

# Enhancer Promoter Interactome and Mendelian Randomization Identify Network of Druggable Vascular Genes in Coronary Artery Disease

**Arnaud Chignon**

IUCPQ

**Samuel Mathieu**

IUCPQ

**Anne Rufiange**

IUCPQ

**Deborah Argaud**

IUCPQ

**Pierre Voisine**

IUCPQ

**Yohan Bossé**

IUCPQ

**Benoit Arsenault**

IUCPQ

**Sébastien Thériault**

IUCPQ

**Patrick Mathieu** (✉ [patrick.mathieu@fmed.ulaval.ca](mailto:patrick.mathieu@fmed.ulaval.ca))

Université Laval <https://orcid.org/0000-0002-3805-2004>

---

## Primary research

**Keywords:** Atherosclerosis, Coronary artery disease, Genes, 3D genome, Mendelian Randomization, Druggable genome, Network

**Posted Date:** November 1st, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-978033/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Human Genomics on March 4th, 2022. See the published version at <https://doi.org/10.1186/s40246-022-00381-4>.

# Abstract

Coronary artery disease (CAD) is a multifactorial disorder, which is partly heritable. Herein, we implemented a mapping of CAD-associated candidate genes by using genome-wide enhancer-promoter conformation (H3K27ac-HiChIP) and expression quantitative trait loci (eQTL). Enhancer-promoter anchor loops from human coronary artery smooth muscle cells (HCASMC) explained 22% of the heritability for CAD. 3D enhancer-promoter genome mapping of CAD-genes in HCASMC was enriched in vascular eQTL genes. By using colocalization and Mendelian randomization analyses, we identified 58 causal candidate vascular genes including some druggable targets (*MAP3K11*, *CAMK1D*, *PDGFD*, *IPO9* and *CETP*). A network analysis of causal candidate genes was enriched in TGF beta and MAPK pathways. The pharmacologic inhibition of causal candidate gene *MAP3K11* in vascular SMC reduced the expression of athero-relevant genes and lowered cell migration, a cardinal process in CAD. Genes connected to enhancers are enriched in vascular eQTL and druggable genes causally associated with CAD.

## Introduction

Coronary artery disease (CAD) is a complex trait disorder and a leading cause of morbidity-mortality. The dysregulated expression of genes and activation of pathways culminate in the development of vascular atheromatous plaque, a hallmark feature of CAD<sup>1</sup>. Functional assessment and cell tracking experiments have shown that vascular smooth muscle cells are recruited to the plaque where they play a significant role in the development of CAD<sup>2</sup>. The identification of molecules involved in the pathophysiology of CAD could lead to the development of novel therapies. However, the discovery of disease-associated drug targets is limited by several factors, which include limited information about the disease process and indirect evidence obtained from animal and cell experiments<sup>3</sup>. Also, epidemiological studies carried out in humans by measuring biomarkers or intermediate molecules is subject to bias and reverse causality<sup>4</sup>. As such, only a small fraction of drug development programs leads to licensed drugs<sup>5</sup>.

Genetic association studies (GWAS) have underlined that CAD is heritable (narrow sense heritability estimated at 40-50%) and involves several loci (~160 identified so far)<sup>6</sup>. Genetic association data provides a rich resource, which may help identify key targets involved in the development of disorders. Gene variants associated with intermediate phenotypes (e.g. gene expression) and the disease can be leveraged to assess causality similarly to a randomized clinical trial<sup>7</sup>. According to Mendel's second law of random allocation of alleles, independent instrumental variables (IVs) (uncorrelated variants) can be assessed to mimic the effect of drugs (exposure, e.g. as determined by gene expression) on the risk of disorders (outcome). The use of multiple genetic variants that are associated with the exposure can be used as IVs in Mendelian randomization (MR) to assess causal associations for the risk of disorders<sup>8</sup>. Since the allocation of alleles is random and occurs before the development of the outcome, MR technique is not prone to bias and reverse causation<sup>9</sup>. Studies have underlined that pharmacological targets supported by genetics have a much higher chance of success during the different phases of drug development<sup>10</sup>.

Despite rigorous preclinical screening, the failure of some compounds is related to drug-related adverse side-effects, which are often discovered in large phase 3 randomized clinical trials (RCTs)<sup>11</sup>. This may result in substantial monetary loss with the attrition of resources for the development of novel drug pipelines<sup>12</sup>. Target-related pleiotropy is one cause of such failure<sup>10</sup>. For instance, a drug target with opposite directional effect on two major outcomes may negate potential benefit. Interrogation of large electronic health record databases in genotyped individuals provides a resource to assess hundreds of traits and disorders. The assessment of risk variants linked to a target in a phenome-wide analysis study (PheWAS) is thus a strategy to evaluate potential side-effects of drug-gene pairs<sup>13</sup>. Also, the assessment of putative causal candidate drug targets for their association with monogenic disorders by using large resources where data are collated, such as in the Human Phenotype Ontology, is another approach to assess drug-related safety issues<sup>14,15</sup>.

Complex systems in which gene expression and interaction of molecules are establishing a trajectory to health or disease are increasingly investigated by the integration of data in network<sup>16</sup>. The network topology is highly modular and provides information for distinctive molecules, which interact together to drive different functions<sup>17</sup>. As such, genes highly connected tend to be enriched in essential functions and in drug targets<sup>18,19</sup>. The implementation of network to assess molecules with impact on the biological pathways relevant to a disease is a useful strategy to prioritize genes and to address the function of the whole system and its components (modules).

In order to identify causal genes by using genome-wide association studies (GWAS), we need to map variants to potential gene targets. Mapping of genes is compounded by several factors, which includes the linkage disequilibrium (LD) between the variants (the extent to which variants are correlated) and the fact that the vast majority of gene variants associated with complex traits and disorders reside in the noncoding genome<sup>20</sup>. Variants within the noncoding genome are enriched in active regulatory regions such as distant acting enhancers and may impact on the expression in *cis* (locally) through the conformation of chromatin<sup>21</sup>. Genome-wide assessment of chromatin conformation based on Hi-C and its derivatives (e.g. HiChIP) has revealed that chromatin looping between enhancers and promoters provides regulatory mechanisms to control gene expression<sup>22</sup>. Growing evidence suggests that a hierarchy among the distant regulatory regions provides another layer of control on gene expression<sup>23</sup>. To this effect, highly connected hub enhancers in 3D communities regulate lineage-specific genes<sup>24</sup>. In addition, the architecture of chromatin conformation is largely cell specific<sup>25</sup>. Hence, GWAS mapping using genome-wide 3D data provides an additional layer of information to identify putative causal genes in disease-relevant cells<sup>26</sup>. Another strategy to map the genetic variants to potential targets is to assess the associations with expression quantitative trait loci (eQTL) derived from disease-relevant tissue(s)<sup>27</sup>. Herein, we implemented an integrative approach including enhancer-promoter chromatin conformation and eQTL mapping, causal inference, interrogation of the druggable genome and network biology to capture GWAS-associated molecules and pathways, which could be targeted pharmacologically. Specifically, we assessed whether enhancer-promoter chromatin looping in vascular smooth muscle cells

explained a significant proportion of the heritability for CAD and if it was enriched in vascular eQTL genes. Genome-wide mapping using enhancer-promoter conformation and eQTL data were integrated in a framework to assess causal associations and to prioritize druggable genes with the help of networks.

## Results

### Tissue enrichment and pathway analyses

Summary-level data of meta-analysis from UK Biobank (UKB) and CARDIoGRAMplusC4D including 123,733 CAD cases and 424,528 controls<sup>28</sup> was leveraged to assess tissue and pathway enrichment. We implemented Data-driven Expression Prioritized Integration for Complex Traits (DEPICT) to document the enrichment of tissue and pathways for genetic association data in CAD. By using GWAS summary statistics, DEPICT provides an analysis in which risk loci mapped genes are integrated to a vast collection of human tissue gene expression to prioritize highly expressed genes and their relevant pathways. According to DEPICT, genetic association for CAD was enriched in arteries ( $P=7.38 \times 10^{-9}$ ) and smooth muscle ( $P=4.33 \times 10^{-6}$ ) (Figure 1A) (**Suppl. Table 1**). We next evaluated the enrichment of genetic association in pathways. DEPICT showed an enrichment for abnormal cell adhesion ( $P=5.04 \times 10^{-14}$ ), abnormal vitelline vasculature morphology ( $P=5.64 \times 10^{-13}$ ) and abnormal cell migration ( $P=3.67 \times 10^{-11}$ ) (Figure 1B) (**Suppl. Table 2**). These data thus indicated that the vasculature and smooth muscle cells as well as their related functions such as adhesion and migration are key features associated with CAD-gene variants.

### Heritability for CAD explained by the vascular enhancer-promoter connectome

Considering the high enrichment of genetic association data for CAD in smooth muscle and arteries, we analyzed publicly available HiChIP for H3K27ac obtained in human coronary artery smooth muscle cells (HCASMC) (GSE101498). H3K27ac-HiChIP provides a high-definition map of chromatin conformation between enhancers and promoters<sup>29</sup>. After a stringent loop call with FitHiChIP ( $FDR < 1 \times 10^{-6}$ ), we identified 224,209 confident loops in HCASMC. The anchor loops were significantly enriched in open chromatin detected by assay for transposase-accessible chromatin and sequencing (ATAC-seq) in HCASMC (fold enrichment=2.9,  $P < 2.2 \times 10^{-16}$ , binomial test) (Figure 1C). By using HOMER, a pathway enrichment of anchor loops in HCASMC using the Kyoto Encyclopedia of Genes and Genomes (KEGG) showed an enrichment for endocytosis ( $P=1.36 \times 10^{-10}$ ), regulation of actin cytoskeleton ( $P=1.70 \times 10^{-8}$ ) and autophagy ( $P=5.54 \times 10^{-8}$ ) (**Suppl. Table 3**). We next wondered whether enhancer-promoter contacts in HCASMC explained a significant part of the heritability in CAD risk. For that purpose, we partitioned the heritability of *cis*-anchor loops in HCASMC by using the full baseline model of 53 annotations in the stratified LD score regression framework (Methods)<sup>30</sup>. This analysis revealed that chromosomal interactions in HCASMC were enriched for regions that explained the heritability of CAD

(enrichment=2.38,  $P=5.90 \times 10^{-5}$ ) (**Suppl. Table 4**). Despite representing only 9% of CAD-single nucleotide polymorphisms (SNPs), variants within anchor loops in HCASMC explained 22% of the heritability for CAD (**Suppl. Table 4**).

## Identification of gene promoters connected to CAD associated variants within distant enhancers

To identify gene promoters mapped by chromatin looping in H3K27ac-HiChIP in HCASMC, we identified lead and individual significant SNPs ( $P_{\text{GWAS}} < 5 \times 10^{-8}$ ,  $r^2 < 0.6$ ) that overlapped with enhancer loops (Methods). In total, 353 individual significant SNP-enhancer-promoter loop pairs tagging 228 gene promoters were identified (**Suppl. Table 5**). The mean number of loops per gene promoter was 1.5 and the average distance between enhancers and gene promoters was  $\sim 168\text{kb}$ . For instance, at 11q13.1, rs3741380, a CAD individual significant SNP intronic to *EHBP1L1*, is localized within an interacting loop with the promoter of *MAP3K11* located  $\sim 32\text{kb}$  downstream (Figure 1D). We next hypothesized that CAD gene variants may be enriched in hub enhancers having a high level of chromatin contacts (in network with a degree  $\geq 90\text{th}$  percentile) (Methods). In HCASMC, we identified 5,455 highly connected hub enhancers involved in chromatin looping. This analysis showed a significant enrichment of CAD individual significant SNPs associated with hub enhancers (observed/expected=3.4,  $P=1.9 \times 10^{-6}$ , binomial test). Among the different genes mapped by CAD variants and enhancer-promoter looping, *SH2B3* is a gene connected to a hub enhancer. *SH2B3* has been previously identified as a causal candidate gene for CAD<sup>31</sup>. On the other hand, some genes identified by chromatin looping were not previously mapped in CAD. For instance, at 11q13.4, rs590121 is located in a distant hub enhancer intronic to *SERPINH1* and having contacts with multiple enhancers-promoters including with the promoter of *ARRB1* located  $\sim 298\text{kb}$  upstream (Figure 1E).

