portal and PheWeb, and variant-level results were recorded. The results were then filtered for genetic variants associated with the vascular disease of interest, at the genome-wide significance level ($P < 1 \times 10^{-8.7}$ in PheWAS and $P < 5 \times 10^{-8}$ in PheWeb), including traits of CAD (atherosclerotic heart disease, coronary atherosclerosis, ischaemic heart disease, angina pectoris, myocardial infarction, major coronary heart disease event, coronary revascularization, coronary angioplasty, coronary artery bypass grafting), blood pressure/hypertension (systolic blood pressure, diastolic blood press, hypertension, hypertension diseases), stroke, and/or AAA.

To investigate the potential functional consequences of the disease-associated variants identified in the abovementioned search in the AstraZeneca PheWAS Portal and PheWeb, we looked up the variants in the Ensembl Genome Browser 110 (www.ensembl.org) to gather information on functional consequence predictions by SIFT (sorts intolerant from tolerant substitutions)⁸³, PolyPhen (polymorphism phenotyping)⁸⁴, CADD (combined annotation dependent depletion)⁸⁵, REVEL (rare exome variant ensemble learner)⁸⁶, MetaLR⁸⁷, and Mutation Assessor⁸⁸. Each variant was individually interrogated in the Ensembl Genome Browser, and the scores of SIFT, PolyPhen, CADD, REVEL, MetaLR and Mutation Assessor were recorded. A SIFT score ≤ 0.05 , a PolyPhen score ≥ 0.85 , a CADD score ≥ 30 , a REVEL score ≥ 0.75 , and/or a MetaLR ≥ 3.5 score was considered to indicate that the variant was deleterious.

Druggability and drug-gene interaction analyses

To investigate whether any of the candidate causal genes identified in this study were potential therapeutic targets, we interrogated these genes in the Drug-Gene Interaction Database (DGIdb)²⁶ which includes information from >30 trusted sources on the druggable genome with potentially druggable genes (genes or gene products that are predicted to be druggable, e.g., kinases) and genes with evidence of drug-gene interaction (genes or gene products that have been tested and reported to be druggable). For each disease, the names of all of the likely causal genes were entered in the search box in the search engine interface (<https://dgidb.org/>), and all genes indicated to be druggable in DGIdb were recorded.

Pooled CRISPR-Cas9-nuclease knockout screens

Pooled CRISPR-Cas9-nuclease knockout screens were performed on the genes that were identified in the colocalization analyses to be likely causal for CAD, hypertension, stroke, and/or AAA, and that were protein-coding genes because the pooled CRISPR-Cas9-nuclease screening technique is unable to screen non-coding RNA genes. A single-guide RNA (sgRNA) library was constructed, which included 4 sgRNAs for each of the target genes and 1000 non-targeting control sgRNAs. The sequences of the sgRNAs were selected from the Human CRISPR knockout Pooled Library (GeCKO v2 or Brunello). Corresponding oligonucleotides were synthesized by GenScript, and cloned into the lentiCRISPRv2 plasmid (Addgene, #52961) as described⁸⁹. The quality of the library was analysed by next-generation sequencing and MAGeCK (Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout)⁹⁰.

Human aortic smooth muscle cells (Lonza, CC-2571) were cultured in 10 plates (15 cm diameter) for the proliferation experiment and 5 plates for the migration experiment, with 3 technical replicates in each experiment. At 80% confluency, cells were transduced with the above-mentioned sgRNA library at a multiplicity of infection (MOI) of 0.2–0.3 as determined by titration. Two days post-transduction, cells were treated with 1 µg/mL puromycin and selected over the course of 8 days to enrich for transduced cells. After selection, the media was replaced with Human Vascular Smooth Muscle Cell Basal Medium (ThermoFisher Scientific, M231500) with Smooth Muscle Growth Supplement (ThermoFisher Scientific, S00725), 2% foetal bovine serum, and 1 × penicillin-streptomycin.

For proliferation experiments, genomic DNA was prepared from 5 plates on day 5 and from another 5 plates on day 28, post-puromycin treatment. Genomic DNA was prepared using the PureLink Genomic DNA Mini Kit (ThermoFisher Scientific, K182002).

For migration experiments, on day 12 post-puromycin treatment, 8×10^6 cells in 1 mL Human Vascular Smooth Muscle Cell Basal Medium (ThermoFisher Scientific, M231500) without supplements, were seeded at 2.5×10^6 per 8 µm pore polycarbonate insert in 6-well (24 mm) transwell plates (Corning, 3428). The lower wells contained 2 mL Human Vascular Smooth Muscle Cell Basal Medium (ThermoFisher Scientific, M231500) with 100 ng/mL platelet-derived growth factor-BB (Sigma-Aldrich, P3201-50UG). After 24 h, cells from both the top and bottom surfaces of the insert were harvested separately with trypsinization and then subjected to genomic DNA extraction with the use of the PureLink Genomic DNA Mini Kit (ThermoFisher Scientific, K182002).

