Defining the Functional Mechanism of *FHL5* that Underlies Coronary Artery Disease Risk

Doris Wong
Quincy, MA

BS, Biochemistry, Boston College, 2015

A Dissertation presented to the Graduate Faculty of the
University of Virginia in Candidacy for the
Degree of Doctor of Philosophy

Department of Biochemistry & Molecular Genetics

University of Virginia
December 2022

Table of Contents

List of Abbreviations

ABC	Activity by Contact
AOR	Aorta
ATAC	Assay for Transposase-Accessible Chromatin
BETA	Binding and Expression Target Analysis
BMP	Bone Morphogenetic Proteins
CAC	Coronary Artery Calcification
CAD	Coronary Artery Disease
CGRP	Calcitonin Gene-Related Peptide
cIMT	Carotid Intimal to Medial Thickness
CRE	cAMP Response Element
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CUT&RUN	Cleavage Under Targets and Release Using Nuclease
dCas9	Cas9 Endonuclease Dead
DEG	Differentially Expressed Genes
ECM	Extracellular Matrix
ENCODE	Encyclopedia of DNA Elements
eQTL	Expression Quantitative Trait Locus
FHL	Four and a Half LIM
FHS	Framingham Heart Study
GO	Gene Ontology
GREAT	Genomics Regions Enrichment of Annotations Tool
GREGOR	Genomic Regulatory Elements and GWAS Overlap Algorithm
GRN	Gene Regulatory Network
gRNA	Guide-RNA
GTEx	Genotype-Tissue Expression Project
GWAS	Genome-wide Association Study
HA	Hemagglutinin
HCASMC	Human Coronary Artery Smooth Muscle Cell
H&E	Hematoxylin and Eosin
HEIDI	Heterogeneity in Dependent Instruments
HDACs	Histone Deacetylase
iPSC	Induced Pluripotent Stem Cells
KDA	Key Driver Analysis
LD	Linkage Disequilibrium
LDL	Low Density Lipoprotein
LIM	Lin11, Isl-1, and Mec-3 Domain
LINC	Linker of Nucleoskeleton and Cytoskeleton
LOLA	Locus Overlap Analysis
MAM	Mammary Artery
MAPCA	Major Aorto-Pulmonary Collateral Arteries
MI	Myocardial Infarction
MLC	Myosin Light Chain
MLCK	Myosin Light Chain Kinase
NLS	Nuclear Localization Signal

NO	Nitric Oxide
NT	Non-Targeting
PAINTOR	Probabilistic Annotation Integrator
PDGF	Platelet Derived Growth Factor
PRS	Polygenic Risk Score
SCENIC	Single-Cell Regulatory Network Inference and Clustering
SAM	Synergistic Activation Mediator
scRNA-seq	Single-Cell RNA-sequencing
SMC	Smooth Muscle Cells
SMR	Summary-Level Mendelian Randomization
SNP	Single Nucleotide Polymorphism
snATAC-seq	Single Nuclei Assay for Transposase-Accessible Chromatin with High Throughput sequencing
STAGE	Stockholm Atherosclerosis Gene Expression
STARNET	Stockholm-Tartu Atherosclerosis Reverse Networks Engineering Task
TF	Transcription Factor
TGFβ	Transforming Growth Factor Beta
TOF	Tetralogy of Fallot
UKBB	United Kingdom Bio Bank
WCGNA	Weighted Gene Co-expression Network Analysis

List of Figures

Dedications

To my parents and my sister,
Thank you for coming on this journey with me even if you never left Boston. Although you may not know exactly what this dissertation is about, it is as much mine as it is yours. Mom and Dad, Thank you for giving up your dreams so I could pursue mine. I hope I have made you proud. Bonnie, thank you for always picking up the phone and reminding me that the world is bigger place that is worth exploring.

To Collin,
Thank you for always believing in me, my science, and that everything will be okay. Even when I could not get bands on a gel. I would not have made it to the finish line without your unwavering support.

To Randy,
Thank you for all your smiles that remind me that there was more to graduate school than science.

To Drs. Rich, Civelek, Garrett-Bakelman and Kucenas,
Thank you for all of our helpful discussions throughout this journey about science and beyond. Your suggestions and feedback helped shape this project.

To Dr. Miller,
Thank you for giving me the space and guidance to grow scientifically and personally. Thank you for believing in me and this project even as I couldn't get bands on a gel. Thank for challenging me to develop the expertise needed to lead both experimental and computational projects. I will without a doubt take this skillset forward. Lastly, thank you for showing me the type of mentor I aspire to be. I hope to one day pay this mentorship forward.

To Adam, Daud, Nelson, Yipei, Chani, Wei, Gaelle, and Jose,
Thank you for listening to this story 100 times. I hope it actually was more interesting the 100^{th} time. Thank you for all our chats over veggie pizza, quesadillas, and coffee with less milk. I hope we will stay connected for a long time.

Chapter 1: Introduction

I: Vascular Physiology

The human body depends on a system of rapid nutrient delivery and waste removal to maintain homeostasis. The cardiovascular system, which includes the heart and a vast network of interconnected blood vessels, evolved to fulfill this critical role. Oxygenated blood flows through blood vessels to reach every organ in the human body, while removing carbon dioxide and metabolic waste products from active tissues.

Each circulatory circuit is composed of a system of blood vessels, which includes arteries, arterioles, capillaries, and veins (**Fig. 1.1**). Arteries are blood vessels that carry oxygenated blood to tissues. In contrast, veins carry deoxygenated blood back to the heart and lungs for oxygenation. The exception to these generalizations is the pulmonary circuit. The pulmonary artery carries deoxygenated blood to the lungs and the pulmonary vein delivers oxygenated blood to the heart. To coordinate the transportation of nutrients to the entire human body, the vasculature is organized in an arborized network. Conduit arteries carrying oxygenated blood from the heart branch into smaller diameter arterioles which connect to an extensive network of capillaries. This system maximizes surface area for efficient gas exchange at the tissue level. Despite structural differences that reflect functional and physiological demands, all blood vessels, aside from the capillaries, are composed of three layers: tunica interna

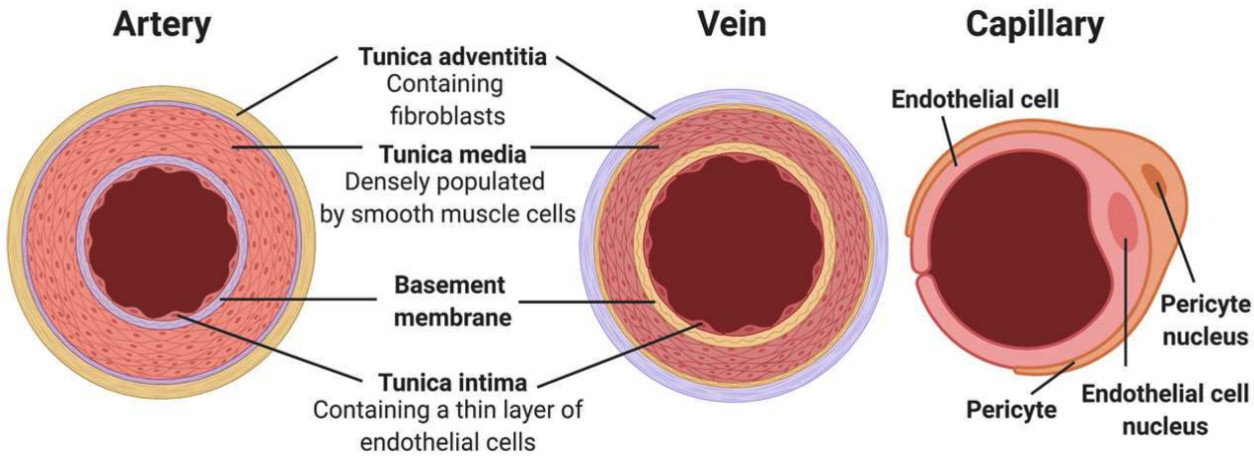

Figure 1.1: Structural Comparison Between the Major Types of Blood Vessels.

Created at BioRender.com

(intima), tunica media (media), and tunica externa (adventitia) [1] (**Fig 1.1**). Capillaries are the smallest blood vessels that are made up of an endothelial cell tube that is covered by pericytes embedded in a basement membrane (**Fig. 1.1**). The density of pericytes is very heterogeneous across different tissues and organs, with the brain and eye possessing the greatest coverage. Pericytes are specialized contractile mural cells that stabilize and regulate vascular tone of these microvessels. Other roles of pericytes in maintaining vascular integrity and angiogenesis, particularly during injury are areas of active research [2] [3] .

II. Cell-Type Composition of the Vessel Wall

The tunica interna consists of the endothelial cell layer lining the lumen that senses shear stress from blood flow [1]. Endothelial cells sit on a basement membrane of extracellular matrix fibers that help establish a tight selective barrier [3] [4]. In response to environmental conditions and mechanical strain, endothelial cells release local cytokines that act on other cells in the vessel wall. For example, endothelin and nitric oxide released by endothelial cells act on SMCs to modulate vascular tone [5] [6]. This is a primary mechanism of blood pressure regulation that is perturbed in hypertension [7].

The middle layer of the blood vessel, below the internal elastic lamina is known as the tunica media. This layer is the thickest portion of muscular and elastic arteries. Elastic arteries represent the largest class of arteries in the circulatory system and are named according to their extensive network of elastin fibers that is needed to withstand the pulsatile strain from the left ventricle [1]. Muscular arteries branch off of elastic arteries into smaller resistance arteries and arterioles and possess a higher proportion of smooth muscle cells (SMC) in the medial layer [1] [8]. SMCs are specialized spindle-shaped cells that maintain a quiescent phenotype, characterized by low proliferation and migration rate. These cells express a collection of contractile markers, ion channels, and transcription factors that enable SMCs to regulate vascular tone through coordinated contraction and relaxation [9]. The medial layer of healthy arteries is supported by an array of extracellular matrix (ECM) fibers. The ECM composition is often altered with age or during the course of vascular diseases (e.g. aneurysm,

hypertension, CAD). Remodeling events such as fragmentation of elastin, increased deposition and crosslinking of collagens, induce structural changes, such as increased intimal to medial thickness, arterial stiffness and reduced compliance [10] [11] [12].

The adventitial layer is the outermost layer composed mainly of fibroblasts and perivascular nerves embedded in a collagen rich extracellular matrix. In large conduit arteries, the adventitia contains the vasa vasorum, a network of microvessels that provide a blood supply to this region [13]. The adventitia has recently gained interest as a source of macrophages, progenitor cells and others that migrate into the vessel wall, under disease settings [14] [15]. Dysfunction in these adventitial cells contributes to disturbed blood flow to critical organs, either by blockage of the lumen or weakening of the vessel wall, with great clinical consequences [16].

Veins are composed of the same three layers as arteries, with a few key differences that accommodate the high capacitance of the venous system (**Fig. 1.1**) [1]. Since veins are subject to much less pressure, the medial layer is relatively thin, compared to the arterial medial layer. After blood exchange at the tissue levels, blood flows into venules, the smallest veins, which branch into larger vessels until the blood reaches the heart. Similar to the contrast of the pulmonary circuit, pulmonary veins transport freshly oxygenated blood from the lungs back to the heart for dissemination. Veins often have large diameters and given the low venous pressure, require the presence of valves to prevent back-flow.

III. Function of Smooth Muscle Cells in the Vessel Wall

The studies of this dissertation focus primarily on SMCs. Therefore, much of this introduction is dedicated to describing the function and upstream regulatory mechanisms of SMC differentiation, although it is noteworthy to mention that other cell types also contribute to maintaining vascular homeostasis. One of the primary functions of differentiated SMCs in the medial layer is contraction, a molecular process that regulates vascular tone and blood pressure [17]. Similar to the mechanisms in skeletal and cardiac myoblasts, SMC contraction involves the formation of actin-myosin cross

bridges. Although there are electrical, chemical and physical contractile stimuli, this process is largely dependent on calcium influx and calcium ion release from the sarcoplasmic reticulum. Free calcium ions in the cytoplasm bind calmodulin, an enzyme that activates the enzyme myosin light chain kinase (MLCK). MLCK-mediated phosphorylation of the myosin regulatory light chain (MLC) is necessary to induce the actin-myosin interactions that generate the necessary tension for cell shortening [17]. To halt the contractile cycle, myosin is dephosphorylated by myosin light chain phosphatase to disrupt the actin-myosin cross bridge cycle.

The dynamic regulation of MLC phosphorylation is also mediated by G-protein coupled receptor and nitric oxide signaling. Contractile agonists, such as phenylephrine, endothelin, and norepinephrine induce the release of intracellular calcium stores and subsequent activation of protein kinase C [18]. Conversely, vasorelaxation stimulants (e.g. adenosine, nitric oxide (NO), prostacyclin) function by increasing cyclic AMP levels to inhibit MLCK activity or stimulating soluble guanylyl cyclase activity to activate MLC phosphatase [19]. Upstream and downstream perturbations of these intricately regulated processes are implicated in various vascular diseases. For example, elevated plasma levels of the vasoconstrictor, endothelin, are linked to hypertension, restonsis and atherosclerosis [20] To underscore this role, endothelin receptor antagonists used in the clinic to treat pulmonary hypertension [21] .

IV. Epigenetic and Transcriptional Regulation of SMC Differentiation

The ability for SMCs to regulate vascular tone is due to its unique gene expression profile which includes high expression of the contractile genes, *MYH11*, *CNN1*, and *TAGLN* [22]*.* SMC differentiation is a complex process that is regulated by the combinatorial interactions between many transcriptional regulators and epigenetic mechanisms, such as the ubiquitously expressed MADS-box transcription factor, serum response factor (SRF) [23,24] [25] (**Fig. 1.2A**). SRF interactions with SMC-enriched transcriptional regulators, such as myocardin (MYOCD) drive expression of SMC marker genes. MYOCD is regarded as the master regulator of SMC differentiation with overexpression sufficient to induce expression of SMC marker genes and activate the

contractile phenotype [26] [27]. The Myocardin-Related Transcription Factor (MRTF) proteins, MRTFA and MRTFB promote similar differentiated SMC phenotypic states [28]. SRF activity is governed by the cooperative interactions of multiple CArG boxes and their spatial distribution [29] (**Fig. 1.2A**). A majority of SMC markers (e.g. *CNN1, MYH11, LMOD1)* harbor at least 2 highly conserved CArG or CArG-like sequences in their promoter and intronic regions [30] [31] [32]. Point mutations in one or more of these critical regulatory elements that abolishes SRF binding is sufficient to diminish gene expression [30].

Given the dynamics of chromatin organization, SMC differentiation is also regulated by chromatin accessibility, epigenetic marks, and DNA methylation [23]. Relative to other cell types in the vessel wall, enhancers functioning upstream of these cell specific markers, many of which harbor CArG boxes, are enriched in active histone marks [33] . Of particular interest, acetylation of H4 (H4ac), a feature which is enriched along SMC selective genes, facilitates the recruitment of SRF [33]. The chromatin mark, H3K4me2, enriched in similar regulatory elements help recruit DNA demethylation complexes (e.g. TET2) and stabilizes the SRF/MYOCD interaction (**Fig. 1.2A**) [33] [34]. This histone modification, in particular, is retained even following SMC phenotypic modulation. *In situ* proximity ligation assays pioneered by the Owens lab have leveraged this histone modification to perform rigorous lineage tracing in SMCs of human and mouse artery tissues throughout injury and disease [35]. DNA methylation is linked to gene repression and involves the addition of a methyl group to a cytosine residue at CpG islands. Relative to fibroblasts and non-SMCs, the promoters of SMC marker genes, such as *TAGLN* and *MYH11* are hypomethylated correlating with increased expression in SMCs [36]. This pattern of DNA methylation is modified by the activity of ten-eleven translocation (TET) methylcytosine dioxygenase family that is essential for DNA demethylation (**Fig. 1.2B**). TET2, an enzyme highly upregulated in SMCs, oxidizes the 5-methylcytosine to 5-hydroxymethylcytosine (5hmC) which is eventually converted to an unmethylated cytosine [37]. Recent studies show that 5hmC may also be dynamic epigenetic modification with crucial regulatory roles in transcription [38]. This modification serves as a

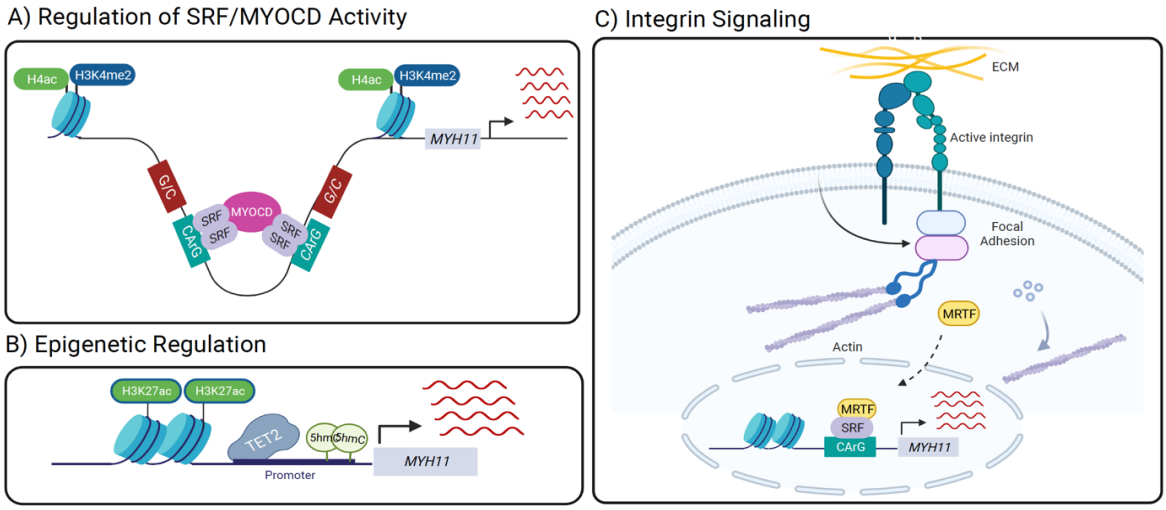

Figure 1.2: Mechanisms Regulating the Expression of SMC Marker Genes.

The expression of differentiated SMC markers is dependent on cis-regulatory elements, combinatorial interactions of transcription factors, and microenvironment cues. A) The expression of SMC markers, such as *MYH11*, is dependent on the cooperative interaction of the SRF/MYOCD transcriptional complex at multiple upstream CArG boxes. SRF/MYOCD binding at these cis-regulatory elements promotes the deposition of activating histone marks in the region to further stimulate gene expression. B) The DNA demethylase enzyme, TET2 oxidizes the repressive DNA methylation mark at promoter regions to hydroxymethycytosine (5-hmC). This modification is eventually removed inducing gene expression of these SMC marker genes. C) Interactions with the ECM activates focal adhesion signaling to induce the nuclear translocation of mechanosensitive transcription factors, such as MRTF. MRTF serves as a potent coactivator for SRF to induce expression of SMC contractile marker genes. Created at BioRender.com

binding motif for some transcription factors to facilitate chromatin opening throughout development, differentiation, and disease [39] [40].

This activity increases the deposition of active histone marks at SMC target gene loci and induces their gene expression [41]. Ectopic expression of TET2 is sufficient to induce SMC differentiation in non-SMCs [42]. The microenvironment plays a pivotal role in SMC phenotypic regulation as well (**Fig. 1.2C**). The ECM in the medial layer is composed of various collagen isoforms, elastin, proteoglycans, and glycoproteins, the composition of which is tightly regulated by a balance of the metalloproteinases (MMPs), disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS), and tissue inhibitors of metalloproteinases (TIMPs) activity [43] [44]. In addition to structural support, the ECM binds to and sequesters paracrine growth factors, such as TGFβ [45]. SMCs interact with the

underlying matrisome to sense changes in the stiffness of the environment and modulate their behavior accordingly largely through integrin mediated signaling [46] [47]. Integrins are a class of transmembrane receptors made up of alpha and beta subunits that recognize laminin, collagens, or the Arg-Gly-Asp (RBD) motif found on many ECM molecules [48]. This class of proteins is responsible for relaying extracellular changes to alter downstream signaling events. The affinity and sequence recognition of these receptors are largely dependent on its subunit makeup. The cytoplasmic portion of the receptor serves as a docking site for many adapter molecules at the focal adhesion complex that link receptor activity to the actin cytoskeleton [49,50] [51]. The nuclear translocation of mechanosensitive transcription factors (TFs), such as YAP/TAZ [52] and MRTFA/B [53], the latter of which is a potent coactivator of SRF and SMC differentiation, is dependent on the dynamics of cytoskeletal rearrangements (**Fig. 1.2C**). These TFs are sequestered in the cytoplasm through interactions with monomeric globular actin that mask a nuclear localization signal. Thus, polymerization of the monomeric actin subunits promotes nuclear import and induces gene expression [54] [55]. Perturbations in the ability of SMCs to sense and respond to the environment are associated with vascular remodeling, SMC migration, and arterial stiffness that occurs with aging. Unlike other terminally differentiated cells, SMCs maintain a degree of phenotypic plasticity in response to injury and environmental stimuli. Following injury or during disease, SMCs undergo a phenotypic modulation process to more closely resemble its precursor cell with reduced expression of SMC contractile markers [56] (**Fig 1.3**). This cell state is also characterized by increased cell proliferation, migration and ECM secretion. At the

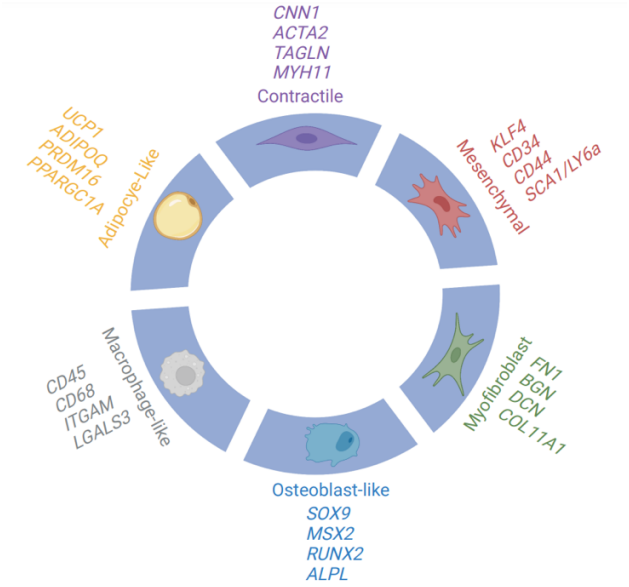

Figure 1.3: SMC Phenotypic Diversity

Due to their epigenetic plasticity, SMC can adopt characteristics of different cell types. Along the edges of the circle are cell markers for each state. Created at BioRender.com

morphological level, disruptions of myofilament organization in modulated SMCs impairs their normal contractile function [57]. *In vitro* studies of this phenomenon have implicated environmental factors, such as platelet derived growth factor (PDGF-BB) [58] [59], inflammatory cytokines (e.g. IL1-β) [60], and oxidized lipids [61]. In addition to resembling an earlier state in differentiation, a spectrum of SMC phenotypic states can be observed following injury or during vascular diseases. For example, in atherosclerotic lesions, SMCs assume phenotypes characteristic of other cell types, such as osteoblasts and fibroblasts (**Fig. 1.3**). This observation has been confirmed through both lineage tracing studies as well as single-cell RNA-sequencing studies (scRNA-seq) of human tissues [62] [63] [64]

II. Coronary artery disease Epidemiology

Coronary artery disease (CAD), one of the leading causes of death worldwide, is a disease that affects the blood vessels supplying blood to the heart muscle or myocardium [65]. Poor blood flow is often caused by the build-up of lipid laden plaques in the vessel wall, although there are also forms of disease that do not involve plaque obstruction. CAD encompasses a broad spectrum of diseases that can be further subdivided into three categories, obstructive, non-obstructive, or spontaneous coronary artery dissection (SCAD). Obstructive CAD is the most common form of ischemic heart disease with the presence of a plaque blocking blood flow. In contrast, non-obstructive CAD is caused by vasomotor defects in SMCs that constricts the vessel [66]. In contrast to the other two forms, the etiology of SCAD is distinct and involves a tear in the vessel wall of the coronary artery. This eventually occludes the artery causing myocardial injury [67].

Patients with late-stage CAD are at high risk for severe clinical endpoints, such as myocardial infarction (MI) and stroke. The American Heart Association (AHA) estimates that 13% of deaths in the United States are due to these complications with related healthcare costs surpassing 200 billion dollars [65]. Despite public health initiatives to increase cardiovascular health, over 1 million people will die from a MI this year [65]. In contrast to past years with industrialized nations accounting for a majority of these

deaths, the CAD mortality rate is rapidly rising in developing nations [68]. This epidemiological transition is fueled in part by economic, scientific, and technological advances that have lengthened life spans and elevated standards of living in these developing areas. Due to its high prevalence throughout the globe and debilitating clinical consequences, effective therapeutics, management strategies, and preventative measures are urgently needed.

Similar to other complex traits, an individual's lifetime CAD susceptibility involves the interplay of genetic, environmental/lifestyle factors, with each accounting for ~50% of an individual's lifetime risk [69]. Since the beginning of the century, physicians have acknowledged the contribution of genetics to atherosclerosis risk, with a family history of fatal CAD serving as an independent risk factor [70] [71]. Large epidemiological cohort studies, such as the Framingham Heart Study (FHS) have linked classic risk factors, such as hypertension, and dyslipidemia to cardiovascular disease risk [72]. These and other studies have also highlighted modifiable behavioral risk factors, such as smoking, sedentary lifestyle, and unhealthy diet to higher incidences of MI and strokes [73]. Non-modifiable risk factors for these conditions include age and male sex. Findings from these studies have motivated the use of a healthy lifestyle as a therapeutic option for patients at high risk for these complications.

Since early intervention is one of the most effective measures to reduce morbidity and mortality, high risk patients are often screened for intermediate vascular phenotypes, indicative of disease progression. Measurements of carotid intimal-medial thickness (cIMT), arterial stiffness, and coronary calcium levels, which can each be assessed noninvasively, serve as surrogate markers for subclinical vascular disease. cIMT is measured using B mode ultrasound and refers to the relative thickness of the intimal to medial layer of the artery [74]. Increases in cIMT correlates with increased risk for adverse events [75] [74]. Arterial stiffness reflects the remodeling processes in the vessel wall, which involves loss of elastic fibers and increased fibrosis. Pulse wave velocity (PWV), the most common metric of arterial stiffness, is correlated with increased severity of CAD and is predictive of future strokes and MI in cohort studies [76] [77]. Levels of coronary

calcium correlate with atherosclerotic plaque burden and function as a stronger predictor than cIMT, C-reactive protein levels, and Framingham risk score for adverse events [78] [79]. These vascular intermediate phenotypes are heritable with a number of overlapping loci with CAD, MI, and stroke, suggesting some shared genetic mechanisms.

In addition to reducing exposure to modifiable CAD risk factors, pharmacological intervention remains a key preventative strategy in the general population. The current standard of care involves a regimen of cholesterol-lowering (e.g. statins) and blood pressure lowering medications (e.g. beta blockers, calcium channel blockers, and Angiotensin-Converting Enzyme inhibitors) [80]. Many of these drug targets have also been identified through GWAS for CAD and vascular intermediate phenotypes. To address more severe presentations of disease, more invasive measures are needed to restore blood flow. Procedures, such as coronary angioplasty and stent placement treat CAD symptoms by widening the occluded vessel, whereas coronary artery bypass graft surgery (CABG) creates new blood vessels from existing vessels to bypass the occlusion [80]. This aggressive treatment plan to combat severe disease is implemented to prevent heart failure, MI, and stroke. Despite the efficacy of these therapeutics, patient adherence to the prescribed regimen reduces their impact. In an effort to reduce the therapeutic burden while leveraging the advances of the CRISPR-Cas9 system, gene editing of some lipid metabolic genes (e.g. *PCSK9*, *LDLR*) is currently tested in clinical trials as a lifetime 'cure' to CAD [81].

II. Genetics of CAD and Mechanisms to Resolve Causal Genes

The heritability of CAD has been recognized through observations in the clinic since the 1950s. Family linkage studies have identified loss of function variants in many lipid metabolism genes, such as *LDLR* and *APOB* that cause familial hypercholesterolemia and early onset CAD [82] [83]. As opposed to these rare variant associations, genome wide association studies (GWAS) have highlighted common genetic variants that confer disease risk. GWAS assesses the frequency of the tested variant among cases relative to controls. For each locus, the lead variant, which is often noncoding, tags a set of

variants in high linkage disequilibrium which can span nearly one hundred kilobases. One or more of these variants may be the causal SNP and affect nearby genes in cis or have distal effects genes in trans to mediate its effect on disease risk [84] (**Fig. 1.4**).

In contrast to coding SNPs that implicate the causal gene, elucidating the mechanisms for non-coding variants is more challenging, due to the complex linkage disequilibrium at most loci. A majority of GWAS SNPs reside in non-coding regions and are predicted to regulate gene expression [85]. The lack of mechanistic characterization for a majority of these GWAS loci limits the clinical translation of these studies.

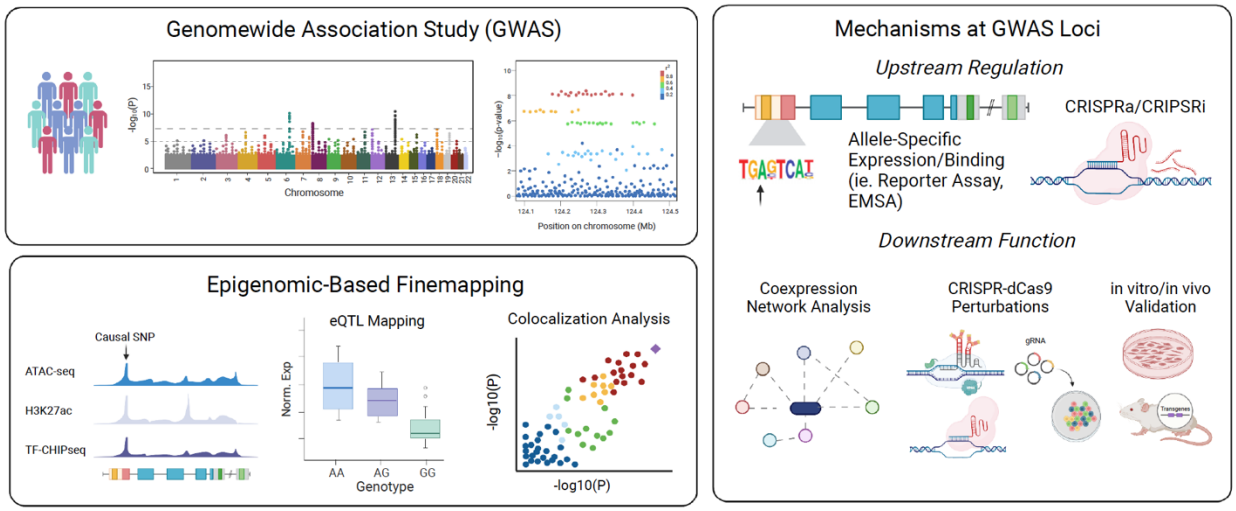

Figure 1.4: Overview of Approaches to Dissect Functional Mechanisms at GWAS Loci

A) Genome-wide Association Studies assess the frequency of the risk allele among individuals with the disease (cases) and healthy controls. Data from these studies are often presented as a Manhattan plot, which shows the lead variants across the genome that surpass the genome wide significance threshold (P<5E-8) or locuszoom plot. A locuszoom plot shows the lead variant at one locus with the other variants at the locus colored according to the linkage disequilibrium with the lead SNP. B) Epigenomic data (ATAC-seq, H3K27ac CHIP-seq) in disease relevant tissue can be used to prioritize candidate causal SNPs. Integration of eQTL and GWAS signals using statistical fine-mapping tools can be performed prioritize candidate causal genes. C) Defining causal mechanisms at GWAS loci require functional follow-up studies. Allele specific expression/binding studies or CRISPR-dCas9 mediated epigenetic perturbations are needed to validate the function of top candidate causal variants. To propose a functional downstream function of candidate genes, construction of coexpression networks provides unbiased support for functional mechanisms. Functional assays following CRISPR-dCas9 mediated epigenetic editing or loss/gain of function will validate the causal association. Created at BioRender.com.