## Identification of vascular eQTLs

We next examined if CAD-associated genes mapped by enhancer-promoter chromatin looping were also significant vascular expression quantitative trait loci (eQTL). CAD-individual significant SNPs were mapped to eQTLs of the aorta in GTEx v8. In total, 15,516 CAD SNP-eQTL gene pairs were significant at FDR 5% in the aorta and tagged 202 genes (**Suppl. Table 6**). Among the 228 genes mapped by enhancer-promoter chromatin looping there were 41 genes (18%), which were also CAD-associated eQTL genes (fold enrichment observed/expected=21.4,  $P < 2.2 \times 10^{-16}$ , binomial test). Hence, these findings highlight that gene mapping of CAD genetic association data with enhancer-promoter chromatin conformation in HCASMC is enriched in vascular eQTL genes.

## Genetic colocalization analyses

By combining enhancer-promoter chromatin looping and eQTL data, there were 383 genes mapped by CAD-associated SNPs. We performed Bayesian colocalization analyses between the eQTL signal (GTEx v8) of the 383 mapped genes in the aorta with the GWAS signal for CAD. We identified 35 genes with a high posterior probability (PP) ( $PP > 0.8$ ) of shared signal between eQTLs in the aorta and CAD-GWAS

(*CDH13*, *PHACTR1*, *TCF21*, *N4BP2L2*, *SYPL2*, *TWIST1*, *PDGFD*, *IPO9*, *FHL3*, *UTP11*, *GGCX*, *SEMA5A*, *DMPK*, *MIA3*, *TMEM133*, *CAMK1D*, *ARHGAP42*, *DAGLB*, *DMWD*, *CETP*, *MORF4L1*, *JCAD*, *MFGE8*, *HAPLN3*, *HHIPL1*, *DHX36*, *B3GNT8*, *BMP1*, *LMOD1*, *FAM117B*, *MAT2A*, *ATP2B1*, *EXOSC5*, *EIF2B2*, *ZEB2*) (**Suppl. Table 7**). Genes with a shared signal were enriched in gene ontology (GO) for negative regulation of cell adhesion ( $P=1.99 \times 10^{-4}$ ) and positive regulation of endothelial cell proliferation ( $P=2.18 \times 10^{-4}$ ) (**Suppl. Table 8**). Figure 2A shows a LocusCompare plot for the GWAS associations at 6p24.1 where there is a shared signal ( $PP=1$ ) with the expression of *PHACTR1*, a gene previously identified as a causal candidate at this locus<sup>32</sup>. Rs9349379 is the strongest SNP for both GWAS and eQTL in the aorta for the expression of *PHACTR1*, a gene encoding for a regulator of actin polymerization<sup>33</sup>. Allele G-rs9349379 increases the risk of CAD (OR: 1.11, 95%CI: 1.10-1.12,  $P_{\text{GWAS}}=2.71 \times 10^{-76}$ ) and is associated with a decreased expression of *PHACTR1* in the aorta (Figure 2B). This analysis underlined some targets identified in previous screen and functional assays such as *CDH13*, *TCF21*, *LMOD1* and *JCAD*<sup>34-37</sup>, but also identified novel potential causal candidates such as *EXOSC5* and *B3GNT8*, which encode for a RNA exosome component and a galactosyltransferase respectively.

## Mendelian Randomization

In order to further evaluate causal associations, we implemented a 2-sample Mendelian randomization (MR) between gene expression in the aorta (GTExV8) by using a minimum of three independent ( $r^2 < 0.1$ ) *cis*-variants as IVs and genetic associations for CAD as the outcome. For that purpose 3D and eQTL mapped genes were evaluated in MR. In inverse variance weighted (IVW) MR, we identified 74 vascular genes expressed in the aorta that were significantly associated with CAD (FDR < 0.05). Among these genes, 33 did not show heterogeneity (Cochran's Q test > 0.05) and were considered as causally associated with CAD (*MRAS*, *HHIPL1*, *CDH13*, *JCAD*, *MFGE8*, *BMP1*, *FGD6*, *CTSK*, *MAP3K11*, *TMEM133*, *CAMK1D*, *DMPK*, *ZEB2*, *EIF2B2*, *HSD17B12*, *CDC25A*, *ARRB1*, *SFMBT1*, *TRIP4*, *KCNH2*, *NME7*, *ATP1B1*, *MRPL35*, *CCDC181*, *AGPAT4*, *RNF123*, *ANKDD1A*, *BEND6*, *CTSH*, *NPHP3*, *PIF1*, *ALKBH5*, *MEAF6*) (Figure 3) (**Suppl. Table 9**). These 33 vascular genes were enriched for GO in cardiac muscle cell repolarization ( $P=1.44 \times 10^{-4}$ ) and proteolysis involved in cellular catabolic process ( $P=4.80 \times 10^{-4}$ ) (**Suppl. Table 10**). Among the causally associated vascular genes, there were 10 genes also identified in colocalization analyses ( $PP > 0.8$ ) (*HHIPL1*, *CDH13*, *JCAD*, *MFGE8*, *BMP1*, *TMEM133*, *CAMK1D*, *DMPK*, *ZEB2*, *EIF2B2*). In total, by using colocalization and MR approaches, we identified 58 potential causal vascular genes for CAD (**Suppl. Table 11**). Among the causal candidate genes, 9 were identified by both eQTL and enhancer-promoter conformation mapping (*FHL3*, *MIA3*, *MAT2A*, *GGCX*, *DAGLB*, *MAP3K11*, *ATP2B1*, *EXOSC5*, *B3GNT8*), whereas 6 genes (*MEAF6*, *AGPAT4*, *KCNH2*, *ARRB1*, *PIF1*, *UTP11*) were mapped only by chromatin conformation. For instance, *AGPAT4*, a gene encoding for a lysophosphatidic acid acyltransferase, is negatively associated with the CAD risk in MR (OR: 0.94, 95%CI: 0.91-0.97,  $P_{\text{causal}}=1.5 \times 10^{-3}$ ). At this locus, two CAD independent significant SNPs (rs9295142, rs142697177) located in an intronic enhancer of *AGPAT4* are within a chromatin loop with the promoter of the same gene (**Suppl. Figure 1**).

In sensitivity analysis, we assessed causal associations by using weighted median MR, which provides a robust assessment with up to 50% of invalid instrumental variables (variants with horizontal pleiotropy)<sup>38</sup>. Among the 33 vascular genes associated with CAD in IVW MR, we found that 29 genes (*MFGE8, CAMK1D, CTSK, JCAD, CDH13, EIF2B2, MAP3K11, FGD6, RNF123, ATP1B1, SFMBT1, MRPL35, MRAS, KCNH2, NPHP3, ANKDD1A, BEND6, CTSH, NME7, CCDC181, HSD17B12, TRIP4, CDC25A, TMEM133, DMPK, ZEB2, ARRB1, BMP1, HHIPL1*) were also significant in weighted median MR analysis (**Suppl. Table 12**). The directional effects were concordant between IVW and weighted median MR analyses.

## Network and prioritization

We aimed to characterize putative causal vascular genes singled out by the colocalization and MR screens by using a network approach with the objectives to identify: 1) key driver genes and 2) pathways of connected genes with functional relevance. A network was constructed from the DifferentialNet dataset, which includes 134,223 protein interactions, and inferred to the aorta from the gene expression profile (Methods)<sup>39</sup>. Candidate causally associated vascular genes identified by MR and colocalization analyses were used as seeds to generate the network, which encompassed 681 nodes (genes) and 763 edges (connections) (Figure 4A). A pathway enrichment analysis showed that genes within the network were enriched in TGF-beta signaling pathway ( $P=1.87 \times 10^{-20}$ ), signaling events mediated by HDAC class I ( $P=3.46 \times 10^{-17}$ ), and MAPK signaling pathway ( $P=1.62 \times 10^{-16}$ ) (**Suppl. Table 13**). Causally associated genes were enriched in nodes with a high degree (degree >90th percentile) (fold enrichment 3.83,  $P=2.70 \times 10^{-25}$ , hypergeometric test). Genes with a high degree and acting as hub in networks are enriched in drug targets and are involved in key biological functions. Predicted CAD causally associated vascular genes in MR and acting as hub gene (high degree in PPI network) include among others *ARRB1* (OR: 0.84, 95%CI: 0.78-0.91,  $P_{\text{causal}}=9.46 \times 10^{-6}$ ), *CDC25A* (OR: 0.96, 95%CI: 0.95-0.98,  $P_{\text{causal}}=5.33 \times 10^{-6}$ ), *KCNH2* (OR: 1.03, 95%CI: 1.01-1.04,  $P_{\text{causal}}=7.14 \times 10^{-5}$ ) and *MAP3K11* (OR: 1.07, 95%CI: 1.05-1.10,  $P_{\text{causal}}=8.46 \times 10^{-9}$ ).

## Druggability of candidate causal genes

We next assessed whether causally associated vascular genes in CAD were potentially druggable by using The Drug Gene Interaction Database (DGIdb)<sup>40</sup>. DGIdb includes an exhaustive list of drug-gene pairs, which are collated from different resources. In total 383 compounds targeting 13 predicted causally associated vascular genes were identified by DGIdb (*CTSK, MAP3K11, CDC25A, ARRB1, KCNH2, ATP1B1, PDGFD, IPO9, GGCX, DAGLB, CETP, DHX36, CAMK1D*) (**Suppl. Table 14**). Considering the directional effects in MR, 5 vascular genes are potential targets for drug inhibition (*CTSK, MAP3K11, KCNH2, ATP1B1, CAMK1D*). For instance, the vascular expression of *CAMK1D* which encodes for a calcium calmodulin dependent protein kinase, is positively associated with the risk of CAD in MR (OR: 1.11, 95%CI: 1.07-1.16,  $P_{\text{causal}}=5.12 \times 10^{-8}$ ) (Figure 4B). Several drugs under investigation are kinase inhibitors, which are reported to target CAMK1D. By using colocalization analyses, the directional effects of the

candidate causal SNP between vascular gene expression and CAD-risk suggest that the inhibition of 4 genes may lower disease-associated risk (*PDGFD*, *IPO9*, *GGCX*, *CETP*). As an example, among the potential drug targets, *PDGFD* encodes for platelet derived growth factor D. In the aorta, colocalization between eQTL for *PDGFD* and CAD-GWAS (PP=0.99) suggests a causal association and prioritizes gene variant rs974819 (Figure 4C). In the aorta, T-rs974819 is associated with a higher expression of *PDGFD* (Figure 4D) and an elevated CAD-risk (OR: 1.06, 95%CI: 1.05-1.07,  $P_{\text{GWAS}}=1.11 \times 10^{-28}$ ). According to the directional effects in MR and colocalization data, 4 genes (*CDC25A*, *ARRB1*, *DAGLB*, *DHX36*) could be targeted by agonist-based therapy. However, for these targets there is no agonist compound reported in DGIdb.