PCR with the use of KAPA HiFi HotStart Ready Mix (Roche, 07958927001) was performed on the genomic DNA samples (32 µg genomic DNA per sample) from the above-described proliferation and migration experiments. The sequences of the PCR primers were as described by Joung et al.⁸⁹. (Supplementary Data 33). The PCR cycling conditions were as follows: 95 °C for 5 min, then 24–26 cycles of 98 °C for 20 s, 60 °C for 15 s and 72 °C for 15 s, and finally 72 °C for 1 min.

The PCR products were pooled and purified using KAPA HyperPure Beads (Roche, 08963843001) with a PCR product-to-bead ratio of 1:0.55. After collecting the supernatant an additional 0.35 volume of beads were added to achieve a final ratio of 1:0.9. Each purified sample was dissolved in 20 µL nuclease-free water, and the concentrations were quantified with the Qubit dsDNA HS Assay Kit (ThermoFisher Scientific, Q32851) and the quality assessed with DNA 1000 kit (Agilent, 5067-1504) on the 2100 Bioanalyzer System (Agilent). Subsequently, libraries with barcoded Rev-1 to Rev-4 were pooled (500 ng each sample, combined into a total volume of 40 µL) and further quantified by quantitative PCR using standards from the NEBNext Library Quant Kit for Illumina (New England Biolabs, E7630S). The combined libraries were sequenced on the Illumina NovaSeq X platform, at 6 million 2×150 bp paired-end reads per sample.

Raw reads files were processed with Cutadapt⁹¹ for adaptor trimming with custom scripts. Subsequently, MAGeCK⁹⁰ (version 0.5.9.5) count and test algorithms were applied to the migration and proliferation datasets to identify positively-selected sgRNAs and negatively-selected sgRNAs, through paired analyses with control sgRNAs normalisation. For the proliferation experiment, data from day 28 were compared with data from day 5, post-puromycin treatment. For the migration experiment, data from the bottom surface of the insert were compared with data from the top surface of the insert. Data from the 3 replicates of the proliferation experiment had similar distributions and corrections, so did data from the 3 replicates of the migration experiment, indicating high reproductivity (Supplementary Fig. 13). Volcano plots were generated with the use of SRplot⁸¹.

Vascular smooth muscle cell proliferation assay

Human coronary artery SMCs (Cell Applications, 350K-05a) were transfected with either *CARF* siRNA (ThermoFisher Scientific, HSS128878) or control siRNA (ThermoFisher Scientific, 4390843), and then a proliferation assay was performed using Cell Counting Kit-8 (MedChemExpress, HY-K0301), with 4 technical replicates. In brief, cells were cultured in 96-well plates (5×10^4 cells/well) for 48 h, and a 2-(2-methoxy-4-nitrophenyl)-3-(4-nitrophenyl)-5-(2,4-disulphophenyl)-2H-tetrazolium solution was added into each well. At 2 h after adding the reagent, a colorimetric assay was carried out using a microplate reader which measured the absorbance at 450 nm.

Human aortic smooth muscle cells (Lonza, CC-2571) were transfected with either *BCAR1* DsiRNA (Integrated DNA Technologies, sequence described in Supplementary 33) or negative control DsiRNA (Integrated DNA Technologies, 51-01-14-04), with 3 technical replicates. At 0, 24, and 48 hours post-transfection, cells were subjected to a proliferation assay with the use of Cell Counting Kit-8 (Abcam, ab228554).

Human aortic smooth muscle cells (Lonza, CC-2571) were transfected with either *SMARCA4* siRNA (ThermoFisher Scientific, HSS110005) or non-targeting control siRNA (ThermoFisher Scientific, 12935300), with 8 technical replicates. At 0, 24, 48, 72, and 96 hours post-transfection, cells were subjected to a proliferation assay using Cell Counting Kit-8 (Abcam, ab228554). A serial dilution starting from 25,000 cells was generated to serve as a standard curve for analysing the results of the proliferation assay.

Vascular smooth muscle cell migration assay

Human aortic smooth muscle cells (Lonza, CC-2571) were transfected with either *FES* siRNA (ThermoFisher Scientific, HSS103636), control siRNA (ThermoFisher Scientific, 12935300), *CSNK2B* DsiRNA (Integrated DNA Technologies, sequence described in Supplementary 33) or negative control DsiRNA (Integrated DNA Technologies, 51-01-14-04), with the use of Lipofectamine RNAiMAX Transfection Reagent (ThermoFisher Scientific, 13778150). At 47 hours post-transfection, cells were incubated with 1 µg/mL mitomycin C (Sigma-Aldrich, M7949-2MG) for 1 h and then 4×10^4 cells were seeded in each well of a Corning Transwell 24-well plate with 8 µm pores (Sigma-Aldrich, CLS3422-48EA), with 6 technical replicates per condition. 24 h later, cells were incubated with 4% formaldehyde (Sigma-Aldrich, F8775-500ML) for 10 min, 5% Triton X-100 (Sigma-Aldrich, T8787) for 5 minutes, and then 1 µg/mL DAPI (4',6-diamidino-2-phenylindole) (ThermoFisher Scientific, 62248) for 10 minutes. Subsequently, 3 images of each well were taken using a fluorescence microscope with a $10 \times$ objective lens and a digital camera. Thereafter, cells on the upper aspect of the transwell were wiped off, and 3 images of the lower aspect were taken. Cells in each image were counted with the use of ImageJ (version 2.14.0) software, and the cell counts in each set of 3 images were averaged.