CAD/MI GWAS have identified over 200 loci associated with disease, which account for ~15% of its heritability [86] [87]. These loci harbor genes related to a broad list of pathways, including traditional risk factors, such as low-density lipoprotein cholesterol and insulin resistance. Genes related to cell adhesion, inflammation and vascular remodeling were also enriched among CAD/MI GWAS loci. These pathways were not previously associated with CAD heritability and underscore the potential of these studies to identify novel biomarkers and therapeutic targets [88]. It is worth acknowledging the various obstacles in conducting genetic-based association studies for CAD/MI that may confound biological interpretation of the results. For example, quantitation of stable and unstable CAD using angiography still remains a challenge and could produce different results influencing the case definitions of these studies [89] [90]. Also, MI is an equally heterogeneous condition with ≈70% of cases due to plaque rupture, ≈30% due to plaque erosion and a minority due to calcified nodules. Due to the distinct causes of these pathological phenotypes, the clinical end points may only provide genetic associations for the most prevalent pathology (e.g. rupture) while obscuring the others [90].

Approaches to Prioritize GWAS Mechanisms

Although CAD GWAS have revealed tremendous insights into vessel wall biology, it has been challenging to define causal genes and mechanisms at many of these loci. One powerful strategy to refine these candidate genes and mechanisms is to leverage molecular quantitative trait locus mapping, such as expression quantitative trait locus (eQTL) and splicing quantitative trait locus (sQTL) in disease relevant tissues [91] [92] (**Fig 1.4**). To this end, the Genotype-Tissue Expression (GTEX) project has performed large-scale mapping of eQTLs and sQTLs across 53 human tissues in an effort to define tissue-specific and broad mechanisms of gene regulation [93]. However, the majority of samples profiled in GTEx are heterogeneous, bulk tissues from non-diseased individuals making it difficult to delineate cell-type specific regulatory mechanisms. The Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task (STARNET) study is another large-scale eQTL resource of 7 cardiometabolic tissues (atherosclerotic-lesion-free internal mammary artery, atherosclerotic aortic root, blood, subcutaneous

fat, visceral abdominal fat, skeletal muscle, and liver) isolated from ≈600 individuals undergoing coronary artery bypass graft procedures [94 95]. This dataset identified 2047 cardiometabolic GWAS SNPs with corresponding cis-eQTLs (≈61%; cis-eQTLs alter gene expression of nearby genes, while trans-eQTLs affect expression of distant genes, often located on different chromosomes) [94]. This is several folds more than the original GTEx dataset, emphasizing that not all disease regulatory variants can be detected in healthy tissues. Importantly this dataset was used to provide functional evidence for top candidate CAD regulatory variants identified from epigenomic profiling of HCASMCs and coronary artery tissues [96].

Fine Mapping Approaches

Integration of these large-scale datasets with GWAS holds great promise for prioritizing candidate causal genes at genome-wide significant loci (**Fig. 1.4**). Early efforts overlapping GWAS and eQTLs showed a significant overlap between the two datasets. However, these early analyses did not take account for the underlying genetic architecture and linkage disequilibrium resulting in many false positives [97]. These challenges motivated the development of statistical colocalization tools that quantified the confidence in the shared signal between a molecular trait and GWAS signal. The fine-mapping tool, *coloc*, developed by Chris Wallace and her team is one of the most popular tools for this work. *coloc* leverages the public availability of summary statistics to quantify the odds of colocalization of 2 signals relative to the odds of the null hypothesis using a Bayesian framework [98 99]. Transcriptome-wide association study (TWAS) approaches are commonly employed as well. TWAS leverages the availability of large bulk tissue eQTL datasets with matched genotype data to create a predictive model to impute gene expression given the genotype of an independent cohort. The predicted expression level of each gene is then used to estimate the association of the phenotype and gene expression [100]. Since this approach aggregates the effects of all eQTLs for each gene, the multiple testing correction burden is lessened to account for the number of genes (~20,000) rather than number of variants. Since the conception of the first TWAS, PrediXcan [101], multiple tools (e.g. S-PrediXCan [102], EpiXcan [102] have extended this method to leverage summary statistics rather than individual-level data and incorporate epigenetic data into the predictive model. Summary-level Mendelian

Randomization (SMR), leverages the Mendelian Randomization (MR) framework to assess the association of gene expression and phenotype, using genetic variants as an instrumental variable [103] [104]. SMR estimates the gene-trait effect size by calculating the ratio of the GWAS effect size and eQTL effect size for each variant. The subsequent HEIDI test is used to distinguish models of linkage where the gene-trait association is driven by variants in linkage disequilibrium or pleiotropy/causality.

To prioritize CAD loci that may function through SMCs, our lab completed eQTL mapping in 52 unrelated, multiethnic human coronary artery smooth muscle cells (HCASMCs) donors. This analysis uncovered 1220 significant loci that were enriched in regions of open chromatin in HCASMCs [105]. Colocalization of eQTL and CAD GWAS signals using 2 methods revealed 5 significant loci, including *FES*, *SMAD3*, *TCF21*, *PDGFRA*, and *SIPA1* [105]. This study also quantitated the contribution of HCASMC to CAD risk and provided evidence that these cells capture a substantial portion of CAD heritability. Overall, these findings highlighted CAD risk genes related to vascular remodeling and phenotypic modulation in HCASMC. These data suggest that genetic upregulation of SMC genes (e.g. *SMAD3*, *PDGFRA*, and *SIPA1*) may increase CAD risk, whereas upregulation of other SMC genes (e.g. *TCF21* and *FES*) may be atheroprotective [105].

Gene Regulatory Networks

Many studies have leveraged unbiased systems genetics and network approaches to predict the function of CAD vascular wall related genes. Networks are graphical representations of molecules, represented as nodes and the interactions between the molecules depicted as edges. Edges may represent functional relationships, such as physical interactions, metabolic pathways and cell signaling cascades. Integration of eQTL information in network construction has been used to construct gene regulatory networks (GRN) and predict key drivers driving disease processes (**Fig. 1.4**) [106]. The STARNET study constructed gene regulatory networks across 7 cardiometabolic tissues to highlight cross-tissue interactions that influence CAD risk [94] [107]. This system biology approach identified 224 GRNs that account for the majority (~54%) of CAD heritability. Moreover, the key drivers of 218 modules were causal for cardiometabolic

traits or CAD in Mendelian Randomization analyses [107]. Using a similar Bayesian network approach, the transcription factor, MAFF, was identified as a key driver in a liver specific subnetwork enriched in CAD genes. Through interactions with BACH1, another CAD transcription factor, MAFF repressed *LDLR* gene expression to perturb cholesterol metabolism [108].

CRISPR-Cas9/dCas9 Based Perturbations

The development of genome editing technologies, such as CRISPR, enables perturbations of CAD-associated risk variants or risk haplotypes (**Fig. 1.4**). Despite these technological advancements, these studies are challenging in HCASMCs due to their phenotypic heterogeneity across individuals and disease states as well as their tendency to undergo replicative senescence. An alternative approach to overcome SMC limitations is to disrupt the region of interest in induced pluripotent stem cells (iPSCs) and differentiate the edited cells into SMCs. Deletion of the CAD associated 9p21 risk haplotype in iPSCs identified a mechanism underlying the most reproducible CAD locus. SMCs harboring the risk haplotype exhibited reduced adhesion, contraction and increased proliferation, relative to isogenic SMCs harboring the non-risk haplotype. These studies suggest that 9p21 risk alleles promote SMC dedifferentiation to modulate CAD risk [109]. The use of endonuclease dead Cas9 (dCas9) fused to transcriptional activation (e.g.p300, VP64, VPR) and repression domains (e.g. KRAB) enables the direct perturbation of regulatory elements in SMCs to support the function of candidate causal genes and enhancers [110] [111]. A recent study utilized a CRISPR-dCas9-KRAB mediated epigenetic editing strategy in SMCs to provide insights into the upstream regulatory mechanisms governing the CAD 2q22.3 locus. Epigenomic repression of candidate enhancers harboring CAD variants in tight LD with the lead SNP, prioritized *ZEB2* as the top candidate causal gene [112].

III. Molecular Mechanisms of Atherosclerosis Progression in the Vessel Wall

CAD is a chronic inflammatory disorder in the vessel wall that progresses over an individual's lifetime. The progression of a healthy artery to a late-stage lesion indicative of future MI is portrayed in **Fig 1.5**. Various stimuli, such as high levels of cholesterol, oxidative stress, and pro-inflammatory cytokines disrupt the integrity of the endothelial

monolayer to form an early-stage lesion [113]. This proatherogenic signaling increases the expression of adhesion molecules that help retain lipoproteins in the intimal space where these molecules undergo oxidation and other modifications [114] [115] Activated endothelial cells also recruit circulating immune cells, particularly monocytes and

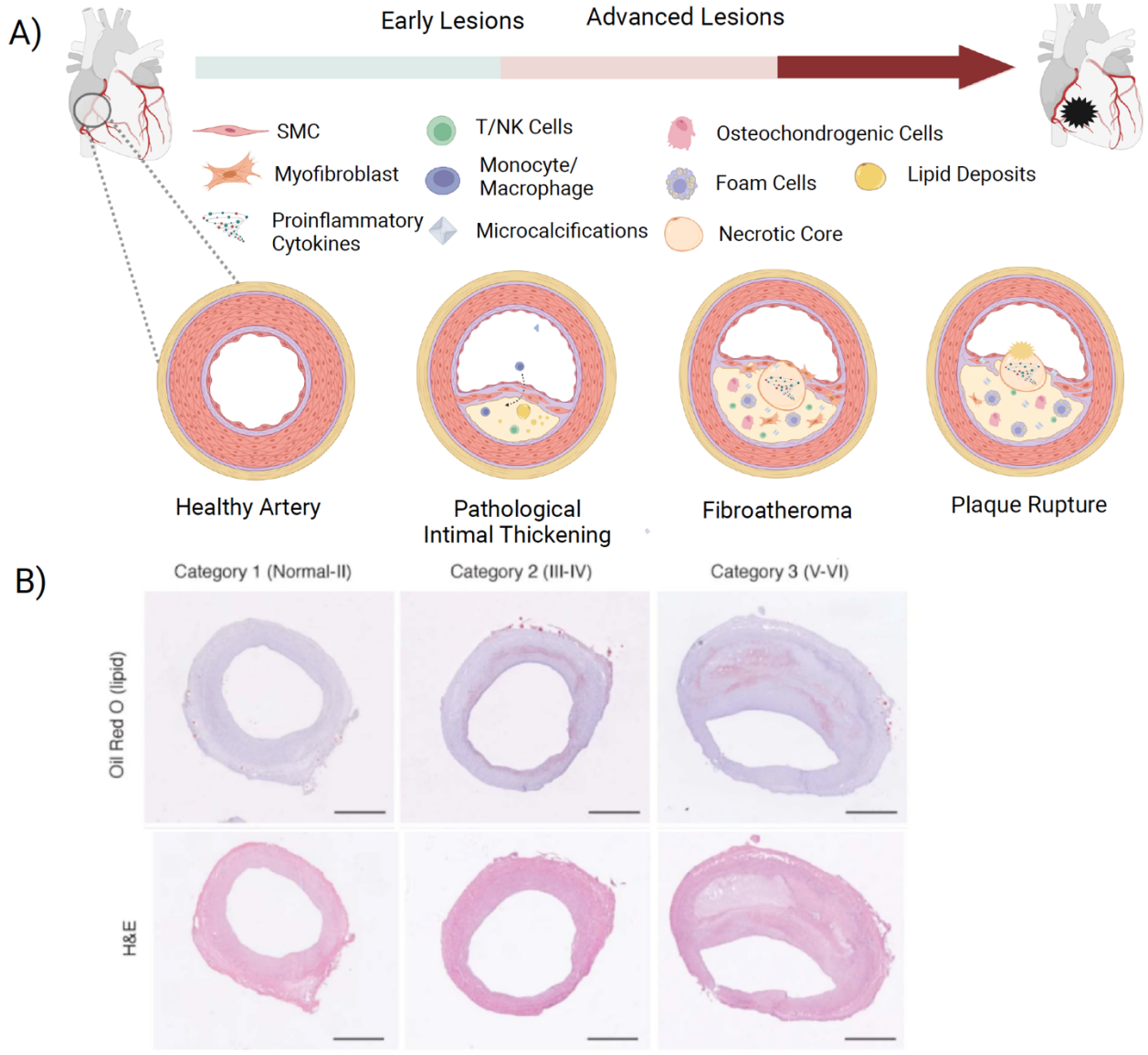

Figure 1.5: Stages of Atherosclerosis in Human Coronary Arteries

A) Overview of the stages as an early-stage atherosclerotic lesion progresses to a late stage lesion. The primary cause of thrombosis leading to adverse cardiovascular events, such as myocardial infarctions and strokes is plaque rupture. Created at BioRender.com
B) Oil Red O and H&E Staining of the different types of atherosclerotic lesions to visualize the accumulation of lipid droplets and structure of the artery respectively. These lesions are categorized as normal, early, and late-stage lesions based on This image was adapted from Turner et al. Single-nucleus chromatin accessibility profiling highlights regulatory mechanisms of coronary artery disease risk. *Nature Genetics.* 2022.

macrophages into the lesion, where they may engulf LDL to transition into macrophage-derived foam cells. Increased proportion of the M1-like proinflammatory macrophages and foam cells in the plaque correlates with disease progression [116]. These molecular processes together contribute to the fatty streaks observed during early stages of CAD. In some individuals, these lesions can be visualized during childhood.

SMCs also play a key role during the formation of these early atherosclerotic lesions. Medial SMCs dedifferentiate to a more proliferative state, where they migrate to populate the intimal lesion. Modulated SMCs downregulate traditional markers of contraction, (e.g. *LMOD1*, *MYH11*, *CNN1*) and secrete high amounts of ECM, such as lipid-retaining proteoglycans [23 61]. SMCs can also adopt a foam-cell like phenotype, although there is no consensus in the field regarding the proportion of these cells in human atherosclerotic plaques [117 118] . Nonetheless, internalization of LDL by SMCs increases apoptosis [119]. These apoptotic cells are normally removed by macrophages but those macrophages in the lesion have reduced phagocytic capabilities. This defect contributes to the overall expansion of the lesion, formation of the necrotic core, and progression towards a late-stage lesion or fibroatheroma. [9].

A common feature of fibroatheromas is the presence of calcification deposits, a feature that is generally not detected in early lesions [120]. The proportion of the plaque covered by calcification differs depending on the plaque type but generally correlates with plaque complexity and percent stenosis [120]. At the molecular level, the deposition of calcium is an active process potentiated by the proinflammatory microenvironment and the release of microvesicles and apoptotic bodies [121 122]. The incompletely cleared debris serves as a nucleation site to form microcalcifications on the collagen rich matrix [123]. With time, these microcalcifications coalesce into large speckles or sheets of macrocalcifications [124]. Vascular calcification is driven by the osteo-chondrogenic transition of SMCs and mirrors many of the molecular mechanisms governing osteogenesis during bone formation. Many of the same genes, such as *BMP2*, *OPN*, *MGP*, and *ALPL,* involved in ossification are also involved in the calcification process of atherosclerotic plaques. The osteoblast enriched transcription factors, Runt-related transcription factor 2 (RUNX2)

and Msh Homeobox 2 (MSX2) are upregulated in calcifying SMCs to activate downstream pathways, such as bone morphogenetic proteins (BMP) and Wnt signaling [125] [126] [127]. These modulated SMCs remodel the existing matrix through degradation and deposition of a new pro-calcifying collagenous matrix.

Fibroatheromas with spotty calcifications near the fibrous cap are often thin-cap atheromas. This plaque type is prone to rupture with a thin fibrous cap and large necrotic core [124] [128]. In contrast to thin cap atheromas, the fibrocalcific plaque is characterized by a thick fibrous cap, often covered by sheets of calcifications [120]. This plaque type is 'stable' and is less prone to rupture. However, these sheets of calcification can fracture leaving a protruding nodule that can penetrate the endothelial cell lining. This mechanical disruption often leads to an MI and accounts for ~5% of coronary artery thrombosis [89]. Although once believed to be detrimental to plaque stability, recent studies of patients with acute coronary syndrome, showed a lesser degree of calcification in thin cap fibroatheromas relative to stable fibrocalcific plaques [120]. However, additional prospective studies are needed to clarify the complex role of calcification and reconcile these observations.

The integrity of the fibrous cap that protects the plaque from rupture, is influenced by both medial and intimal SMCs that migrate into the lesion in response to growth factors and cytokines [9]. Given the correlations between SMC content and fibrous cap thickness and colocalization between SMC and the type I and III collagen-rich matrix, modulated SMCs are key contributors to the stability of the fibrous cap [129] [130]. This ECM-secreting fibromyocyte-like transition is driven by the transcription programs regulated by the transcription factor, Transcription Factor 21 (TCF21) [62]. The function of 2 embryonic stem cell pluripotency factors, Kruppel-like Factor 4 (KLF4) and Octamer-Binding Transcription Factor 4 (OCT4), regulate opposing of transcriptomic profiles that increase and decrease the thickness of the fibrous cap respectively [131] [130] [63].

Transcription Factors Relevant to SMC Phenotypic Modulation in Atherosclerotic Lesions

Given the significant contribution of SMC phenotypic modulation to CAD etiology as described above, it is not surprising that genes related to this process contribute to CAD heritability. Notably, KLF4 and TCF21, 2 key transcription factors regulating SMC phenotypic states in atherosclerotic lesions, are CAD causal genes (**Fig 6**). Mechanistic insights into the function of these genes are supported by a combination of *in vitro* functional studies, murine lineage tracing, and scRNA-seq studies of human artery tissues [63] [62] [130].

KLF4

KLF4 is a zinc finger transcription factor that downregulates the expression of SMC marker genes by binding to the G/C repressor element located near the CArG box. This interaction impairs activation by the SRF/MYOCD complex, resulting in lower expression. [132]. *KLF4* overexpression also induced epigenetic changes recruited histone deacetylases (HDACs) to SMC contractile genes to reduce gene expression[132]. Although not robustly expressed in SMCs or in elastic arteries, *KLF4* expression is upregulated following vascular injury, coincident with the downregulation of SMC marker genes [133].

More recently, in murine models of atherosclerosis, SMC-specific *Klf4* knock-out reduced lesion size and increased indices of plaque stability [130]. The proportion of osteoblast-like SMC-derived cells, characterized by high expression of *Sox9, Runx2, Cytl1, Ibsp,* and *Alpl* was reduced approximately 4 fold and replaced with *Acta2*+, *Phactr1*+ SMC derived cells near the fibrous cap [63]. These studies reinforce the detrimental role of *Klf4* mediated SMC transitions in atherosclerotic lesions. Finally, of major interest, KLF4 binding in murine atherosclerotic plaques altered the expression of about ⅓ of CAD GWAS target genes [63]. Additional follow up studies will be needed to ascertain the causal roles of KLF4 target gene to CAD and SMC phenotypic modulation.

TCF21

TCF21 was first identified as a regulator of cell fate decisions. Cardiac fibroblasts that downregulate its expression are primed to become coronary artery SMCs [134]. Consistent with this role in development, reduced *TCF21* expression upregulated

markers of SMC differentiation [135]. At the molecular level, TCF21 interferes with the SRF/MYCOD signaling cascade to promote SMC dedifferentiation. TCF21 binds to CArG boxes of contractile markers to reduce SRF/MYOCD mediated transactivation. In this way, the positive feedback loop driving expression of SRF and MYOCD is inhibited [136]. As a transcription factor, *TCF21* regulates genes related to migration and proliferation, with binding sites overlapping those of the pro-mitogenic AP-1 TF family. These gene set enrichments mirrored the functional effects of *TCF21* knockdown in reducing proliferation and migration [137].

In regards to atherosclerosis, *TCF21* plays a protective role by regulating the SMC fibromyocyte transition. Fibromyocytes are modulated SMCs that proliferate and migrate, and secrete abundant levels of ECM to stabilize the fibrous cap. Accordingly, SMC specific knock-out of *Tcf21* was linked to a thin fibrous cap in mouse models [62] .This protective effect was confirmed by human genetics where the CAD risk allele was associated with reduced *TCF21* expression. TCF21 target genes were also highly enriched in lead CAD GWAS variants or those in high linkage disequilibrium with CAD SNPs, suggesting that TCF21 acts as a master regulator of downstream CAD-associated genes to influence disease risk [137].

As exemplified by the mechanisms of both *KLF4* and *TCF21*, many signaling pathways regulating SMC phenotypic switching converge on the activity of the SRF/MYOCD transcriptional complex. SRF also interacts with other cofactors to regulate downstream signaling in SMCs. Of particular interest to this project are those cofactors that contain the highly conserved Lin11, Isl-1, and Mec-3 (LIM) domain [138]. The LIM domain is made up of ~60 amino acids that form a cysteine-rich double zinc finger to facilitate protein-protein or protein-DNA interactions [139]. Proteins that contain LIM domains are involved in focal adhesion signaling, cytoskeletal rearrangements and transcriptional regulation [140].

LIM Domain Containing Cofactors Regulate SRF/MYOCD Activity

The LIM containing proteins, Cysteine and Glycine Rich Protein 2 (CSRP2), Leupaxin (LXPN), and Four and a Half LIM domain 2 (FHL2) modulate SMC phenotypic states

through direct interactions with SRF. *CSRP2* is coexpressed with differentiated SMC markers and transactivates SRF target genes by forming a scaffold with SRF and GATA6 near the regulatory element of contractile genes [141] [142]. Loss of *CSRP2* increased SMC migration and neointima formation after vascular injury, consistent with the direction of other SMC markers [143]. The cytoskeletal adapter protein LXPN, was first identified in a screen for focal adhesion kinase with expression enriched in vascular SMCs [144]. In the nucleus, LXPN forms a complex with SRF at CArG box and induces expression of contractile genes in response to focal adhesion kinase activity [144].

Four and Half LIM Domain Family of Cofactors

FHL2 functions as an SRF cofactor to inhibit SMC differentiation by competing with the SMC cofactors, MYOCD and MRTFA/B for SRF binding [145]. Similar to other LIM domain proteins, FHL2 translocates to the nucleus following disengagement of focal adhesions and phosphorylation by focal adhesion kinase on soft substrates. In the nucleus, FHL2 localizes to regions of active transcription to modulate genes related to cell proliferation [146]. Loss of *Fhl2* in mouse models of atherosclerosis was associated with increased stability and reduced plaque size [147].

In the most recent CAD/MI meta-analysis, *FHL3* and *FHL5*, which belong to the same family as *FHL2*, surpassed the genome-wide significance threshold [86]. These proteins, along with *FHL1*, form the FHL family that is named according to its structure. Each member of the FHL family contains four LIM domains and one N-terminal half LIM domain. The expression of the FHL family is enriched in skeletal, cardiac, and smooth muscle [148]. At the molecular level, FHL1, FHL2, and FHL3, interact with the actin cytoskeleton and shuttle into the nucleus to regulate transcription [149] [150]. *FHL1* is involved in proper sarcomere assembly with loss of *FHL1* linked to human muscular dystrophies and cardiomyopathies [151] [152]. *FHL3*, another family member that is enriched in skeletal muscle, negatively regulates muscle differentiation, in part through antagonism of a critical regulator myoblast differentiation, MYOD. Through direct interactions with MYOD, FHL3 downregulates the expression of myogenic genes and myotube formation [153] [154].

FHL5 is the most understudied member with limited expression *in vitro* and in mouse models. *FHL5* was first identified as a cofactor during sperm cell maturation by transactivating CREM target genes [155]. *Fhl5* knock-out mice were fertile but exhibited abnormal sperm morphology [156]. More recently, *FHL5* was identified as one of the most upregulated genes in stenotic post grafted veins relative to normal veins. Follow up *in vitro* studies uncovered a role of FHL5 in regulating cell proliferation through activation of cAMP Response Element-Binding Protein target genes [157] [158].

Motivation for Thesis

CAD GWAS have made remarkable progress in highlighting over 200 loci that contribute to its heritability. These studies hold great promise to identify novel therapeutic targets linked to disease risk. To date, however, findings from these large-scale population genetic studies have not been very successful in therapeutic development, with the few bona fide targets (*PCSK9*, *LDLR*), associated with traditional risk factors. One possible interpretation of this disconnect is the limited number of studies that annotate causal gene(s) at GWAS loci and investigate their functional role in disease relevant tissues. These functional follow-up studies are critical for assessing the druggability of top candidate causal genes. Given the challenges of these types of studies, a majority of CAD loci lack defined mechanisms.

This dissertation addresses this clinical gap in the field by defining the molecular mechanism at the *UFL1-FHL5* locus. This locus was among the novel CAD/MI loci reported in the combined United Kingdom Biobank and CARDIoGRAMC4D meta-analysis. In this dissertation, we characterize the upstream and downstream function of *UFL1-FHL5* to investigate how genetic variants modulate CAD risk in the general population. In chapter 2, we use an epigenomic based fine-mapping strategy to investigate the upstream regulatory mechanism perturbed by the CAD/MI risk allele. We implicate *FHL5* as the top candidate casual gene at this locus, with increased expression associated with increased risk of CAD/MI. We also characterize the functional consequences of modulated *FHL5* expression in SMCs under both basal and more atherogenic conditions. We investigate how these phenotypic, transcriptomic, and epigenomic changes mediated by FHL5 increases CAD risk. In the last chapter, I discuss the implications of these findings in the field of CAD and SMC biology as well as highlight future directions to move towards clinical translation.

Chapter 2

Manuscript is in Revisions at Circulation Research.

FHL5 controls vascular disease-associated gene programs in smooth muscle cells

Doris Wong[1,2,3], Gaëlle Auguste[2], Christian L. Lino Cardenas[4], Adam W. Turner[2], Yixuan Chen[2], Lijiang Ma[5], R. Noah Perry[2,3,6], Redouane Aherrahrou[2], Maniselvan Kuppusamy[3], Chaojie Yang[1,2], Jose Verdezoto Mosquera[1,2], Collin J. Dube[7], Mohammad Daud Khan[2], Meredith Palmore[2], Maryam Kavousi[8], Patricia A. Peyser[9], Ljubica Matic[10], Ulf Hedin[10], Ani Manichaikul[1,2,11], Swapnil K. Sonkusare[3,12], Mete Civelek[2,3,6] Jason C. Kovacic[13,14,15], Johan L.M. Björkegren[5,16], Rajeev Malhotra[4], Clint L. Miller[1,2,3,11]

Affiliations

[1]Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, Virginia, USA.

[2]Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia, USA.

[3]Robert M. Berne Cardiovascular Research Center, University of Virginia, Charlottesville, Virginia, USA.

[4]Cardiovascular Research Center, Cardiology Division, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA.

[5]Department of Genetics and Genomic Sciences, Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, USA.

[6]Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia, USA.

[7]Department of Microbiology, Immunology, and Cancer Biology, University of Virginia, Charlottesville, Virginia, USA.

[8]Department of Epidemiology, Erasmus University Medical Center, The Netherlands.

[9]Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA.
[10]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden.
[11]Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia, USA.
[12]Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia, USA.
[13]Cardiovascular Research Institute, Icahn School of Medicine at Mount Sinai, New York, USA.
[14]Victor Chang Cardiac Research Institute, Darlinghurst, New South Wales, Australia.
[15]St. Vincent's Clinical School, University of New South Wales, Sydney, Australia.
[16]Integrated Cardio Metabolic Centre, Department of Medicine, Karolinska Institutet, Huddinge, Sweden.

Corresponding author
Clint L. Miller, PhD
University of Virginia
Center for Public Health Genomics, MSB 3231
PO Box 800717
Charlottesville, VA 22908 USA
clintm@virginia.edu

Abstract

Background: Genome-wide association studies (GWAS) have identified hundreds of loci associated with common vascular diseases such as coronary artery disease (CAD), myocardial infarction (MI), and hypertension. However, the lack of mechanistic insights for a majority of these loci limits translation of these findings into the clinic. Among these loci with unknown functions is *UFL1-FHL5* (chr6q16.1), a locus that reached genome-wide significance in a recent CAD/MI GWAS meta-analysis. In addition to CAD/MI,

UFL1-FHL5 is also implicated to coronary calcium, intracranial aneurysm and migraine risk, consistent with the widespread pleiotropy observed among other GWAS loci.

Methods: We apply a multimodal approach leveraging statistical fine-mapping, epigenomic profiling, and imaging of human coronary artery tissues to implicate Four-and-a-half LIM domain 5 (*FHL5)* as the top candidate causal gene. We unravel the molecular mechanisms of the cross-phenotype genetic associations through *in vitro* functional analyses and epigenomic profiling experiments.

Results: We prioritized *FHL5* as the top candidate causal gene at the *UFL1-FHL5* locus through eQTL colocalization methods. *FHL5* gene expression was enriched in the SMC and pericyte population in human artery tissues with coexpression network analyses supporting a functional role in regulating SMC contraction. Unexpectedly, under procalcifying conditions, *FHL5* overexpression promoted vascular calcification and dysregulated processes related to extracellular matrix organization and calcium handling. Lastly, by mapping FHL5 binding sites and inferring FHL5 target genes function using artery tissue gene regulatory network analyses, we highlight regulatory interactions between FHL5 and downstream CAD/MI loci, such as *FOXL1* and *FN1* that have roles in vascular remodeling.

Conclusion: Taken together, these studies provide mechanistic insights into the pleiotropic genetic associations of *UFL1-FHL5.* We show that *FHL5* mediates vascular disease risk through transcriptional regulation of downstream vascular remodeling loci. These *trans*-acting mechanisms may account for a portion of the heritable risk for complex vascular diseases.

Introduction

Vascular diseases, such as atherosclerosis and aneurysm, encompass a wide range of disorders that affect blood vessels and perturb blood flow to critical tissues. Despite the heritability of these common pathologies, many of the associated genetic risk factors are unknown. To address this clinical gap, genome-wide association studies (GWAS)

have been used to uncover the genetic basis for many complex traits. In the case of coronary artery disease (CAD), which is among the leading causes of death worldwide, GWAS have identified over 200 loci[159]. A majority of these loci harbor genes that function independently of lipid metabolism and other classic risk factors. These studies highlight additional pathways, such as vascular remodeling and inflammation that directly contribute to CAD pathogenesis and prioritize novel therapeutic targets[160,88].

As demonstrated by lineage tracing and single-cell RNA sequencing (RNA-seq) studies in mice and humans, smooth muscle cells (SMCs) play a key role in both the development and progression of atherosclerosis[130,63]. SMCs undergo phenotypic transitions to give rise to diverse cell populations that drive atherosclerosis progression in early stages and contribute to fibrous cap stability in advanced lesions[62]. Intimal SMCs that acquire mesenchymal and ultimately osteoblast-like phenotypes deposit calcium-mineral in the collagenous matrix[161]. These remodeling events coupled with arterial calcification serve as strong predictors of adverse cardiovascular events[120]. Untangling the complex contribution of SMCs to CAD using human genetics may inform novel drug candidates targeting these primary disease processes in the vessel wall.

The most recent CAD and myocardial infarction (MI) meta-analysis of the combined UK Biobank and CARDIoGRAMplusC4D cohorts identified a genome-wide significant association of the *UFL1-FHL5* locus (chr6q16.1) [86]. In addition to CAD/MI, *UFL1-FHL5* is also associated with multiple vascular pathologies, including hypertension[162], intracranial aneurysm[163], and migraines[164], similar to another pleiotropic locus *PHACTR1-EDN1*[165]. The lead CAD/MI single nucleotide polymorphism (SNP), rs9486719 is located in the first intron of *FHL5*. FHL5 is a member of the four-and-a-half LIM (FHL) domain family of cofactors, which also includes structurally similar proteins, FHL1, FHL2, and FHL3[149]. Among the FHL family, FHL5 is the most understudied, which may be attributed to its low expression *in vitro* and limited availability of suitable animal models. Early functional studies have focused on its expression in germ cells[166,167], however FHL5 was recently linked to intimal hyperplasia in aortic SMCs by activating CREB target genes[157,158].