As some potential disease-associated targets may be linked with adverse side-effects, we undertook a phenome-wide analysis (PheWAS) by using Gene ATLAS, which includes 778 diseases-traits from the UK Biobank<sup>41</sup>. Genes identified as candidates for drug inhibitors were included in this analysis (*CTSK*, *MAP3K11*, *KCNH2*, *ATP1B1*, *CAMK1D*, *PDGFD*, *IPO9*, *GGCX*, *CETP*). The strongest instrumental variable (lowest P-value) in MR or the variant prioritized in colocalization analyses for each of the potential drug target gene was assessed in the cross-phenotype association analysis. Diseases-traits were deemed significantly associated with the variant at FDR 5% and data are presented in **Suppl. Table 15**. In GTEx, gene variant A-rs12042263 (strongest IV in MR for the expression) is associated with a lower vascular expression of *CTSK* ( $P_{\text{eQTL}}=1.69 \times 10^{-10}$ ) and a decreased risk for CAD (OR: 0.96, 95%CI: 0.95-0.98,  $P_{\text{GWAS}}=1.18 \times 10^{-3}$ ). In PheWAS, A-rs12042263 is positively associated with the risk of cerebrovascular disease ( $P_{\text{PheWAS}}=3.15 \times 10^{-4}$ ) (167 other cerebrovascular disease) (**Suppl. Table 15**). These data suggest that the inhibition of *CTSK* may lower the risk of CAD (data in MR are concordant as the vascular expression of *CTSK* is positively associated with CAD-risk), but at the expense of increasing cerebrovascular events. Several drugs under investigation are reported in DGIdb to inhibit *CTSK* (**Suppl. Table 14**). A recent randomized clinical trial evaluating odanacatib, a *CTSK* inhibitor developed for postmenopausal osteoporosis, has found that inhibition of the target lowered primary endpoints, but increased the risk of stroke (HR 1.32,  $p=0.03$ )<sup>42</sup>. Also, inhibition of *ATP1B1*, a gene positively associated with the CAD-risk in MR, is predicted to increase the risk of venous thromboembolic disease (**Suppl. Table 15**). Prediction based on colocalization analysis suggests that inhibition of *GGCX* may lower CAD-risk, but according to the PheWAS it is associated with an increase risk of malignant neoplasm of prostate ( $P_{\text{PheWAS}}=3.06 \times 10^{-9}$ ) (**Suppl. Table 15**). We next interrogated the Human Phenotype Ontology database<sup>43</sup> to identify disease-trait associations for the genes deemed druggable (**Suppl. Table 16**). In Human Phenotype Ontology, *CTSK*, *GGCX* and *KCNH2*, have been linked to bone-related conditions, coagulation defects and ventricular arrhythmia respectively. *KCNH2* encodes for the Ether-A-Go-Go-Related Protein 1, a potassium voltage-gated channel, involved in arrhythmia. Consistently, in the Side Effect Resource (SIDER)<sup>44</sup>, several drugs targeting *KCNH2* in DGIdb are reported to induce ventricular arrhythmia. After a comprehensive filtering based on multiple resources including a PheWAS analysis and interrogation of Human Phenotype Ontology as well as SIDER databases, potential vascular drug targets such as *MAP3K11*, *CAMK1D*, *PDGFD*, *IPO9* and *CETP* were not predicted to be associated with

major adverse side effects (cardiovascular, neurologic, metabolic, cancer) that would result from drug inhibition. Thus, some of these genes may represent suitable potential drug targets for follow-up studies.

## Impact of targeting MAP3K11 in vascular smooth muscle cells

By using the present framework, we showed that some causally associated vascular genes were central in a network and were potentially druggable. Among those genes, *MAP3K11* (also known as *MLK3*) is positively associated with the CAD risk (OR: 1.07, 95%CI: 1.05-1.10,  $P_{\text{causal}} = 8.46 \times 10^{-9}$ ) and is a target of the experimental compound URM-099<sup>45</sup>. Community network analysis using random walks showed that a module including *MAP3K11* ( $P = 1.94 \times 10^{-6}$ ) (**Suppl. Figure 2**) was enriched in GO for the regulation of JNK cascade ( $P = 1.81 \times 10^{-10}$ ) (**Suppl. Table 17**), which is involved in inflammation and cell migration<sup>46</sup>. URM-099 is an experimental compound that inhibits MAP3K11. We hypothesized that URM-099 may modulate the expression of key cytokines and transcription factors involved in the development of atherosclerosis. We observed that vascular smooth muscle cells (VSMCs) treated with URM-099 (100nM) for 6 hours had lower expression of transcripts encoding for interleukin1 beta (*IL1B*), C-C motif chemokine ligand 2 (*CCL2* and also known as *MCPT1*) and plasminogen activator urokinase (*PLAU*) (Figure 5A-C). IL1B, CCL2 and PLAU are key mediators involved in plaque activity<sup>47-49</sup>. Early growth response 1 (*EGR1*) is a transcription factor known to enhance the expression of IL1B and chemokines as well as to promote the development of atherosclerosis in mice<sup>50</sup>. In isolated VSMCs, URM-099 reduced the transcript level encoding for *EGR1* by 30% (Figure 5D). Considering the pathway enrichment of CAD gene variants in cell migration (**Suppl. Table 2**), we assessed the impact of URM-099 on VSMC transmigration in a Boyden chamber. Consistent with the impact of URM-099 on genes having a role on cell migration such as *CCL2* and *EGR1*<sup>51</sup>, we observed a reduction of cell migration by 40% in VSMC treated with the inhibitor for 6 hours (Figure 5E). As a proof of concept, these data provide further evidence that the pharmacological inhibition of predicted causally associated genes such as *MAP3K11* in VSMC impacts athero-relevant gene expression and cell phenotype.

## Discussion

In this work, we provide evidence that promoter-enhancer anchor loops in HCASMC explain 22% of the heritability for CAD. 3D mapped vascular smooth muscle genes were enriched in eQTL genes and in predicted causal targets. After filtering by using cross-phenotype analyses and curation of Human Phenotype Ontology and SIDER databases, we identified a set of druggable vascular genes, which are not predicted to be associated with major adverse events. Network of CAD-causally associated vascular genes was enriched in TGF-beta signaling, HDAC class I and MAPK pathways. As an illustrative case, the pharmacological inhibition of causally associated gene *MAP3K11* in vascular smooth muscle cells reduced cell migration, a key process involved in the development of plaque<sup>52</sup>.

Among the causal candidate vascular genes identified by MR, *ARRB1* (OR: 0.84, 95%CI: 0.78-0.91,  $P_{\text{causal}}=9.45 \times 10^{-6}$ ) and *MFGE8* (OR: 1.13, 95%CI: 1.09-1.17,  $P_{\text{causal}}=2.08 \times 10^{-12}$ ) were the genes with the highest effect size on the risk of CAD. *ARRB1* encodes for arrestin beta 1, which is involved in the regulation of G-protein coupled receptor (GPCR) signaling. As illustrated by the network approach, *ARRB1* interacts with a large number of proteins and modulates, in a context specific manner, a myriad of signaling events related, among others, to inflammation<sup>53</sup>. At the 15q26.1 locus, we found a shared signal between the vascular expression of *MFGE8* and the risk of CAD (PP=0.90). Taken together, colocalization and MR analyses strongly militate for a role of *MFGE8*, which encodes for Milk Fat Globule EGF And Factor V/VIII Domain Containing (also known as lactadherin), on the risk of CAD. A previous study conducted in VSMCs showed that the knockdown of *MFGE8* reduced the proliferation rate of cells<sup>54</sup>. Similarly, in a mouse model of vascular injury, the silencing of *MFGE8* decreased the migration of VSMCs and the formation of neointima<sup>55</sup>. Taken together, these data are consistent with an implication of *MFGE8* on the proliferation/migration of VSMC, a key process in atherogenesis.

Among the causal associations with CAD, 15 genes were identified by enhancer-promoter looping including 6 genes that were mapped only by using 3D genome data. Among those genes, several novel causal candidates, which have not been investigated in the context of atherosclerosis, were identified and include *EXOSC5*, *B3GNT8*, *ARRB1*, *PIF1*, *UTP11*, *AGPAT4* and *MAP3K11*. For instance, MR data indicate that the vascular expression of *AGPAT4* is negatively associated with CAD-risk. *AGPAT4* encodes for a membrane enzyme that inactivates lysophosphatidic acid, a small lipid mediator with pro-atherogenic activity<sup>56</sup>. Also, *PIF1*, which encodes for a helicase involved in the activity of telomerase<sup>57</sup>, is a candidate gene for further exploration as the telomere length has been linked to the CAD-risk<sup>58</sup>.

Network provides a holistic approach in identifying pathways and gene modules with specialized functions in chronic disorders<sup>59</sup>. A pathway analysis of the network constructed by using the candidate causal genes showed an enrichment for TGF-beta and MAPK pathways. These data are consistent with a role of the TGF-beta pathway in atherogenesis<sup>60</sup>. Drug targets are enriched in hub molecules<sup>61</sup> and the present data are in line with this notion. In this regard, we found that causally associated vascular genes were acting as hub in a network and were also deemed druggable according to DGIdb. After a comprehensive assessment of potential side effects, we narrowed down the number of druggable causal candidate genes to 5 targets (*MAP3K11*, *CAMK1D*, *PDGFD*, *IPO9* and *CETP*). *PDGFD* encodes for a platelet derived growth factor with implication in atherosclerosis<sup>62</sup>. *CETP*, which has been identified by colocalization analysis, encodes for cholesteryl ester transfer protein. *CETP* is involved in the metabolism of high-density lipoprotein (HDL) and low-density lipoprotein (LDL). Genetic signal at the *CETP* locus suggested that inhibition of this enzyme may lower the risk of CAD<sup>63</sup>. Drugs targeting *CETP* were evaluated in 4 different randomized clinical trials with inconsistent results, which are possibly linked to drug-related off-target effects on the blood pressure and the design of studies (reviewed in<sup>64</sup>). In the present work, we identified by using MR that two druggable kinases (*MAP3K11*, *CAMK1D*) were positively associated with CAD. A previous analysis using MR showed that the expression of *CAMK1D* in blood cells

was positively associated with the risk of CAD (OR: 1.05)<sup>16</sup>. These data are concordant with the present findings, which demonstrate a causal association for the vascular expression of *CAMK1D* on the risk of CAD (OR: 1.11). As a proof of concept, we evaluated a drug under development, which targets MAP3K11, on the expression of key genes involved in atherosclerosis and on the migration of VSMCs. In VSMCs, drug-induced inhibition of MAP3K11 reduced the expression of *IL1B* and *CCL2*, two genes causally associated with CAD<sup>65,66</sup>. Also, the pharmacological inhibition of MAP3K11 decreased substantially the migration of VSMC, a cardinal process in the development of atheromatous plaques.

The present work has some limitations as causal inference using MR is subject to horizontal pleiotropy (i.e. gene variants that affect the outcome through an alternative pathway)<sup>67</sup>. However, the assessment of pleiotropy with the Cochran's Q test and also the implementation of sensitivity analyses with the median weighted analyses minimize this risk. Also, the assessment of both MR and colocalization provides robustness to the findings. As such, the combined evidence derived from colocalization and MR for a drug target (gene) increased substantially the likelihood for a drug to be licensed<sup>68</sup>.