RNA-seq

Human coronary artery smooth muscle cells (Cell Applications, 350K-05a) transfected with either *FES* siRNA (ThermoFisher Scientific, HSS103636) or control siRNA (ThermoFisher Scientific, 12935300), with 5 technical replicates per condition, were subjected to RNA-seq. A strand-specific library with rRNA removal was prepared from total RNA of each sample, and 150 bp paired-end sequencing at a >30 million read depth (12 G raw data per sample) was performed with the Illumina NovaSeq 6000 Sequencing System. Raw data (raw reads) in FASTQ format were processed as follows: (1) removing reads with 5' adaptors, reads without 3' adaptor or insert sequence, reads with >10% undetermined nucleotides, reads with >50% nucleotides with Qphred \leq 20, and reads with ploy A/T/G/C; (2) removing adaptor sequences from the 3' ends of reads. Thereafter, the processed clean reads from each sample were mapped to the human reference genome with HISAT2 software, and subsequently transcripts were assembled with the StringTie programme. For quantification of the transcripts and genes, RPKM (Reads Per Kilobase of transcript per Million mapped reads) values were determined with the use of StringTie. A differential expression analysis was performed with the use of Cuffdiff, and *P*-values were adjusted using the Benjamini and Hochberg method. A volcano plot of differentially-expressed genes (adjusted *P* < 0.05) was generated with the use of SPplot⁸¹.

Proteomics and phosphoproteomics analyses

Human coronary artery smooth muscle cells (Cell Applications, 350K-05a) transfected with either *FES* siRNA (ThermoFisher Scientific, HSS103636) or control siRNA (ThermoFisher Scientific, 12935300), with 5 technical replicates per condition, were subjected to quantitative proteomics and phosphoproteomics analysis. In brief, protein extracts from cells were reduced with dithiothreitol, alkylated with iodoacetamide, washed with acetone, and then dissolved in a solution containing 0.1M triethylammonium bicarbonate (pH 8.5) and 6M urea. The concentration of total protein in the samples was quantified using the Bradford method, and protein quality was assessed using sodium dodecyl sulfate-polyacrylamide gel electrophoresis with Coomassie Brilliant Blue R-250 staining. The protein samples were digested by incubation with trypsin and subjected to tandem mass tag (TMT) labelling using a mass tagging kit (ThermoFisher Scientific, 90113). Phosphopeptides in the samples were enriched using PHOS-Select Iron Affinity Gel (Sigma-Aldrich, P9740). Subsequently, liquid chromatography-mass spectrometry was performed with the use of an EASY-nLCTM 1200 System (ThermoFisher Scientific) and a Q Exactive HF-X mass spectrometer (ThermoFisher Scientific).

The liquid chromatography-mass spectrometry data were analysed using Proteome Discoverer Software 2.4 (ThermoFisher Scientific) with the following parameters: a mass tolerance of 10 ppm for precursor ion scans and a mass tolerance of 0.02 Da for the product ion scans. Phosphorylation of serine (S)/threonine (T)/tyrosine (Y), oxidation of methionine (M), and tandem mass tags (TMT)-plex of lysine (K) were specified as dynamic modifications. Carbamidomethyl was set as a fixed modification. TMT-plex, acetylation, and loss of methylation were specified as N8 terminal modifications. The motif-x algorithm was used to identify motifs enriched within a set of phosphosites. All enrichments were carried out for 7 residues surrounding the central residue with occurrences > 20 and *P* < 10⁻⁶. WebLogo was used to build the motif figures. To predict kinase-substrate relationships, all identified serine, threonine and tyrosine phosphorylation sites (pS/T/Y) were scored with the NetPhorest algorithm.

Each protein was identified with at least 1 unique peptide at a false discovery rate (FDR) < 1%. Proteins with similar peptides that could not be distinguished based on the mass spectrometry data were combined as protein groups. Quantification based on intensity was performed.

A GO enrichment analysis was conducted on proteins that showed suggestive differences in modification between cells with siRNA-

induced *FES* knockdown and cells transfected with the control siRNA (*P* < 0.05 by *t* test), using the default setting of Gene Ontology (<http://geneontology.org>). A bubble plot of enriched pathways was generated with the use of SPplot⁸¹.

Any difference in the quantity of each protein between cells with *FES* knockdown and cells transfected with the control siRNA was ascertained by *t* test. A volcano plot of differentially expressed proteins (*P*-value < 0.05 and > 1.2fold) was generated with the use of SPplot⁸¹.