Here, we present a comprehensive study that investigates the molecular underpinnings of the genetic association of the *UFL1-FHL5* locus with CAD/MI and other vascular pathologies. We identify an FHL5-regulated transcriptional network that contributes to the maladaptive extracellular matrix (ECM) remodeling and vascular tone defects associated with clinical disease risk.

Methods

A detailed and expanded methods section can be found in the Supplemental Material. A list of key reagents can be found in the Major Resources Table located in the Supplementary Materials section.

Statistical analysis

Data in bar graphs are presented with mean+/- standard error of mean (SEM) with each point represented as an individual replicate. Data in box plots are presented with lines denoting the 25th, median and 75th percentile with each point representing an individual donor. Pairwise comparisons were made using the student's t test or Wilcoxon rank test as appropriate. Comparisons between more than two groups were assessed using a one-way ANOVA test or Kruskal-Wallis test. The normality of the data was assessed using the Shapiro-Wilkes test, with $P > 0.05$, supporting a normal distribution. For each of these analyses, we considered $P < 0.05$ as significant.

To identify differentially expressed genes, we used the false discovery rate (FDR) adjusted $P< 0.05$ threshold and log2FoldChange > 0.6 or log2FoldChange < 0.6. Heatmaps were created using the pheatmap package and represent normalized expression (Z-score) for genes, scaled across each row. Gene ontology enrichment analyses were performed relative to all expressed genes using Fisher’s Exact Test, with a significant threshold of 5% FDR.

Data Availability

All raw and processed CUT&RUN and RNA-seq datasets are made available on the Gene Expression Omnibus (GEO) database (accession: GSE201572). All other

datasets are publicly available and detailed in the Major Resources Table Section or Supplemental Materials section. All custom scripts used are available at https://github.com/MillerLab-CPHG/FHL5_Manuscript. All software tools used in this study are publicly available and full names and versions are provided in Supplementary Materials.

Ethics statement

All research described herein complies with ethical guidelines for human subjects research under approved Institutional Review Board (IRB) protocols at Stanford University (#4237) and the University of Virginia (#20008), for the procurement and use of human tissues and information, respectively.

Results

FHL5 is the top candidate causal gene associated with CAD/MI risk

The lead variant, rs9486719, tagging the *UFL1-FHL5* locus (chr6q16.1) is associated with CAD (*P* = 1.1E-8) and MI (*P* = 6.8E-10) risk (**Figure 2.1A** and **Figure S2.1A**), as reported in the combined genome-wide association study (GWAS) meta-analysis of CARDIoGRAMplusC4D and UK Biobank (UKBB) data [159 86]. The lead MI SNP, rs9486719 is also associated with intermediate traits predictive of cardiovascular adverse events, such as blood pressure [168 169] and hypertension [162] **(Figure S2.1B)**. We did not observe genetic associations of rs9486719 with lipid metabolism or other traditional CAD risk factors (e.g. type 2 diabetes or obesity) in the PhenoScanner database [170 171] (**Figure S2.1C)**, implicating heritable CAD risk at this locus acting primarily through the vessel wall.

rs9486719 is located in the fifth intron of *FHL5*, and similar to the majority of GWAS variants likely modulates gene expression to influence disease risk [172 173] (**Figure 2.1A)**. To prioritize the most likely target gene(s) at this locus, which includes 7 genes, we leveraged *cis*-expression quantitative trait loci (*cis*-eQTLs) in cardiometabolic tissues from Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task (STARNET)

[94]. This SNP was strongly associated with *FHL5* gene expression in both mammary arteries (MAM) with subclinical atherosclerotic disease, and aortic root

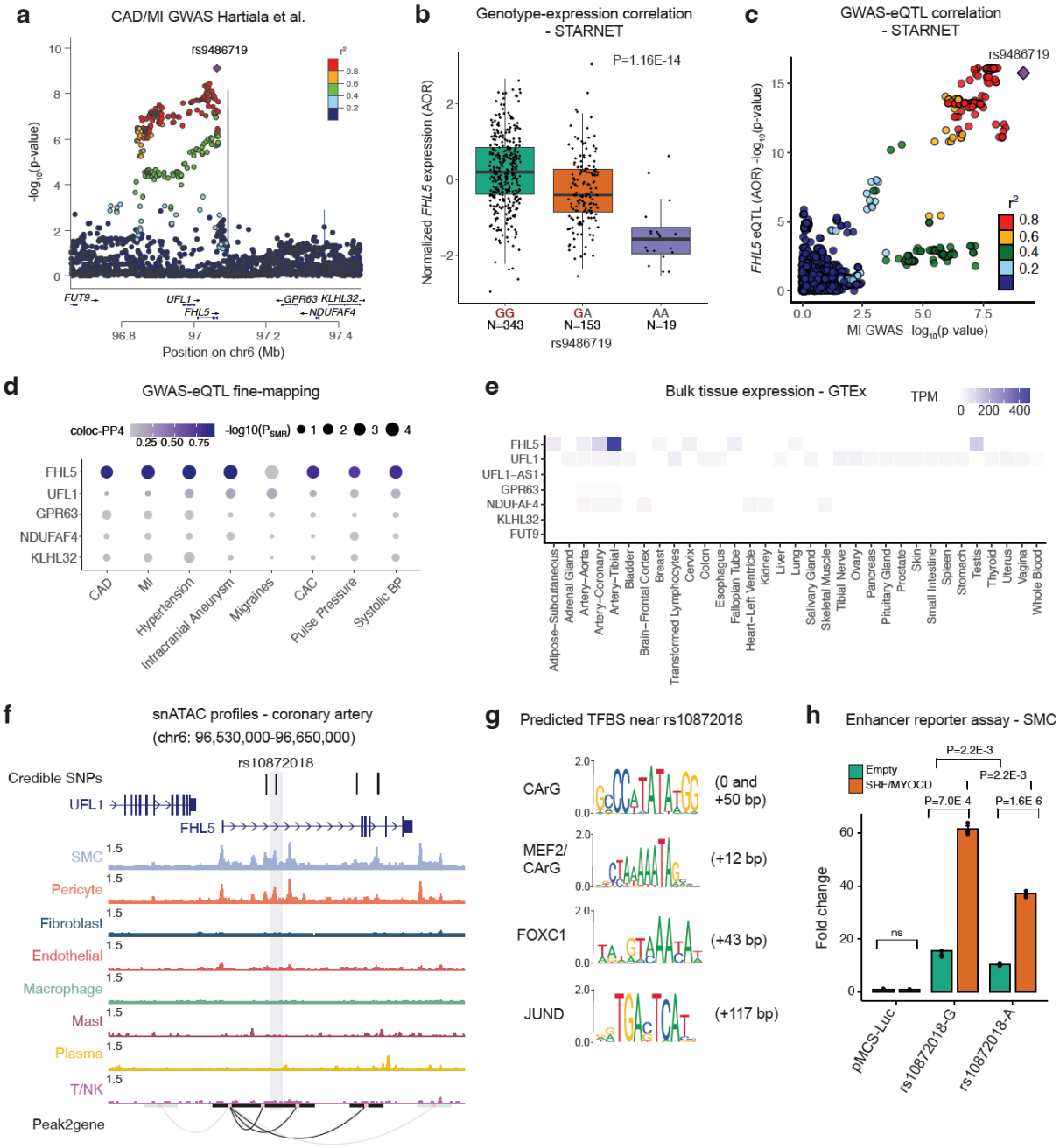

Figure 2.1. *FHL5* is the top candidate causal gene at the *UFL1-FHL5* locus associated with increased CAD/MI risk. (a) LocusZoom plot showing the association of *UFL1-FHL5* locus with myocardial infarction (MI) in the combined UKBB and CARDIoGRAMplusC4D meta-analysis. Data previously published. Figure created by DW. **(b)** Association of rs9486719 with *FHL5* gene expression in STARNET aortic tissue with the CAD risk allele (rs9486719-G) associated with increased FHL5 gene expression. Box plots represent the median with the box

spanning the first and third quartiles and whiskers as 1.5 times IQR. Data previously published. Figure created by DW. **(c)** Locuscompare plot showing the associations of top *FHL5* STARNET aortic tissue *cis*-eQTLs with MI. Data previously published. Figure created by DW. **(d)** Summary of SMR and coloc analyses nominating *FHL5* as top candidate causal gene for vascular diseases and cardiovascular risk factors. The size of the dot reflects the -log10(pSMR) and the intensity of purple color represents the posterior probability of colocalization. Analysis done by DW. **(e)** Normalized gene expression as transcripts per million (TPM) for all genes at the *UFL1-FHL5* locus across GTEx tissues. Data previously published. Figure created by DW. **(f)** Genome browser view of human coronary artery snATAC-seq peaks showing overlap between CAD 95% credible set of SNPs (highlighting top candidate rs10872018) with putative enhancers correlated with the *FHL5* promoter through peak2Gene analyses. Data previously published. Figure created by AT. **(g)** Predicted transcription factor binding sites (TFBS) at or around rs10872018 determined from JASPAR 2022. Distance (bp) of motif sequence is also shown relative to the rs10872018 SNP. Analysis done by CLM **(h)** Luciferase reporter assay in A7r5 SMCs comparing rs10872018 allele-specific enhancer activity co-transfected with empty vector (control) or *SRF* and *MYOCD* expression constructs. Experiments done by CLM.

tissues (AOR) with atherosclerosis (**Figure 2.1B**). These results were consistent with GTEx eQTLs, in which rs9486719 was significantly associated with *FHL5* expression in aorta (P=2.3E-7) (**Figure S2.2A**). Importantly, *FHL5* eQTLs colocalized with CAD SNPs in aortic tissues, compared to eQTLs for neighboring genes (e.g. *UFL1*) (**Figure 2.1C**). In order to systematically prioritize the GWAS effector genes at the *UFL1-FHL5* locus, we employed two complementary and independent fine-mapping methods, *coloc* [98 99] and Summary-level based Mendelian Randomization (SMR) [103 104]. By integrating GTEx and STARNET artery tissue eQTLs for all of the genes at the *UFL1-FHL5* locus with GWAS summary statistics, we consistently identified *FHL5* as the top candidate causal gene for CAD/MI and other common vascular traits (**Figure 2.1D**). This was consistent with the tissue distribution of *FHL5* gene expression, which was highly enriched in artery tissues in GTEx (**Figure 2.1E**) and STARNET (**Figure S2.2B**) compared to other genes at the locus. Together, these analyses suggest that the *FHL5* risk allele (rs9486719-G) increases CAD/MI risk by increasing *FHL5* gene expression in artery tissue (**Figure S2.2C).**

Epigenomic based fine-mapping of UFL1-FHL5 locus in human artery tissue

Next, we leveraged human coronary artery epigenomic profiles to resolve candidate causal variants at the *UFL1-FHL5* locus. The MI 95% credible set determined from Probabilistic Annotation Integrator v3 (PAINTOR) [174] using human coronary artery SMC

and pericyte open chromatin profiles (snATAC-seq) as prior functional annotation weights contained 4 SNPs. By overlapping variants in this credible set with SMC Peak2Gene linkages [175], we prioritized rs10872018, a SNP located in an active *FHL5* regulatory element *in vivo*. (**Figure 2.1F, Figure S2.3A**). These interactions were supported by the Activity-by-Contact Model (ABC) [176] generated using ENCODE human coronary artery tissue ATAC-seq, H3K27ac, and HiChIP data, which further prioritized rs10872018 as the most likely causal variant. This variant was identified as one of the top eQTLs in STARNET human vascular tissues (**Figure S2.3B**).

Since trait-associated SNPs are predicted to alter regulatory elements [85], we scanned the genomic sequence +/- 100bp of rs10872018 for putative transcription factor binding sites and identified multiple conserved CArG boxes [177] (**Figure 2.1G** and **Figure S2.3C**). The alternate allele (rs10872018-A) disrupts a non-canonical CArG box motif, which likely perturbs binding of the SRF-MYOCD transcriptional complex, a well-characterized regulator of SMC differentiation [178 179]. We validated this putative upstream regulatory mechanism using allele-specific enhancer luciferase reporter assays in A7r5 aortic SMCs. The risk allele (rs10872018-G) increased luciferase activity relative to the non-risk allele (rs10872018-A) consistent with the *FHL5* eQTL direction in human artery tissues (**Figure 2.1H** and **Figure S2.3B)**. SRF and MYOCD overexpression further potentiated the luciferase activity in an allele-specific manner (**Figure 2.1H**). Together, these results demonstrate that altered SRF-MYOCD binding due to rs10872018 may partially explain the observed eQTL effects in human artery tissues during disease.

FHL5 is highly enriched in contractile mural cells in coronary arteries

To confirm expression of FHL5 protein in human artery tissues and identify its endogenous localization, we performed immunofluorescence on sections of human coronary arteries with subclinical atherosclerosis (**Figure 2.2A**). FHL5 protein was enriched in the SMC-containing medial layer as well as the intima, colocalizing with F-actin positive cells (**Figure 2.2A**). Consistent with previous reports as a transcriptional

regulator [180], we observed nuclear localization of FHL5 in both the medial and intimal layers, but also perinuclear/cytoplasmic localization. To identify the cell-type specific

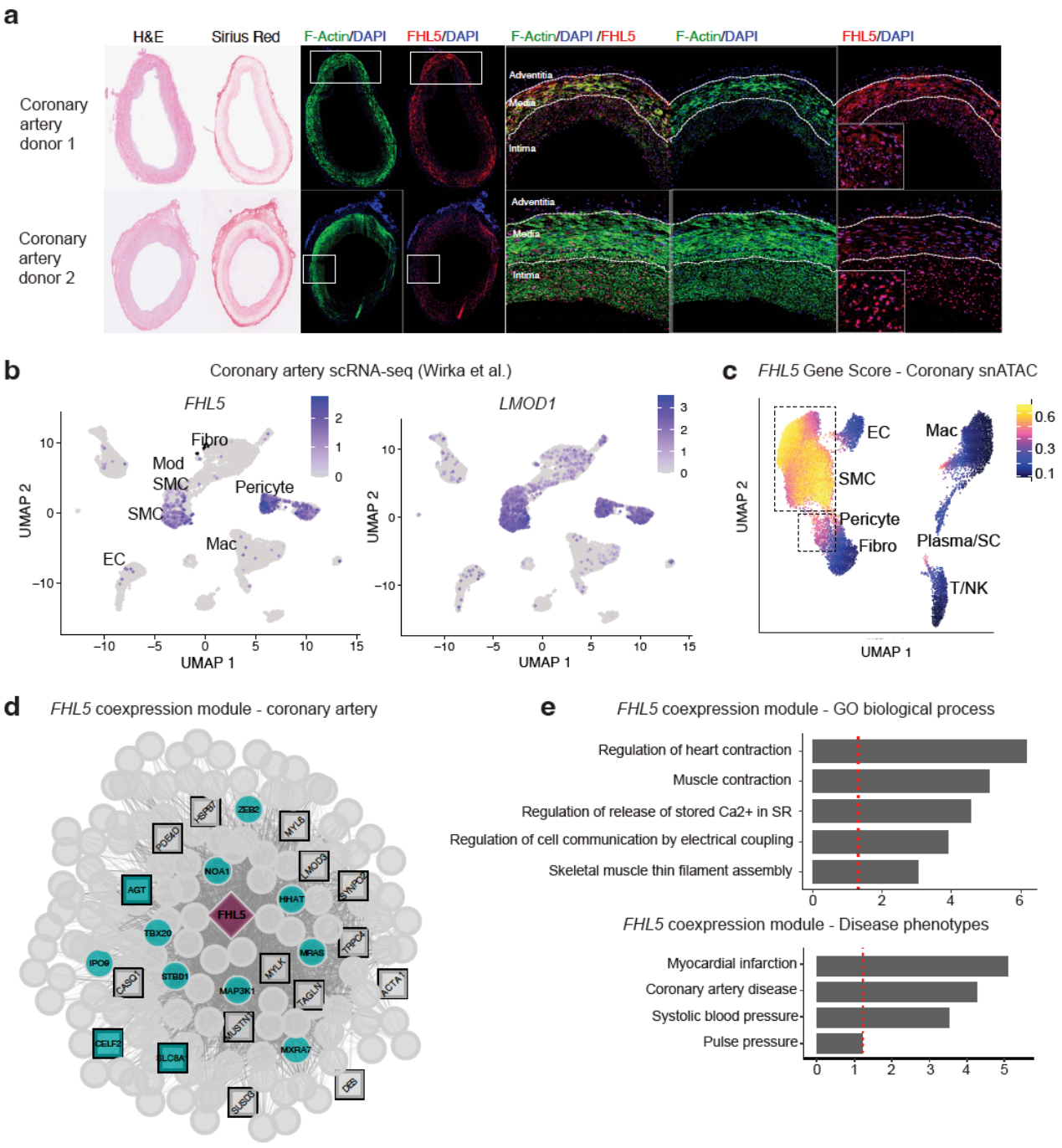

Figure 2.2. *FHL5* expression is enriched in SMCs and pericytes in human coronary arteries.

(a) Representative Immunofluorescence (IF) of FHL5 protein (red) colocalizing with F-actin (green) immunopositive SMCs in the medial and intimal layers of human subclinical atherosclerotic coronary arteries (n=8 unique donors). Experiments done by CLC. **(b)** UMAP

visualization of *FHL5* and *LMOD1* gene expression in different human coronary artery cell types from Wirka et al [62]. Data Previously Published. Figure Created by DW. **(c)** UMAP visualization of human coronary artery snATAC-seq cell clusters colored according to *FHL5* gene score calculated in ArchR. Data Previously Published. Figure Created by AT. **(d)** Undirected *FHL5* coexpression network identified in human coronary arteries using weighted gene co-expression network analysis (WGCNA). Each node represents a gene and the lines connecting 2 nodes are weighted according to degree of correlation. The teal color highlights module genes associated with CAD/MI or blood pressure. The square nodes represent module genes that regulate SMC contraction. *FHL5* is placed at the center and depicted as a red diamond. Analysis Done by RNP. **(e)** Enrichment of gene ontology biological processes (GO-BP) and cardiometabolic disease phenotypes in the *FHL5* module protein-coding genes. Analysis Done by RNP.

expression profile, we queried a human coronary artery scRNA-seq dataset from 4 donors [62 181], which revealed *FHL5* gene expression enriched in the mural cells (SMCs and pericytes) (**Figure 2.2B)**. In fact, *FHL5* was one of the most specific markers identified in the SMC cluster along with well-established regulators of SMC contraction, e.g. *LMOD1* and *MYOCD* (**Figure 2.2B** and **Figure S2.4A**). Similar results were observed in an advanced carotid artery single cell RNA-seq dataset [182] (**Figure S2.4B**). Lastly, we corroborated these results in a single-nucleus ATAC-seq (snATAC-seq) dataset of 41 human coronary arteries [183], in which *FHL5* had the highest gene score in the SMC and pericyte clusters, similar to *LMOD1* (**Figure 2.2C** and **Figure S2.4C**).

To gain insights into the functional role of *FHL5* using a systems biology approach, we performed iterative Weighted Gene Co-expression Network Analysis (WGCNA) [184 185] on transcriptomic data from 148 human coronary artery donors. After removing outliers and lowly expressed genes, we identified *FHL5* in the light-green module, which included key SMC contractile mediators, such as *MYLK, ACTA2, and TAGLN* **(Figure 2.2D)***.* This module was enriched in SMC processes, such as muscle contraction (GO:0006936) and regulation of cell communication by electrical coupling (GO:0010649) as well as cardiometabolic GWAS candidate genes annotated in the Cardiovascular Disease Knowledge Portal (**Figure 2.2E**). In further support of this link to vascular disease risk through regulation of SMC functions, *FHL5* was identified as a key driver in module 152, a cross-tissue gene regulatory network enriched in ECM and CAD candidate genes from the Stockholm Atherosclerosis Gene Expression (STAGE)

cohort [95]. The distinct enrichment in pathways among these two vascular tissues may reflect differences in the atherosclerotic lesion stage.

To further characterize *FHL5* regulatory interactions *in vivo*, we constructed a Bayesian GRN incorporating STARNET aortic tissue eQTL data as priors (**Figure S2.5A**). A key driver analysis (KDA) of this network supported the regulatory potential of *FHL5* which was predicted to function upstream of 2 pulse pressure genes, *ACTG* and *MUSTN1* [168] that have established roles in SMC contraction (**Figure S2.5B**). *ITIH3* was the key driver gene in this subnetwork with a reported association with MI in the Japanese population [186]. These integrative analyses suggest that this *FHL5* GRN related to SMC contractility may contribute to pleiotropic association of FHL5 with common vascular diseases.

FHL5 expression increases SMC contraction and calcification

To overcome the limitations in replicative senescence in primary HCASMCs, we generated an immortalized coronary artery SMC line via overexpression of hTERT[187]. This immortalized cell line (HCASMC-hTERT) maintained protein expression of differentiated SMC markers and closely resembled primary HCASMCs at the transcriptomic level (**Figure S2.6A** through **S2.6C**). Despite robust expression in intact arteries, *FHL5* is potently downregulated *in vitro*, similar to the downregulation of other SMC contractile markers (**Figure S2.7A)**. Therefore, to investigate its role in SMCs, we overexpressed either wildtype FHL5 or FHL5-NLS (FHL5 with a C-terminal nuclear localization signal) and confirmed protein expression (**Figure 2.3A**). The expression level of *FHL5* was physiological and comparable to endogenous levels in human coronary arteries (**Figure 2.3B)**. Motivated by our coexpression network analyses, we functionally validated the role of *FHL5* in regulating SMC contractile pathways. Since calcium is a critical initiator of contraction, we also quantitated intracellular calcium levels. We observed increased SMC contraction and elevated calcium levels in the FHL5-NLS cells relative to HA control cells. While not statistically significant, *FHL5* overexpressing cells trended in the same direction (**Figure 2.3C** and **2.3D**).

Next, to investigate the molecular mechanisms contributing to *FHL5* regulation of SMC phenotypes, we performed bulk RNA-seq on HCASMC-hTERT expressing HA, FHL5 or FHL5-NLS constructs. The direction of effect for top differentially expressed genes were

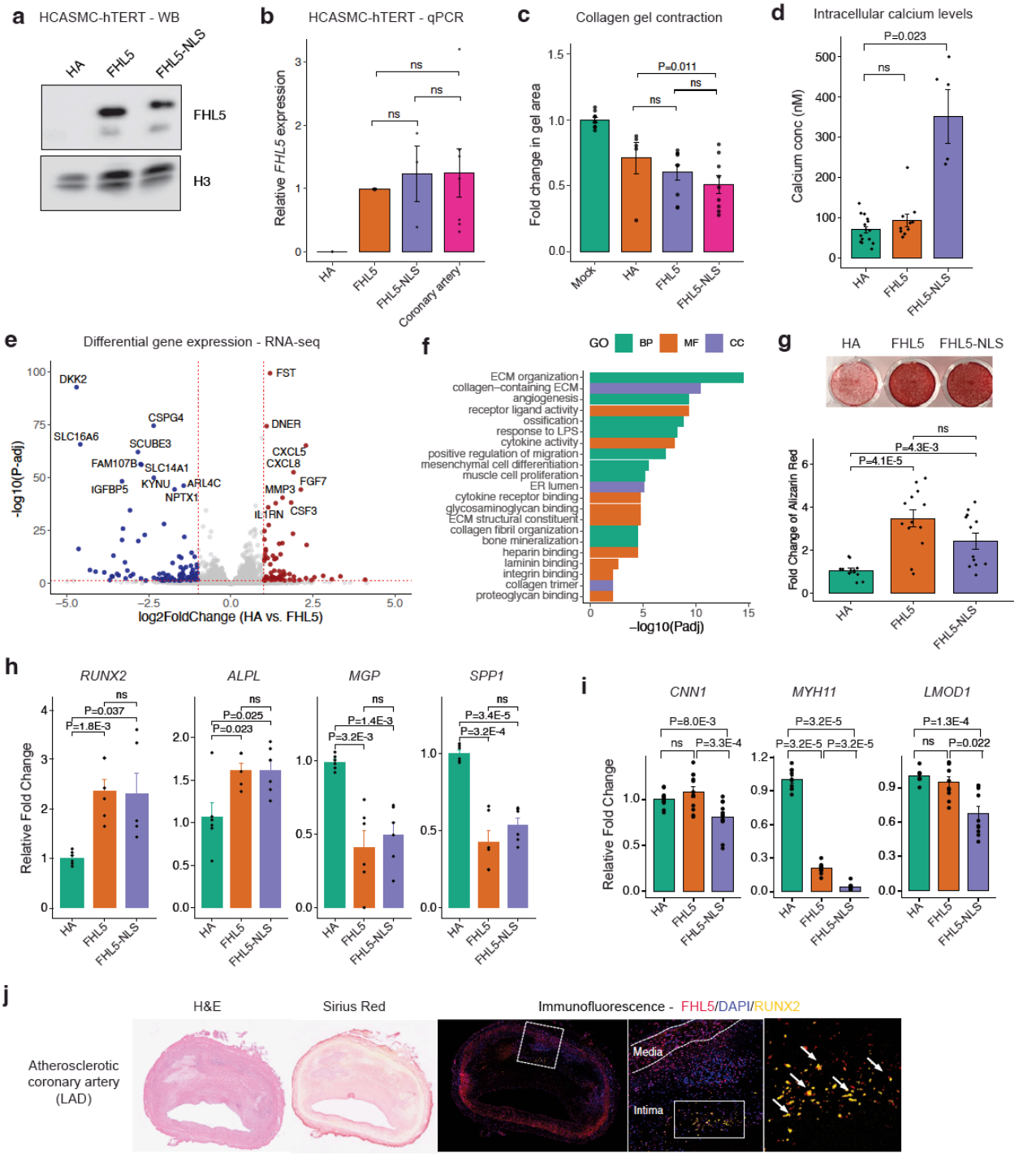

Figure 2.3. *FHL5* regulates SMC contraction and calcification. (a) Western blot confirming overexpression of FHL5 and FHL5-NLS in HCASMC-hTERT cell lines. Experiment Completed

by DW, CD. **(b)** Relative expression of FHL5 in HCASMC-hTERT (HA, FHL5, FHL5-NLS) compared to endogenous levels of FHL5 in human coronary arteries (n=3 donors) determined using qPCR. Expression was normalized to levels of *GAPDH*. Experiment Completed by DW, YC. **(c)** Relative change in collagen gel area (mm²) in SMC contraction assay under basal conditions (n=4 independent biological replicates). Experiment Completed by DW. **(d)** Quantification of intracellular calcium concentrations following stimulation with 10μm phenylephrine (n=4 biological replicates). Experiment Completed by MM. **(e)** Volcano plot showing differentially expressed genes (DEGs) following FHL5 overexpression in HCASMC-hTERT. Red points represent upregulated genes and blue points represent downregulated genes. For clarity, genes with log2FoldChange > 5 and log2Foldchange <-5 are not represented. The full DEG list is provided in Supplementary Table 8. Analysis Completed by DW. **(f)** GO enrichment analysis of FHL5 DEGs showing top over-represented terms for biological process (BP), molecular function (MF), and cell compartment (CC). Analysis Completed by DW. **(g)** Quantification (OD405) and representative image of alizarin red staining 21 days post treatment of HCASMC-hTERT in osteogenic media. Analysis Completed by DW**. (h)** Relative expression of vascular calcification activators (*RUNX2, ALPL)* and inhibitors (*MGP, SPP1)* 14 days post treatment in osteogenic media. Experiments Completed by DW, YC. **(i)** Relative expression of SMC markers 14 days post treatment in osteogenic media. **(j)** Representative immunofluorescence staining of human coronary arteries showing FHL5 (red) colocalization with RUNX2(yellow) in the intima layer near regions of calcium deposition (n=4 independent donors). Adjacent sections subjected to histology staining for hematoxylin & eosin (H&E), collagen deposition (Sirius Red), and calcification (Von Kossa). Experiments Completed by CLC. All error bars represent mean+/-SEM. P-values determined from paired Student's t-test. Individual points reflect replicates from at least n=3 independent experiments. ns: non-significant.

mainly concordant direction between FHL5 and FHL5-NLS samples (**Figure S2.7B** through **S2.7D**). When comparing FHL5 versus HA cells, we identified 377 differentially expressed genes (log2FoldChange > 0.6 or log2FoldChange < 0.6 and FDR < 0.05), of which 191 and 186 genes were upregulated and downregulated respectively (**Figure 2.3E**). Top upregulated genes included various metalloproteinases (*MMP1*, *MMP3*, *MMP10*) and vessel wall matrisome proteins (*DCN*, *COL5A3*, *ANGPTL4*). We noted that differentially expressed genes (DEGs) were overrepresented in vascular rremodeling pathways, such as cytokine activity/inflammation, ECM organization and ossification (**Figure 2.3F**). FHL5 DEGs were also enriched with vascular and inflammatory disease candidate genes (**Figure S2.7E**).

Based on the perturbation of vascular remodeling and ossification pathways supported by RNA-seq, we next hypothesized that FHL5 may regulate vascular calcification to mediate CAD/MI risk. We treated SMCs with an osteogenic cocktail as done previously[188,189]. We observed increased mineral deposition quantified by alizarin red

staining upon *FHL5* and *FHL5-NLS* overexpression (**Figure 2.3G**). This increased calcification correlated with increased expression of osteogenic activators, *RUNX2* and *ALPL* and reduced expression of osteogenic inhibitors, *MGP* and *SPP1* (**Figure 2.3H)**. Consistent with promoting SMC phenotypic transitions toward an osteogenic state, *FHL5* and *FHL5-NLS* overexpression also coincided with downregulated expression of SMC markers, *LMOD1*, *MYH11*, and *CNN1* (**Figure 2.3I).** Lastly, to validate the role of FHL5 in increasing vascular calcification, we immunostained sections of human coronary artery plaques. In accordance with our *in vitro* findings, we observed colocalization between FHL5 and RUNX2, a transcriptional activator of osteogenic differentiation in close proximity to intimal calcium deposits (**Figure 2.3J**). These results were consistent with the direction of effect of the *FHL5* genetic association with coronary artery calcification (CAC), where the CAC risk allele was associated with increased *FHL5* gene expression. Together, these analyses suggest that FHL5 promotes a shift towards the SMC osteogenic phenotypic state in atherosclerotic lesions to increase CAD/MI risk.

FHL5 serves as a SMC cofactor to regulate disease-associated ECM interactions

In order to further decipher the direct transcriptional network regulated by FHL5 in SMCs under basal conditions, we used the Cleavage Under Targets and Release Using Nuclease (CUT&RUN) method to map genome-wide binding sites [190 191]. In total, we identified 17,201 FHL5 binding sites that map to 6,776 unique genes, with a majority of binding sites located in annotated promoters and intronic regions (**Figure S2.4A** and **S2.8A**) [192 193]. The active enhancer H3K27ac and promoter mark, H3K4me3 signals were enriched around the center of FHL5 peaks (**Figure 2.4B, 2.4C** and **S2.8B through S2.8D**), with about 75% of binding sites overlapping primary HCASMC ATAC-seq peaks[96] (**Figure S2.8E**). To examine the functional consequences of FHL5 binding on transcription, we integrated genome-wide FHL5 binding sites with differential expression analysis using the Binding and Expression Target Analysis (BETA) tool [194]. We observed substantial overlap between FHL5 DEGs and gene harboring FHL5 peaks

(P=6.3E-42, Fisher's Exact Test). A majority (51%) of these differentially expressed genes harboring at least 1 FHL5 binding site near its transcription start site.