## Methods

### CAD genetic associations

Summary statistics from GWAS data of a meta-analysis including 122,733 CAD cases and 424,528 controls from the UK Biobank and CARDIoGRAMplusC4D were downloaded for analyses<sup>28</sup>. GWAS meta-analysis was adjusted for age, gender and the first 30 principal components and included 7,947,838 gene variants. Cases from UK Biobank were identified by using ICD codes I21-I25 for ischemic heart disease and the corresponding OPCS-4 codes K40-K46, K49, K50 and K75 including percutaneous angioplasty. Patients reporting cardiovascular events such as myocardial infarction, coronary angioplasty and coronary artery bypass grafting were also included.

### Pathway analysis of genetic association data

Genetic association data for CAD were analyzed by using DEPICT, which provides an integrative approach to assess the most likely causal genes at risk loci and to infer pathway and tissue-cell enrichments<sup>69</sup>. Plink was used to identify independent loci based on GWAS significant SNPs ( $P_{\text{GWAS}} < 5E-8$ ) and DEPICT was used with the prioritization of genes, gene sets and tissue-cells.

### Mapping the GWAS

Genetic association data for CAD was mapped by using the Functional Mapping and Annotation of GWAS (FUMA) framework<sup>70</sup>. Genomic risk loci were defined using a pre-calculated LD structure of the 1000G EUR reference population. SNPs in genomic loci with LD  $r^2 < 0.6$ ,  $P\text{-value} < 5 \times 10^{-8}$  and  $\text{MAF} \geq 0.01$  were identified as independent significant SNPs (IndSigSNPs). IndSigSNPs independent from each other (LD  $r^2 < 0.1$ ) were identified as lead SNPs. Genomic loci closely located ( $< 250\text{kb}$  based on the most right

and left SNPs of each locus) were merged into one genomic locus. Gene annotation was based on Ensembl (build 85) and entrez ID yielding identification of 19,436 protein coding genes. Vascular eQTL (GTEx v8, aorta) were mapped by using IndSigSNPs to genes in *cis* ( $\pm 500\text{kb}$ ). SNP eQTL gene pairs were deemed significant at FDR 5%.

### **Analysis of HiChIP**

H3K27ac-HiChIP FASTQ files (GSE101498) from HCASMC were downloaded and aligned with HiC-Pro using the default settings<sup>71</sup>. Loop call was performed with FitHiChIP<sup>72</sup> at  $\text{FDR} < 1 \times 10^{-6}$  and a resolution of 5kb. Mapping of SNP to gene promoters was performed by using bedtools with the intersect function. Gene promoters were identified as a region  $\pm 2\text{kb}$  from the transcription start site (TSS) by using GENCODE version 35 in build 37. 1D H3K27ac-HiChIP track was generated using deepTools bamCoverage from the sorted BAM file generated by HiC-Pro. Hub enhancers were identified from H3K27ac-HiChIP by generating an interaction matrix of interacting pairs, which was analyzed with Cytoscape (version 3.8.2). Most connected hub enhancers were defined as those with a degree  $\geq 90^{\text{th}}$  percentile.

### **Analysis of ATAC-seq and ChIP-seq**

ATAC-seq and H3K27ac ChIP-seq data from HCASMC were downloaded (GSE124011). FASTQ files were extracted from SRA by using parallel-fastq-dump. Data were aligned on hg19 using Bowtie 2 and converted to bam files with Samtools. Duplicate reads were removed with Picard tools. Peak call was performed with MACS2 for broad peak with cutoff at 0.01. BigWig files were generated with deepTools bamCoverage and the sorted bam files.

### **Enrichment analysis of anchor loops with open chromatin**

Anchor loops and ATAC-seq bed files were analyzed for enrichment of overlap by using the R package GenometriCorr<sup>73</sup>. The projection test, which uses a binomial distribution was used to assess significance of calculated enrichment. Relative distance between the reference and query features were quantified by the density function correlation and illustrated in a heatmap.

### **Pathway analysis of anchor loops**

Significant loop calls of interacting pair regions in HCASMC H3K27ac-HiChIP identified from FitHiChIP were transformed into a bed file, which integrated the interacting regions or anchor loops. Enrichment was performed by using HOMER and the KEGG pathway.

### **Partitioned heritability**

Partitioned heritability of anchor loops in HCASMC was integrated to the full baseline model of 53 annotations in the stratified LD score regression framework<sup>30</sup>. LD score was calculated for each

chromosome. CAD summary statistics was converted using the `munge_sumstats.py` and partitioned heritability calculated using the `-h2` flag.

## Genetic Colocalization

Shared genetic signal between the eQTL and CAD was assessed by using the HyPrColoc package, which provides Bayesian colocalization analysis between traits<sup>74</sup>. Genomic regions were defined as 500kb down- and 500kb up-stream of the TSS of each gene. Shared genetic signal between the expression (eQTL) and CAD was considered if the posterior probability  $PP > 0.8$ . Data were visualized by using LocusCompare.

## Mendelian randomization

Causal inference was evaluated with two-sample MR by selecting independent SNPs (instrumental variables) associated with the expression. SNPs were analyzed within a window of 500kb around the TSS of each gene, then only SNPs strongly associated with the expression ( $P < 0.001$  corresponds to  $\sim F$  statistics  $> 10$ ) and independent ( $r^2 < 0.1$  based on the 1000G EUR reference panel) were selected as instrumental variables. MR was performed by using the Mendelian Randomisation package and a minimum of three instrumental variables was required to perform the analysis. Horizontal pleiotropy was estimated by using the Cochran's Q test and was considered significant when  $P_{\text{heterogeneity}} < 0.05$ . Genes were considered causally associated with CAD at FDR 5% using the IVW MR. Sensitivity analysis was performed with the weighted median MR.

## Network analysis

Candidate causal genes were used to extract a network from the DifferentialNet dataset<sup>39</sup>. Data inferred for the aorta were extracted and analyzed by using NetworkAnalyst<sup>75</sup>. Centrality indices were downloaded from NetworkAnalyst and hub nodes (genes) were identified as those with a degree  $\geq 90^{\text{th}}$  percentile. Community identification was performed by using the random walk algorithm as implemented in NetworkAnalyst. Enrichment analysis was performed in Enrichr with the BioPlanet 2019 resource.

## Druggable genome

The Drug Gene Interaction Database (DGIdb) was leveraged to assess the druggability of candidate causal genes<sup>40</sup>. Drug-gene pairs identified in DGIdb were evaluated as potential candidate by using the directional effect of the expression on the CAD risk and whether the compound was identified as an agonist or antagonist (in case of enzyme: an inhibitor).

## PheWAS analysis

PheWAS analysis was performed in Gene ATLAS, which integrates 778 traits computed from 452,264 individuals in the UK Biobank<sup>41</sup>. Traits and disorders were deemed significantly associated with a SNP at FDR 5%.

## Cell culture experiments

Human aorta smooth muscle cells were obtained from patients undergoing heart transplantation. All donated tissues have been obtained with an explicit written consent approved by the local ethical committee and the investigations conducted in accordance with the Helsinki Declaration. Aortic roots were cut, the tunica adventitia and tunica intima were removed, and tissues were cut into pieces and incubated 8 minutes in Trypsin (Invitrogen, Thermo Fisher Scientific, ON, Canada) at 37°C under agitation. Trypsin was then removed and tissues were resuspended in complete DMEM media (DMEM, 20% FBS with L-glutamine and sodium pyruvate). After cell growth (approximately 1 month), DMEM media was removed and cells were cultured in smooth muscle cell basal media with growth supplements (#310-470 and #311-GS, Cell Application, CA, USA). Cells were used between passages 2 to 5. Cells were treated with 100 nM of URMC-099 (MedKoo Biosciences, NC, USA) as indicated in the result section. The expression of *IL1B*, *CCL2*, *PLAU* and *EGR1* was evaluated by quantitative real-time PCR. RNA from cells was isolated with E.Z.N.A.® Micro RNA kit (Omega Bio-tek, VWR, QC, Canada). One µg of RNA was reverse transcribed using the Qscript cDNA supermix from Quanta (VWR, QC, Canada). qPCRs were performed with perfecta sybr supermix from Quanta on the Rotor-Gene 6000 system (Corbett Robotics Inc, CA, USA). Primers for *CCL2* were obtained from Qiagen (ON, Canada) and *IL1B*, *PLAU* and *EGR1* from IDT (IDT, IL, USA) (**Supplementary Table 18**). Transmigration assay was carried out in Boyden chamber with the QCM Chemotaxis Cell Migration Assay, 24-well (5 µm), fluorimetric (Millipore, Burlington, USA). 10,000 cells in DMEM 0% FBS with 100 nM of URMC-099 or with the control were loaded into the insert, and DMEM 10% FBS was applied outside the insert. Cell migration was quantified after 6 hours according to manufacturer's instructions.

## Statistics

For cell analyses, continuous data were expressed as mean ± SEM. Normality was tested with the Shapiro-Wilk test. Data with normal distribution were compared with Student t-test. For data with non-normal distribution, groups were compared with the Wilcoxon-Mann-Whitney test. Statistical analyses were performed with Prism 8.0.2. Hypergeometric and binomial tests were performed by using R version 4.0. FDR was calculated by using the R package multtest with the Benjamini and Hochberg test.

## Conclusions

In this work we found that the connectome of enhancer-promoter in HCASMC explained a significant proportion of the heritability for CAD. The mapping of genes using 3D enhancer-promoter contacts was enriched in vascular eQTL genes and causally-associated CAD genes. Among the causal candidate vascular genes, some are druggable and need further investigation to assess their role as potential pharmacologic targets for CAD.

## Abbreviations

CAD : coronary artery disease; eQTL: expression quantitative trait loci; GWAS: genetic association studies; HCASMC : human coronary artery smooth muscle cells; IVs: instrumental variables; LD: linkage disequilibrium; MR: Mendelian randomization; PheWAS: phenome-wide analysis study; RCTs: randomized clinical trials.

## **Declarations**

### **Ethics approval and consent to participate**

All donated tissues have been obtained with an explicit written consent approved by the local ethical committee and the investigations were conducted in accordance with the Helsinki Declaration. Datasets from GEO and the GWAS summary statistics were publicly available and did not require ethical approval.

### **Consent for publication**

Not applicable.

### **Availability of data and materials**

All the data are available in the manuscript including in the supplementary tables and figures. GWAS summary statistics for CAD and vascular eQTL (GTEx v8, aorta) as well as FASTQ files from GEO dataset were publicly available. Accession number and URLs are provided in the manuscript methods section.

### **Competing interests**

No competing interest to declare.

### **Funding**

This work was supported by the Canadian Institutes of Health Research grants to P.M. (FRN148778, FRN159697) and the Quebec Heart and Lung Institute Fund. Y.B. holds a Canada Research Chair in Genomics of Heart and Lung Diseases. S.T. and B.J.A. hold a scholarship from Fonds de Recherche du Québec-Santé (FRQS).

### **Authors' contributions**

A.C. and P.M. conceived and designed experiments. A.C. S.M. and P.M. conducted mapping, colocalization and MR analyses. S.M. and P.M. performed statistical analyses, enrichment and network analyses. P.M. analyzed HiChIP data. Functional assays were conducted by A.C. P.M. drafted the manuscript. S.M. drafted the figures. All the authors critically reviewed the manuscript and provided intellectual inputs.