RT-PCR

Cells were lysed using TRIzol reagent (ThermoFisher Scientific, 15596026), and thereafter total RNA was extracted using the Direct-zol RNA MiniPrep kit (Zymo Research, R2050) according to the manufacturer's protocol. RNA was then reverse transcribed into cDNA with the use of iScript Reverse Transcription Supermix (Bio-Rad, 1708841). Subsequently, cDNA of the genes of interest was amplified using iTaq Universal SYBR Green Supermix (Bio-Rad, 1725124) and the primers described in Supplementary Data 35. The PCR thermocycling conditions were as follows: 95 °C for 2 minutes, followed by 40 cycles of 95 °C (denaturation) for 5 seconds and 60 °C (annealing and extension) for 30 s. Differences in expression levels of the genes of interest between cells transfected with the target gene siRNA and cells transfected with the control RNA were determined using the $\Delta\Delta$ Ct method, with the housekeeping gene *ACTB* as a reference.

Western blot analyses

Protein extracts from transfected cells were prepared using RIPA (radioimmunoprecipitation assay) lysis buffer (ThermoFisher Scientific, #89901), and protein concentrations in the extracts were quantified using Pierce BCA Protein Assay Kit (ThermoFisher Scientific, #23327). Equal amounts of protein from each sample were subjected to sodium dodecyl-sulfate polyacrylamide gel electrophoresis and thereafter transferred onto polyvinylidene difluoride membranes. The membranes were then blocked with 5% bovine serum albumin in Tris-buffered saline with Tween-20 for 2 h at room temperature, and subsequently incubated with a primary antibody (Supplementary Data 35) at 4 °C overnight. After washing, the membranes were incubated with a secondary antibody conjugated with horse-radish peroxidase (Supplementary Data 36) for 1 h at room temperature. The bands on the membrane were detected using a ChemiDoc Touch Imaging system (Bio-Rad).

Animal study

The study was conducted in accordance with the UK Animals (Scientific Procedures) Act 1986, under a Home Office Project License (60/4332), and following the guidelines of ARRIVE (Animal Research: Reporting of In Vivo Experiments), the University of Leicester and the University of Oxford.

Animals. The generation of *Fes*^{-/-}/*ApoE*^{-/-} mice and *Fes*^{+/-}/*ApoE*^{-/-} control littermates has been described previously²⁸. In brief, *Fes*^{-/-} mice were backcrossed onto a C57BL6/J background for seven generations, and subsequently intercrossed with *ApoE*^{-/-} mice (B6.129P2-Apoetm1Unc/J purchased from Jackson Laboratory) to generate *Fes*^{-/-}/*ApoE*^{-/-} mice and *Fes*^{+/-}/*ApoE*^{-/-} control littermates. Mice were housed in a pathogen-free environment. Animals were maintained in ventilated cages with a controlled 12:12-hour light/dark cycle (lights on at 07:00 h). The ambient room temperature was maintained at 22 ± 2 °C with a relative humidity of 55 ± 10%.

Study of atherosclerosis. *Fes*^{-/-}/*ApoE*^{-/-} and littermate control (*Fes*^{+/-}/*ApoE*^{-/-}) mice of 6 weeks of age were fed a high fat diet for 12 weeks. The extent of aortic atherosclerotic lesions in the mice was analysed by *en face* staining of aortae with Oil Red O, as previously described⁹². In

brief, the aorta was opened longitudinally along the ventral midline from the iliac arteries to the aortic root, pinned out flat on a surface, and stained *en face* with Oil Red O. The stained aortas were photographed using a digital camera, and the Oil Red O-stained areas analysed by computer-assisted quantification using the Image-Pro software.

Arterial pressure monitoring. *Fes^{+/+}/Apoe^{-/-}* and littermate control (*Fes^{+/+}/Apoe^{-/-}*) mice, at 8 weeks of age, were anaesthetised with 4% isoflurane in medical O₂ and maintained on a nose cone at 1.5–1% isoflurane for measurements of blood pressure and heart rate. The aorta was cannulated via the right common carotid artery using a 1.4 French blood pressure probe and transducer (Millar Instruments Inc. Houston, TX) coupled to a Powerlab/4SP data acquisition system (ADInstruments Pty Ltd., New Castle, Australia). Measurements were obtained after a 15 min equilibrium period; phenylephrine was dosed via a catheter placed in the jugular vein. Blood pressure and heart rate measurements were taken for 30 s and average recordings during this time taken. Increasing doses of phenylephrine (5 µg/kg and 25 µg/kg) were administered with a 5-minute wash out period between doses. After a further 10-minute wash out period, acetylcholine (5 µg/kg) was administered. Delta change in blood pressure was measured by subtracting baseline measurements from blood pressure after drug administration. A total of $n = 7$ *Fes^{+/+}/Apoe^{-/-}* and 5 *Fes^{-/-}/Apoe^{-/-}* mice per group were used in the study, with procedures performed blinded. One *Fes^{-/-}/Apoe^{-/-}* mouse was excluded due to a technical issue with the catheterisation. Two hemodynamic measurements were excluded in the wild type group due to technical problems with the traces (one 5 µg/kg phenylephrine and one acetylcholine), and one wild type measurement at 5 µg/kg was missed due to a problem with the injection.