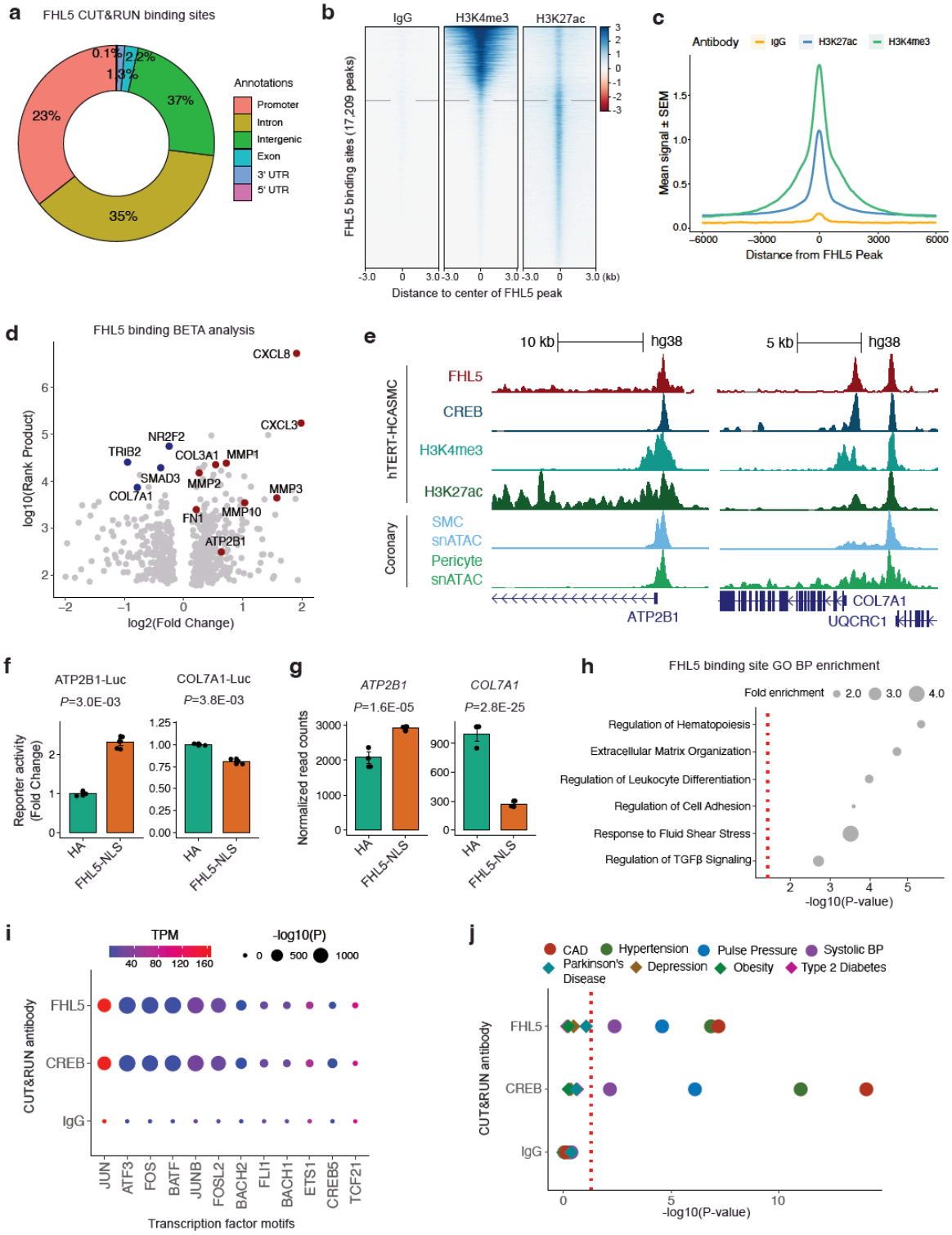

Figure 2.4. FHL5 serves as a cofactor for the transcription factor, CREB, to regulate ECM

organization in SMCs. (a) Overlap of FHL5 peaks identified from CUT&RUN with genomic features. **(b)** Heatmap showing distribution of active chromatin histone marks (H3K4me3 and H3K27ac) +/-3kb from the center of FHL5 peaks, compared to IgG control. Experiment and Analysis Completed by DW. **(c)** Density plot showing genome-wide enrichment of active chromatin histone marks +/- 6kb from the center of FHL5 peaks. Analysis Completed by DW. **(d)** Results of Binding and Expression Target Analysis (BETA) using FHL5 CUT&RUN peaks and FHL5 differentially expressed genes (DEGs) as input. Top candidate direct target genes highlighted in red, or blue based on upregulated and downregulated expression, respectively. Log normalized rank sum product score reflects the likelihood of direct transcriptional regulation for each gene. Analysis Completed by DW. **(e)** UCSC genome browser tracks showing FHL5 binding sites at the *ATP2B1* and *COL7A1* loci, highlighting candidate regulatory elements overlapping CREB binding sites and H3K27ac and H3K4me3 peaks. These regions also overlapped accessible chromatin SMCs and pericytes determined from human coronary artery snATAC-seq. Analysis Completed by DW. **(f)** Relative change in luciferase reporter activity of *ATP2B1* and *COL7A1* enhancer regions in FHL5 overexpressing SMCs. Experiment Completed by DW. **(g)** Relative change in RNA expression of *ATP2B1* and *COL7A1* in SMCs overexpressing FHL5, shown as normalized read counts. Experiment Completed by DW. **(h)** Top GO-BP overrepresented in FHL5 target genes identified from GREAT analysis of FHL5 peaks. Analysis Completed by DW. **(i)** Top transcription factor motifs enriched in FHL5, CREB, and IgG peaks identified from HOMER known motif analysis (dot size) as well as normalized (transcripts per million: TPM) expression level of the corresponding transcription factors in SMCs (color scale). Analysis Completed by DW. **(j)** GREGOR analysis showing enrichment of vascular trait GWAS risk variants in FHL5, CREB, and IgG binding sites. Analysis Completed by DW.

FHL5 directly upregulated key genes involved in vascular remodeling, including contractile genes, metalloproteinases, and ECM components (**Figure 2.4D**). We identified FHL5 binding at the 5' regulatory elements near the *ATP2B1*, a CAD associated gene involved in calcium ion homeostasis, and at the gene promoter of *COL7A1*, a component of the basement membrane (**Figure 2.4E**). We validated the activity of these regulatory elements at the *ATP2B1* and *COL7A1* promoter in SMCs. *FHL5* overexpression increased luciferase activity and correlated with increased *ATP2B1* expression (**Figure 2.4F and 2.4G**). In contrast, *FHL5* overexpression reduced the luciferase activity of the *COL7A1* enhancer and *COL7A1* gene expression, supporting the gene-specific activating or repressive role of FHL5 (**Figure 2.4F and 2.4G**). Gene ontology enrichment analysis of these high confidence target genes supported FHL5 mediated direct regulation of pathways related to ECM organization and cell adhesion (**Figure S2.8F**). We complemented this approach with Genomic Regions Enrichment of Annotations Tool (GREAT)[195] analysis. FHL5 binding sites were proximal to genes related to inflammation, ECM organization and cell adhesion **(Figure**

2.4H). These enrichments mirror the BETA enrichment results and further highlight the regulatory interactions between FHL5 and downstream vascular remodeling genes in mural cells.

Since FHL5 has not been reported to bind DNA directly, we performed motif enrichment analysis on genome-wide FHL5 peaks to identify candidate transcription factor binding partners, which revealed enrichment of AP1 and cAMP-response element (CRE) motifs within these peaks (**Figure 2.4I**). Regulatory elements proximal to upregulated genes after FHL5 overexpression were also enriched in AP1 motifs. Consistent with previous studies, 48% of FHL5 binding sites overlapped CREB binding sites, which was also strikingly enriched in AP1 binding motifs as well (**Figure S2.8E** and **S2.8H**). We further explored additional potential transcription factor binding partners using the Locus Overlap Analysis (LOLA) enrichment tool [196] and assessed the enrichment of FHL5 binding sites in the ENCODE collection of transcription factor ChIP-seq datasets. Relative to the HCASMC ATAC-seq peak set, we observed significant enrichment of Pol2, AP1 members, (c-Fos and c-Jun), CREB and its common cofactor, p300 (**Figure S2.8G**). We validated the role of FHL5 as a cofactor for CREB through a cAMP response element (CRE) luciferase reporter assay. Relative to the HA control, FHL5 overexpression upregulated CRE activity (**Figure S2.8I**).

FHL5 regulates a transcriptional network that contributes to CAD/MI risk by modulating SMC functions and vascular remodeling processes

As a transcriptional regulator of CAD-relevant pathways in the vessel wall, we next hypothesized that FHL5 regulation of downstream target genes contribute to the mechanistic link between *FHL5* and vascular disease risk. We evaluated the enrichment of vascular disease GWAS SNPs in FHL5 binding sites using the Genomic Regulatory Elements and Gwas Overlap algoRithm (GREGOR) software package [197 198]. Relative to a matched set of random SNPs, FHL5 and CREB binding sites were highly enriched for CAD, MI, and blood pressure risk variants (**Figure 2.4J**), compared to non-vascular traits, further emphasizing the specific contribution of *FHL5* to common vascular diseases.

To further support this hypothesis and extend our findings to human artery tissues, we assessed the distal effects of *FHL5* gene expression in STARNET cardiometabolic tissues. rs10872018, a top *FHL5* cis-eQTL was associated with the expression of 743 and 568 eGenes in aortic and mammary artery tissues respectively at nominal significance ($P < 0.05$). These eGenes included key matrisome genes, such as *FN1*, *COL4A4*, and *LUM*. Integration of *trans*-eQTL target genes from STARNET artery tissues and FHL5 binding sites identified a network of downstream genes that harbor CAD risk variants (**Figure 2.5A)**.

We prioritized *FN1 and FOXL1*, two CAD/MI and CAC candidate genes, as likely downstream effectors of FHL5. The rs9486719-G CAD/MI risk allele associated with increased *FHL5* gene expression was associated with both increased *FN1* (**Figure 2.5B**) and *FOXL1* (**Figure 2.5C**) gene expression in *trans*. We identified an FHL5 binding site overlapping active regulatory marks harboring CAD and CAC risk variants at both the *FN1* promoter and upstream regulatory element of *FOXL1* (**Figure 2.5D**). This candidate *FOXL1* enhancer harbors rs423984, a variant in the MI 95% credible set that is also associated with CAC and *FOXL1* gene expression in muscle tissue (**Figure 2.5D** and **S2.9A**). As is common for intergenic variants at GWAS loci, this variant is likely correlated with expression of both *FOXL1* and neighboring gene *FOXC2,* based on coronary artery snATAC peak2gene links.

To provide insights into the function of FHL5 target genes in human arteries, we queried STARNET cross tissue gene regulatory networks. *FN1* is a member of module 39 in AOR tissue, which was enriched in genes predicted to regulate ECM organization and cell adhesion and correlated with clinical traits indicative of CAD severity (DUKE and SYNTAX score). Similarly, we identified *FOXL1* and *FHL5* in module 28, a cross-tissue GRN that also includes CAD-associated genes, *SVEP1* and *MFEG8*. This module was functionally enriched in genes related to ECM organization, SMC contraction and calcium ion regulation (**Figure 2.5E** and **S2.9B**), and correlated with atherosclerotic lesion presence and disease severity (**Figure 2.5E**). These enrichment analyses recapitulate the established role of *FN1* in the vessel wall and its link to disease and highlight *FOXL1* as a candidate effector gene during vascular remodeling.

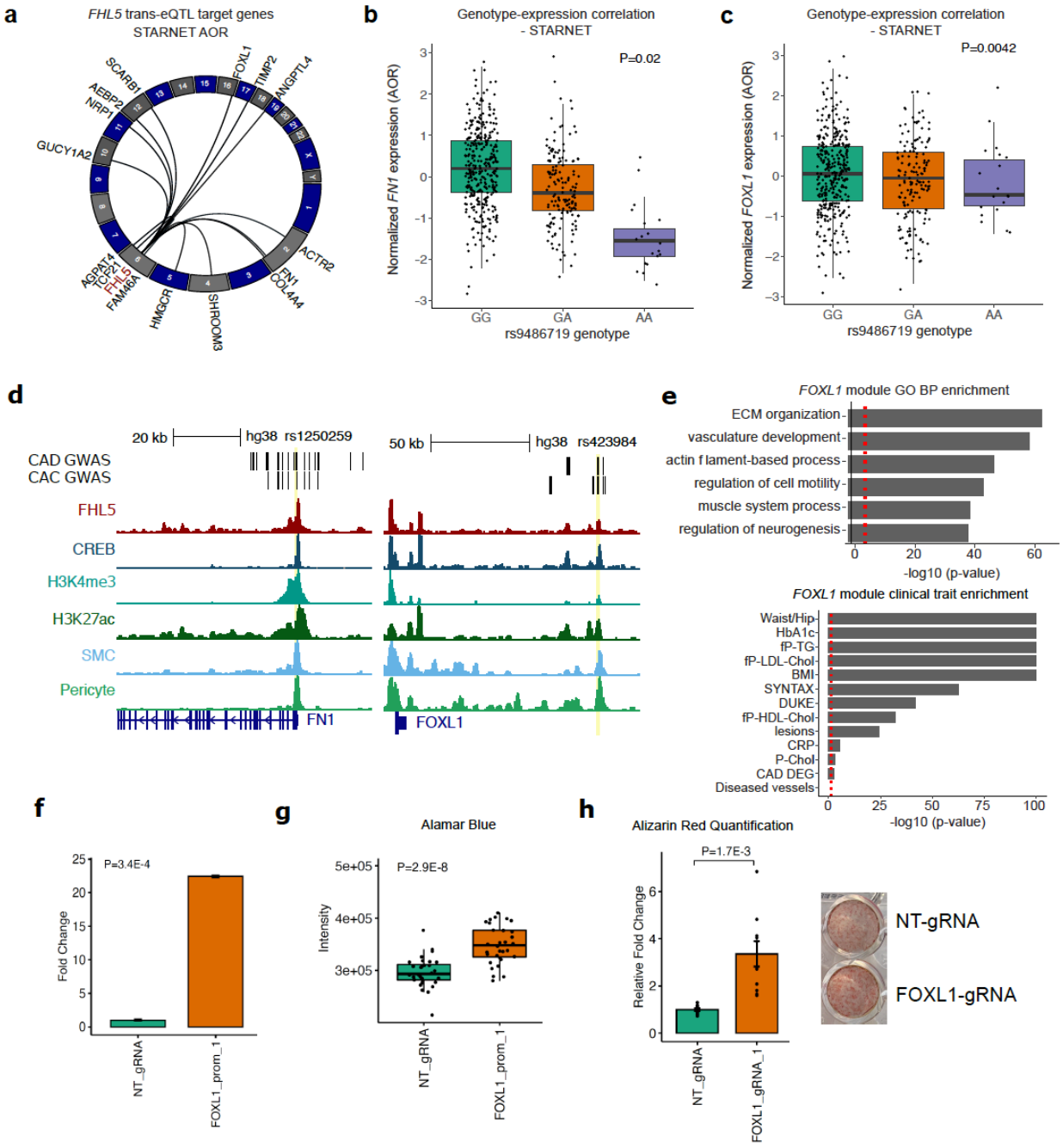

Figure 2.5. FHL5 regulates a network of CAD associated genes in human artery tissues. **(a)** Circos plot showing top *trans*-eQTL target genes for *FHL5* lead SNP rs9486719 from STARNET artery tissues, which also harbor FHL5 binding sites in SMC. Genes are plotted according to genomic position with each black link representing significant *trans*-eQTL effects (P<0.05). Previously Published Data. Figure Created by DW. **(b)** Association of rs9486719 with *FN1* gene expression in atherosclerotic aortic artery (AOR) in STARNET. Previously Published Data. Figure Created by DW. **(c)** Association of rs9486719 with *FOXL1* gene expression in AOR tissue. Previously Published Data. Figure Created by DW. **(d)** UCSC genome browser tracks showing FHL5 binding sites overlapping CREB, H3K4me3, and H3K27ac peaks, and accessible chromatin peaks in native SMC and pericytes (human coronary artery snATAC). The

peak2Gene link indicates significant correlation of the candidate enhancer with the *FOXL1* promoter (left) and *FN1* promoter (right). Previously Published Data. Figure Created by DW. **(e)** Top, enrichment of clinical traits in module 28 containing *FOXL1* and *FHL5* in STARNET cross-tissue networks. The red dotted line corresponds to FDR<0.05. Bottom, top GO-BP enriched in module 28. The red dotted line corresponds to a nominal threshold of FDR<0.05. Previously Published Data. Figure Created by DW. **(f)** Relative expression of *FOXL1* gene expression using dCas9-SAM/MPH lentiviral system. Experiments Performed by DW. **(g)** Quantification of Alamar Blue fluorescent intensity. Experiments Performed by DW **(h)** Quantification of alizarin red staining (OD405) at 21dpt. Experiments Performed by DW.

To further explore the potential link between *FHL5* and *FOXL1*, we observed *FOXL1* was highly correlated with *FHL5* gene expression in human coronary arteries (**Figure S2.9C**). Similar to the expression profile of *FHL5*, *FOXL1* gene expression was highly enriched in SMC and pericytes (**Figure S2.9D**) and in bulk artery tissues in both STARNET and GTEx (**Figure S2.9E** and **S2.9F**). We next experimentally validated the functional significance of *FOXL1* to SMC phenotypic modulation using the lentiviral dCas9 Synergistic Activator Mediator (SAM) system [199]. Relative to a non-targeting gRNA control, the *FOXL1* gRNA targeting the promoter region increased expression over twenty-fold (**Figure 5F**). This upregulation of *FOXL1* increased SMC proliferation (**Figure 5G**) and calcification (**Figure 5H**), mirroring the direction of effect of *FHL5* overexpression. Consistent with promoting maladaptive SMC transitions, *FOXL1* gene expression was elevated in atherosclerotic aortic tissues relative to healthy controls (P=1.4E-6). Taken together, *FHL5* mediated regulation of *FOXL1* and other ECM genes in *trans* may contribute in part to the mechanistic link between regulatory variants and vascular disease risk.

Discussion

Despite the identification of over 200 loci associated with CAD, the mechanisms underlying these associations are largely unknown. Given at least half of these loci are predicted to function independently of traditional risk factors[200], characterization of these novel loci may provide insights into some of the 'hidden' heritable risk factors. Our study addresses this gap by implicating *FHL5* as the top candidate gene at the *UFL1-FHL5* locus associated with vascular diseases. *UFL1-FHL5* is also a well-established risk locus for migraine[201], consistent with the genetic correlation and vascular etiology[202,203]. In this study, we elucidated the upstream regulatory mechanisms and downstream

function of FHL5 in the human coronary artery (**Figure 6**). The CAD/risk allele (rs10872018-G) preserves the SRF/MYOCD binding site to increase *FHL5* gene expression, supporting the detrimental role of FHL5 in regards to CAD/MI risk. At the protein level, FHL5 promotes SMC calcification and contraction through dysregulation of vascular remodeling and calcium handling genes. Lastly, we characterized the FHL5 regulated SMC transcriptional network and highlighted *trans*-acting mechanisms mediated by FHL5 that contribute to the heritability of CAD and other common vascular diseases.

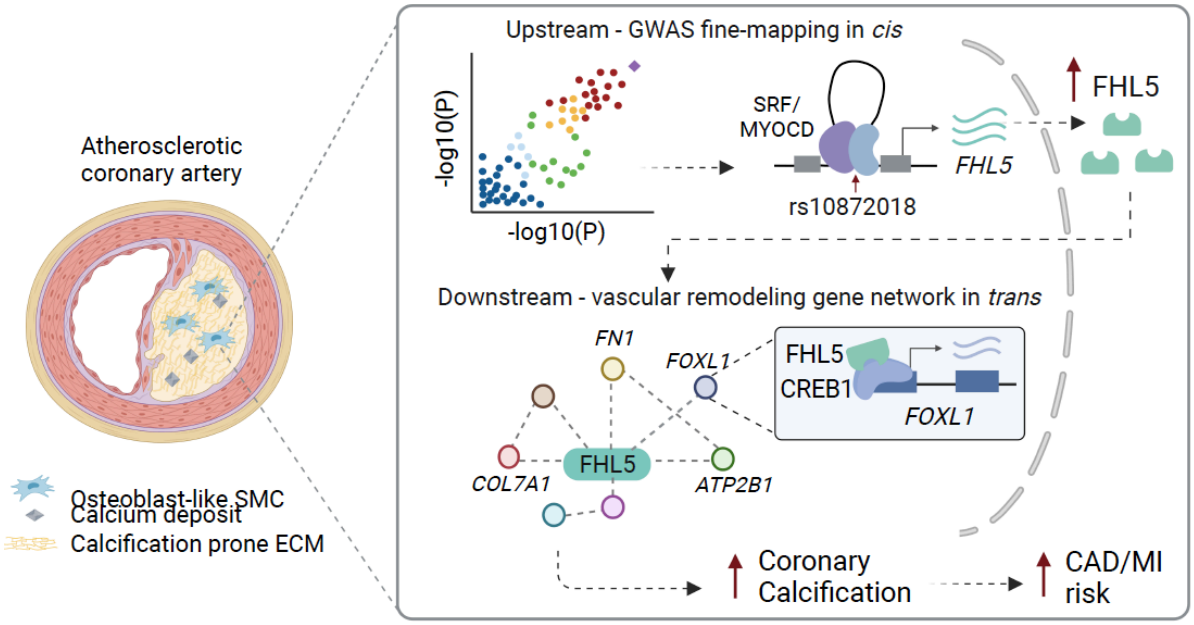

Figure 6: Proposed Model Underlying the *FHL5* Association of with CAD/MI.
Schematic depicting the proposed upstream and downstream mechanisms underlying the *FHL5* genetic association with CAD/MI and other vascular traits. Top, risk alleles for candidate causal variants (e.g. rs10872018) are associated with increased *FHL5* gene expression, which are predicted to function through SRF-MYOCD enhancers in *cis*. Bottom, this results in increased binding of FHL5 cofactor to CREB regulatory elements in vascular remodeling gene network in *trans*, ultimately leading to maladaptive ECM remodeling and vascular calcification in SMC and osteogenic-like SMC, thus increasing disease risk. Created with BioRender.com Figure created by DW.

In atherosclerotic lesions, SMCs adopt phenotypes characteristic of other cell lineages, such as macrophages, fibroblasts, and osteoblasts[62,204]. SMCs that adopt this osteogenic phenotype participate in the deposition of intimal microcalcifications. Our findings support a mechanism where FHL5 mediated dysregulation of various metalloproteinases, collagens, and calcium signaling genes contributes to this pro-

calcifying phenotype. Given the colocalization of RUNX2 and FHL5 in human coronary artery lesions, we postulate that FHL5 expression in this population of dedifferentiated SMCs alters the ECM composition to promote mineral deposition in the lesion. The significance of this calcification phenotype is underscored by the use of calcium scores in the clinic to identify patients at high risk for MI and other adverse cardiovascular events [205 206]. Despite observing increased contractility under basal conditions, which may reflect the homeostatic role of FHL5 in maintaining vascular tone, our studies under osteogenic conditions implicate FHL5 as a pro-calcifying factor as well. In support of a putative role in regulating both phenotypes, recent work by Karlöf et al [207] reported that SMC contractile markers, which included *FHL5, MYOCD and CNN1*, were upregulated ~4-fold in highly calcified carotid artery plaques relative to lowly calcified plaques. This upregulation of SMC markers also correlated with increased expression of recently characterized SMC calcification markers (e.g. proteoglycan 4 (*PRG4*)) [208]. This is further supported by gene ontology enrichment analysis highlighting calcium signaling, cytoskeletal rearrangements, and muscle contraction as overrepresented pathways among differentially expressed genes between highly and lowly calcified lesions[207]. To reconcile these phenotypes, we speculate that FHL5 regulation of intracellular calcium ion homeostasis contributes to both processes, given the critical role of calcium in the initiation of SMC contraction [209 210], cell stiffness [211 212], and vascular calcification [179 213].

This study confirms and extends previous work identifying FHL5 as a cofactor for CREB. Our CUT&RUN data supports a similar mechanism where FHL5, through interactions with CREB, regulates target genes associated with vascular remodeling in SMCs. Previous studies have linked activation of the CREB/ATF3 signaling pathway to increased SMC proliferation, migration, and calcification [214]. In addition to the CRE motif, we also observed strong enrichment of the AP1 binding motif in FHL5 binding sites. This motif was also found to be enriched in accessible regulatory elements throughout the SMC transition to a fibroblast-like state [215]. Although we cannot rule out additional direct interactions with members of the AP1 family, this enrichment may also reflect epigenetic functions, similar to AP1 interactions at downstream CAD-associated loci, *SMAD3* and *CDKN2B-AS1* [216]. Lastly, while we highlight that *FHL5* is regulated by

SRF-MYOCD in *cis*, whether it directly interacts with or competes with MYOCD for binding to SMC gene promoters to regulate gene expression in *trans* [136], requires further investigation.

Dissecting the regulatory network of disease associated transcription factors such as *KLF4* and *TCF21*, have identified regulatory interactions with downstream CAD loci that govern SMC phenotypic modulation [63 62]. In support of the proposed omnigenic model, these studies highlight *trans*-acting mechanisms that may contribute to the majority of CAD heritability [217 218]. Our findings support the role of *FHL5* as a "peripheral" transcriptional regulator of downstream "core" genes that have predicted functions in the vessel wall. By integrating STARNET vascular tissue *trans*-eQTL data, we identified a network of putative core genes that contribute to CAD pathogenesis, which include the well-characterized matrisome gene, *FN1* and the transcription factor, *FOXL1*. *FN1* has already been implicated as a regulator of SMC phenotypic modulation associated with CAD [62 64 219]. In contrast, the precise mechanism of how *FOXL1* functions in artery tissues and atherosclerosis has not been fully characterized. *FOXL1* and neighboring gene *FOXC2* are both associated with CAD and bone mineral density [220]. As supported by our coronary artery SMC and STARNET cross-tissue GRNs, we propose that *FOXL1* may regulate vascular remodeling pathways *in vivo* to impact disease risk. Functional dissection of other *FHL5* target genes may uncover novel mechanisms and candidate genes contributing to heritable risk for CAD and other vascular diseases.

Although our study integrates multiple lines of human evidence to unravel the function of *FHL5* in SMCs, we acknowledge known limitations. First, since we focus primarily on subclinical disease, this study does not address the potential roles for *FHL5* in advanced disease stages. Future studies using more complex *in vitro* and human clinical samples will be needed to establish the causal link between FHL5 and indices of plaque stability and clinical outcomes. Given the lack of *Fhl5* expression in murine artery tissues, investigations utilizing traditional models of atherosclerosis and vascular calcific disease may not be applicable. Second, this study focuses primarily on the transcriptional role of FHL5 in SMCs. Other FHL family members, *FHL1* and *FHL3*, are

known to function as scaffolds to regulate sarcomere formation in myoblasts[149 221 222]. We speculate that FHL5, due to the presence of similar conserved LIM domains, may mediate similar interactions with the actin cytoskeleton [223 224] or focal adhesions[225 146]. While our study provides insights into the FHL5 mediated gene regulatory mechanisms, future studies will be needed to clarify its role in mediating mechanotransduction or other cell signaling events.

In summary, this work reveals a molecular mechanism by which *FHL5*, the top candidate causal gene at the pleiotropic *UFL1-FHL5* locus, modulates vascular disease risk. We propose that FHL5 regulates a network of downstream genes in SMCs that participate in adverse vascular remodeling events driving disease progression. Similar to the effects of modulating other transcriptional regulators, modest increases in FHL5 gene expression may be propagated through its gene regulatory network in the vascular wall, ultimately increasing CAD/MI risk over time. Characterization of these *trans* effects in SMCs also highlights regulatory interactions at the molecular level governing atherogenic SMC phenotypic modulation. Future studies mapping genome-wide interactions of other disease associated transcriptional regulators may uncover new mechanisms and prioritize effector genes that impact primary disease processes, thereby accelerating the development of therapeutics to prevent MI and other vascular related pathologies.

Novelty and Significance:

What is known?

- Common genetic variants at *UFL1-FHL5* contribute to the genetic risk for a number of vascular diseases, including CAD/MI, hypertension, aneurysm, and migraines.
- Expression quantitative trait locus analysis indicates that vascular disease risk variants modulate *FHL5* gene expression in human artery tissues.
- *FHL5* was previously characterized as a potent coactivator during development and more recently was linked to vein graft intimal hyperplasia.

What new information does this article contribute?

- *FHL5* is the top candidate causal gene underlying the pleiotropic genetic associations with CAD/MI and other common vascular diseases.
- *FHL5* promotes the osteogenic SMC transition to increase vascular calcification *in vitro* under calcifying conditions and within human atherosclerotic lesions *ex vivo*.
- *FHL5* regulation of downstream gene programs harboring CAD/MI risk loci (e.g. *FN1*, *FOXL1*) linked to SMC phenotypic modulation and vascular remodeling, contributes to the heritable risk of common vascular diseases.

Supplementary Methods

Coronary artery tissues and human subjects

Freshly explanted hearts from orthotopic heart transplantation recipients were procured at Stanford University under IRB approved protocols and written informed consent. Participants were not compensated for this study. Hearts were arrested in cardioplegic solution and rapidly transported from the operating room to the adjacent lab on ice. The proximal 5-6 cm of three major coronary vessels (left anterior descending (LAD), left circumflex (LCX), and right coronary artery (RCA)) were dissected from the epicardium on ice, trimmed of surrounding adipose (and in some samples the adventitia), rinsed in cold phosphate buffered saline (PBS), and rapidly snap frozen in liquid nitrogen. Coronary artery samples were also obtained at Stanford University (from Donor Network West and California Transplant Donor Network) from non-diseased donor hearts rejected by surgeons for heart transplantation and procured for research studies. All hearts were procured after written informed consent from legal next-of-kin or authorized parties for the donors. Reasons for rejected hearts include size incompatibility, comorbidities or risks for cardiotoxicity. Hearts were arrested in cardioplegic solution and transported on ice following the same protocol as hearts used for transplant. Explanted hearts were generally classified as ischemic or non-ischemic cardiomyopathy and prior ischemic events and evidence of atherosclerosis was obtained through retrospective review of electronic health records at Stanford Hospital

and Clinics. The disease status of coronary segments from donor and explanted hearts was also evaluated by gross inspection at the time of harvest (for presence of lesions), as well as histological analysis of adjacent frozen tissues embedded in Tissue-Tek O.C.T. compound (Sakura) blocks. Frozen tissues were transferred to the University of Virginia through a material transfer agreement and Institutional Review Board approved protocols. All samples were then stored at -80°C until day-of-processing.

GWAS fine-mapping analyses

We used Summary-level Mendelian Randomization (SMR)[103] to prioritize candidate causal genes that underlie the vascular GWAS signals at the *UFL1-FHL5* locus. SMR is a transcriptome wide association study (TWAS) method that tests whether the association of a variant with a phenotype is mediated through changes in gene expression, while accounting for the linkage disequilibrium in the study population. We performed SMR using STARNET and GTEx artery tissue eGenes for all genes +/-500kb of the FHL5 transcription start site (*FHL5*, *UFL1*, *GPR63*, *NDUFAF4*, *KLHL32*, *MMS22L*), using the 1000G European reference panel. We used the HEIDI (heterogeneity in dependent instruments) test to filter associations driven by linkage disequilibrium rather than pleiotropy (pHEIDI > 0.01) as previously done[104]. We considered pSMR < 0.008 (0.05/6) and pHEIDI > 0.01 as evidence of pleiotropy or causality.

In a complementary approach, we performed a Bayesian colocalization analysis using the R package, *coloc 4.0-3* [99,98] to calculate the posterior probability that the GWAS and eQTL data share a common signal. We colocalized GWAS signals with GTEx and STARNET artery tissue eQTLs using the coloc.signals function with default priors (p1=1E-4, p2=1E-4, p12=1E-5). We considered the PP4 > 0.80 as evidence of colocalization.