### **Acknowledgments**

Not applicable.

Summary statistics GWAS CAD: <https://data.mendeley.com/datasets/gbbsrpx6bs/1>

GEO DataSets: <https://www.ncbi.nlm.nih.gov/gds>

GTEx v8: <https://www.gtexportal.org/home/>

DEPICT: <https://data.broadinstitute.org/mpg/depict/>

PLINK: <http://zzz.bwh.harvard.edu/plink/>

FUMA: <https://fuma.ctglab.nl/>

HiC-Pro: <https://github.com/nservant/HiC-Pro>

FitHiChIP: <https://ay-lab.github.io/FitHiChIP/>

parallel-fastq-dump: <https://github.com/rvalieris/parallel-fastq-dump>

Bowtie 2: <http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>

Samtools : <http://www.htslib.org/>

Picard: <https://broadinstitute.github.io/picard/>

MACS2: <https://pypi.org/project/MACS2/>

deepTools: <https://deeptools.readthedocs.io/en/develop/index.html>

GenometriCorr: <http://genometricorr.sourceforge.net/>

HOMER: <http://homer.ucsd.edu/homer/>

Partitioned heritability: <https://github.com/bulik/ldsc>

HyPrColoc: <https://github.com/jrs95/hyprcoloc>

LocusCompare: <http://locuscompare.com/>

MR: <https://cran.r-project.org/web/packages/MendelianRandomization/index.html>

NetworkAnalyst : <https://www.networkanalyst.ca/>

Cytoscape : <https://cytoscape.org/>

DGIdb: <https://www.dgidb.org/>

Gene Atlas: <http://geneatlas.roslin.ed.ac.uk/>

# References

1. Libby P, Buring JE, Badimon L, Hansson GK, Deanfield J, Bittencourt MS, Tokgözoğlu L, Lewis EF. Atherosclerosis. *Nature Reviews Disease Primers*. 2019;5:1–18.
2. Allahverdian S, Chaabane C, Boukais K, Francis GA, Bochaton-Piallat M-L. Smooth muscle cell fate and plasticity in atherosclerosis. *Cardiovasc Res*. 2018;114:540–550.
3. Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock P, Macleod M, Mignini LE, Jayaram P, Khan KS. Comparison of treatment effects between animal experiments and clinical trials: systematic review. *BMJ*. 2007;334:197.
4. Sattar N, Preiss D. Reverse Causality in Cardiovascular Epidemiological Research: More Common Than Imagined? *Circulation*. 2017;135:2369–2372.
5. Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, Karlsson A, Al-Lazikani B, Hersey A, Oprea TI, Overington JP. A comprehensive map of molecular drug targets. *Nat Rev Drug Discov*. 2017;16:19–34.
6. Khera AV, Kathiresan S. Genetics of coronary artery disease: discovery, biology and clinical translation. *Nat Rev Genet*. 2017;18:331–344.
7. Zheng J, Baird D, Borges M-C, Bowden J, Hemani G, Haycock P, Evans DM, Smith GD. Recent Developments in Mendelian Randomization Studies. *Current Epidemiology Reports*. 2017;4:330–345.
8. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37:658–665.
9. Burgess S, Foley CN, Zuber V. Inferring Causal Relationships Between Risk Factors and Outcomes from Genome-Wide Association Study Data. *Annu Rev Genomics Hum Genet*. 2018;19:303–327.
10. Floris M, Olla S, Schlessinger D, Cucca F. Genetic-Driven Druggable Target Identification and Validation. *Trends Genet*. 2018;34:558–570.
11. Hingorani AD, Kuan V, Finan C, Kruger FA, Gaulton A, Chopade S, Sofat R, MacAllister RJ, Overington JP, Hemingway H, Denaxas S, Prieto D, Casas JP. Improving the odds of drug development success through human genomics: modelling study. *Scientific Reports*. 2019;9:1–25.
12. Hinkson IV, Madej B, Stahlberg EA. Accelerating Therapeutics for Opportunities in Medicine: A Paradigm Shift in Drug Discovery. *Front Pharmacol*. 2020;11:770.
13. Diogo D, Tian C, Franklin CS, Alanne-Kinnunen M, March M, Spencer CCA, Vangjeli C, Weale ME, Mattsson H, Kilpeläinen E, Sleiman PMA, Reilly DF, McElwee J, Maranville JC, Chatterjee AK, Bhandari A, Nguyen K-DH, Estrada K, Reeve M-P, Hutz J, Bing N, John S, MacArthur DG, Salomaa V, Ripatti S, Hakonarson H, Daly MJ, Palotie A, Hinds DA, Donnelly P, Fox CS, Day-Williams AG, Plenge RM, Runz H. Phenome-wide association studies across large population cohorts support drug target validation. *Nat Commun*. 2018;9:4285.
14. Nguyen PA, Born DA, Deaton AM, Nioi P, Ward LD. Phenotypes associated with genes encoding drug targets are predictive of clinical trial side effects. *Nat Commun*. 2019;10:1579.

15. Robinson PN, Köhler S, Bauer S, Seelow D, Horn D, Mundlos S. The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. *Am J Hum Genet.* 2008;83:610–615.
16. Bon-Baret V, Chignon A, Boulanger M-C, Li Z, Argaud D, Arsenault BJ, Thériault S, Bossé Y, Mathieu P. System Genetics Including Causal Inference Identify Immune Targets for Coronary Artery Disease and the Lifespan. *Circ Genom Precis Med.* 2021;14:e003196.
17. Mitra K, Carvunis A-R, Ramesh SK, Ideker T. Integrative approaches for finding modular structure in biological networks. *Nat Rev Genet.* 2013;14:719–732.
18. Chignon A, Bon-Baret V, Boulanger M-C, Li Z, Argaud D, Bossé Y, Thériault S, Arsenault BJ, Mathieu P. Single-cell expression and Mendelian randomization analyses identify blood genes associated with lifespan and chronic diseases. *Communications Biology.* 2020;3:1–15.
19. Hu JX, Thomas CE, Brunak S. Network biology concepts in complex disease comorbidities. *Nat Rev Genet.* 2016;17:615–629.
20. Gloss BS, Dinger ME. Realizing the significance of noncoding functionality in clinical genomics. *Exp Mol Med.* 2018;50:1–8.
21. Mishra A, Hawkins RD. Three-dimensional genome architecture and emerging technologies: looping in disease. *Genome Med.* 2017;9:87.
22. Gorkin DU, Leung D, Ren B. The 3D genome in transcriptional regulation and pluripotency. *Cell Stem Cell.* 2014;14:762–775.
23. Huang J, Li K, Cai W, Liu X, Zhang Y, Orkin SH, Xu J, Yuan G-C. Dissecting super-enhancer hierarchy based on chromatin interactions. *Nat Commun.* 2018;9:943.
24. Madsen JGS, Madsen MS, Rauch A, Traynor S, Van Hauwaert EL, Haakonsson AK, Javierre BM, Hyldahl M, Fraser P, Mandrup S. Highly interconnected enhancer communities control lineage-determining genes in human mesenchymal stem cells. *Nat Genet.* 2020;52:1227–1238.
25. Krijger PHL, Di Stefano B, de Wit E, Limone F, van Oevelen C, de Laat W, Graf T. Cell-of-Origin-Specific 3D Genome Structure Acquired during Somatic Cell Reprogramming. *Cell Stem Cell.* 2016;18:597–610.
26. Lu L, Liu X, Huang W-K, Giusti-Rodríguez P, Cui J, Zhang S, Xu W, Wen Z, Ma S, Rosen JD, Xu Z, Bartels CF, Kawaguchi R, Hu M, Scacheri PC, Rong Z, Li Y, Sullivan PF, Song H, Ming G-L, Li Y, Jin F. Robust Hi-C Maps of Enhancer-Promoter Interactions Reveal the Function of Non-coding Genome in Neural Development and Diseases. *Mol Cell.* 2020;79:521-534.e15.
27. Porcu E, Rüeger S, Lepik K, eQTLGen Consortium, BIOS Consortium, Santoni FA, Reymond A, Kutalik Z. Mendelian randomization integrating GWAS and eQTL data reveals genetic determinants of complex and clinical traits. *Nat Commun.* 2019;10:3300.
28. van der Harst P, Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ Res.* 2018;122:433–443.
29. Mumbach MR, Satpathy AT, Boyle EA, Dai C, Gowen BG, Cho SW, Nguyen ML, Rubin AJ, Granja JM, Kazane KR, Wei Y, Nguyen T, Greenside PG, Corces MR, Tycko J, Simeonov DR, Suliman N, Li R, Xu J, Flynn RA, Kundaje A, Khavari PA, Marson A, Corn JE, Quertermous T, Greenleaf WJ, Chang HY.

- Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat Genet.* 2017;49:1602–1612.
30. Finucane HK, Reshef YA, Anttila V, Slowikowski K, Gusev A, Byrnes A, Gazal S, Loh P-R, Lareau C, Shores N, Genovese G, Saunders A, Macosko E, Pollack S, Brainstorm Consortium, Perry JRB, Buenrostro JD, Bernstein BE, Raychaudhuri S, McCarroll S, Neale BM, Price AL. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet.* 2018;50:621–629.
  31. Devallière J, Charreau B. The adaptor Lnk (SH2B3): an emerging regulator in vascular cells and a link between immune and inflammatory signaling. *Biochem Pharmacol.* 2011;82:1391–1402.
  32. Wang X, Musunuru K. Confirmation of Causal rs9349379- PHACTR1 Expression Quantitative Trait Locus in Human-Induced Pluripotent Stem Cell Endothelial Cells. *Circ Genom Precis Med.* 2018;11:e002327.
  33. Kasikara C, Schilperoort M, Gerlach B, Xue C, Wang X, Zheng Z, Kuriakose G, Dorweiler B, Zhang H, Fredman G, Saleheen D, Reilly MP, Tabas I. Deficiency of macrophage PHACTR1 impairs efferocytosis and promotes atherosclerotic plaque necrosis. *J Clin Invest.* 2021;131:145275.
  34. Kyriakakis E, Frisanti A, Dasen B, Pfaff D, Rivero O, Lesch K-P, Erne P, Resink TJ, Philippova M. T-cadherin promotes autophagy and survival in vascular smooth muscle cells through MEK1/2/Erk1/2 axis activation. *Cell Signal.* 2017;35:163–175.
  35. Nanda V, Wang T, Pjanic M, Liu B, Nguyen T, Matic LP, Hedin U, Koplev S, Ma L, Franzén O, Ruusalepp A, Schadt EE, Björkegren JLM, Montgomery SB, Snyder MP, Quertermous T, Leeper NJ, Miller CL. Functional regulatory mechanism of smooth muscle cell-restricted LMOD1 coronary artery disease locus. *PLoS Genet.* 2018;14:e1007755.
  36. Douglas G, Mehta V, Al Haj Zen A, Akoumianakis I, Goel A, Rashbrook VS, Trelfa L, Donovan L, Drydale E, Chuaiphichai S, Antoniadou C, Watkins H, Kyriakou T, Tzima E, Channon KM. A key role for the novel coronary artery disease gene JCAD in atherosclerosis via shear stress mechanotransduction. *Cardiovasc Res.* 2020;116:1863–1874.
  37. Nagao M, Lyu Q, Zhao Q, Wirka RC, Bagga J, Nguyen T, Cheng P, Kim JB, Pjanic M, Miano JM, Quertermous T. Coronary Disease-Associated Gene TCF21 Inhibits Smooth Muscle Cell Differentiation by Blocking the Myocardin-Serum Response Factor Pathway. *Circ Res.* 2020;126:517–529.
  38. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol.* 2016;40:304–314.
  39. Basha O, Shpringer R, Argov CM, Yeger-Lotem E. The DifferentialNet database of differential protein-protein interactions in human tissues. *Nucleic Acids Research.* 2018;46:D522–D526.
  40. Griffith M, Griffith OL, Coffman AC, Weible JV, McMichael JF, Spies NC, Koval J, Das I, Callaway MB, Eldred JM, Miller CA, Subramanian J, Govindan R, Kumar RD, Bose R, Ding L, Walker JR, Larson DE,