Collapsing analyses of human FES protein truncating variants in relation to blood pressure and CAD traits

Through the AstraZeneca PheWAS Portal²⁵, we collated human data of deleterious genetic variants in the protein-coding regions of the human *FES* gene in relation to blood pressure (data fields 4079 and 4080), hypertension (I10), and the CAD traits angina (I20) and myocardial infarction (I21), based on whole genome sequencing and phenotypic measurements of European participants in the UK Biobank. Collapsing analyses of the human *FES* protein truncating variants in relation to the blood pressure and CAD traits were performed.

Statistical analyses

The Mann-Whitney test or *t* test was used to ascertain differences between experimental groups in the VSMC proliferation assays, VSMC migration assays, RT-PCR analyses, standardised band intensity in the Western blot analyses, and atherosclerotic lesion areas and blood pressure in the animal studies, respectively.

Reporting summary

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

Data availability

The data that support the findings of this study are available from the corresponding author. The RNA-Seq data and H3K27ac HiChIP-seq data are available from Gene Expression Omnibus with the accession numbers [GSE189300](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE189300) and [GSE282557](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE282557). The DNA methylation data are available from ArrayExpress with the accession ID [E-MTAB-15426](https://www.ebi.ac.uk/ena/arrayexpress/experiments/E-MTAB-15426). The proteomics data and phosphoproteomics data are available from MassIVE with the accession IDs PXD061984 [<https://massive.ucsd.edu/ProteoSAFe/private-dataset.jsp?task=8660bf01bde54a50899b9126643cf52c>] and PXD061992 [<https://massive.ucsd.edu/ProteoSAFe/private-dataset.jsp?task=>

[c1493e7d255a401b93782a307e5f0b64](https://doi.org/10.1038/s41467-026-69273-8)], respectively. Source data are provided in this paper.

References

- Aragam, K. G. et al. Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nat. Genet.* **54**, 1803–1815 (2022).
- Nelson, C. P. et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat. Genet.* **49**, 1385–1391 (2017).
- van der Harst, P. & Verweij, N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. *Circ Res.* **122**, 433–443 (2018).
- Howson, J. M. M. et al. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. *Nat. Genet.* **49**, 1113–1119 (2017).
- The_Pan-UKBB_team. *Pan UKBB*, <https://pan.ukbb.broadinstitute.org> (2020).
- Malik, R. et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat. Genet.* **50**, 524–537 (2018).
- Mishra, A. et al. Stroke genetics informs drug discovery and risk prediction across ancestries. *Nature* **611**, 115–123 (2022).
- Roychowdhury, T. et al. Genome-wide association meta-analysis identifies risk loci for abdominal aortic aneurysm and highlights PCSK9 as a therapeutic target. *Nat. Genet.* **55**, 1831–1842 (2023).
- Basatemur, G. L., Jorgensen, H. F., Clarke, M. C. H., Bennett, M. R. & Mallat, Z. Vascular smooth muscle cells in atherosclerosis. *Nat. Rev. Cardiol.* **16**, 727–744 (2019).
- Touyz, R. M. et al. Vascular smooth muscle contraction in hypertension. *Cardiovasc. Res.* **114**, 529–539 (2018).
- Nordon, I. M., Hinchliffe, R. J., Loftus, I. M. & Thompson, M. M. Pathophysiology and epidemiology of abdominal aortic aneurysms. *Nat. Rev. Cardiol.* **8**, 92–102 (2011).
- Frosen, J. & Joutel, A. Smooth muscle cells of intracranial vessels: from development to disease. *Cardiovasc. Res.* **114**, 501–512 (2018).
- Khera, A. V. & Kathiresan, S. Genetics of coronary artery disease: discovery, biology and clinical translation. *Nat. Rev. Genet.* **18**, 331–344 (2017).
- Maurano, M. T. et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012).
- Aherrahrou, R. et al. Genetic regulation of atherosclerosis-relevant phenotypes in human vascular smooth muscle cells. *Circ. Res.* **127**, 1552–1565 (2020).
- Solomon, C. U. et al. Effects of coronary artery disease-associated variants on vascular smooth muscle cells. *Circulation* **146**, 917–929 (2022).
- Wirka, R. C. et al. Atheroprotective roles of smooth muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell analysis. *Nat. Med.* **25**, 1280–1289 (2019).
- Aherrahrou, R. et al. Genetic regulation of SMC gene expression and splicing predict causal CAD genes. *Circ. Res.* **132**, 323–338 (2023).
- Evangelou, E. et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat. Genet.* **50**, 1412–1425 (2018).
- Hormozdiari, F. et al. Colocalization of GWAS and eQTL Signals Detects Target Genes. *Am. J. Hum. Genet.* **99**, 1245–1260 (2016).
- Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
- Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).