We used the Bayesian fine-mapping tool, PAINTOR_V3[174] to prioritize CAD causal variants at the *UFL1-FHL5* locus. We included 1Mb around the lead CAD GWAS SNP, rs9486719 and computed Pairwise Pearson correlations for all bi-allelic SNPs based on

the 1000G European reference panel. We incorporated human coronary artery ATAC-seq peaks from Turner et al[183] as functional annotations to calculate prior probabilities for the model. Following calculation of the posterior probability, we ranked SNPs based on this metric and identified the 95% credible set.

Human coronary artery smooth muscle cell immortalization and characterization

Primary human coronary artery SMCs (HCASMCs) were purchased from Cell Applications and immortalized by lentiviral transduction of the hTERT-IRES-hygro construct (Addgene Plasmid #85140)[187] in 10ng/ml polybrene. Transduced cells were selected using 400μg/ml of hygromycin (Gibco) and maintained in SMC basal medium (SmBM) supplemented with SmGM-2 SingleQuots kit (PromoCell). The maintenance of similar levels of SMC marker gene expression (e.g. SM22-alpha) was confirmed via western blot. The transcriptomes of these immortalized SMCs (HCASMC-hTERT) were characterized via RNA-seq and correlated with RNA-seq data of the parental cell line[105]. A principal component analysis was performed incorporating primary HCASMCs, HCASMC-hTERT, HEPG2, K562, and human umbilical vein endothelial cell (HUVEC) RNA-seq data (**Supplementary Fig 6**). Fastq files for the primary HCASMCs were extracted from Sequence Read Archive (SRA) (SRR7064063, SRR7058289, SRR7058290). Fastq files for HEPG2, K562, and HUVECs were downloaded from ENCODE.

Cell culture conditions

Primary HCASMCs and HCASMC-hTERT were maintained in SMC basal medium (SmBM) supplemented with SmGM-2 SingleQuots kit (PromoCell #C-22162), which includes 5% Fetal Bovine Serum (FBS), insulin, Fibroblast Growth Factor (FGF), and Epidermal Growth Factor (EGF). HEK293T and A7r5 cells were cultured in 10% FBS (Gibco) supplemented DMEM (Sigma-Aldrich). All cell cultures were maintained at 37°C and 5% CO_2.

Lentivirus generation and transduction

FHL5 cDNA was amplified from genomic DNA and cloned into the lentiviral vector, Lenti-III-EF1α (ABM #LV043) using EcoR1 and BamH1 restriction sites. The transfer vector, psPAX, (Addgene#12260), and pMD2.G (Addgene #12259) were transfected into HEK293T cells using Lipofectamine 3000 (Invitrogen) according to the manufacturer's protocol. The lentiviral containing supernatant was harvested at 24 and 48 hours post transfection and frozen at -80°C until use.

HCASMCs/HCASMC-hTERT were transduced at 60% confluence with 500µl of lentiviral containing supernatant and 10ng/ml polybrene (EMD-Millipore). The viral containing supernatant was removed the following day. 1.0µg/ml puromycin (Gibco) was added to the culture medium for antibiotic selection 48 hours post transduction. Puromycin resistant cells were expanded, and *FHL5* overexpression was validated via qPCR and/or western blot.

CRISPR-dCas9 SAM/MPH

gRNAs targeting the *FOXL1* promoter was designed using CRISPick algorithm (https://portals.broadinstitute.org/gppx/crispick/public) and cloned into the lentiSAMv2 (Addgene #75112) using the BsmBI restriction enzyme sites. HCASMC-hTERT were transduced with lentiMPHv2 (Addgene #89308) and lentiSAMv2 with gRNA insert. Cell lines were maintained in SMC media with 400ug/ml hygromycin and 5ug/ml blasticidin.

Immunofluorescence

Coronary arteries were embedded in OCT and cryosectioned at 8µm. Frozen slides were washed with sterile PBS twice for 2min followed by fixation with formaldehyde at 4% for 10 min. Then slides were washed with PBS twice for 2 min and tissue were permeabilized with Triton X at 0.05% for 10 min. Coronary artery tissues were blocked with donkey serum at 10% for 1 hour followed by incubation overnight at 4°C with anti-rabbit-FHL5 (Novus, NBP-32600) at 1:100 dilution and anti-mouse-α-SMA (Santa Cruz Biotechnology, SC-53142) at 1:100 dilution. Slides were washed with PBS-tween at 0.1% 3 times for 3 min each followed by incubation with secondary antibody (1:400) and F-actin (ActinGreen™ 488 ReadyProbes, ThermoFisher Scientific, USA) for 1 hr at

room temperature. Then slides were washed with PBS-t (PBS, 0.05% Tween 20), 4 times for 3 min each and slides were mounted with diamond mounting medium containing DAPI. Slides were visualized with the Leica TCS SP8 confocal microscopy station and micrographs were digitized with the Leica Application Suite X software.

Luciferase reporter assay

The enhancer sequence (180bp) harboring the rs10872018-G or rs10872018-A was cloned into the pLUC-MCS (Agilent) vector. 400ng of the luciferase reporter constructs (pLuc-MCS, pLuc-MCS-rs10872018-G, pLuc-MCS-rs10872018-A), 100ng of each expression plasmid (pCGN-SRF, pCDNA6-MYOCD, or pCMV6-empty), and 5ng of Renilla-luciferase were transfected into A7r5 cells using Lipofectamine 3000. (Invitrogen). The media was replaced 6 hours post transfection. Samples were prepared using the Dual Luciferase Assay Kit (Promega) according to the manufacturer's protocol 24 hours post transfection. The luciferase activity was measured using the SpectraMax L luminometer (Molecular Devices). The Firefly luciferase signal was normalized to the Renilla luciferase signal for each sample.

The FHL5 bound enhancer sequence was amplified from HCASMC-hTERT genomic DNA and cloned into the lentiviral luciferase reporter plasmid, pLS-MP-Luc (Addgene #106253). HCASMC-hTERT were transduced with 500µl of lentiviral-containing supernatant and the media was replaced the next day. Samples were prepared using the Dual-Glo Luciferase Assay system (Promega) 48 hours post transduction and the Firefly luciferase signal was normalized to total protein concentration determined using a Bradford colorimetric assay (Thermo Scientific 23225).

Collagen gel contraction assay

250,000 HCASMCs were embedded in a 2mg/ml collagen solution (Advanced BioMatrix #5074). Following incubation at 37°C for 90 minutes, the solidified gels were detached and 1ml of media was added on top of each collagen lattice. After 18-24 hours at 37°C, the diameter of the collagen gels was measured and quantified using ImageJ.

SMC calcification and alizarin red staining

30,000 HCASMC-hTERT were maintained in DMEM (Sigma #D6429) supplemented with 10% FBS (Gibco #26140079), 10mM β-glycerophosphate disodium salt hydrate (Sigma #G9422), 50µg/mL ascorbic acid (Sigma, #A8960), and 10nM dexamethasone (Sigma, #D2915) for 21 days, with media replaced every other day. Following 14 and 21 days of treatment, RNA was extracted from cells for qPCR. Following 21 days of treatment, cells were washed 2x with 500µL of PBS and fixed with 4% paraformaldehyde (Thermo Scientific #28908) at room temperature for 15min. Wells were rinsed twice with water and incubated with 500µL of alizarin red staining solution (Sigma, #TMS-008-C) for 20min at room temperature. Excess dye was removed by rinsing with water. Alizarin red staining was quantitated by extraction of stained calcified material at low pH. 400µL of 10% acetic acid (Ricca Chemical #R0135000) was added to each well and incubated at room temperature for 30 min. The monolayer and acetic acid slurry was vortexed, heated to 85°C for 10 min, cooled on ice for 5 min, and centrifuged at 4°C at 20,000xg for 15 min. The supernatant was neutralized with 150µl of 10% ammonium hydroxide (Alfa Aesar 35575-AP) and OD405 was measured. Alizarin red concentration was determined relative to known concentrations of Alizarin Red standards.

SMC calcium quantification

SMCs were incubated with Fluo-4 AM (2.5 µM; ThermoFisher #F14201) and pluronic acid (0.004%) at 37°C for 20 minutes[226]. Images were then acquired at 30 frames per second using an Andor Revolution WD (with Borealis) spinning-disk confocal imaging system (Andor Technology) comprising an upright Nikon microscope with a 40X water-dipping objective (numerical aperture, 0.8) and an electron-multiplying CCD camera. Fluo-4 was excited using a 488 nm solid-state laser, and emitted fluorescence was captured using a 525/36 nm band-pass filter. Following baseline measurements, SMCs were treated with phenylephrine (5 µM; Sigma). Ca^{2+} ionophore, ionomycin (10 µM; Sigma #I0634) was used at the end of the experiment to calculate the total maximum calcium concentration[227].

Images were analyzed using custom-designed SparkAn software[228,229]. Fractional fluorescence traces (F/F_0) were obtained by region of interest (ROI) using polygon drawing on each cell, excluding the nucleus.

Estimates of $[Ca^{2+}]_i$ were made using the following equation,

Ca2+= KdF/Fmax-1/Rf1-F/Fmax

where F is fluorescence measured within an ROI, Fmax is the fluorescence intensity of Fluo-4 at a saturating maximum Ca^{2+} concentration, K_d is the dissociation constant of Fluo-4 (340 nM)[230], and R*f* (= 100) is the Fluo-4 AM maximum:minimum ratio measured at saturating and zero Ca^{2+} concentrations[231]. Fmax was obtained individually for each culture dish by adding the Ca^{2+} ionophore, ionomycin (10 μM), and 20 mM external Ca^{2+} at the end of the experiment. Fractional fluorescence (F/F_0) was determined by dividing the fluorescence intensity (F) within an ROI by the mean fluorescence value (F_0), determined from images collected before stimulation.

Alamar Blue Assay

2,000 HCASMC-hTERT were seeded into each well of 96-well plate in 100ul SMC media. After 4 days, 10u of almar Blue reagent (Thermo Fisher) was added to each well and incubated for 1 hour at 37C. Fluorescence intensity was measured using excitation wavelength of 540 nm and emission wavelength 585 nm.

Western blot

Cells were lysed in 1X RIPA buffer (Millipore Sigma #20-188) supplemented with a protease inhibitor cocktail (Roche #11697498001). Protein concentration was determined by a BCA Protein Assay (Thermo Scientific #PI23225). 20ug of protein lysate per sample was fractionated through 4-20% Tris-Glycine gel (Invitrogen #XP04205BOX) under denaturing conditions. Gels were transferred at 20V for 1hr to 0.2um pore size PVDF membrane (Invitrogen #LC2002) using the Invitrogen Mini Gel Tank (Invitrogen #A25977). The PVDF membrane was blocked for 1hr at RT in 5% non-fat dry milk (Lab Scientific #M0841) in PBS with 0.1% tween-20 (PBS-t) and incubated

overnight at 4°C with primary antibody diluted in 5% non-fat dry milk. The following primary antibodies were used in western blots: anti-FHL5 (Abnova #H00009457-M01) at 1:1000 dilution, anti-HA (Abcam #ab9110) at 1:5000 dilution, anti-H3 (Cell Signaling #9715S) at 1:5000 dilution, anti-alpha smooth muscle Actin (Abcam #ab5694) at 1:2500 dilution, and anti-SM22 (Abcam #ab14106). The following day the membrane was washed 3x with PBS-T, incubated with a HRP conjugated secondary antibody (Abcam #ab205718) for 1 hour diluted 1:10000 in 5% non-fat dry milk. The signal was detected using chemiluminescent substrate (Thermo Scientific #34580) and visualized in a linear range on the Amersham Imager 600 (GE Healthcare).

Quantitative PCR analysis

RNA was extracted using the Quick-RNA Minoprep kit (Zymo Research # R1055). Equal amounts of RNA, quantified by the Qubit Fluorometer (Invitrogen #Q33326), was reverse-transcribed into cDNA the high-capacity RNA-to-cDNA kit (Applied Biosystems #4387406) according to the manufacturer's instructions. Taqman qPCR was performed in duplicate for each sample using the QuantStudio 5 qPCR instrument (ThermoFisher), normalized to GAPDH expression levels and analyzed via the standard $2^{-\Delta\Delta Ct}$ method. Alternatively, SYBR qPCR was performed in duplicate for each sample, normalized to U6 or GAPDH expression levels, and analyzed using the standard $2^{-\Delta\Delta Ct}$ method. All SYBR oligos were pre-validated for single amplicons using high-resolution melt curve analyses in QuantStudio.

RNA-seq library preparation, sequencing, and analysis

HCASMC-hTERT

RNA was extracted from HCASMC-hTERT cells (HA, FHL5 or FHL5-NLS) using the Quick-RNA Miniprep kit (Zymo Research #R1054). The quality of RNA was assessed on the Agilent 4200 TapeStation. High quality samples, (RIN score > 8) were sequenced at the University of Virginia Genome Analysis and Technology Core in triplicate. Salmon [232] was used to quantify transcripts from demultiplexed fastq files. Data normalization and differential expression (DE) analysis was performed using DESeq2 (v1.30.1). The Wald Test as implemented in the standard DESeq2 pipeline

was used to determine differentially expressed genes. Genes with FDR < 0.05 and log2 fold change > 0.6 were considered significant. Gene ontology enrichment and pathway analyses were performed using the EnrichR web-server [233 234] or ClusterProfiler (v4.1.4) [235 236].

Coronary artery tissues

RNA was extracted from ~50mg of frozen human coronary artery tissue using the Qiagen miRNeasy Mini RNA extraction kit (Qiagen #217004). Prior to column-based RNA isolation, the tissue was pulverized using a mortar and pestle and then homogenized in Qiazol lysis buffer using stainless steel beads in a Bullet Blender (Next Advances). RNA concentration was determined using Qubit 3.0 and RNA quality was determined using the Agilent 4200 TapeStation. Libraries were prepared from high quality RNA samples (RNA Integrity Number (RIN) > 5.5 and Illumina DV_{200} > 75) using the Illumina TruSeq Stranded Total RNA Gold kit (catalog #20020599) and barcoded with TruSeq RNA unique dual indexes (catalog # 20022371). 150bp paired end sequencing was performed at Novogene on an Illumina NovaSeq S4 Flowcell to a medial depth of 100 million total reads per library. The raw passed filter sequencing reads obtained from Novogene were demultiplexed using the bcl2fastq script. The quality of the reads was assessed using FASTQC and the adapter sequences were trimmed using trimgalore. Trimmed reads were aligned to the hg38 human reference genome using STAR v2.7.3a according to the GATK Best Practices for RNA-seq. To increase mapping efficiency and sensitivity, novel splice junctions discovered in a first alignment pass with high stringency, were used as annotation in a second pass to permit lower stringency alignment and therefore increase sensitivity. PCR duplicates were marked using Picard and WASP was used to filter reads prone to mapping bias. Total read counts and RPKM were calculated with RNA-SeQC v1.1.8 using default parameters and additional flags “-n 1000 -noDoC -strictMode” and GENCODE v30 reference annotation. The transcript and isoform expression levels were estimated using the RSEM package[237].

CUT&RUN assay, library preparation and sequencing

CUT&RUN assay was performed as previously described with few modifications[238]. The detailed protocol can be found at https://www.protocols.io/view/cut-amp-run-targeted-in-situ-genome-wide-profiling-14egnr4ql5dy/v3. Briefly, 250,000 HCASMC-hTERT were washed twice with wash buffer 20mM HEPES pH 7.5, 150mM NaCl, 0.5mM spermidine, 1x protease inhibitor). 10µl conA beads (Bang Laboratories #BP531) were washed three times with Binding Buffer and incubated on ice until use. Cells were incubated at RT for 10min on an orbital shaker (nutator, VWR) with the activated conA beads. The mixture was incubated overnight at 4°C with primary antibody (rabbit anti-FHL5 or anti-HA or IgG) diluted in Antibody Buffer (1x Wash Buffer, 0.05% Digitonin, 2.5mM EDTA). All primary antibodies and secondary antibodies are described in Table S2. Next day, the unbound antibody was washed away with 1mL Dig-Wash Buffer (0.05% digitonin) twice. Samples were resuspended in 150µl Dig-Wash Buffer and incubated with diluted secondary antibody for 1 hour at RT. Unbound secondary antibody was washed away with three washes of Dig-Wash Buffer and then resuspended in 50µl Dig-Wash buffer. 2.5µl of pAG-MNase (Epicypher # 15-1016) was added to each sample and incubated for 1 hour at 4°C on a nutator. After washing away unbound pAG-MNase three times with Dig-Wash buffer and resuspension in 150µl of Dig-Wash buffer, 1mM $CaCl_2$ was added to activate MNase cleavage. Samples were incubated at 0°C for 30min. 150µl STOP Buffer (200mM NaCl, 20mM EDTA, 50ug/mL RNASE A, 40ug/mL glycogen) was added to halt digestion and the resulting DNA fragments was released into solution following Proteinase K digestion and purified via phenol-chloroform extraction. DNA pellets were dissolved in 30µl of TE Buffer (1mM Tris-HCl pH8.0, 0.1mM EDTA) and stored at 4°C overnight.

The CUT&RUN libraries were prepared as previously described[190,191]. 10µl of ERA buffer (4X T4 Ligase Buffer, 2mM dNTP, 1mM ATP, 10% PEG4000, 0.5U/µl T4 PNK, 0.05U/µl T4 DNA polymerase, 0.05U/µl of Taq DNA polymerase) was added to each sample. The samples were placed in a thermocycler with the following program: Cycle 1: 12°C /15min, Cycle 2: 37°C/15min, Cycle 3: 58°C/ 45min, Cycle 4: 8°C Hold. 5µl of 0.15uM of annealed Illumina Truseq adapters and 40µl of 2x Ligation Buffer (2x Rapid Ligase Buffer (Enzymatics #B101L) and 4µl T4 Ligase (Enzymatics L6030-HC-L) were

added. Samples were then incubated at 20°C for 25min. Following adapter ligation, the libraries were amplified 14 cycles using NEBNext Q5 Ultra Master Mix (NEB #M0544L) and purified with 1.2X volume of Ampure Beads (Beckman CoulterCatalog #A63880). DNA was eluted into 20µl of Tris-EDTA Buffer and stored at -20°C until sequencing. DNA concentration was determined using Qubit dsDNA High Sensitivity Assay (Invitrogen #Q32851) and library size was assessed using DNA High Sensitivity Tape-Station kit (Agilent #5067-5584). Libraries were pooled and paired-end sequencing (2x42bp, 8bp index) was performed using the NextSeq 2000 instrument with the NextSeq 2000 P2 kit at the UVA Genomics Core.

CUT&RUN analysis

CUT&RUN libraries were analyzed using the CUT&RUN Tools pipeline[192,239]. Briefly, adapter sequences were trimmed from demultiplexed reads using Trimmomatic[240] and filtered reads were aligned to the hg38 genome with bowtie2[241]. After removal of unmapped and duplicate reads using Picard[242], peaks were called using MACS2[243] with a threshold qvalue < 0.01. The pooled IgG sample was used as the control in peak calling. The global settings in the pipeline were as described: https://github.com/fl-yu/CUT-RUNTools-2.0, aside from organism_build=hg38. The parameters for the individual software called by the pipeline were also left unchanged. Bigwig files for each sample were created using deeptools[244] and visualized on the UCSC genome browser. Gene ontology analysis was performed through GREAT[195] using the basal-plus extension to define putative target genes of FHL5 peak set, with the whole genome used as background. Known motif enrichment in FHL5, CREB, and IgG binding sites was determined in HOMER (v4.11) using the findMotifsGenome.pl command. FHL5 peaks were annotated according to the nearest protein-coding gene in CHIPseeker[193]. Differential expression analysis and FHL5 biding site data was integrated using the Binding and Expression Target Analysis (BETA) software package[194]. BETA calculates a rank sum product score to reflect the likelihood of direct transcriptional regulation by incorporating both DEGs and epigenomic profiles. We used the following thresholds for expression changes: log2 Fold Change > 0.6, and for binding: FDR < 0.05 and Q < 0.01

and signal value > 5. We considered the BETA rank sum product < 1E-3 as evidence of direct transcriptional regulation.

GWAS SNP enrichment

We used the GREGOR software tool[197] to determine the enrichment of GWAS risk variants in FHL5, CREB, and IgG binding sites. GWAS summary statistics were filtered to include SNPs P<5E-5. This list of suggestive SNPs were pruned using PLINK (v1.9) to retain the most significant SNPs with pairwise LD (r2) threshold < 0.2 in the 1000G European reference panel. We then used the default GREGOR parameters as described in https://genome.sph.umich.edu/wiki/GREGOR.

STARNET analyses

eQTL analysis

The STARNET cohort and datasets were described previously[94,107]. The STARNET cohort includes 600 individuals with CAD and 250 control samples. RNA-seq libraries from cardiometabolic tissues (aorta, mammary artery, liver, subcutaneous fat, visceral fat, blood, and skeletal muscle) were prepared using the polyA and Ribo-Zero library preparation protocols. 50-100bp single-end sequencing was performed using the Illumina HiSeq sequencer to a depth of 20-30million reads per sample. *Cis*-eQTLs and *trans*-eQTLs were identified using the Matrix QTL package.

Gene regulatory network analysis

Normalized gene expression for all STARNET tissues were used to construct co-expression modules using block-wise Weighted Gene Co-expression Network Analysis (WGCNA). The modules were annotated based on gene ontology enrichment analyses using the Fisher's exact test employed in the WGCNA package. Phenotypic correlations with clinical traits were calculated by aggregating Pearson correlation *P*-values for genes in the module by Fisher's method. Gene regulatory networks were inferred using the GENIE3 package with edges constrained by eQTL genes and transcription factor annotations. Key driver analyses were performed in the

Mergenomics R package. The co-expression and gene regulatory network data can be accessed at http://starnet.mssm.edu/.

Coronary artery gene regulatory network analysis

WGCNA

Total RNA-seq data was filtered to exclude genes not present in greater than 80% of the samples. Hierarchical clustering was used to identify sample outliers (UVA047, UVA125). The remaining 15,720 genes and 148 samples were used as input into WGCNA[184] to detect gene modules. WGCNA calculates the co-expression of genes through an adjacency matrix based on co-expression similarity between the i-th gene and the j-th gene. Hierarchical clustering of these gene co-expression values was used to determine gene modules. Module detection was performed multiple times using iterativeWGCNA[185] to prune poorly fitting genes and generate more robust gene modules. Genes not assigned to any of the modules were designated to the gray module. Coexpression modules were visualized using Cytoscape.

Bayesian network construction

Bayesian networks for the *FHL5*-containing module were constructed using RIMBANET (Reconstructing Integrative Molecular Bayesian Networks)[245]. STARNET aortic tissue eQTL data were also used as genetic priors such that genes with *cis*-eQTLs are allowed to be parent nodes of genes with coincident trans eQTLs, but not vice versa. One thousand Bayesian networks were reconstructed using different starting random seeds. Edges that appeared in greater than 30% of the networks were used to define a consensus network.

Key driver analysis

Given a set of genes (G) and directed gene network (N), Key driver analysis (KDA) generates a subnetwork, N_G, defined as the set of nodes in N that are no more than h-layers from the nodes of G and subsequently computes the size of the h-layer neighborhood (HLN) for each node. The key driver score increases if the HLN was greater than $\mu + \sigma(\mu)$, where μ is average size of HLN. Total key driver scores for each node were computed by summing all scores at each h-layer scaled according to h, Σ

n=1h(1/n). Genes with key driver scores in the top 5% were annotated as module key drivers.

Statistical analyses

Data in bar graphs are presented with mean+/- standard error of mean (SEM) with each point represented as an individual replicate. Data in box plots are presented with lines denoting the 25th, median and 75th percentile with each point representing an individual donor. Pairwise comparisons were made using the student's t test or Wilcoxon rank test as appropriate. Comparisons between more than two groups were assessed using a one-way ANOVA test or Kruskal-Wallis test. The normality of the data was assessed using the Shapiro-Wilkes test, with $P > 0.05$, supporting a normal distribution. For each of these analyses, we considered $P < 0.05$ as significant.

To identify differentially expressed genes, we used the false discovery rate (FDR) adjusted $P< 0.05$ threshold and log2FoldChange > 0.6. Heatmaps were created using the pheatmap package and represent normalized expression (Z-score) for genes, scaled across each row. Gene ontology enrichment analyses were performed relative to all expressed genes using Fisher's Exact Test, with a significant threshold of 5% FDR.

Acknowledgements

We thank Dr. Pat Pramoonjago and Sheri VanHoose and members of the histology core facilities for tissue sectioning and histological staining assistance and Dr. Katia Sol-Church and members of the Genomics Core facility for assistance with library construction and sequencing. We also thank Dr. R. Kirk Riemer and Dr. Xiaoyuan Ma at Stanford University for helpful discussions.

Sources of Funding

This work was supported by grants from: the National Institutes of Health (R01HL148239 and R00HL125912 to C.L.M.; F31 HL156463 to D.W.) and the Fondation Leducq ('PlaqOmics' 18CVD02 to J.LM.B. and C.L.M.)

Author Contributions

C.L.M supervised research primarily related to the study. R.M., J.LM.B., J.C.K., M.C., S.K.S. and A.M. supervised research secondarily related to the study. C.L.M. and D.W. conceived and designed the experiments. D.W., G.A., C.L.L.C, A.W.T., Y.C., M.Ku., C.J.D. and M.P. performed the experiments. D.W. and C.Y. performed the statistical analyses. D.W., G.A., Li.M., R.N.P., R.A., J.V.M., M.Ku. and M.D.K. analyzed the data. M.Ka., P.P., Lj.M., U.H., J.C.K. and J.LM.B. contributed reagents/materials/analysis tools. D.W., G.A. and C.L.M. wrote the paper.

Disclosures

Johan Björkegren is a shareholder in Clinical Gene Network AB and has an invested interest in STARNET. Jason Kovacic is the recipient of an Agilent Thought Leader Award (January 2022), which includes funding for research that is unrelated to the current manuscript. All other authors declare that they have no competing interests relevant to the contents of this paper to disclose.

Data Availability

All raw and processed CUT&RUN and RNA-seq datasets are made available on the Gene Expression Omnibus (GEO) database (accession: GSE201572).The following publicly available datasets were used in this study: human coronary artery scATAC-seq generated by Turner et al. (accession: GSE175621), bulk human coronary ATAC-seq generated by Miller et al. (accession: GSE72696), bulk left ventricle and liver tissue ATAC-seq generated by the ENCODE project (https://www.encodeproject.org/). Bulk RNA-seq fastq files for human coronary artery SMCs (SRR705828, SRR705829, SRR706406) were downloaded from sequence read archive (SRA) using sratools. ENCODE bulk RNA-seq fastq files for HUVECs (ENCFF000DUK, ENCFF000DUH), HEPG2 (ENCFF982FAM, ENCFF564BSM) and K562 (ENCFF104ZSG, ENCFF695XOC) cells were downloaded from the ENCODE project (https://www.encodeproject.org/). Processed scRNA-seq data for human coronary arteries generated by Wirka et al. and carotid arteries generated by Alsaigh et al. are accessible at the PlaqView single-cell data portal (https://www.plaqview.com). Normalized gene expression levels and expression quantitative trait loci (eQTL) data

are available at the Genotype-Tissue Expression (GTEx) portal website (https://www.gtexportal.org). Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task (STARNET) eQTL data and GRNs are available at http://starnet.mssm.edu. Summary statistics and gene annotations for cardiometabolic GWAS (hypertension, diastolic blood pressure, CAD, systolic blood pressure, BMI, and pulse pressure) were accessed through the Cardiovascular Disease Knowledge Portal (https://cvd.hugeamp.org/).

Code Availability

All custom scripts used are available at https://github.com/MillerLab-CPHG/FHL5_Manuscript. All software tools used in this study are publicly available and full names and versions are provided in the reporting summary.

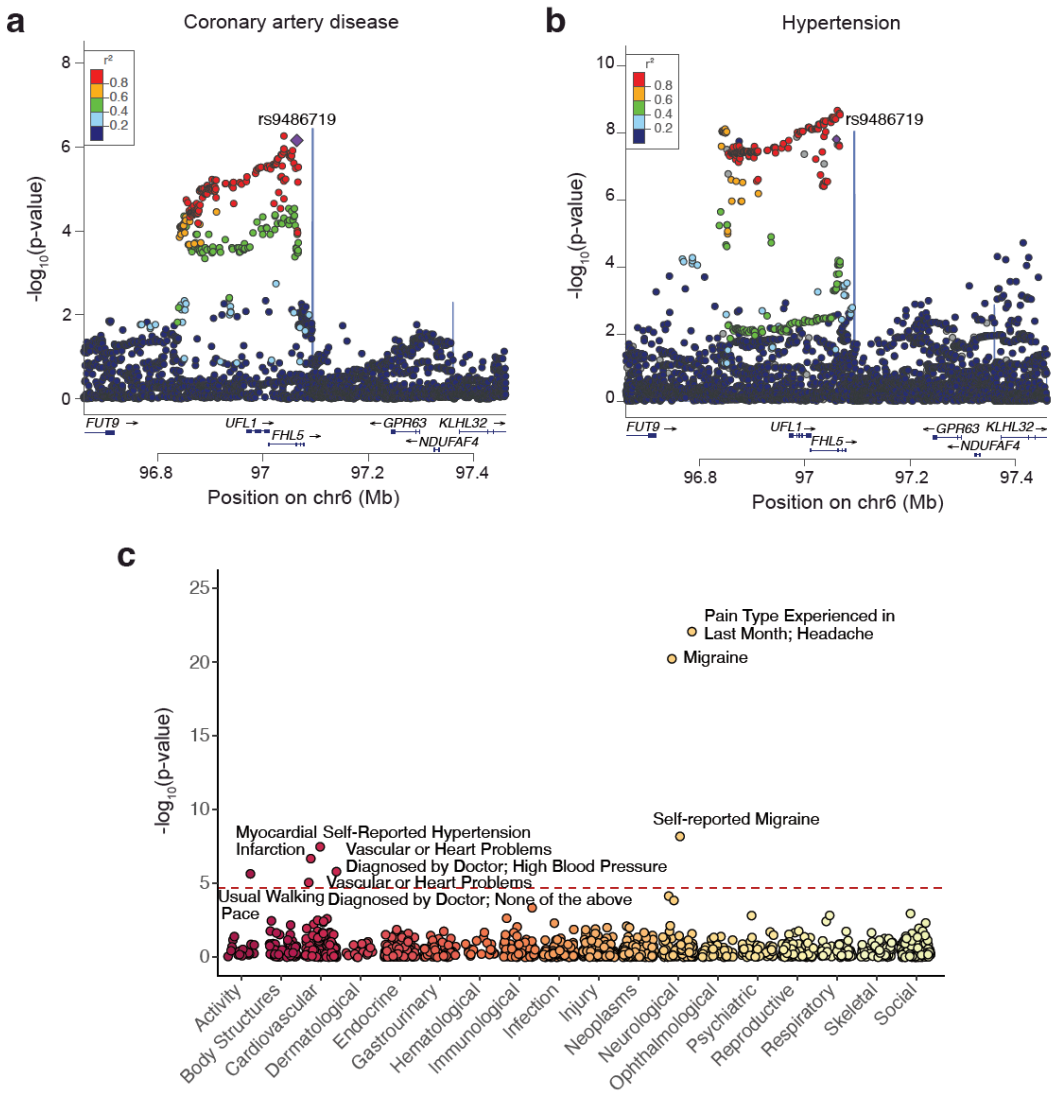

Fig. S2.1. UFL1-FHL5 locus is associated with vascular diseases.

LocusZoom plot highlighting the genetic association of the UFL1-FHL5 locus with coronary artery disease (CAD) (a) as determined from the CARDIoGRAMplusC4D and UK Biobank meta-analysis, and hypertension (b) as determined from the UK Biobank. Data was Previously Published. Figure Created by DW. (c) Genetic association of rs9486719 across multiple phenotypes using GWAS studies reported in Phenoscanner, and traits organized in physiological categories on x-axis. Analysis by DW.