- Dooling DJ, Smith SM, Ley TJ, Mardis ER, Wilson RK. DGldb: mining the druggable genome. *Nat Methods*. 2013;10:1209–1210.
41. Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. *Nat Genet*. 2018;50:1593–1599.
42. McClung MR, O'Donoghue ML, Papapoulos SE, Bone H, Langdahl B, Saag KG, Reid IR, Kiel DP, Cavallari I, Bonaca MP, Wiviott SD, de Villiers T, Ling X, Lippuner K, Nakamura T, Reginster J-Y, Rodriguez-Portales JA, Roux C, Zanchetta J, Zerbini CAF, Park J-G, Im K, Cange A, Grip LT, Heyden N, DaSilva C, Cohn D, Massaad R, Scott BB, Verbruggen N, Gurner D, Miller DL, Blair ML, Polis AB, Stoch SA, Santora A, Lombardi A, Leung AT, Kaufman KD, Sabatine MS, LOFT Investigators. Odanacatib for the treatment of postmenopausal osteoporosis: results of the LOFT multicentre, randomised, double-blind, placebo-controlled trial and LOFT Extension study. *Lancet Diabetes Endocrinol*. 2019;7:899–911.
43. Köhler S, Vasilevsky NA, Engelstad M, Foster E, McMurry J, Aymé S, Baynam G, Bello SM, Boerkoel CF, Boycott KM, Brudno M, Buske OJ, Chinnery PF, Cipriani V, Connell LE, Dawkins HJS, DeMare LE, Devereau AD, de Vries BBA, Firth HV, Freson K, Greene D, Hamosh A, Helbig I, Hum C, Jähn JA, James R, Krause R, F Laulederkind SJ, Lochmüller H, Lyon GJ, Ogishima S, Olry A, Ouwehand WH, Pontikos N, Rath A, Schaefer F, Scott RH, Segal M, Sergouniotis PI, Sever R, Smith CL, Straub V, Thompson R, Turner C, Turro E, Veltman MWM, Vulliamy T, Yu J, von Ziegenweidt J, Zankl A, Züchner S, Zemojtel T, Jacobsen JOB, Groza T, Smedley D, Mungall CJ, Haendel M, Robinson PN. The Human Phenotype Ontology in 2017. *Nucleic Acids Res*. 2017;45:D865–D876.
44. Kuhn M, Letunic I, Jensen LJ, Bork P. The SIDER database of drugs and side effects. *Nucleic Acids Res*. 2016;44:D1075-1079.
45. Marker DF, Tremblay M-È, Puccini JM, Barbieri J, Gantz Marker MA, Loweth CJ, Muly EC, Lu S-M, Goodfellow VS, Dewhurst S, Gelbard HA. The new small-molecule mixed-lineage kinase 3 inhibitor URM-099 is neuroprotective and anti-inflammatory in models of human immunodeficiency virus-associated neurocognitive disorders. *J Neurosci*. 2013;33:9998–10010.
46. Huang Z, Yan D-P, Ge B-X. JNK regulates cell migration through promotion of tyrosine phosphorylation of paxillin. *Cell Signal*. 2008;20:2002–2012.
47. Libby P. Interleukin-1 Beta as a Target for Atherosclerosis Therapy: Biological Basis of CANTOS and Beyond. *J Am Coll Cardiol*. 2017;70:2278–2289.
48. Lin J, Kakkar V, Lu X. Impact of MCP-1 in atherosclerosis. *Curr Pharm Des*. 2014;20:4580–4588.
49. Fuhrman B. The urokinase system in the pathogenesis of atherosclerosis. *Atherosclerosis*. 2012;222:8–14.
50. McCaffrey TA, Fu C, Du B, Eksinar S, Kent KC, Bush H, Kreiger K, Rosengart T, Cybulsky MI, Silverman ES, Collins T. High-level expression of Egr-1 and Egr-1-inducible genes in mouse and human atherosclerosis. *J Clin Invest*. 2000;105:653–662.
51. Tarcic G, Avraham R, Pines G, Amit I, Shay T, Lu Y, Zwang Y, Katz M, Ben-Chetrit N, Jacob-Hirsch J, Virgilio L, Rechavi G, Mavrothalassitis G, Mills GB, Domany E, Yarden Y. EGR1 and the ERK-ERF axis

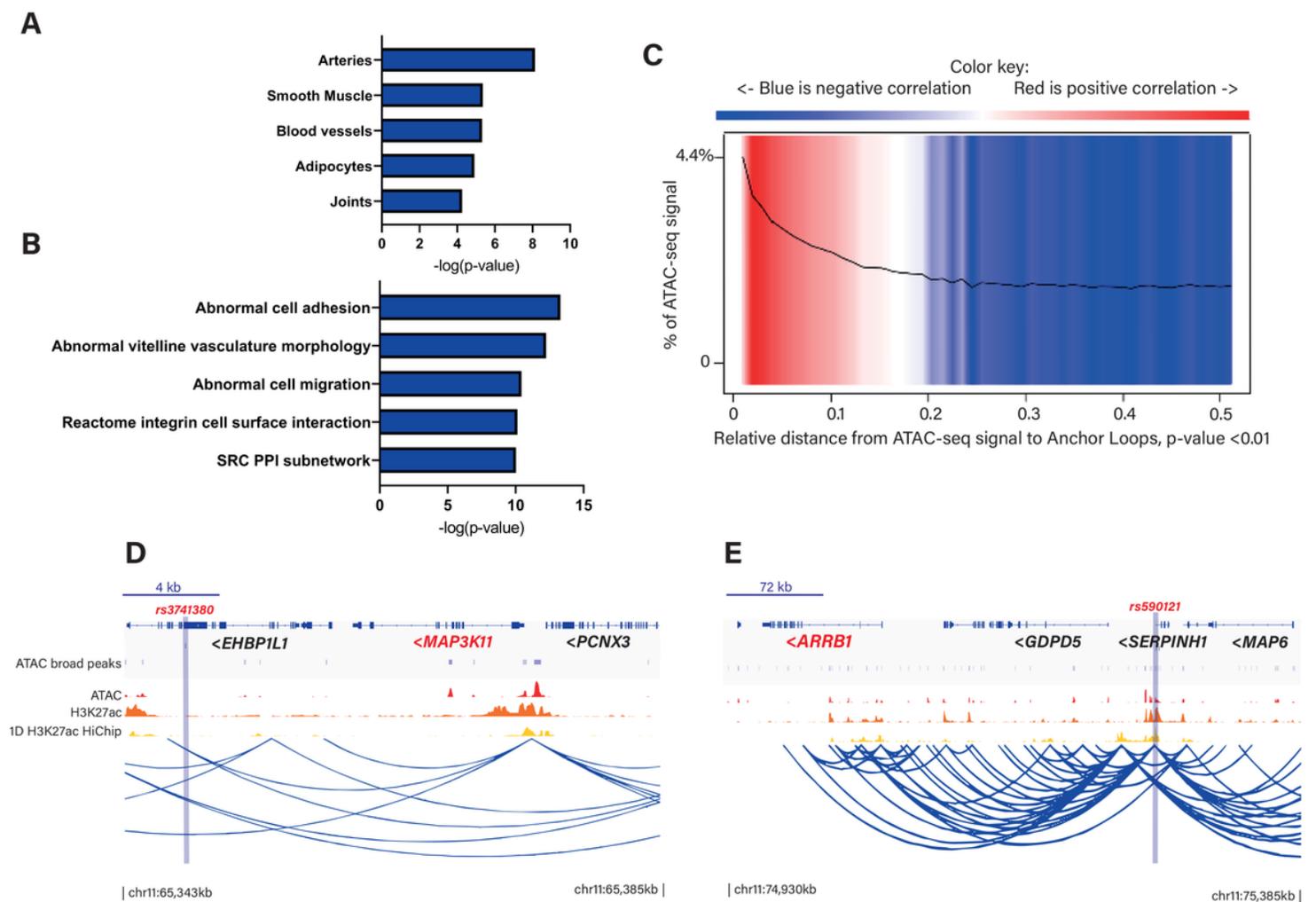
- drive mammary cell migration in response to EGF. *FASEB J.* 2012;26:1582–1592.
52. Grootaert MOJ, Bennett MR. Vascular smooth muscle cells in atherosclerosis: Time for a reassessment. *Cardiovasc Res.* 2021;cvab046.
53. Freedman NJ, Shenoy SK. Regulation of inflammation by  $\beta$ -arrestins: Not just receptor tales. *Cell Signal.* 2018;41:41–45.
54. Soubeyrand S, Nikpay M, Turner A, Dang A-T, Herfkens M, Lau P, McPherson R. Regulation of MFGE8 by the intergenic coronary artery disease locus on 15q26.1. *Atherosclerosis.* 2019;284:11–17.
55. Viola JR, Lemnitzer P, Paulin N, Drechsler M, Nazari-Jahantigh M, Maas S, De Jong RJ, Winter J, Schober A, Weber C, Atabai K, Soehnlein O. Deletion of MFGE8 Inhibits Neointima Formation upon Arterial Damage. *Thromb Haemost.* 2018;118:1340–1342.
56. Schober A, Siess W. Lysophosphatidic acid in atherosclerotic diseases. *Br J Pharmacol.* 2012;167:465–482.
57. Stinus S, Paeschke K, Chang M. Telomerase regulation by the Pif1 helicase: a length-dependent effect? *Curr Genet.* 2018;64:509–513.
58. Farzaneh-Far R, Cawthon RM, Na B, Browner WS, Schiller NB, Whooley MA. Prognostic value of leukocyte telomere length in patients with stable coronary artery disease: data from the Heart and Soul Study. *Arterioscler Thromb Vasc Biol.* 2008;28:1379–1384.
59. Barabási A-L, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet.* 2011;12:56–68.
60. Chen P-Y, Qin L, Li G, Wang Z, Dahlman JE, Malagon-Lopez J, Gujja S, Cilfone NA, Kauffman KJ, Sun L, Sun H, Zhang X, Aryal B, Canfran-Duque A, Liu R, Kusters P, Sehgal A, Jiao Y, Anderson DG, Gulcher J, Fernandez-Hernando C, Lutgens E, Schwartz MA, Pober JS, Chittenden TW, Tellides G, Simons M. Endothelial TGF- $\beta$  signalling drives vascular inflammation and atherosclerosis. *Nature Metabolism.* 2019;1:912–926.
61. Cheng F, Desai RJ, Handy DE, Wang R, Schneeweiss S, Barabási A-L, Loscalzo J. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat Commun.* 2018;9:2691.
62. Folestad E, Kunath A, Wågsäter D. PDGF-C and PDGF-D signaling in vascular diseases and animal models. *Mol Aspects Med.* 2018;62:1–11.
63. Webb TR, Erdmann J, Stirrups KE, Stitzel NO, Masca NGD, Jansen H, Kanoni S, Nelson CP, Ferrario PG, König IR, Eicher JD, Johnson AD, Hamby SE, Betsholtz C, Ruusalepp A, Franzén O, Schadt EE, Björkegren JLM, Weeke PE, Auer PL, Schick UM, Lu Y, Zhang H, Dube M-P, Goel A, Farrall M, Peloso GM, Won H-H, Do R, van Iperen E, Kruppa J, Mahajan A, Scott RA, Willenborg C, Braund PS, van Capelleveen JC, Doney ASF, Donnelly LA, Asselta R, Merlini PA, Duga S, Marziliano N, Denny JC, Shaffer C, El-Mokhtari NE, Franke A, Heilmann S, Hengstenberg C, Hoffmann P, Holmen OL, Hveem K, Jansson J-H, Jöckel K-H, Kessler T, Kriebel J, Laugwitz KL, Marouli E, Martinelli N, McCarthy MI, Van Zuydam NR, Meisinger C, Esko T, Mihailov E, Escher SA, Alver M, Moebus S, Morris AD, Virtamo J, Nikpay M, Olivieri O, Provost S, AlQarawi A, Robertson NR, Akinsansya KO, Reilly DF, Vogt TF, Yin W,