23. Szustakowski, J. D. et al. Advancing human genetics research and drug discovery through exome sequencing of the UK Biobank. *Nat. Genet.* **53**, 942–948 (2021).
24. Kurki, M. I. et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* **613**, 508–518 (2023).
25. Wang, Q. et al. Rare variant contribution to human disease in 281,104 UK Biobank exomes. *Nature* **597**, 527–532 (2021).
26. Freshour, S. L. et al. Integration of the Drug-Gene Interaction Database (DGIdb 4.0) with open crowdsourcing efforts. *Nucleic Acids Res.* **49**, D1144–D1151 (2021).
27. Greer, P. Closing in on the biological functions of Fps/Fes and Fer. *Nat. Rev. Mol. Cell Biol.* **3**, 278–289 (2002).
28. Karamanavi, E. et al. The FES gene at the 15q26 coronary artery disease locus inhibits atherosclerosis. *Circ. Res.* **131**, 1004–1017 (2022).
29. Bauer, R. C. et al. Knockout of Adamts7, a novel coronary artery disease locus in humans, reduces atherosclerosis in mice. *Circulation* **131**, 1202–1213 (2015).
30. Douglas, G. et al. A key role for the novel coronary artery disease gene JCAD in atherosclerosis via shear stress mechanotransduction. *Cardiovasc. Res.* **116**, 1863–1874 (2020).
31. Cheng, P. et al. Smad3 regulates smooth muscle cell fate and mediates adverse remodeling and calcification of the atherosclerotic plaque. *Nat. Cardiovasc. Res.* **1**, 322–333 (2022).
32. Woon, P. Y. et al. Aryl hydrocarbon receptor nuclear translocator-like (BMAL1) is associated with susceptibility to hypertension and type 2 diabetes. *Proc. Natl. Acad. Sci. USA* **104**, 14412–14417 (2007).
33. Klemens, C. A. et al. Loss of chloride channel 6 (CLC-6) affects vascular smooth muscle contractility and arterial stiffness via alterations to golgi calcium stores. *Hypertension* **77**, 582–593 (2021).
34. Jin, Y. J. et al. Protein kinase N2 mediates flow-induced endothelial NOS activation and vascular tone regulation. *J. Clin. Invest.* **131**, e145734 (2021).
35. Steffensen, L. B. et al. Basement membrane collagen IV deficiency promotes abdominal aortic aneurysm formation. *Sci. Rep.* **11**, 12903 (2021).
36. Zhang, X. et al. Brg1 trans-activates endothelium-derived colony stimulating factor to promote calcium chloride induced abdominal aortic aneurysm in mice. *J. Mol. Cell Cardiol.* **125**, 6–17 (2018).
37. Zhao, Q. et al. Molecular mechanisms of coronary disease revealed using quantitative trait loci for TCF21 binding, chromatin accessibility, and chromosomal looping. *Genome Biol.* **21**, 135 (2020).
38. Tsai, S. et al. TGF-beta through Smad3 signaling stimulates vascular smooth muscle cell proliferation and neointimal formation. *Am. J. Physiol. Heart Circ. Physiol.* **297**, H540–H549 (2009).
39. Lacolley, P., Regnault, V., Nicoletti, A., Li, Z. & Michel, J. B. The vascular smooth muscle cell in arterial pathology: a cell that can take on multiple roles. *Cardiovasc. Res.* **95**, 194–204 (2012).
40. Maki-Petaja, K. M. et al. Mechanisms underlying vascular endothelial growth factor receptor inhibition-induced hypertension: the HYPAN trial. *Hypertension* **77**, 1591–1599 (2021).
41. Wang, H. & Keiser, J. A. Vascular endothelial growth factor upregulates the expression of matrix metalloproteinases in vascular smooth muscle cells: role of flt-1. *Circ. Res.* **83**, 832–840 (1998).
42. Raines, E. W. PDGF and cardiovascular disease. *Cytokine Growth Factor Rev.* **15**, 237–254 (2004).
43. Jang, Y., Lincoff, A. M., Plow, E. F. & Topol, E. J. Cell adhesion molecules in coronary artery disease. *J. Am. Coll. Cardiol.* **24**, 1591–1601 (1994).
44. Jones, R. B., Gordus, A., Krall, J. A. & MacBeath, G. A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature* **439**, 168–174 (2006).
45. Hause, R. J. Jr. et al. Comprehensive binary interaction mapping of SH2 domains via fluorescence polarization reveals novel functional diversification of ErbB receptors. *PLoS ONE* **7**, e44471 (2012).
46. Yamanaka, Y. et al. EGF family ligand-dependent phenotypic modulation of smooth muscle cells through EGF receptor. *Biochem. Biophys. Res. Commun.* **281**, 373–377 (2001).
47. Schreier, B. et al. Substance-specific importance of EGFR for vascular smooth muscle cells motility in primary culture. *Biochim. Biophys. Acta* **1863**, 1519–1533 (2016).
48. Nutt, J. E., Mellon, J. K., Qureshi, K. & Lunec, J. Matrix metalloproteinase-1 is induced by epidermal growth factor in human bladder tumour cell lines and is detectable in urine of patients with bladder tumours. *Br. J. Cancer* **78**, 215–220 (1998).
49. Cho, A., Graves, J. & Reidy, M. A. Mitogen-activated protein kinases mediate matrix metalloproteinase-9 expression in vascular smooth muscle cells. *Arterioscler. Thromb. Vasc. Biol.* **20**, 2527–2532 (2000).
50. Park, H. J. et al. Gastrin-releasing peptide promotes the migration of vascular smooth muscle cells through upregulation of matrix metalloproteinase-2 and -9. *BMB Rep.* **50**, 628–633 (2017).
51. Azar, P. et al. Smooth muscle cells in atherosclerosis: essential but overlooked translational perspectives. *Eur. Heart J.* **46**, 4862–4875 (2025).
52. Zaman, S. et al. The Lancet Commission on rethinking coronary artery disease: moving from ischaemia to atheroma. *Lancet* **405**, 1264–1312 (2025).
53. Newby, A. C. Dual role of matrix metalloproteinases (matrixins) in intimal thickening and atherosclerotic plaque rupture. *Physiol. Rev.* **85**, 1–31 (2005).
54. Renna, N. F., de Las Heras, N. & Miatello, R. M. Pathophysiology of vascular remodeling in hypertension. *Int. J. Hypertens.* **2013**, 808353 (2013).
55. Intengan, H. D. & Schiffrin, E. L. Vascular remodeling in hypertension: roles of apoptosis, inflammation, and fibrosis. *Hypertension* **38**, 581–587 (2001).
56. Nelson, M. R. et al. The support of human genetic evidence for approved drug indications. *Nat. Genet.* **47**, 856–860 (2015).
57. Plenge, R. M., Scolnick, E. M. & Altshuler, D. Validating therapeutic targets through human genetics. *Nat. Rev. Drug Discov.* **12**, 581–594 (2013).
58. Finan, C. et al. The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* **9**, eaag1166 (2017).
59. Liu, B. et al. Genetic regulatory mechanisms of smooth muscle cells map to coronary artery disease risk loci. *Am. J. Hum. Genet.* **103**, 377–388 (2018).
60. Kim-Hellmuth, S. et al. Cell type-specific genetic regulation of gene expression across human tissues. *Science* **369**, eaaz8528 (2020).
61. Leik, C. E., Willey, A., Graham, M. F. & Walsh, S. W. Isolation and culture of arterial smooth muscle cells from human placenta. *Hypertension* **43**, 837–840 (2004).
62. Huang, Q. Q., Ritchie, S. C., Brozynska, M. & Inouye, M. Power, false discovery rate and Winner’s Curse in eQTL studies. *Nucleic Acids Res.* **46**, e133 (2018).
63. Erdmann, J., Kessler, T., Munoz Venegas, L. & Schunkert, H. A decade of genome-wide association studies for coronary artery disease: the challenges ahead. *Cardiovasc. Res.* **114**, 1241–1257 (2018).
64. Buniello, A. et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
65. Zhao, H. et al. CrossMap: a versatile tool for coordinate conversion between genome assemblies. *Bioinformatics* **30**, 1006–1007 (2014).
66. Aherrahrou, R., Lue, D. & Civelek, M. Genetic regulation of circular RNA expression in human aortic smooth muscle cells and vascular traits. *HGG Adv.* **4**, 100164 (2023).
67. Ardlie, K. G., Kruglyak, L. & Seielstad, M. Patterns of linkage disequilibrium in the human genome. *Nat. Rev. Genet.* **3**, 299–309 (2002).