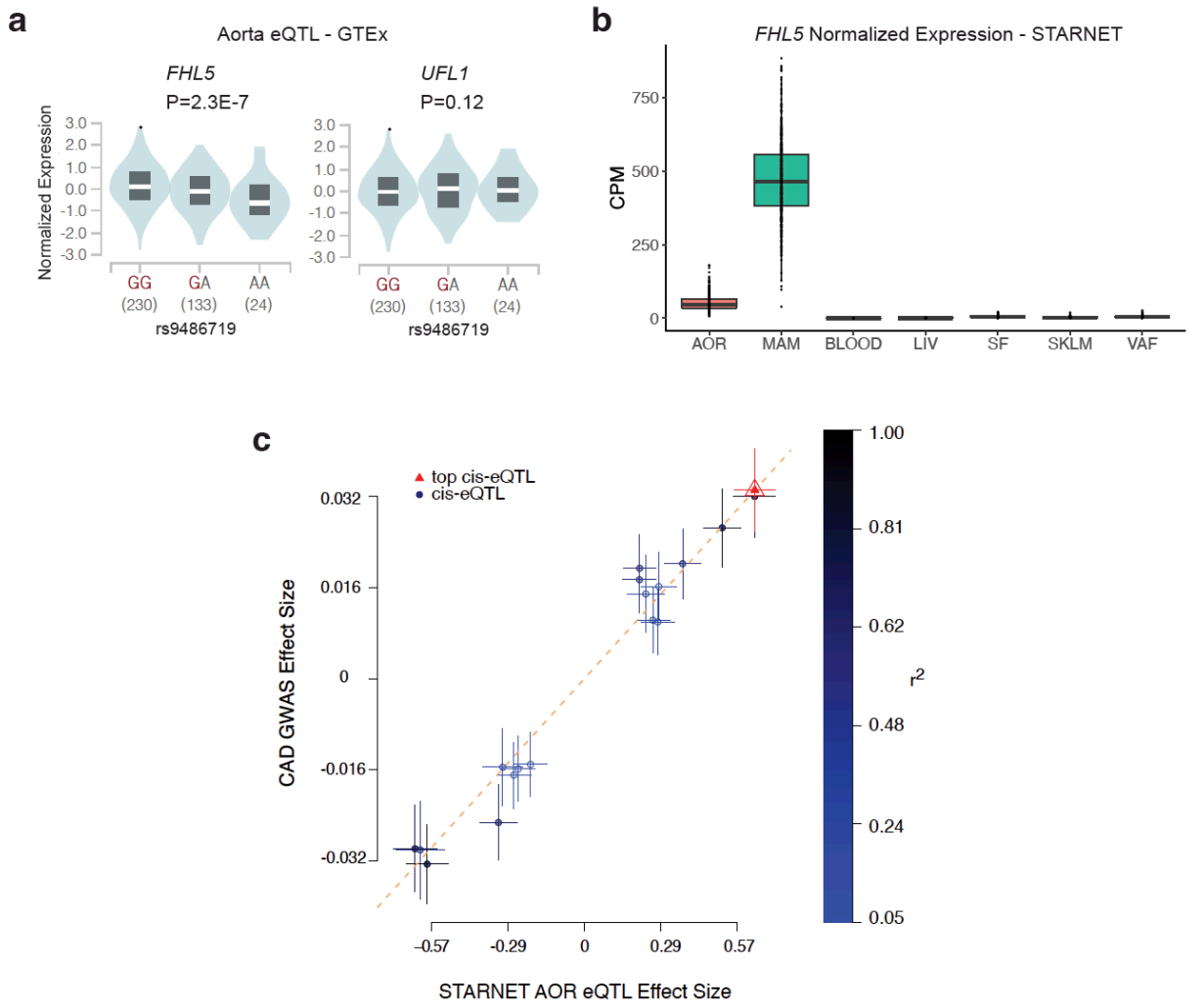

Fig. S2.2. FHL5 is the top candidate causal gene at the CAD/MI *ULF1-FHL5* locus.

(a) Violin plots showing the association of rs9486719 with FHL5 gene expression (left) and lack of association with UFL1 gene expression (right) in GTEx (Aorta). Data was Previously Published. Figure Created by DW. (b) Normalized FHL5 gene expression in counts per million (CPM) across STARNET cardiometabolic tissues: AOR: atherosclerotic aorta, MAM: mammary artery, LIV: liver, SF: subcutaneous fat, SKLM: skeletal muscle: VAF: visceral adipose fat. Data was Previously Published. Figure Created by DW. (c) Correlation of the effect size of MI risk alleles with FHL5 eQTL effect size in STARNET AOR. Color scale depicts Pearson's r-squared values for the correlation of individual alleles. Analysis Done by DW.

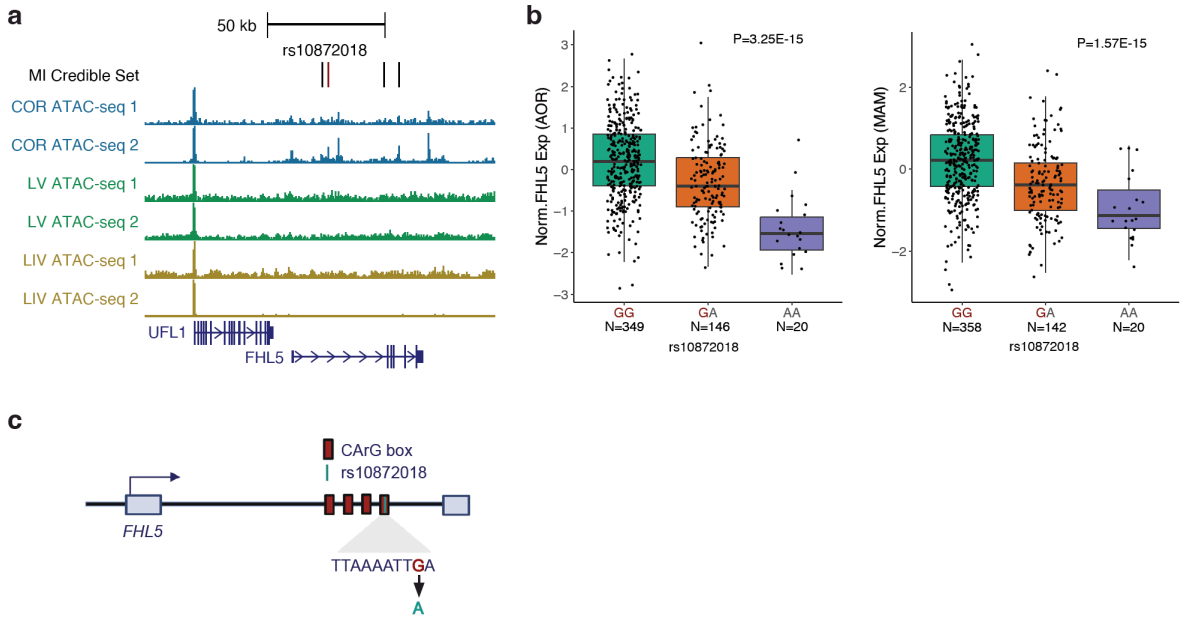

Fig. S2.3. rs10872018 is the top candidate causal variant at underlying the MI *UFL1-FHL5* locus

(a) UCSC genome browser screenshot showing the overlap of *FHL5* MI GWAS credible set of variants with ATAC-seq peaks in coronary artery (COR), left ventricle (LV), and liver (LIV), from two independent donors per tissue. Top candidate SNP rs10872018 highlighted in red. Analysis Done by DW. (b) Association of rs10872018 with *FHL5* gene expression in STARNET atherosclerotic aorta (AOR) (left) and mammary artery (MAM) (right). Data Previously Published. Figure Created by DW. (c) Schematic highlighting the 4 CArG boxes in intron 1 of *FHL5*, with the last motif disrupted by rs10872018-A. Created at Biorender.com. Figure Created by DW.

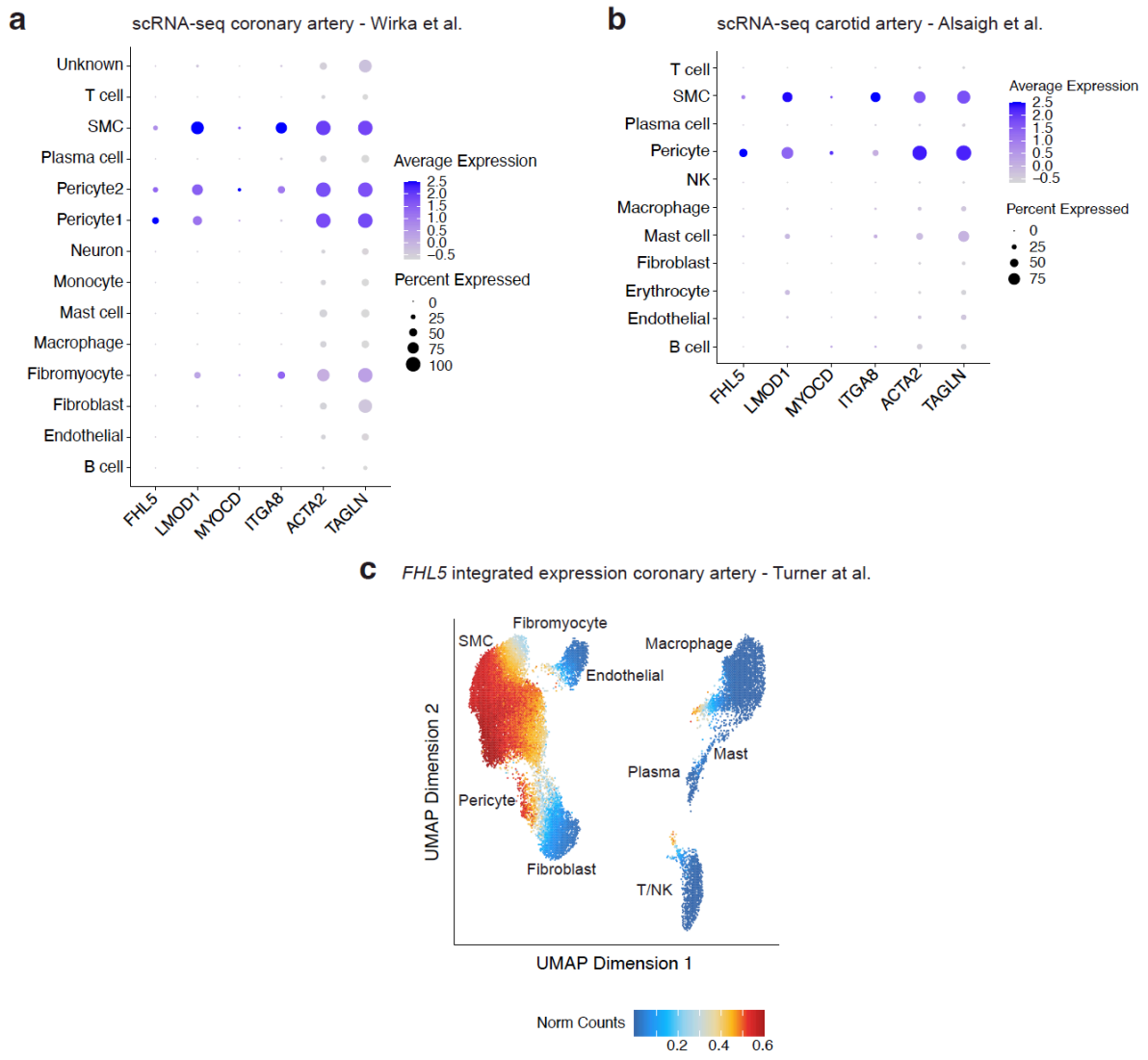

Fig. S2.4. *FHL5* gene expression is enriched in SMC and pericytes in human coronary arteries. (a) Dot plot showing average expression and percentage of cells expressing *FHL5* and other SMC markers in scRNAseq of subclinical atherosclerotic human coronary arteries (Wirka et al.). Data was Previously Published. Figure Created by DW. (b) Dot plot showing average expression and percentage of cells expressing FHL5 and other markers in scRNAseq of healthy and atherosclerotic human carotid arteries (Alsaigh et al.). Data was Previously Published. Figure Created by DW. (c) UMAP plot showing *FHL5* integrated expression in SMC and pericytes using integrated coronary artery snATAC-seq and scRNA-seq dataset from Turner et al. Clusters are labeled by the main cell-type annotations after integration of the two datasets. Data was Previously Published. Figure Created by AT and JVM.

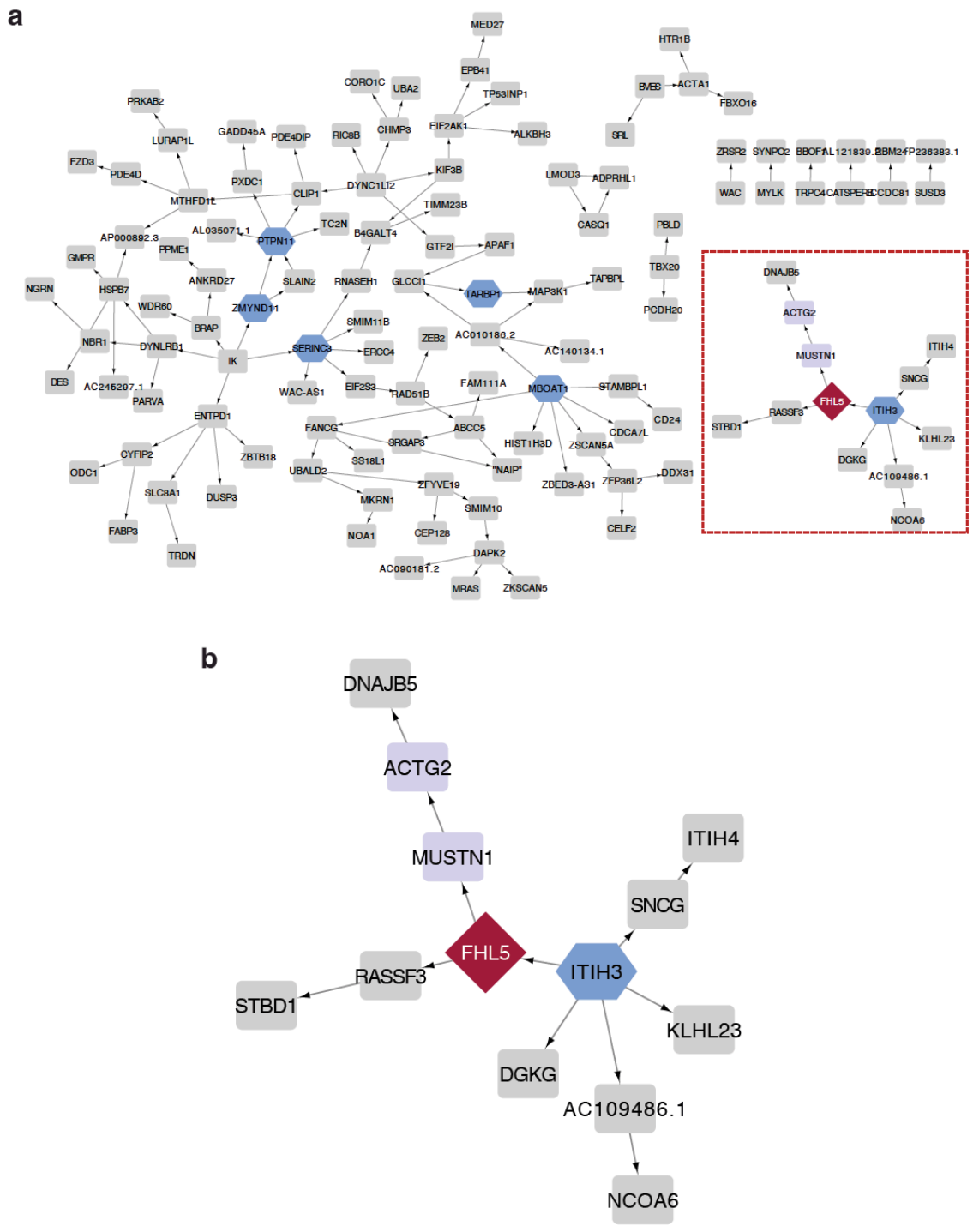

Fig. S2.5. *FHL5* gene regulatory network (GRN) in human coronary arteries.

(a) Directed network depicting the *FHL5* GRN in human coronary arteries. The key drivers in the network are highlighted in blue hexagons. The dotted red box denotes the *FHL5* subnetwork. Analysis Done by RNP. Figure Created by DW. (b) Zoomed in view of the *FHL5* subnetwork. The key driver of this subnetwork, *ITIH3* is depicted in the blue hexagon. Downstream FHL5 genes that function in SMC contraction biological process are highlighted in the purple rectangles. Figure Created by DW. (b

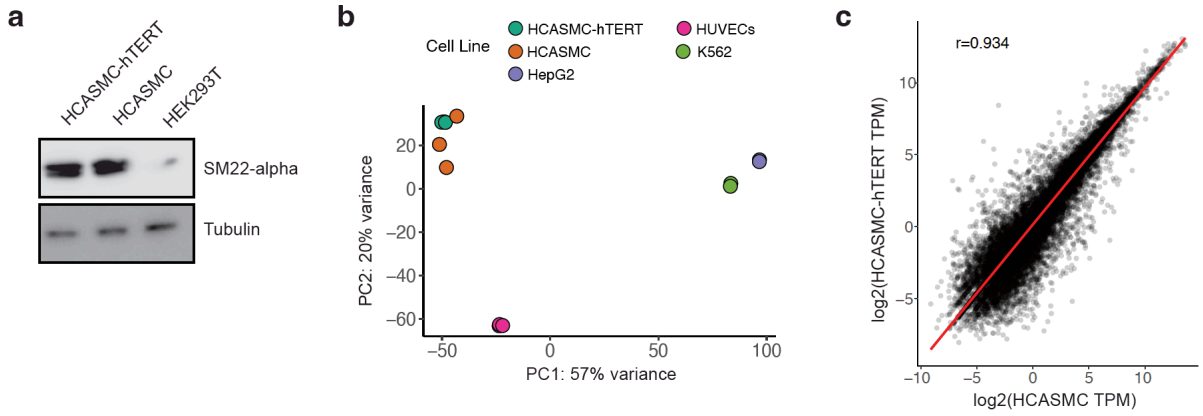

Fig. S2.6. Characterization of HCASMC-hTERT.

(a) Western blot showing expression levels of SMC marker, SM22- alpha, in HCASMC-hTERT and the parental HCASMC (Cell Applications #2105). Tubulin is shown as a loading control. Experiment Done by DW. (b) Principal component analysis (PCA) showing similarity in the transcriptomes of HCASMC-hTERT and primary HCASMC relative to HEPG2, HUVEC, and K562 cells. Analysis Done by DW. (c) Pearson correlation of HCASMC-hTERT and parental primary HCASMC transcriptomes shown as log2 normalized gene expression (transcripts per million (TPM)). Analysis Done by DW.

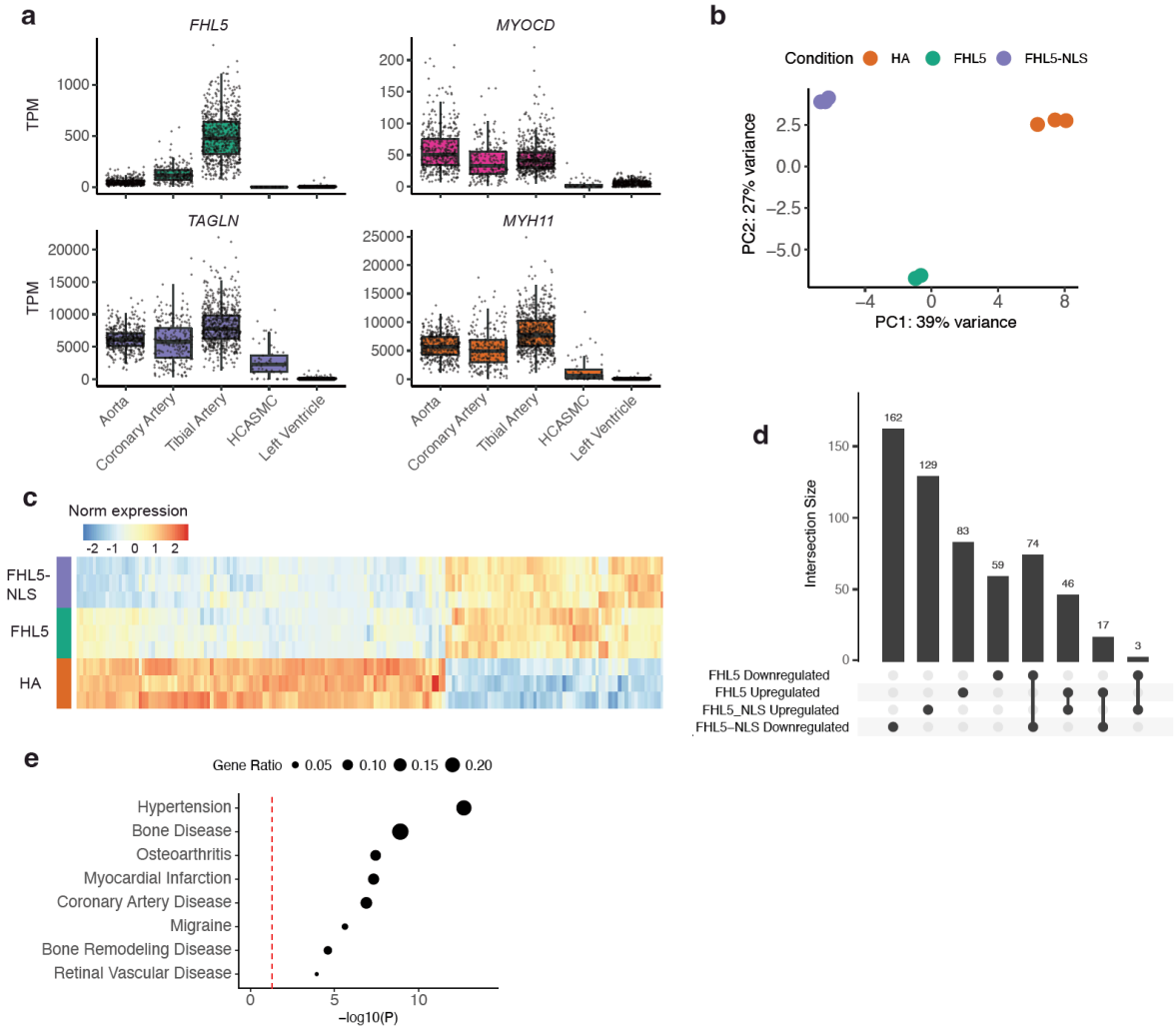

Fig. S2.7. Comparison of FHL5 and FHL5-NLS differentially expressed genes.

(a) Normalized expression level in transcripts per million (TPM) of *FHL5* and SMC markers in human aorta (N=432), coronary artery (N=240), tibial artery (N=663), left ventricle (N=432) and HCASMC (N=61). (b) PCA of normalized HA, FHL5, and FHL5-NLS using RNAseq transcriptomes (n=3 per group). Data Previously Published. Figure Created by DW. (c) Heatmap showing log2 normalized expression of 168 DEGs derived from the union of FHL5 and FHL5-NLS significantly upregulated and downregulated DEGs and clustered using n=3 biological replicates per group. Analysis Done by DW. (d) Upset plot showing overlap between FHL5 and FHL5-NLS upregulated and downregulated genes. Analysis Done by DW. (e) Disease enrichment analysis of the FHL5 DEGs showing enrichment of disease ontology (DO) terms from 381 FHL5 DEGs against the whole transcriptome as a background. Analysis Done by DW.

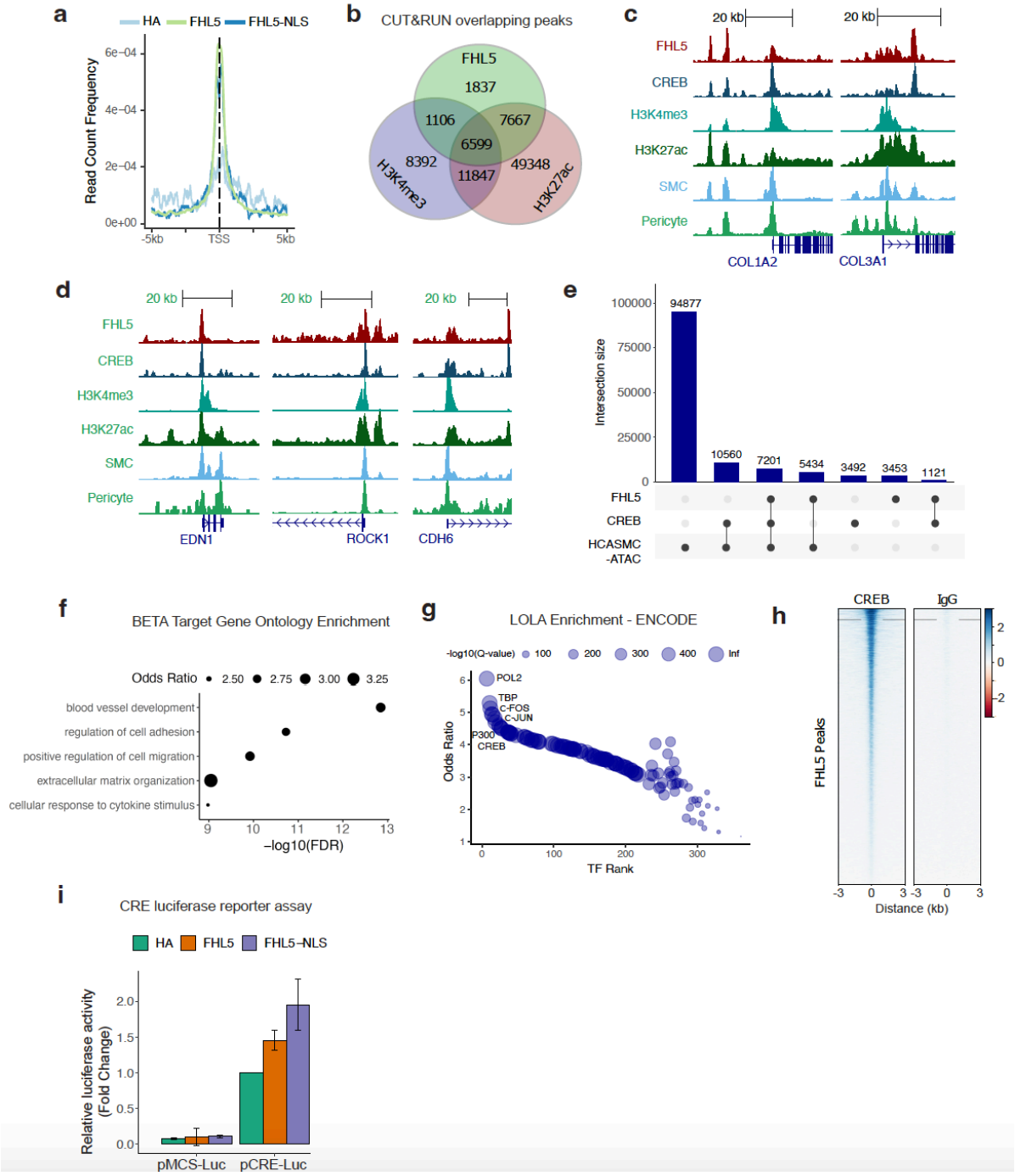

Fig. S2.8. FHL5 interacts with CREB to transcriptionally regulate extracellular matrix and cell adhesion genes.
(a) Density plot showing similar enrichment profiles of FHL5 (light green) and FHL5-NLS (dark blue) CUT&RUN peaks centered on transcription start sites (TSS). Analysis performed by DW. (b) Overlap of FHL5, H3K27ac, and H3K4me3 CUT&RUN peaks in SMCs. Analysis performed by DW. (c) UCSC genome browser tracks showing FHL5 and CREB binding sites near the promoters of extracellular matrix genes, *COL1A2* and *COL3A1*. Analysis performed by DW. (d) UCSC genome browser tracks showing FHL5 and CREB binding sites near the promoters of cell adhesion genes, *EDN1*, *ROCK1*, and *CDH6*.

Analysis performed by DW. (e) Gene ontology (Biological Process) enrichment of FHL5 target genes identified from BETA analysis. Analysis performed by DW. (f) Overlap of FHL5 and CREB bidding sites with primary coronary artery SMC ATAC-seq peaks. (g) LOLA enrichment of FHL5 binding sites in ENOCDE transcription factor (TF) ChIP-seq datasets highlighting AP-1 family TFs, CREB, and transcriptional machinery proteins, RNA polymerase II (POL2) and TATA Binding Protein (TBP), and cofactor, p300. Analysis performed by DW. (h) Heatmap visualization of the distance of CREB binding sites to the center of FHL5 binding sites, compared with non-specific IgG binding sites. Analysis performed by DW. (i) Relative fold change in CRE luciferase reporter activity after FHL5 overexpression. Analysis performed by DW.

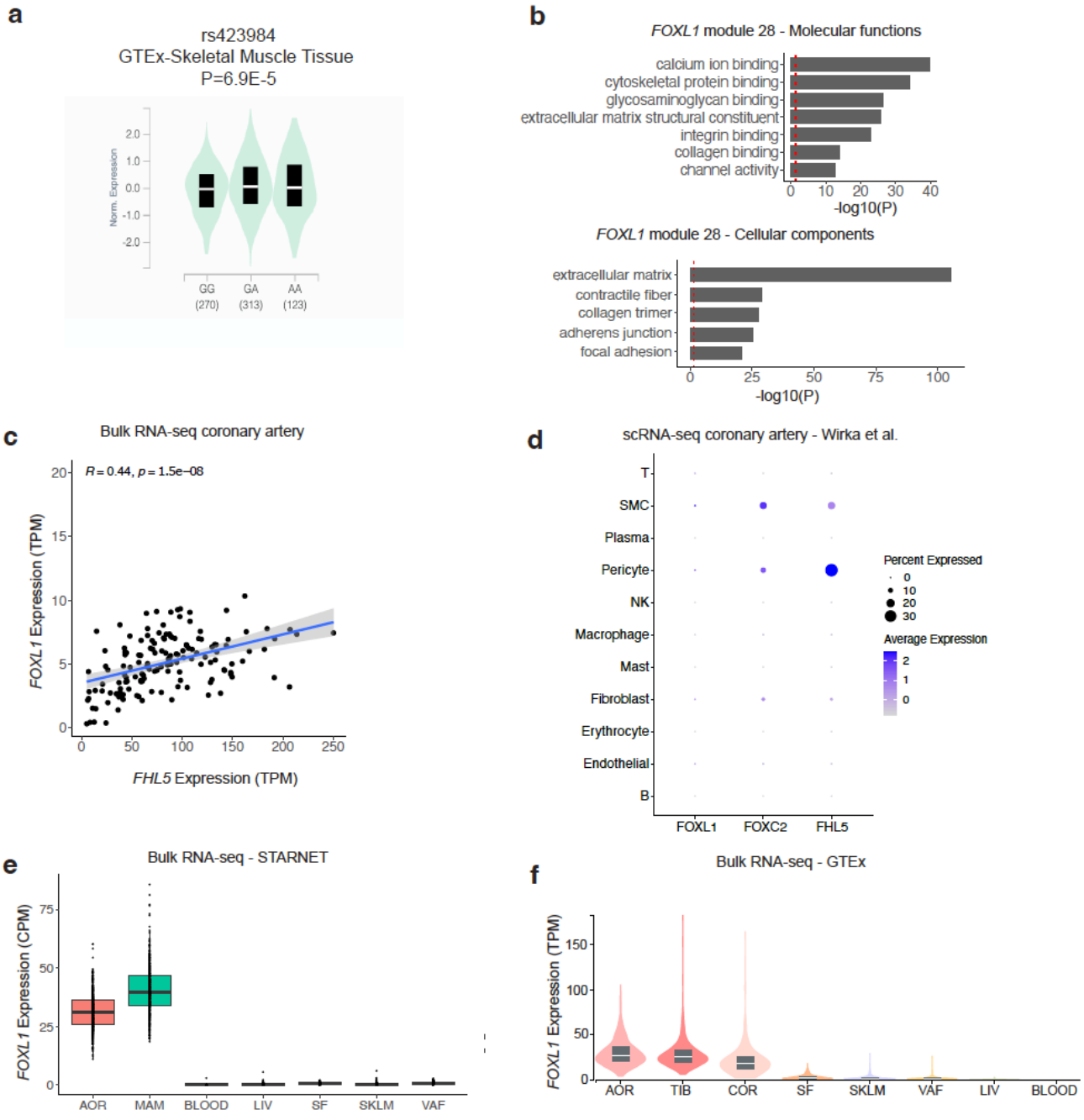

Fig. S2.9. FHL5 regulation of *FOXL1* may contribute to the underlying mechanism of its association with CAD/MI.