- Asselbergs FW, Kooperberg C, Jackson RD, Stahl E, Müller-Nurasyid M, Strauch K, Varga TV, Waldenberger M, Wellcome Trust Case Control Consortium, Zeng L, Chowdhury R, Salomaa V, Ford I, Jukema JW, Amouyel P, Kontto J, MORGAM Investigators, Nordestgaard BG, Ferrières J, Saleheen D, Sattar N, et al. Systematic Evaluation of Pleiotropy Identifies 6 Further Loci Associated With Coronary Artery Disease. *J Am Coll Cardiol*. 2017;69:823–836.
64. Holmes MV, Smith GD. Dyslipidaemia: Revealing the effect of CETP inhibition in cardiovascular disease. *Nat Rev Cardiol*. 2017;14:635–636.
65. Ridker PM, Everett BM, Thuren T, MacFadyen JG, Chang WH, Ballantyne C, Fonseca F, Nicolau J, Koenig W, Anker SD, Kastelein JJP, Cornel JH, Pais P, Pella D, Genest J, Cifkova R, Lorenzatti A, Forster T, Kobalava Z, Vida-Simiti L, Flather M, Shimokawa H, Ogawa H, Dellborg M, Rossi PRF, Troquay RPT, Libby P, Glynn RJ, CANTOS Trial Group. Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N Engl J Med*. 2017;377:1119–1131.
66. Georgakis MK, Gill D, Rannikmäe K, Traylor M, Anderson CD, Lee J-M, Kamatani Y, Hopewell JC, Worrall BB, Bernhagen J, Sudlow CLM, Malik R, Dichgans M. Genetically Determined Levels of Circulating Cytokines and Risk of Stroke. *Circulation*. 2019;139:256–268.
67. Cho Y, Haycock PC, Sanderson E, Gaunt TR, Zheng J, Morris AP, Davey Smith G, Hemani G. Exploiting horizontal pleiotropy to search for causal pathways within a Mendelian randomization framework. *Nat Commun*. 2020;11:1010.
68. Zheng J, Haberland V, Baird D, Walker V, Haycock PC, Hurle MR, Gutteridge A, Erola P, Liu Y, Luo S, Robinson J, Richardson TG, Staley JR, Elsworth B, Burgess S, Sun BB, Danesh J, Runz H, Maranville JC, Martin HM, Yarmolinsky J, Laurin C, Holmes MV, Liu JZ, Estrada K, Santos R, McCarthy L, Waterworth D, Nelson MR, Smith GD, Butterworth AS, Hemani G, Scott RA, Gaunt TR. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat Genet*. 2020;52:1122–1131.
69. Pers TH, Karjalainen JM, Chan Y, Westra H-J, Wood AR, Yang J, Lui JC, Vedantam S, Gustafsson S, Esko T, Frayling T, Speliotes EK, Genetic Investigation of ANthropometric Traits (GIANT) Consortium, Boehnke M, Raychaudhuri S, Fehrmann RSN, Hirschhorn JN, Franke L. Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun*. 2015;6:5890.
70. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8:1826.
71. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen C-J, Vert J-P, Heard E, Dekker J, Barillot E. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol*. 2015;16:259.
72. Bhattacharyya S, Chandra V, Vijayanand P, Ay F. Identification of significant chromatin contacts from HiChIP data by FitHiChIP. *Nat Commun*. 2019;10:4221.
73. Favorov A, Mularoni L, Cope LM, Medvedeva Y, Mironov AA, Makeev VJ, Wheelan SJ. Exploring massive, genome scale datasets with the GenometriCorr package. *PLoS Comput Biol*. 2012;8:e1002529.

74. Foley CN, Staley JR, Breen PG, Sun BB, Kirk PDW, Burgess S, Howson JMM. A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nat Commun.* 2021;12:764.
75. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.* 2019;47:W234–W241.

## Figures

**Figure 1**

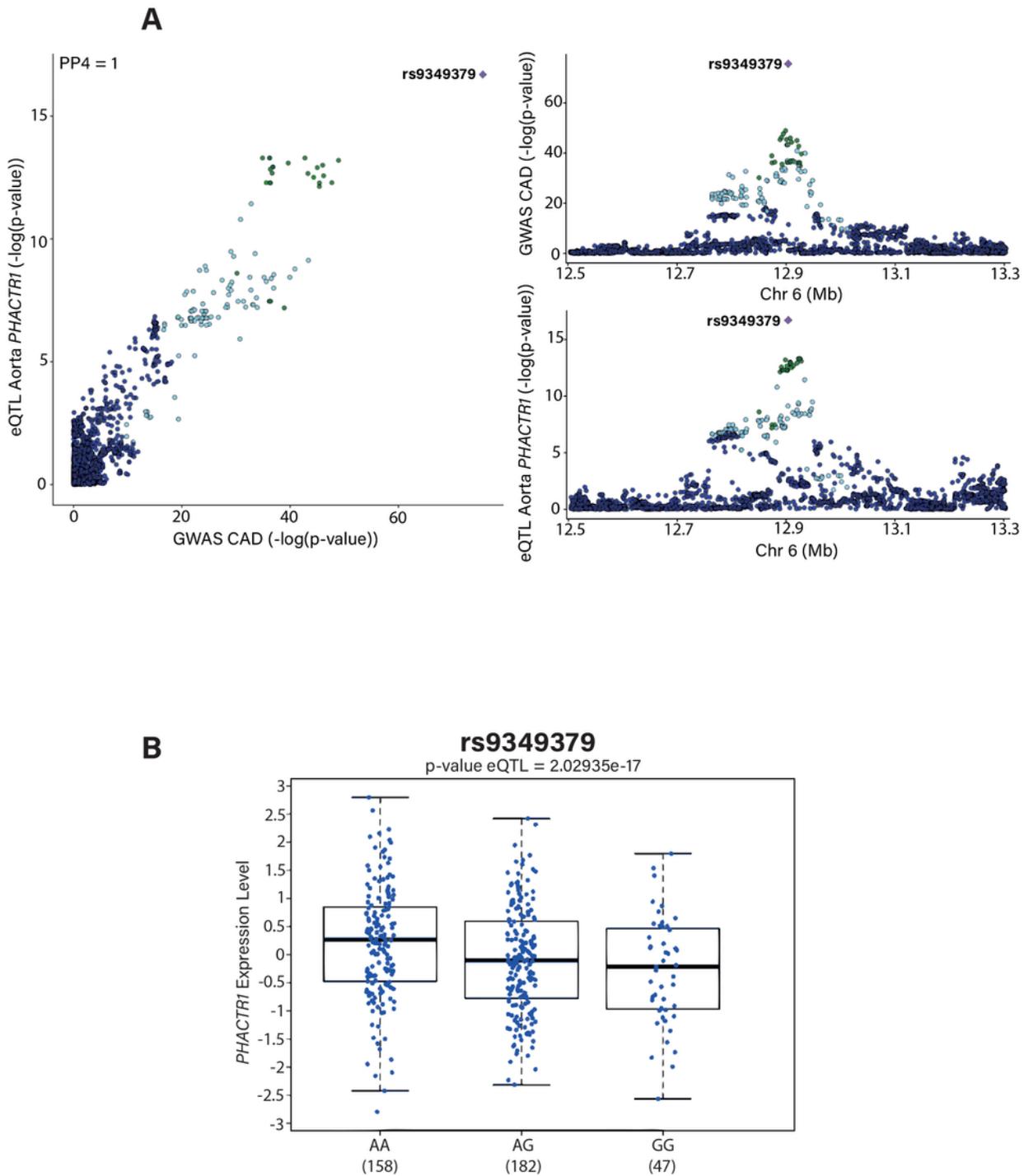


**Figure 1**

Enrichment and enhancer-promoter mapping. A-B) Tissue and pathway enrichment of CAD GWAS. C) Heatmap representing the relative distance of ATAC-seq signal to enhancer-promoter anchor loops; correlation is the measure of the observed to expected relative distance to the query point. D-E) Individual independent significant SNPs and 3D mapping in enhancer-promoter HiChIP; tracks represent genes,

ATAC-seq, H3K27ac ChIP, 1D H3K27ac-HiChIP and arcs of significant loops; vertical bar represents the SNP of interest and mapped genes are in red.