68. Pritchard, J. K. & Przeworski, M. Linkage disequilibrium in humans: models and data. *Am. J. Hum. Genet.* **69**, 1–14 (2001).
69. Varilo, T. et al. The interval of linkage disequilibrium (LD) detected with microsatellite and SNP markers in chromosomes of Finnish populations with different histories. *Hum. Mol. Genet.* **12**, 51–59 (2003).
70. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
71. McVey, D. G. et al. Genetic influence on vascular smooth muscle cell apoptosis. *Cell Death Dis.* **15**, 402 (2024).
72. Batut, B. et al. Community-Driven Data Analysis Training for Biology. *Cell Syst.* **6**, 752–758 (2018).
73. Rickner H. D., Niu S. Y. & Cheng C. S. ATAC-seq Assay with low mitochondrial DNA contamination from primary human CD4+ T lymphocytes. *J. Vis. Exp.* <https://doi.org/10.3791/59120> (2019).
74. Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).
75. Anene-Nzelu, C. G. et al. Assigning Distal Genomic Enhancers to Cardiac Disease-Causing Genes. *Circulation* **142**, 910–912 (2020).
76. Open2C, et al. Pairtools: From sequencing data to chromosome contacts. *PLoS Comput. Biol.* **20**, e1012164 (2024).
77. Hahne F. & Ivanek R. Visualizing genomic data using Gviz and Bioconductor. in *Statistical Genomics: Methods and Protocols* (eds Mathé E&Davis S). (Springer New York, 2016).
78. Morris, T. J. et al. ChAMP: 450k Chip analysis methylation pipeline. *Bioinformatics* **30**, 428–430 (2014).
79. Tian, Y. et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics* **33**, 3982–3984 (2017).
80. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
81. Tang, D. et al. SRplot: A free online platform for data visualization and graphing. *PLoS ONE* **18**, e0294236 (2023).
82. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
83. Ng, P. C. & Henikoff, S. Predicting deleterious amino acid substitutions. *Genome Res.* **11**, 863–874 (2001).
84. Ramensky, V., Bork, P. & Sunyaev, S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res.* **30**, 3894–3900 (2002).
85. Kircher, M. et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
86. Ioannidis, N. M. et al. REVEL: An ensemble method for predicting the pathogenicity of rare missense variants. *Am. J. Hum. Genet.* **99**, 877–885 (2016).
87. Dong, C. et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum. Mol. Genet.* **24**, 2125–2137 (2015).
88. Reva, B., Antipin, Y. & Sander, C. Determinants of protein function revealed by combinatorial entropy optimization. *Genome Biol.* **8**, R232 (2007).
89. Joung, J. et al. Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **12**, 828–863 (2017).
90. Li, W. et al. MAGECK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).
91. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **17**, 10–12 (2011).
92. Laxton, R. C. et al. A role of matrix metalloproteinase-8 in atherosclerosis. *Circ. Res.* **105**, 921–929 (2009).
- to S.Y.), The National Medical Research Council of Singapore (MOH-001229 and MOH-001479 to S.Y.), and the National University of Singapore/National University Health System (NUHSRO/2022/004/Startup/01 to S.Y.). C.U.S. is a Leicester British Heart Foundation Accelerator Award (AA/18/3/34220) Research Fellow. D.G.M. is supported by the British Heart Foundation Research Excellence Award (RE/24/130031), the van Geest Foundation Heart and Cardiovascular Diseases Research Fund, and was awarded a BHF Accelerator Early Careers Researcher Interdisciplinary Fellowship and pump-priming funding from the Leicester British Heart Foundation Accelerator Award (AA/18/3/34220). This work falls under the portfolio of research conducted within the National Institute for Health Research Leicester Biomedical Research Centre and the Leicester BHF Centre of Research Excellence. This research used the ALICE High Performance Computing Facility at the University of Leicester.