(a) Association of rs423984 with FOXL1 gene expression in skeletal muscle tissue (GTEx-v8) (b, upper) Top molecular functions enriched in module 28. Analysis Done by DW. (b, lower) Top cell compartments enriched in module 28 genes. The red dotted line corresponds to a nominal threshold of FDR<0.05. Data Previously Published. Figure Created by DW. (c) Pearson correlation of *FHL5* and *FOXL1* gene expression from bulk RNA-seq of human coronary arteries (N=148). Analysis Done by DW. (d) Dot plot showing enrichment of *FOXL1*, *FOXC2* and *FHL5* gene expression in SMC and pericytes in human coronary artery single-cell RNA-seq (Wirka et al.). Data Previously Published. Figure Created by CLM. (e)

Normalized expression level in counts per million (CPM) of *FOXL1* in cardiometabolic tissues profiled in STARNET bulk RNA-seq analysis. Data Previously Published. Figure Created by DW. (f) Normalized expression level in transcripts per million (TPM) of *FOXL1* in bulk RNA-seq of GTEx cardiometabolic tissues. AOR: aorta, MAM: mammary artery, BLOOD: whole blood, LIV: liver, SF: subcutaneuous adipose fat, SKLM: skeletal muscle, VAF: visceral adipose fat, TIB: tibial artery, COR: coronary artery. Data Previously Published. Figure Created by CLM.

Chapter 3: Discussion and Future Directions

Summary:

Throughout the course of this project, we aimed to uncover the molecular mechanisms linking the pleiotropic *UFL1-FHL5* locus with multiple common vascular diseases and traits, including CAD/MI, hypertension, and coronary artery calcium. Comprehensive characterization of the molecular basis of these genetic associations highlights overlapping processes that may be amenable for therapeutic intervention for co-occurring diseases. We used a multimodal approach integrating human genetics, genomics, and *in vitro* studies to implicate *FHL5* as the top candidate causal gene contributing to common vascular disease risk.

In chapter 2, we elucidated the upstream mechanisms regulating *FHL5* gene expression that is perturbed by the CAD/MI risk alleles. We nominated rs10872018, a top *FHL5* eQTL, as the top candidate causal variant, with the alternate allele impairing transactivation by the SRF/MYOCD transcriptional complex.

We also characterize the downstream function and transcriptional role of *FHL5* in the vessel wall. We propose that *FHL5* regulates a network of downstream genes in SMCs that participate in adverse vascular remodeling events that are critical for deposition of calcification in the vessel wall. We identify downstream interactions at the CAD/MI loci, *FN1* and *FOXL1* that have putative roles in vascular remodeling. These gene regulatory effects may magnify the modest changes in *FHL5* expression and account for a portion of the missing 'hidden' heritability in the vessel wall.

Genetic Association of *FHL5* with Migraine

Although this story begins with genetic association of *UFL1-FHL5* with CAD/MI, this locus did not surpass the genome-wide significance threshold until 2021 in the combined UK Biobank and CARDIoGRAMplusC4D consortium meta-analysis [86]. Our interest in this locus was not piqued by this suggestive association alone but also by *UFL1-FHL5* also harboring variants associated with hypertension [162], blood pressure [169],

migraine [201], and aneurysm [163]. Although pleiotropy is common across GWAS SNPs with an estimated 90% of SNPs associated with more than one ICD-10 code [246], the specific enrichment of vascular disease genetic associations caught our attention. This collection of genetic associations mirrored those of the well-established *PHACTR1/EDN1* locus [165]. We observed the same paradoxical associations for *FHL5*, where the direction of effect in regards to CAD/MI risk [86] was not concordant with the direction of effect for hypertension [162], blood pressure [169], and fibromuscular dysplasia [247], albeit the last below genome-wide significance threshold (P=2E-05). We focus on the contribution of *FHL5* to atherosclerotic diseases and intermediate phenotypes, leaving questions regarding non-atherosclerotic vascular diseases outstanding.

The molecular mechanisms linking *FHL5* function to migraine risk is especially intriguing, given that migraines are thought to be a neurological based abnormality. Although migraines is a leading cause of 'years lived with a disability' metric [248 249], the etiology of this complex disease is poorly characterized with much debate regarding its primary origin in vascular or neuronal dysfunction [250]. Migraine GWAS, which have uncovered over 100 loci, support the vascular contribution with genetic signals enriched near active genes in artery tissues and SMCs in culture [201 251]. One hypothesis reconciling these observations is the causative link between the vasodilation of the cerebral arteries and migraine onset, as hemodynamic changes in blood flow often precedes migraine attacks [252 253]. This hypothesis is further supported by the use of the monoclonal antibodies targeting the calcitonin gene-related peptide (CGRP) or the CGRP receptor as a prophylactic treatment to reduce the prevalence and severity of migraines [254 255]. CGRP is involved in the multiple pathophysiological functions, among which includes vascular tone effects. It is noteworthy to mention that this mechanism has human genetics support, as the loci harboring the genes encoding the CGRP subunits (*CALCA* and *CALCB*) surpassed the genome wide significance threshold in the most recent migraine GWAS meta-analysis [201].

Based on our in vitro experiments and *in vivo* coexpression analysis, we suspect that FHL5 increases vascular tone through increased SMC contractility, which may mitigate

the degree of arterial dilation. Consistent with this potential contribution, the migraine risk allele is associated with reduced *FHL5* gene expression, suggesting a protective effect. The exacerbation of migraine symptoms through CGRP related signaling is another interesting angle to consider in regards to *FHL5* function. CGRP functions to induce SMC relaxation through a cyclic adenosine monophosphate dependent pathway that likely converges on CREB transcriptional activation [256] . As we and others [157] show that *FHL5* functions as a transcriptional cofactor for CREB, *FHL5* may mediate some of the transcriptional effects of CGRP signaling. A clinical trial (NCT01924052) assessed whether individuals harboring the migraine risk variant, rs13208321 near *FHL5*, affected the susceptibility or severity of CGRP-induced migraine attacks. Although the results were not statistically significant, with increasing sample sizes, these epistatic interactions contributing to migraine risk could be significant and help guide CGRP antagonist-based therapies.

Contribution of *FHL5* to Congenital Heart Disease Pathology

In addition to contributing to the heritability of a handful of common vascular diseases, *FHL5* may also contribute to the pathology of major aorto-pulmonary collateral arteries (MAPCA), given its dramatic upregulation in these vessels [257]. MAPCAs are tortuous collateral vessels present in some infants with a severe form of tetralogy of fallot (TOF). Afflicted patients use MAPCAs to compensate for the hypoplastic or nonexistent pulmonary vascular system [258] [259]. However, MAPCAs are not suited to withstand the increased shear stress from blood flow after birth and often become stenotic over time [260] [261]. Surgical intervention to construct a new pulmonary vascular system from existing MAPCAs is the standard of care for these patients[262] [263].

Relative to pulmonary arteries isolated from infants without TOF, MAPCAs upregulate genes related to extracellular matrix, contraction and cell differentiation. *FHL5* was the most upregulated gene in this differential expression analysis [257]. Although there is no evidence yet supporting the association of *FHL5* genetic variants with TOF or the MAPCA phenotype, the upregulation of *FHL5* in these vessels is striking. Considering the epigenomic and transcriptomic data in SMCs presented in chapter 2, *FHL5* likely

contributes to these vessel wall structural changes. However, it is unclear if *FHL5* functions upstream to orchestrate the gene expression program to form these collateral vessels or functions downstream as a consequence of the increased shear stress to drive SMC hypertrophy and vessel stenosis.

One approach to overcome the lack of suitable mouse models recapitulating the MAPCA phenotype is to perform transcriptomic profiling on human MAPCA tissues directly. Single cell transcriptomic analyses of MAPCAs will provide a global view of the different cell types and cell states in these tissues and enable comparisons with pulmonary arteries from age-matched healthy infants. Further, to investigate *FHL5*-dependent mechanisms of disease in particular, differential expression analysis between SMCs expressing high and low levels of *FHL5* can be performed to help prioritize target genes and pathways that contribute to disease pathology. Since classic *in vitro* and *in vivo* perturbation experiments face clear limitations, the construction of gene regulatory networks from scRNA-seq data can be used to identify functions of uncharacterized genes in an unbiased manner [264] [265] [266]. Gene ontology analysis of the *FHL5*-containing module will provide insights into its function in the SMCs of the most disease-relevant tissues.

To address questions regarding the upstream transcriptional regulators driving different SMC states in the MAPCAs, the SCENIC pipeline can be run. [264] [267] [266]. This well-established pipeline has been used to define candidate transcription factors driving gene programs within immune cells of COVID patients [268], tumor microenvironment [269], and liver cirrhosis [270]. The identification of disease relevant regulons and upstream drivers of SMC cell states linked to MAPCAs will generate new hypotheses that can be validated using genetic and pharmacological perturbations in TOF patient derived iPSCs or vascular organoid models.

Defining the *FHL5 Interactome*

Throughout the course of this project, we highlight contributions of *FHL5* to CAD/MI heritability through gene regulation. However, we would be remiss to not consider mechanisms independent of transcription in these future directions, especially given the

presence of LIM domains that often function as scaffolds to facilitate the formation of large complexes [150] [271] [140]. Other members of the FHL family, which share structural similarities, have crucial roles in myoblast and cardiomyocytes. FHL1 [272] [273] and FHL3 [224] function as scaffolds to regulate sarcomere assembly, loss of which is associated with myopathies [274] [152]. FHL2 regulates the clustering of focal adhesion complexes in response to matrix stiffness [146]. In line with these functions in the cytoplasm of its family members, we speculate that FHL5 may mediate similar interactions with the actin cytoskeleton and/or focal adhesions, a function which may contribute to the increased contractility we observed *in vitro*. Consistent with this hypothesis as well as its role in transcription, we observed both cytoplasmic and nuclear localization in human coronary arteries. Unexpectedly, we also observed perinuclear localization in the same tissue samples, which may reflect involvement in nuclear organization, perhaps through interactions with the Linker of Nucleoskeleton and Cytoskeleton (LINC) complex. This complex serves as a bridge between the nuclear envelope and cytoskeleton to receive and transduce mechanical signals at the plasma membrane. It is possible that FHL5 localizes to the nuclear envelope to facilitate interactions with the LINC complex and/or actin cytoskeleton, given its predicted cytoskeletal interactions. Previous work has shown actin and its adapter proteins to be involved in genomic compartment organization and RNA polymerase II clustering to regulate gene expression [275] [276] [277]. As we only observed ~50% of differentially expressed genes harboring FHL5 binding sites, it is tempting to speculate whether some of these indirect effects are attributed to effects on 3D genome organization. Further work is necessary to characterize these perinuclear interactions.

High throughput proteomic screens in yeast hint at potential roles in focal adhesion signaling and cytoskeletal rearrangements [278] [279]. FHL5 was reported to interact with other LIM domain proteins, such as PDLIM7, a SMC marker gene localized to actin stress fibers. Similar to the function of FHL5 in SMC contraction, loss of PDLIM7 downregulated the expression of surrogate contraction markers, *LMOD1* and *PLN* [280]. Since these proteomic screens lack physiological context, constructing a more comprehensive view of this interactome in SMCs would reveal additional insights. As it

is a unifying function of proteins harboring LIM domains, we anticipate that FHL5 functions as the hub central to the organization of the actin cytoskeleton in SMCs. Recent advances in the proximity labeling techniques using the biotin ligase-based suite of tools (e.g. BioID [281] [282], TurboID [283]) enables the detection of transient protein interactions that may be more reflective of physiological dynamics. In order to maximize the interpretability of the interactome, we can create a stable SMC cell line overexpressing the TurboID-FHL5 fusion protein and perform these experiments under both basal and osteogenic conditions. Characterizing the subcellular FHL5 proximity interactome will validate FHL5 interactions with its dominant binding partners that we presented in chapter 2 and reveal other transcription factor binding partners with less defined recognition motifs. Similarly, identifying the FHL5 network of interaction in the cytoplasm and subsequent gene ontology enrichment analysis will provide unbiased support for possible roles in mechanotransduction and focal adhesion signaling. Finally, integration of this proximity interactome with vascular disease GWAS candidate genes may also generate hypotheses regarding the mechanism of uncharacterized candidate genes. Similar analyses deciphering the interactome of the hallmark congenital heart disease transcription factors, *GATA4* and *TBX2* revealed an enrichment of genes harboring disease *de novo* variants [284].

FHL5-Based Polygenic Risk Score

In this project, we characterized the *FHL5* gene regulatory network in SMCs, which revealed regulatory interactions with downstream CAD loci to modulate disease risk. In particular, we highlight two downstream CAD and CAC loci, *FN1* and *FOXL1* with putative functions in extracellular matrix remodeling in the vessel wall. Notably, we recapitulated the established role of *FN1* and supported a lesser-known gene in *FOXL1* in this process. To extend these findings to clinical translation, it would be interesting to assess the utility of *FHL5* target genes in the clinic through the development of a gene set CAD polygenic risk score (PRS). Gene set PRS have been developed by combining the effect of risk alleles mapping to specific gene sets, modules, or pathways [285] [286] [287]. Recent work constructed gene set PRS based on SMC cell state signature genes [288]. Although these analyses did not yet improve the predictive performance of the

traditional PRS, they highlight biological processes that account for much of the disease risk that may be exploited for therapeutic development. These few studies to date represent proof of concept studies and highlight some challenges due to limited biological insights of GWAS loci, further underscoring the significance of mechanistic follow-up studies. It would also be intriguing to estimate the amount of CAD heritability attributed to SNPs near the *FHL5* target gene set. Given the enrichment analysis of FHL5 binding sites, differentially expressed genes, and coexpression analysis presented in chapter 2, we expect that these regulatory interactions *in trans* account for a portion of the missing heritability.

In Vivo Models

One obvious future direction of this study is the development of an *in vivo* model that will enable the direct study of vascular disease in a physiologically relevant system. We attempt to address this limitation by coupling *in vitro* studies with human genetics and the epigenomics of human vascular tissues. However, future studies demonstrating the function of FHL5 directly in physiological contexts of atherosclerosis, calcific disease, or blood pressure regulation will complement the studies presented here and strengthen the causative link to these diseases.

Despite robust and specific expression in human artery tissues, *FHL5* is not expressed in the vasculature of traditional mouse models. It is noteworthy to mention that *FHL5* is robustly expressed in mouse adult testis, although loss of *Fhl5* was not associated with infertility [156 166]. This tissue specific gene expression profile may reflect the great diversity of the testis transcriptome, since germ cells are estimated to express 90% of protein coding genes at least at the transcript level [289]. Since *Fhl5* is not endogenously expressed in a cardiometabolic tissues, classic loss of function mouse studies may not be applicable. This limitation necessitates the generation of a novel SMC specific *Fhl5* overexpression model in a Myh11-CreERT2 ROSA floxed STOP eYFP Apoe−/− background [130]. Use of this mouse model under a high fat diet can validate the causal link between FHL5 and CAD, as predicted by human genetics. Similar studies using a SMC specific knock-out strategy of other CAD GWAS transcription factors (e.g.

Klf4 [63], *Tcf21* [62] , *Smad3* [290], *Zeb2* [112]) uncovered regulators of SMC phenotypic modulation using single cell transcriptomic profiling. Similarly, scRNA-seq analysis of plaques in the SMC specific *Fhl5* overexpression mouse model will address the role of FHL5 in SMC modulation under atherogenic conditions. Coupling this analysis with classic histology stains, such as von Kossa, alizarin red, and Sirius red staining will connect the pathology of the plaque with transcriptomic changes. Although the human genetics data suggest that *FHL5* increases CAD/MI risk, these studies will confirm the true direction of effect of *FHL5*, given the direct manipulation. We speculate that *FHL5* may play a detrimental role in reducing plaque stability due to increased incidences of microcalcifications in the lesion.

Given the strong association with hypertension and blood pressure regulation, *in vivo* and *ex vivo* murine studies using the SMC specific *Fhl5* overexpression mouse model described above can also address the function of *FHL5* in hypertension pathogenesis. More complex models will be needed to test this hypothesis and validate the causal effect between *FHL5* expression levels and hypertension risk. One possible mechanism contributing to its effect on blood pressure regulation is myogenic tone, a hallmark feature of resistance arteries [291] [292]. In response to increased pressure, SMCs activate voltage-gated calcium channels after depolarization to increase intracellular calcium and constrict the vessel [293]. Perturbations in cytoskeletal remodeling and ECM organization and interactions contribute to the regulation of myogenic tone [294]. *Ex vivo* resistance artery preparations from the SMC specific *Fhl5* overexpression mouse model for pressure myography will determine the effect of *Fhl5* overexpression on vasoreactivity under near-physiological conditions. These *ex-vivo* studies will clarify the detrimental or protective effect of FHL5 on blood pressure. Although human genetics suggest that *FHL5* plays a protective role, this direction of effect needs to be validated experimentally. *FHL5* also increases CAD/MI risk, suggesting distinct contributions to these co-occurring common diseases.

Zebrafish Model

Zebrafish represent an emerging model system to interrogate mechanisms of vascular diseases, such as atherosclerosis [295] [296], arterial-venous malformation [297] [298] , ischemic stroke [299], and age-associated vascular calcification [300]. Despite a more primitive vasculature, zebrafish vessels are composed of a similar structure with an endothelial cell tube surrounded by mural cells that act synergistically to regulate vascular tone [301] [302] [303]. Mural cells in both vertebrates express a similar set of contractile markers and use conserved vasoactive agents to constrict and dilate the vessel [304] .

In our preliminary studies to assess the suitability of the zebrafish model to study *fhl5* function, we first characterized its expression profile. We performed i*n situ hybridization (*ISH*)* to localize fhl5 expression to the pharynx and eye region at 48-96 hours post fertilization (hpf), recapitulating the published ZFIN ISH data (**Fig. 3.1**). Consistent with the ZFIN data we did not observe *fhl5* expression at 24hpf. To determine the *fhl5* expression profile across embryonic, larval, and adult zebrafish, we also generated a reporter zebrafish line (Tg(*fhl5:mCherry*)) which labeled specific regions in the eye and pharynx region (**Fig. 3.2**). These specific localizations may reflect *fhl5* expression in the

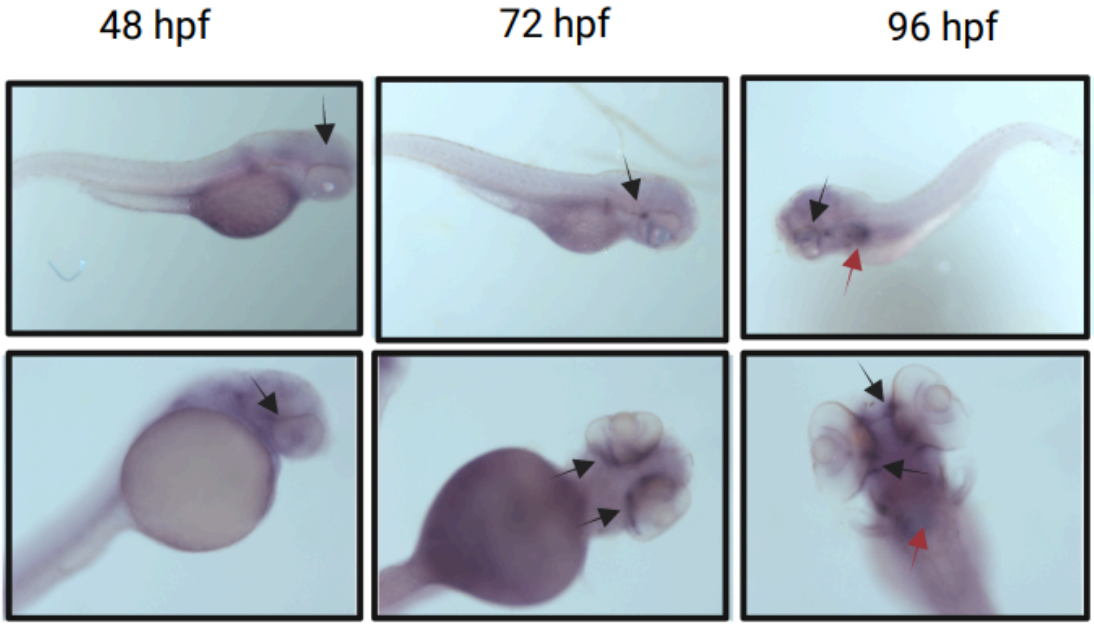

Figure 3.1: *Fhl5* gene expression in zebrafish embryos.
Whole mount in situ hybridization (ISH) across zebrafish embryos, 48-96hpf identifying *fhl5* near the eye and pharynx region. Red and black arrows highlight signal near the pharynx and eye respectively.

skeletal muscle controlling respiration and ocular movements. Although FHL5 is enriched in artery tissues in humans, we did not observe expression in the dorsal aorta, the major blood vessel in zebrafish. It is possible that fhl5 may be expressed in a subset of blood vessels in these tissues. However, that hypothesis is difficult to address without a double transgenic model labeling both fhl5 expressing cells and the endothelial cells or SMCs.

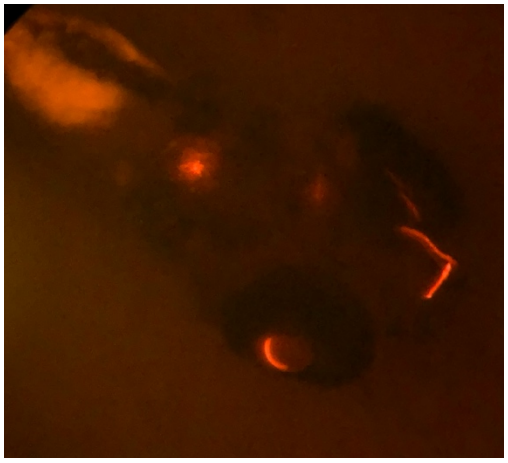

Figure 3.2: Localization of *fhl5* gene expression in transgenic zebrafish model.
Tg(*fhl5:mCherry*) zebrafish model show fhl5 expression in the eye and pharynx region at 120hpf.

In order to define the function of fhl5 in vascular physiology and angiogenesis, we aimed to generate a mural cell specific *Fhl5* overexpression transgenic zebrafish in a Tg(*tagln-eGFP*; *kdrl:mCherry*) background [303]. This model labels endothelial cells and SMCs with mCherry and eGFP respectively. In our preliminary studies, we constructed a Tol2 plasmid containing the *Tagln* promoter driving fhl5 expression and microinjected zebrafish embryos immediately following fertilization. Given this new zebrafish model, we can address the role of *Fhl5* on blood vessel morphology throughout development across different vascular beds, such as the dorsal aorta and cerebral vasculature. Live cell microscopy can be performed on the transparent embryos to visualize the recruitment of mural cells to the nascent blood vessels, a process which is conserved among vertebrates. Microangiography will highlight vascular leakages and disruptions to the structural integrity due to mural cell dysfunction from *Fhl5* overexpression. Lastly, to address the vascular tone hypothesis, vasoreactivity defects can also be interrogated with this model. Following administration of vasoactive factors, changes in vessel diameter and blood flow velocity will reflect deficits in vascular tone regulation [305] [306]. Together, these studies would support the functional mechanism linking perturbation of *Fhl5* gene expression to common vascular disease.

Therapeutic Potential

As we showed in chapter 2, *FHL5* target genes are enriched in CAD candidate genes, many of which are predicted to function in vascular remodeling. Given the limited FDA approved drugs targeting these maladaptive pathways in the vessel wall, there is an opportunity to leverage this molecular characterization to identify existing therapeutics that may be effective in treating cardiovascular disease. As a proof of concept, preliminary analyses intersecting high confidence *FHL5* target genes with the druggable genome identified a number of hits, including *IL6*. The proatherogenic role of IL6 has already been validated through Mendelian Randomization studies that causally link IL6 levels to adverse cardiovascular events [307] [308]. As zebrafish is an already established model for high throughput drug screening [309], this model system can be used to prioritize other *FHL5* downstream genes. The emerging popularity of zebrafish spurred the development of the atherosclerosis zebrafish model, which mirrors many of the features of early-stage human lesions [310] [296] [311]. Moreover, zebrafish developed obstructive plaques after 45 days of atherogenic stimulation. The accumulation of macrophages, neutrophils, and accumulation of cholesterol in the vessel wall can be monitored through living imaging to screen for effective compounds. Validation of these hits in a mouse model will further support the clinical translation of this work.

Conclusion:

GWAS have provided tremendous insights into CAD heritability. However, a complete understanding of its genetic basis requires additional follow-up studies to tease apart the disease mechanisms at GWAS loci. Despite the significance of such studies to the field, the disease mechanisms at a majority of CAD loci remain unknown. Through the work presented in this dissertation, we address this gap and characterize the upstream and downstream disease mechanisms at the pleiotropic *UFL1-FHL5* locus. This project defines the function of FHL5 in SMCs with implications not only for CAD but also for other common non-atherosclerotic vascular diseases. We provide a framework to dissect the molecular mechanisms of other CAD associated transcriptional regulators with predicted functions in the vessel wall.

WorksIN Cited

1. Tucker, W. D., Arora, Y. & Mahajan, K. Anatomy, Blood Vessels. in *StatPearls* (StatPearls Publishing, 2022).

2. Pham, T. T. D. *et al.* Heart and Brain Pericytes Exhibit a Pro-Fibrotic Response After Vascular Injury. *Circ. Res.* **129**, e141–e143 (2021).

3. Armulik, A., Abramsson, A. & Betsholtz, C. Endothelial/pericyte interactions. *Circ. Res.* **97**, 512–523 (2005).

4. Eelen, G. *et al.* Endothelial Cell Metabolism. *Physiol. Rev.* **98**, 3–58 (2018).

5. Yanagisawa, M. *et al.* A novel potent vasoconstrictor peptide produced by vascular endothelial cells. *Nature* **332**, 411–415 (1988).

6. Lamas, S., Marsden, P. A., Li, G. K., Tempst, P. & Michel, T. Endothelial nitric oxide synthase: molecular cloning and characterization of a distinct constitutive enzyme isoform. *Proc Natl Acad Sci USA* **89**, 6348–6352 (1992).

7. Galley, H. F. & Webster, N. R. Physiology of the endothelium. *Br. J. Anaesth.* **93**, 105–113 (2004).

8. Leloup, A. J. A. *et al.* Elastic and Muscular Arteries Differ in Structure, Basal NO Production and Voltage-Gated Ca(2+)-Channels. *Front. Physiol.* **6**, 375 (2015).

9. Basatemur, G. L., Jørgensen, H. F., Clarke, M. C. H., Bennett, M. R. & Mallat, Z. Vascular smooth muscle cells in atherosclerosis. *Nat. Rev. Cardiol.* **16**, 727–744 (2019).

10. Kohn, J. C., Lampi, M. C. & Reinhart-King, C. A. Age-related vascular stiffening: causes and consequences. *Front. Genet.* **6**, 112 (2015).

11. Barallobre-Barreiro, J. *et al.* Extracellular matrix in vascular disease, part 2/4: JACC focus seminar. *J. Am. Coll. Cardiol.* **75**, 2189–2203 (2020).

12. Ma, Z., Mao, C., Jia, Y., Fu, Y. & Kong, W. Extracellular matrix dynamics in vascular remodeling. *Am J Physiol, Cell Physiol* **319**, C481–C499 (2020).

13. Phillippi, J. A. On vasa vasorum: A history of advances in understanding the vessels of vessels. *Sci. Adv.* **8**, eabl6364 (2022).

14. Kokkinopoulos, I. *et al.* Adventitial SCA-1+ Progenitor Cell Gene Sequencing Reveals the Mechanisms of Cell Migration in Response to Hyperlipidemia. *Stem Cell Reports* **9**, 681–696 (2017).

15. Michelis, K. C. *et al.* CD90 Identifies Adventitial Mesenchymal Progenitor Cells in Adult Human Medium- and Large-Sized Arteries. *Stem Cell Reports* **11**, 242–257 (2018).

16. Stenmark, K. R. *et al.* The adventitia: essential regulator of vascular wall structure and function. *Annu. Rev. Physiol.* **75**, 23–47 (2013).

17. Webb, R. C. Smooth muscle contraction and relaxation. *Adv. Physiol. Educ.* **27**, 201–206 (2003).

18. Wynne, B. M., Chiao, C.-W. & Webb, R. C. Vascular Smooth Muscle Cell Signaling Mechanisms for Contraction to Angiotensin II and Endothelin-1. *J. Am. Soc. Hypertens.* **3**, 84–95 (2009).

19. Kuo, I. Y. & Ehrlich, B. E. Signaling in muscle contraction. *Cold Spring Harb. Perspect. Biol.* **7**, a006023 (2015).

20. Schiffrin, E. L. Role of endothelin-1 in hypertension and vascular disease. *Am. J. Hypertens.* **14**, 83S-89S (2001).

21. Rich, S. & McLaughlin, V. V. Endothelin receptor blockers in cardiovascular disease. *Circulation* **108**, 2184–2190 (2003).

22. Owens, G. K. Regulation of differentiation of vascular smooth muscle cells. *Physiol. Rev.* **75**, 487–517 (1995).

23. Alexander, M. R. & Owens, G. K. Epigenetic control of smooth muscle cell differentiation and phenotypic switching in vascular development and disease. *Annu. Rev. Physiol.* **74**, 13–40 (2012).

24. Mack, C. P., Thompson, M. M., Lawrenz-Smith, S. & Owens, G. K. Smooth muscle alpha-actin CArG elements coordinate formation of a smooth muscle cell-selective, serum response factor-containing activation complex. *Circ. Res.* **86**, 221–232 (2000).

25. Kumar, M. S. & Owens, G. K. Combinatorial control of smooth muscle-specific gene expression. *Arterioscler. Thromb. Vasc. Biol.* **23**, 737–747 (2003).

26. Wang, D. *et al.* Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. *Cell* **105**, 851–862 (2001).

27. Yoshida, T. *et al.* Myocardin is a key regulator of CArG-dependent transcription of multiple smooth muscle marker genes. *Circ. Res.* **92**, 856–864 (2003).

28. Parmacek, M. S. Myocardin-related transcription factors: critical coactivators regulating cardiovascular development and adaptation. *Circ. Res.* **100**, 633–644 (2007).

29. Mack, C. P. & Owens, G. K. Regulation of smooth muscle alpha-actin expression in vivo is dependent on CArG elements within the 5' and first intron promoter regions. *Circ. Res.* **84**, 852–861 (1999).

30. Miano, J. M., Carlson, M. J., Spencer, J. A. & Misra, R. P. Serum response factor-dependent regulation of the smooth muscle calponin gene. *J. Biol. Chem.* **275**, 9814–9822 (2000).

31. Strobeck, M. *et al.* Binding of serum response factor to CArG box sequences is necessary but not sufficient to restrict gene expression to arterial smooth muscle cells. *J. Biol. Chem.* **276**, 16418–16424 (2001).

32. Nanda, V. & Miano, J. M. Leiomodin 1, a new serum response factor-dependent target gene expressed preferentially in differentiated smooth muscle cells. *J. Biol. Chem.* **287**, 2459–2467 (2012).