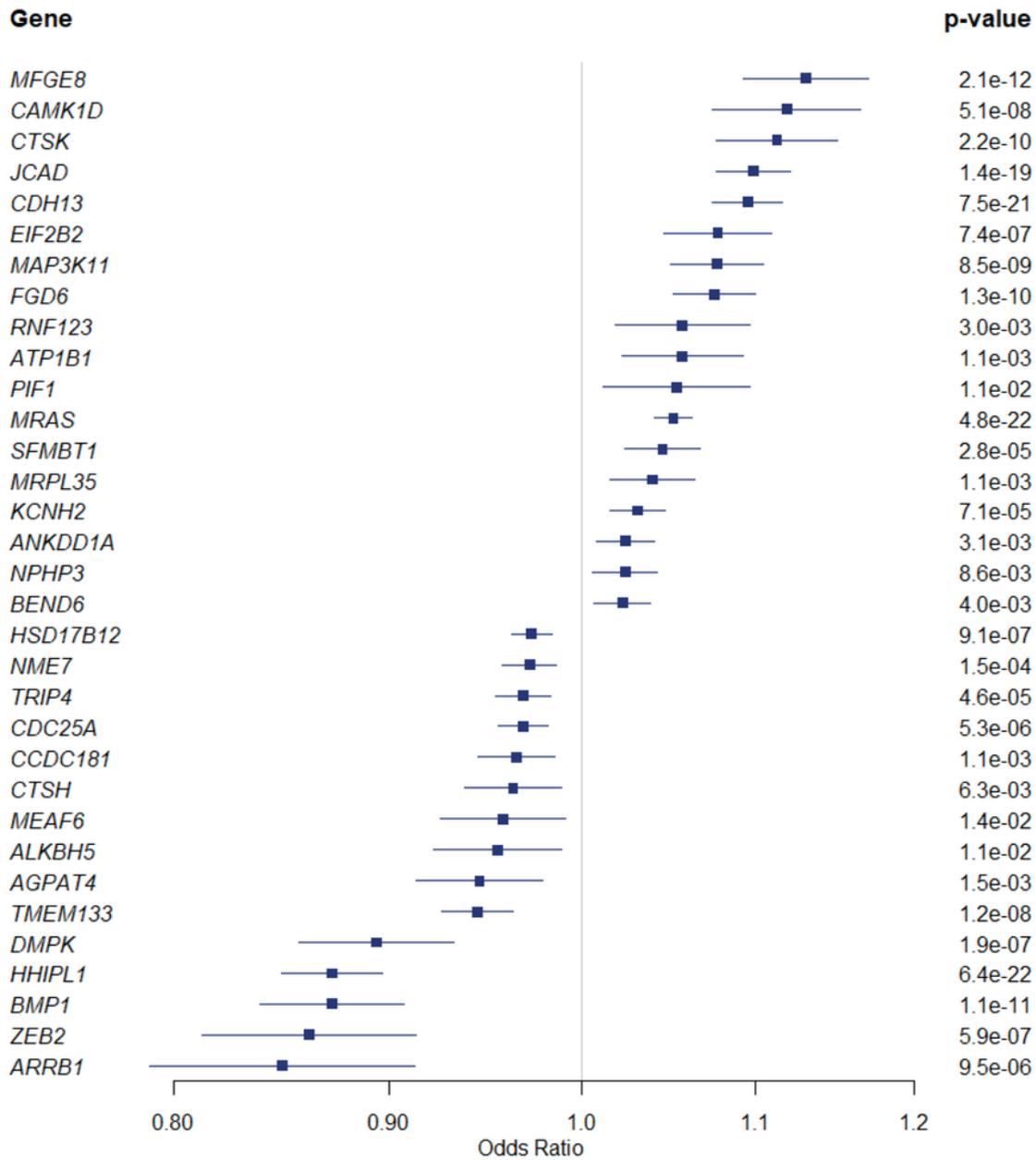
**Figure 2**



**Figure 2**

Colocalization and vascular eQTL. A) LocusCompare plot showing  $-\log p$ -values for CAD-GWAS and eQTL in the aorta (GTEx v8) at the PHACTR1 locus. B) eQTL in the aorta (GTEx v8) for rs9349379.

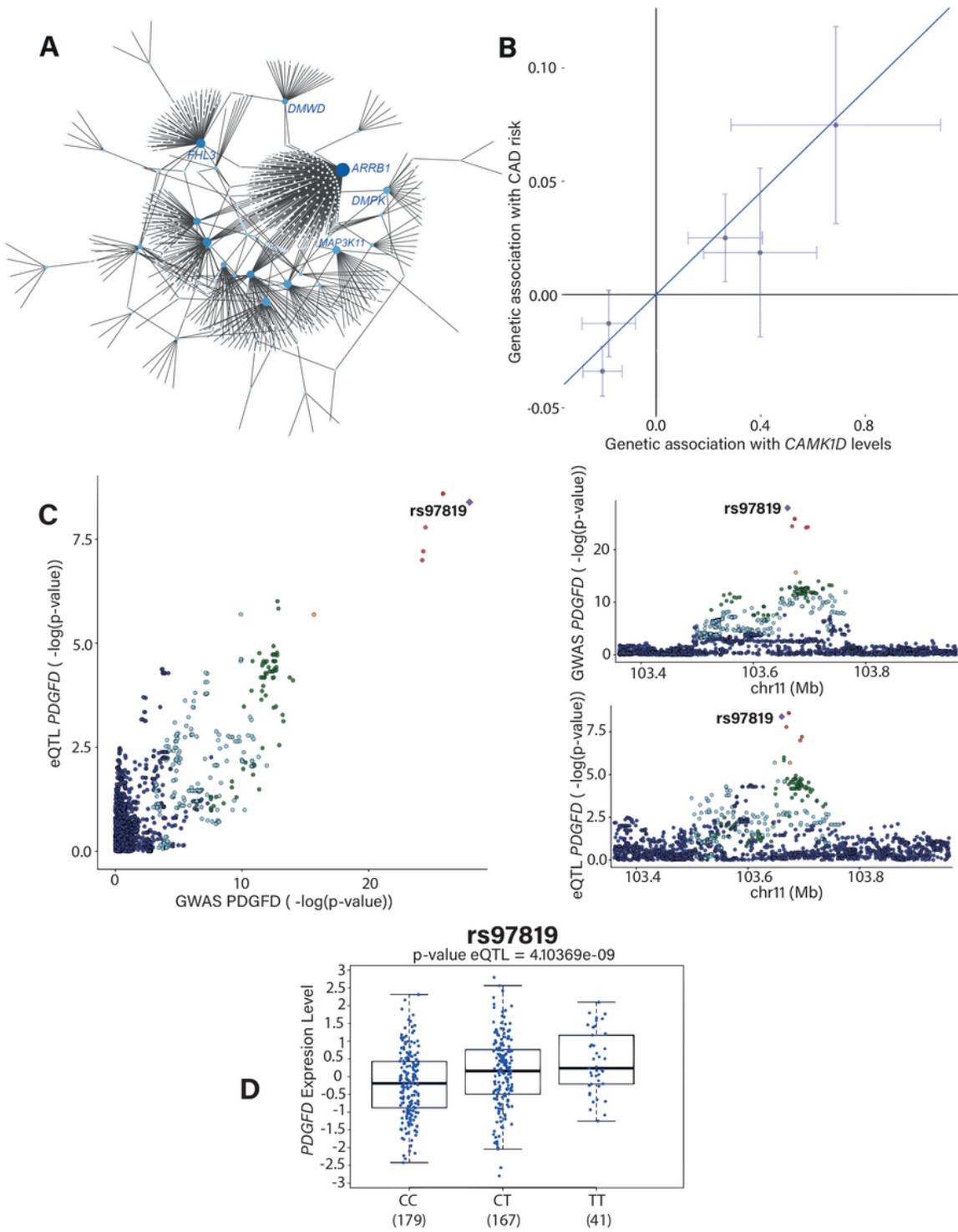
**Figure 3**



**Figure 3**

Mendelian randomization. Odds ratio and 95% CI for vascular genes (aorta) and CAD risk. Data are for 1SD.

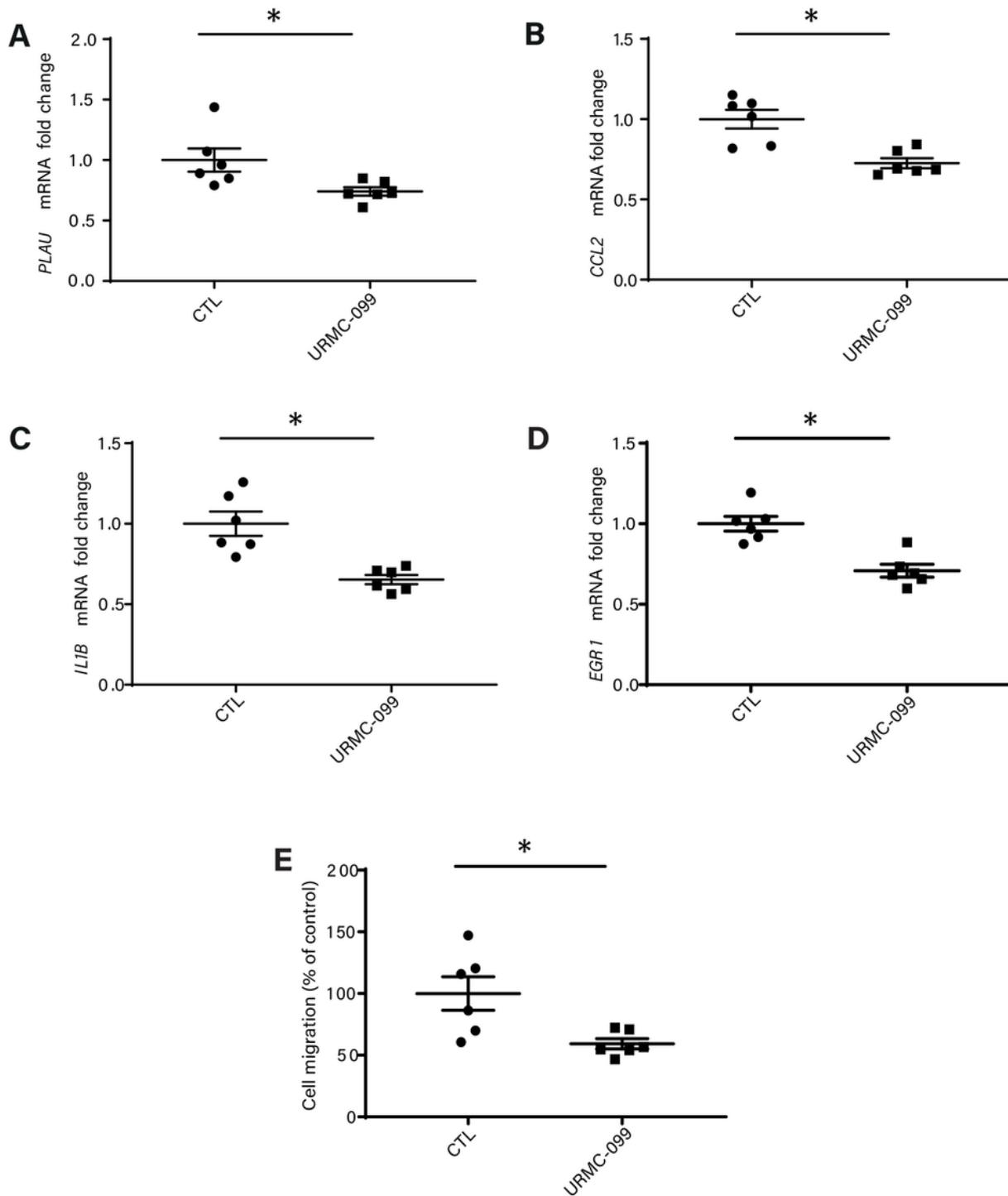
**Figure 4**



**Figure 4**

Network and causally associated vascular genes. A) Protein interactions derived network extracted from causally associated vascular genes. B) IVW MR for CAMK1D. C) LocusCompare plot showing  $-\log p$ -values for CAD-GWAS and eQTL in the aorta (GTEx v8) at the PDGFD locus. D) eQTL in the aorta (GTEx v8) for rs97819.

**Figure 5**



**Figure 5**

Functional impact of MAP3K11 inhibition on HCASMC. A-D) VSMC treated with an inhibitor of MAP3K11 (URM-099, 100 nM for 6 hours) and expression mRNAs encoding for PLAU, CCL2, IL1B and EGR1. E) Cell transmigration assay with URM-099 (100 nM for 6 hours).

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ST1.xlsx](#)
- [ST2.xlsx](#)
- [ST3.xlsx](#)
- [ST4.xlsx](#)
- [ST5.xlsx](#)
- [ST6.xlsx](#)
- [ST7.xlsx](#)
- [ST8.xlsx](#)
- [ST9.xlsx](#)
- [ST10.xlsx](#)
- [ST11.xlsx](#)
- [ST12.xlsx](#)
- [ST13.xlsx](#)
- [ST14.xlsx](#)
- [ST15.xlsx](#)
- [ST16.xlsx](#)
- [ST17.xlsx](#)
- [ST18.xlsx](#)
- [SUPPL1.pdf](#)
- [SUPPL2.pdf](#)