Author contributions

C.U.S., D.G.M., H.Z., and Y.Z. performed the analyses. D.G.M., C.A., P.G., L.T., D.S.S.S., H.Z., D.P.L., E.K., W.Y., J.C., R.C., K.E.H., C.G.A-N., H.L., M.J.D., P.Y.L, X.H., G.E.M., E.J.S., and G.D. performed the experiments. P.A.G. provided the mouse line. H.Y., R.S.Y.F., N.J.S., T.R.W., and S.Y. supervised the work. S.Y. wrote the paper. All the authors revised the manuscript, contributed with discussions and revisions and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-69273-8>.

Correspondence and requests for materials should be addressed to Shu Ye.

Peer review information *Nature Communications* thanks Bo Zhang and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

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

© The Author(s) 2026

Acknowledgements

This work was funded by the British Heart Foundation [RG/16/13/32609, RG/19/9/34655, PG/16/9/31995, PG/18/73/34059, and SP/19/2/344612

¹Division of Cardiovascular Sciences, University of Leicester, Leicester, UK. ²National Institute for Health Research Leicester Biomedical Research Centre, Leicester, UK. ³Leicester British Heart Foundation Centre of Research Excellence, University of Leicester, Leicester, UK. ⁴Cardiovascular-Metabolic Disease Translational Research Programme, Department of Medicine, National University of Singapore, Singapore, Singapore. ⁵Division of Basic Medicine, Shantou University Medical College, Shantou, China. ⁶Department of Pathophysiology, Gannan Medical University, Ganzhou, Jiangxi, China. ⁷The First Affiliated Hospital of Shantou University Medical College, Shantou, China. ⁸Montreal Heart Institute, Montreal, Canada. ⁹Department of Pathology and Molecular Medicine, Queen's University, Kingston, Canada. ¹⁰Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK. ¹¹Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. ¹²These authors contributed equally: Charles U. Solomon, David G. McVey. ¹³These authors jointly supervised this work; Nilesh J. Samani, Tom R. Webb, Shu Ye. ✉ e-mail: shuye68@nus.edu.sg