33. McDonald, O. G., Wamhoff, B. R., Hoofnagle, M. H. & Owens, G. K. Control of SRF binding to CArG box chromatin regulates smooth muscle gene expression in vivo. *J. Clin. Invest.* **116**, 36–48 (2006).

34. Liu, M. *et al.* H3K4 di-methylation governs smooth muscle lineage identity and promotes vascular homeostasis by restraining plasticity. *Dev. Cell* **56**, 2765-2782.e10 (2021).

35. Gomez, D., Shankman, L. S., Nguyen, A. T. & Owens, G. K. Detection of histone modifications at specific gene loci in single cells in histological sections. *Nat. Methods* **10**, 171–177 (2013).

36. Gomez, D., Swiatlowska, P. & Owens, G. K. Epigenetic control of smooth muscle cell identity and lineage memory. *Arterioscler. Thromb. Vasc. Biol.* **35**, 2508–2516 (2015).

37. Ito, S. *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303 (2011).

38. Bachman, M. *et al.* 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat. Chem.* **6**, 1049–1055 (2014).

39. Tsagaratou, A. *et al.* Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation. *Proc Natl Acad Sci USA* **111**, E3306-15 (2014).

40. Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S. & Heintz, N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151**, 1417–1430 (2012).

41. Liu, R. *et al.* Ten-eleven translocation-2 (TET2) is a master regulator of smooth muscle cell plasticity. *Circulation* **128**, 2047–2057 (2013).

42. Ostriker, A. C. *et al.* TET2 protects against vascular smooth muscle cell apoptosis and intimal thickening in transplant vasculopathy. *Circulation* **144**, 455–470 (2021).

43. Wagenseil, J. E. & Mecham, R. P. Vascular extracellular matrix and arterial mechanics. *Physiol. Rev.* **89**, 957–989 (2009).

44. Wang, X. & Khalil, R. A. Matrix metalloproteinases, vascular remodeling, and vascular disease. *Adv. Pharmacol.* **81**, 241–330 (2018).

45. Wijelath, E. S. *et al.* Heparin-II domain of fibronectin is a vascular endothelial growth factor-binding domain: enhancement of VEGF biological activity by a singular growth factor/matrix protein synergism. *Circ. Res.* **99**, 853–860 (2006).

46. Misra, A. *et al.* Integrin beta3 regulates clonality and fate of smooth muscle-derived atherosclerotic plaque cells. *Nat. Commun.* **9**, 2073 (2018).

47. Jain, M. *et al.* Integrin α9 regulates smooth muscle cell phenotype switching and vascular remodeling. *JCI Insight* (2021).

48. Moiseeva, E. P. Adhesion receptors of vascular smooth muscle cells and their functions. *Cardiovasc. Res.* **52**, 372–386 (2001).

49. Jeong, K. *et al.* Nuclear Focal Adhesion Kinase Controls Vascular Smooth Muscle Cell Proliferation and Neointimal Hyperplasia Through GATA4-Mediated Cyclin D1 Transcription. *Circ. Res.* **125**, 152–166 (2019).

50. Geiger, B., Spatz, J. P. & Bershadsky, A. D. Environmental sensing through focal adhesions. *Nat. Rev. Mol. Cell Biol.* **10**, 21–33 (2009).

51. Slepian, M. J., Massia, S. P., Dehdashti, B., Fritz, A. & Whitesell, L. Beta3-integrins rather than beta1-integrins dominate integrin-matrix interactions involved in postinjury smooth muscle cell migration. *Circulation* **97**, 1818–1827 (1998).

52. Dupont, S. *et al.* Role of YAP/TAZ in mechanotransduction. *Nature* **474**, 179–183 (2011).

53. Staus, D. P. *et al.* Nuclear RhoA signaling regulates MRTF-dependent SMC-specific transcription. *Am. J. Physiol. Heart Circ. Physiol.* **307**, H379-90 (2014).

54. Sotiropoulos, A., Gineitis, D., Copeland, J. & Treisman, R. Signal-regulated activation of serum response factor is mediated by changes in actin dynamics. *Cell* **98**, 159–169 (1999).

55. Olson, E. N. & Nordheim, A. Linking actin dynamics and gene transcription to drive cellular motile functions. *Nat. Rev. Mol. Cell Biol.* **11**, 353–365 (2010).

56. Owens, G. K., Kumar, M. S. & Wamhoff, B. R. Molecular regulation of vascular smooth muscle cell differentiation in development and disease. *Physiol. Rev.* **84**, 767–801 (2004).

57. Worth, N. F., Rolfe, B. E., Song, J. & Campbell, G. R. Vascular smooth muscle cell phenotypic modulation in culture is associated with reorganisation of contractile and cytoskeletal proteins. *Cell Motil. Cytoskeleton* **49**, 130–145 (2001).

58. Dandré, F. & Owens, G. K. Platelet-derived growth factor-BB and Ets-1 transcription factor negatively regulate transcription of multiple smooth muscle cell differentiation marker genes. *Am. J. Physiol. Heart Circ. Physiol.* **286**, H2042-51 (2004).

59. Blank, R. S. & Owens, G. K. Platelet-derived growth factor regulates actin isoform expression and growth state in cultured rat aortic smooth muscle cells. *J. Cell. Physiol.* **142**, 635–642 (1990).

60. Alexander, M. R., Murgai, M., Moehle, C. W. & Owens, G. K. Interleukin-1β modulates smooth muscle cell phenotype to a distinct inflammatory state relative to PDGF-DD via NF-κB-dependent mechanisms. *Physiol. Genomics* **44**, 417–429 (2012).

61. Pidkovka, N. A. *et al.* Oxidized phospholipids induce phenotypic switching of vascular smooth muscle cells in vivo and in vitro. *Circ. Res.* **101**, 792–801 (2007).

62. 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).

63. Alencar, G. F. *et al.* Stem Cell Pluripotency Genes Klf4 and Oct4 Regulate Complex SMC Phenotypic Changes Critical in Late-Stage Atherosclerotic Lesion Pathogenesis. *Circulation* **142**, 2045–2059 (2020).

64. Pan, H. *et al.* Single-Cell Genomics Reveals a Novel Cell State During Smooth Muscle Cell Phenotypic Switching and Potential Therapeutic Targets for Atherosclerosis in Mouse and Human. *Circulation* **142**, 2060–2075 (2020).

65. Tsao, C. W. *et al.* Heart Disease and Stroke Statistics-2022 Update: A Report From the American Heart Association. *Circulation* **145**, e153–e639 (2022).

66. Pizzi, C. *et al.* Nonobstructive Versus Obstructive Coronary Artery Disease in Acute Coronary Syndrome: A Meta-Analysis. *J. Am. Heart Assoc.* **5**, (2016).

67. Hayes, S. N. *et al.* Spontaneous coronary artery dissection: current state of the science: A scientific statement from the american heart association. *Circulation* **137**, e523–e557 (2018).

68. Okrainec, K., Banerjee, D. K. & Eisenberg, M. J. Coronary artery disease in the developing world. *Am. Heart J.* **148**, 7–15 (2004).

69. Zdravkovic, S. *et al.* Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. *J. Intern. Med.* **252**, 247–254 (2002).

70. Lloyd-Jones, D. M. *et al.* Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA* **291**, 2204–2211 (2004).

71. Murabito, J. M. *et al.* Sibling cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults. *JAMA* **294**, 3117–3123 (2005).

72. Wilson, P. W. *et al.* Prediction of coronary heart disease using risk factor categories. *Circulation* **97**, 1837–1847 (1998).

73. Khera, A. V. & Kathiresan, S. Genetics of coronary artery disease: discovery, biology and clinical translation. *Nat. Rev. Genet.* **18**, 331–344 (2017).

74. Granér, M. *et al.* Association of carotid intima-media thickness with angiographic severity and extent of coronary artery disease. *Am. J. Cardiol.* **97**, 624–629 (2006).

75. Kablak-Ziembicka, A. *et al.* Association of increased carotid intima-media thickness with the extent of coronary artery disease. *Heart* **90**, 1286–1290 (2004).

76. Feistritzer, H.-J. *et al.* Prognostic Value of Aortic Stiffness in Patients After ST-Elevation Myocardial Infarction. *J. Am. Heart Assoc.* **6**, (2017).

77. van Sloten, T. T. *et al.* Carotid stiffness is associated with incident stroke: a systematic review and individual participant data meta-analysis. *J. Am. Coll. Cardiol.* **66**, 2116–2125 (2015).

78. Greenland, P., LaBree, L., Azen, S. P., Doherty, T. M. & Detrano, R. C. Coronary artery calcium score combined with Framingham score for risk prediction in asymptomatic individuals. *JAMA* **291**, 210–215 (2004).

79. Yeboah, J. *et al.* Comparison of novel risk markers for improvement in cardiovascular risk assessment in intermediate-risk individuals. *JAMA* **308**, 788–795 (2012).

80. Writing Committee Members *et al.* 2021 ACC/AHA/SCAI guideline for coronary artery revascularization: A report of the american college of cardiology/american heart association joint committee on clinical practice guidelines. *J. Am. Coll. Cardiol.* **79**, e21–e129 (2022).

81. Musunuru, K. *et al.* In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* **593**, 429–434 (2021).

82. Lehrman, M. A. *et al.* Mutation in LDL receptor: Alu-Alu recombination deletes exons encoding transmembrane and cytoplasmic domains. *Science* **227**, 140–146 (1985).

83. Soria, L. F. *et al.* Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. *Proc Natl Acad Sci USA* **86**, 587–591 (1989).

84. Uffelmann, E. *et al.* Genome-wide association studies. *Nat. Rev. Methods Primers* **1**, 59 (2021).

85. Maurano, M. T. *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012).

86. Hartiala, J. A. *et al.* Genome-wide analysis identifies novel susceptibility loci for myocardial infarction. *Eur. Heart J.* **42**, 919–933 (2021).

87. Tcheandjieu, C. *et al.* Large-scale genome-wide association study of coronary artery disease in genetically diverse populations. *Nat. Med.* **28**, 1679–1692 (2022).

88. 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).

89. Jia, H. *et al.* In vivo diagnosis of plaque erosion and calcified nodule in patients with acute coronary syndrome by intravascular optical coherence tomography. *J. Am. Coll. Cardiol.* **62**, 1748–1758 (2013).

90. Guagliumi, G. *et al.* Mechanisms of atherothrombosis and vascular response to primary percutaneous coronary intervention in women versus men with acute myocardial infarction: results of the OCTAVIA study. *JACC Cardiovasc. Interv.* **7**, 958–968 (2014).

91. Gilad, Y., Rifkin, S. A. & Pritchard, J. K. Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends Genet.* **24**, 408–415 (2008).

92. Cookson, W., Liang, L., Abecasis, G., Moffatt, M. & Lathrop, M. Mapping complex disease traits with global gene expression. *Nat. Rev. Genet.* **10**, 184–194 (2009).

93. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).

94. Franzén, O. *et al.* Cardiometabolic risk loci share downstream cis- and trans-gene regulation across tissues and diseases. *Science* **353**, 827–830 (2016).

95. Talukdar, H. A. *et al.* Cross-Tissue Regulatory Gene Networks in Coronary Artery Disease. *Cell Syst.* **2**, 196–208 (2016).

96. Miller, C. L. *et al.* Integrative functional genomics identifies regulatory mechanisms at coronary artery disease loci. *Nat. Commun.* **7**, 12092 (2016).

97. Lappalainen, T. *et al.* Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* **501**, 506–511 (2013).

98. Wallace, C. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genet.* **16**, e1008720 (2020).

99. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).

100. Li, B. & Ritchie, M. D. From GWAS to Gene: Transcriptome-Wide Association Studies and Other Methods to Functionally Understand GWAS Discoveries. *Front. Genet.* **12**, 713230 (2021).

101. Gamazon, E. R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nat. Genet.* **47**, 1091–1098 (2015).

102. Barbeira, A. N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, 1825 (2018).

103. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).

104. Wu, Y. *et al.* Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat. Commun.* **9**, 918 (2018).

105. 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).

106. Civelek, M. & Lusis, A. J. Systems genetics approaches to understand complex traits. *Nat. Rev. Genet.* **15**, 34–48 (2014).

107. Koplev, S. *et al.* A mechanistic framework for cardiometabolic and coronary artery diseases. *Nat. Cardiovasc. Res.* **1**, 85–100 (2022).

108. von Scheidt, M. *et al.* Transcription factor MAFF (MAF basic leucine zipper transcription factor F) regulates an atherosclerosis relevant network connecting inflammation and cholesterol metabolism. *Circulation* **143**, 1809–1823 (2021).

109. Lo Sardo, V. *et al.* Unveiling the Role of the Most Impactful Cardiovascular Risk Locus through Haplotype Editing. *Cell* **175**, 1796-1810.e20 (2018).

110. Hilton, I. B. *et al.* Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33**, 510–517 (2015).

111. Joung, J. *et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **12**, 828–863 (2017).

112. Cheng, P. *et al.* ZEB2 shapes the epigenetic landscape of atherosclerosis. *Circulation* **145**, 469–485 (2022).

113. Libby, P. The changing landscape of atherosclerosis. *Nature* **592**, 524–533 (2021).

114. Li, D., Liu, L., Chen, H., Sawamura, T. & Mehta, J. L. LOX-1, an oxidized LDL endothelial receptor, induces CD40/CD40L signaling in human coronary artery endothelial cells. *Arterioscler. Thromb. Vasc. Biol.* **23**, 816–821 (2003).

115. Li, D. *et al.* LOX-1 mediates oxidized low-density lipoprotein-induced expression of matrix metalloproteinases in human coronary artery endothelial cells. *Circulation* **107**, 612–617 (2003).

116. Li, D. & Mehta, J. L. Antisense to LOX-1 inhibits oxidized LDL-mediated upregulation of monocyte chemoattractant protein-1 and monocyte adhesion to human coronary artery endothelial cells. *Circulation* **101**, 2889–2895 (2000).

117. Wang, Y. *et al.* Smooth Muscle Cells Contribute the Majority of Foam Cells in ApoE (Apolipoprotein E)-Deficient Mouse Atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **39**, 876–887 (2019).

118. Allahverdian, S., Chehroudi, A. C., McManus, B. M., Abraham, T. & Francis, G. A. Contribution of intimal smooth muscle cells to cholesterol accumulation and macrophage-like cells in human atherosclerosis. *Circulation* **129**, 1551–1559 (2014).

119. Chattopadhyay, A. *et al.* Cholesterol-Induced Phenotypic Modulation of Smooth Muscle Cells to Macrophage/Fibroblast-like Cells Is Driven by an Unfolded Protein Response. *Arterioscler. Thromb. Vasc. Biol.* **41**, 302–316 (2021).

120. Mori, H. *et al.* Coronary Artery Calcification and its Progression: What Does it Really Mean? *JACC Cardiovasc. Imaging* **11**, 127–142 (2018).

121. Kockx, M. M. *et al.* Apoptosis and related proteins in different stages of human atherosclerotic plaques. *Circulation* **97**, 2307–2315 (1998).

122. New, S. E. P. *et al.* Macrophage-derived matrix vesicles: an alternative novel mechanism for microcalcification in atherosclerotic plaques. *Circ. Res.* **113**, 72–77 (2013).

123. Roijers, R. B. *et al.* Microcalcifications in early intimal lesions of atherosclerotic human coronary arteries. *Am. J. Pathol.* **178**, 2879–2887 (2011).

124. Hutcheson, J. D. *et al.* Genesis and growth of extracellular-vesicle-derived microcalcification in atherosclerotic plaques. *Nat. Mater.* **15**, 335–343 (2016).

125. Leopold, J. A. Vascular calcification: Mechanisms of vascular smooth muscle cell calcification. *Trends Cardiovasc. Med.* **25**, 267–274 (2015).

126. Shao, J.-S. *et al.* Msx2 promotes cardiovascular calcification by activating paracrine Wnt signals. *J. Clin. Invest.* **115**, 1210–1220 (2005).

127. Hill, T. P., Später, D., Taketo, M. M., Birchmeier, W. & Hartmann, C. Canonical Wnt/beta-catenin signaling prevents osteoblasts from differentiating into chondrocytes. *Dev. Cell* **8**, 727–738 (2005).

128. Jinnouchi, H. *et al.* Calcium deposition within coronary atherosclerotic lesion: Implications for plaque stability. *Atherosclerosis* **306**, 85–95 (2020).

129. Nadkarni, S. K. *et al.* Measurement of collagen and smooth muscle cell content in atherosclerotic plaques using polarization-sensitive optical coherence tomography. *J. Am. Coll. Cardiol.* **49**, 1474–1481 (2007).

130. Shankman, L. S. *et al.* KLF4-dependent phenotypic modulation of smooth muscle cells has a key role in atherosclerotic plaque pathogenesis. *Nat. Med.* **21**, 628–637 (2015).

131. Cherepanova, O. A. *et al.* Activation of the pluripotency factor OCT4 in smooth muscle cells is atheroprotective. *Nat. Med.* **22**, 657–665 (2016).

132. Salmon, M., Gomez, D., Greene, E., Shankman, L. & Owens, G. K. Cooperative binding of KLF4, pELK-1, and HDAC2 to a G/C repressor element in the SM22α promoter mediates transcriptional silencing during SMC phenotypic switching in vivo. *Circ. Res.* **111**, 685–696 (2012).

133. Yoshida, T., Kaestner, K. H. & Owens, G. K. Conditional deletion of Krüppel-like factor 4 delays downregulation of smooth muscle cell differentiation markers but accelerates neointimal formation following vascular injury. *Circ. Res.* **102**, 1548–1557 (2008).

134. Acharya, A. *et al.* The bHLH transcription factor Tcf21 is required for lineage-specific EMT of cardiac fibroblast progenitors. *Development* **139**, 2139–2149 (2012).

135. Nurnberg, S. T. *et al.* Coronary artery disease associated transcription factor TCF21 regulates smooth muscle precursor cells that contribute to the fibrous cap. *PLoS Genet.* **11**, e1005155 (2015).

136. Nagao, M. *et al.* Coronary Disease-Associated Gene TCF21 Inhibits Smooth Muscle Cell Differentiation by Blocking the Myocardin-Serum Response Factor Pathway. *Circ. Res.* **126**, 517–529 (2020).

137. Sazonova, O. *et al.* Characterization of TCF21 downstream target regions identifies a transcriptional network linking multiple independent coronary artery disease loci. *PLoS Genet.* **11**, e1005202 (2015).

138. Bach, I. The LIM domain: regulation by association. *Mech. Dev.* **91**, 5–17 (2000).

139. Schmeichel, K. L. & Beckerle, M. C. The LIM domain is a modular protein-binding interface. *Cell* **79**, 211–219 (1994).

140. Kadrmas, J. L. & Beckerle, M. C. The LIM domain: from the cytoskeleton to the nucleus. *Nat. Rev. Mol. Cell Biol.* **5**, 920–931 (2004).

141. Kihara, T. *et al.* Regulation of cysteine-rich protein 2 localization by the development of actin fibers during smooth muscle cell differentiation. *Biochem. Biophys. Res. Commun.* **411**, 96–101 (2011).

142. Chang, D. F. *et al.* Cysteine-rich LIM-only proteins CRP1 and CRP2 are potent smooth muscle differentiation cofactors. *Dev. Cell* **4**, 107–118 (2003).

143. Wei, J. *et al.* Increased neointima formation in cysteine-rich protein 2-deficient mice in response to vascular injury. *Circ. Res.* **97**, 1323–1331 (2005).

144. Sundberg-Smith, L. J., DiMichele, L. A., Sayers, R. L., Mack, C. P. & Taylor, J. M. The LIM protein leupaxin is enriched in smooth muscle and functions as an serum response factor cofactor to induce smooth muscle cell gene transcription. *Circ. Res.* **102**, 1502–1511 (2008).

145. Philippar, U. *et al.* The SRF target gene Fhl2 antagonizes RhoA/MAL-dependent activation of SRF. *Mol. Cell* **16**, 867–880 (2004).

146. Nakazawa, N., Sathe, A. R., Shivashankar, G. V. & Sheetz, M. P. Matrix mechanics controls FHL2 movement to the nucleus to activate p21 expression. *Proc Natl Acad Sci USA* **113**, E6813–E6822 (2016).

147. Ebrahimian, T. *et al.* Absence of Four-and-a-Half LIM Domain Protein 2 Decreases Atherosclerosis in ApoE-/- Mice. *Arterioscler. Thromb. Vasc. Biol.* **35**, 1190–1197 (2015).

148. Chu, P. H., Ruiz-Lozano, P., Zhou, Q., Cai, C. & Chen, J. Expression patterns of FHL/SLIM family members suggest important functional roles in skeletal muscle and cardiovascular system. *Mech. Dev.* **95**, 259–265 (2000).

149. Shathasivam, T., Kislinger, T. & Gramolini, A. O. Genes, proteins and complexes: the multifaceted nature of FHL family proteins in diverse tissues. *J. Cell. Mol. Med.* **14**, 2702–2720 (2010).

150. Winkelman, J. D., Anderson, C. A., Suarez, C., Kovar, D. R. & Gardel, M. L. Evolutionarily diverse LIM domain-containing proteins bind stressed actin filaments through a conserved mechanism. *Proc Natl Acad Sci USA* **117**, 25532–25542 (2020).

151. Keßler, M. *et al.* A zebrafish model for FHL1-opathy reveals loss-of-function effects of human FHL1 mutations. *Neuromuscul. Disord.* **28**, 521–531 (2018).

152. Domenighetti, A. A. *et al.* Loss of FHL1 induces an age-dependent skeletal muscle myopathy associated with myofibrillar and intermyofibrillar disorganization in mice. *Hum. Mol. Genet.* **23**, 209–225 (2014).

153. Cottle, D. L. *et al.* FHL3 binds MyoD and negatively regulates myotube formation. *J. Cell Sci.* **120**, 1423–1435 (2007).

154. Zhang, Y. *et al.* FHL3 differentially regulates the expression of MyHC isoforms through interactions with MyoD and pCREB. *Cell. Signal.* **28**, 60–73 (2016).

155. Fimia, G. M., De Cesare, D. & Sassone-Corsi, P. A family of LIM-only transcriptional coactivators: tissue-specific expression and selective activation of CREB and CREM. *Mol. Cell. Biol.* **20**, 8613–8622 (2000).

156. Lardenois, A. *et al.* Fhl5/Act, a CREM-binding transcriptional activator required for normal sperm maturation and morphology, is not essential for testicular gene expression. *Reprod. Biol. Endocrinol.* **7**, 133 (2009).

157. Nakanishi, K., Saito, Y., Azuma, N. & Sasajima, T. Cyclic adenosine monophosphate response-element binding protein activation by mitogen-activated protein kinase-activated protein kinase 3 and four-and-a-half LIM domains 5 plays a key role for vein graft intimal hyperplasia. *J. Vasc. Surg.* **57**, 182–93, 193.e1 (2013).

158. Uchida, D. *et al.* Development of gene therapy with a cyclic adenosine monophosphate response element decoy oligodeoxynucleotide to prevent vascular intimal hyperplasia. *J. Vasc. Surg.* **71**, 229–241 (2020).

159. 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).

160. Schunkert, H. *et al.* Genetics of coronary artery disease in the light of genome-wide association studies. *Clin. Res. Cardiol.* **107**, 2–9 (2018).

161. Speer, M. Y. *et al.* Smooth muscle cells give rise to osteochondrogenic precursors and chondrocytes in calcifying arteries. *Circ. Res.* **104**, 733–741 (2009).

162. Canela-Xandri, O., Rawlik, K. & Tenesa, A. An atlas of genetic associations in UK Biobank. *Nat. Genet.* **50**, 1593–1599 (2018).

163. Bakker, M. K. *et al.* Genome-wide association study of intracranial aneurysms identifies 17 risk loci and genetic overlap with clinical risk factors. *Nat. Genet.* **52**, 1303–1313 (2020).

164. Gormley, P. *et al.* Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nat. Genet.* **48**, 856–866 (2016).

165. Gupta, R. M. *et al.* A Genetic Variant Associated with Five Vascular Diseases Is a Distal Regulator of Endothelin-1 Gene Expression. *Cell* **170**, 522-533.e15 (2017).

166. Kotaja, N. *et al.* Abnormal sperm in mice with targeted deletion of the act (activator of cAMP-responsive element modulator in testis) gene. *Proc Natl Acad Sci USA* **101**, 10620–10625 (2004).

167. Macho, B. *et al.* CREM-dependent transcription in male germ cells controlled by a kinesin. *Science* **298**, 2388–2390 (2002).

168. 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).

169. Giri, A. *et al.* Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat. Genet.* **51**, 51–62 (2019).

170. Staley, J. R. *et al.* PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* **32**, 3207–3209 (2016).

171. Kamat, M. A. *et al.* PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* **35**, 4851–4853 (2019).

172. Cano-Gamez, E. & Trynka, G. From GWAS to function: using functional genomics to identify the mechanisms underlying complex diseases. *Front. Genet.* **11**, 424 (2020).

173. Hao, K. *et al.* Integrative prioritization of causal genes for coronary artery disease. *Circ. Genom. Precis. Med.* **15**, e003365 (2022).

174. Kichaev, G. *et al.* Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLoS Genet.* **10**, e1004722 (2014).

175. Granja, J. M. *et al.* ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat. Genet.* **53**, 403–411 (2021).

176. Nasser, J. *et al.* Genome-wide enhancer maps link risk variants to disease genes. *Nature* **593**, 238–243 (2021).

177. Benson, C. C., Zhou, Q., Long, X. & Miano, J. M. Identifying functional single nucleotide polymorphisms in the human CArGome. *Physiol. Genomics* **43**, 1038–1048 (2011).

178. Long, X., Bell, R. D., Gerthoffer, W. T., Zlokovic, B. V. & Miano, J. M. Myocardin is sufficient for a smooth muscle-like contractile phenotype. *Arterioscler. Thromb. Vasc. Biol.* **28**, 1505–1510 (2008).

179. Yang, H., Curinga, G. & Giachelli, C. M. Elevated extracellular calcium levels induce smooth muscle cell matrix mineralization in vitro. *Kidney Int.* **66**, 2293–2299 (2004).

180. Fimia, G. M., Morlon, A., Macho, B., De Cesare, D. & Sassone-Corsi, P. Transcriptional cascades during spermatogenesis: pivotal role of CREM and ACT. *Mol. Cell. Endocrinol.* **179**, 17–23 (2001).

181. Ma, W. F. *et al.* Enhanced single-cell RNA-seq workflow reveals coronary artery disease cellular cross-talk and candidate drug targets. *Atherosclerosis* **340**, 12–22 (2022).

182. Alsaigh, T., Evans, D., Frankel, D. & Torkamani, A. Decoding the transcriptome of atherosclerotic plaque at single-cell resolution. *BioRxiv* (2020) doi:10.1101/2020.03.03.968123.

183. Turner, A. W. *et al.* Cell-specific chromatin landscape of human coronary artery resolves regulatory mechanisms of disease risk. *BioRxiv* (2021) doi:10.1101/2021.06.07.447388.

184. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).

185. Greenfest-Allen, E., Cartailler, J.-P., Magnuson, M. A. & Stoeckert, C. J. iterativeWGCNA: iterative refinement to improve module detection from WGCNA co-expression networks. *BioRxiv* (2017) doi:10.1101/234062.

186. Ebana, Y. *et al.* A functional SNP in ITIH3 is associated with susceptibility to myocardial infarction. *J. Hum. Genet.* **52**, 220–229 (2007).

187. Hayer, A. *et al.* Engulfed cadherin fingers are polarized junctional structures between collectively migrating endothelial cells. *Nat. Cell Biol.* **18**, 1311–1323 (2016).

188. Pustlauk, W. *et al.* Induced osteogenic differentiation of human smooth muscle cells as a model of vascular calcification. *Sci. Rep.* **10**, 5951 (2020).

189. Malhotra, R. *et al.* HDAC9 is implicated in atherosclerotic aortic calcification and affects vascular smooth muscle cell phenotype. *Nat. Genet.* **51**, 1580–1587 (2019).

190. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **6**, (2017).

191. Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **13**, 1006–1019 (2018).

192. Zhu, Q., Liu, N., Orkin, S. H. & Yuan, G.-C. CUT&RUNTools: a flexible pipeline for CUT&RUN processing and footprint analysis. *Genome Biol.* **20**, 192 (2019).

193. Yu, G., Wang, L.-G. & He, Q.-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).

194. Wang, S. *et al.* Target analysis by integration of transcriptome and ChIP-seq data with BETA. *Nat. Protoc.* **8**, 2502–2515 (2013).

195. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).

196. Sheffield, N. C. & Bock, C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. *Bioinformatics* **32**, 587–589 (2016).

197. Schmidt, E. M. *et al.* GREGOR: evaluating global enrichment of trait-associated variants in epigenomic features using a systematic, data-driven approach. *Bioinformatics* **31**, 2601–2606 (2015).

198. Khetan, S. *et al.* Type 2 Diabetes-Associated Genetic Variants Regulate Chromatin Accessibility in Human Islets. *Diabetes* **67**, 2466–2477 (2018).

199. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2015).

200. Webb, T. R. *et al.* Systematic evaluation of pleiotropy identifies 6 further loci associated with coronary artery disease. *J. Am. Coll. Cardiol.* **69**, 823–836 (2017).

201. Hautakangas, H. *et al.* Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. *Nat. Genet.* **54**, 152–160 (2022).

202. Winsvold, B. S. *et al.* Genetic analysis for a shared biological basis between migraine and coronary artery disease. *Neurol. Genet.* **1**, e10 (2015).

203. Kurth, T. *et al.* Migraine and risk of cardiovascular disease in women. *JAMA* **296**, 283–291 (2006).

204. Miano, J. M., Fisher, E. A. & Majesky, M. W. Fate and state of vascular smooth muscle cells in atherosclerosis. *Circulation* **143**, 2110–2116 (2021).

205. Budoff, M. J. *et al.* Progression of coronary calcium and incident coronary heart disease events: MESA (Multi-Ethnic Study of Atherosclerosis). *J. Am. Coll. Cardiol.* **61**, 1231–1239 (2013).

206. Detrano, R. *et al.* Coronary calcium as a predictor of coronary events in four racial or ethnic groups. *N. Engl. J. Med.* **358**, 1336–1345 (2008).

207. Karlöf, E. *et al.* Correlation of computed tomography with carotid plaque transcriptomes associates calcification with lesion-stabilization. *Atherosclerosis* **288**, 175–185 (2019).

208. Seime, T. *et al.* Proteoglycan 4 Modulates Osteogenic Smooth Muscle Cell Differentiation during Vascular Remodeling and Intimal Calcification. *Cells* **10**, (2021).

209. Adelstein, R. S. & Sellers, J. R. Effects of calcium on vascular smooth muscle contraction. *Am. J. Cardiol.* **59**, 4B-10B (1987).

210. Waugh, W. H. Role of calcium in contractile excitation of vascular smooth muscle by epinephrine and potassium. *Circ. Res.* **11**, 927–940 (1962).

211. Huang, H., Sun, Z., Hill, M. A. & Meininger, G. A. A calcium mediated mechanism coordinating vascular smooth muscle cell adhesion during kcl activation. *Front. Physiol.* **9**, 1810 (2018).

212. Shen, X. *et al.* Extracellular calcium ion concentration regulates chondrocyte elastic modulus and adhesion behavior. *Int. J. Mol. Sci.* **22**, (2021).

213. Kapustin, A. N. *et al.* Calcium regulates key components of vascular smooth muscle cell-derived matrix vesicles to enhance mineralization. *Circ. Res.* **109**, e1-12 (2011).

214. Choe, N. *et al.* miR-27a-3p Targets ATF3 to Reduce Calcium Deposition in Vascular Smooth Muscle Cells. *Mol. Ther. Nucleic Acids* **22**, 627–639 (2020).

215. Örd, T. *et al.* Single-Cell Epigenomics and Functional Fine-Mapping of Atherosclerosis GWAS Loci. *Circ. Res.* **129**, 240–258 (2021).

216. Zhao, Q. *et al.* TCF21 and AP-1 interact through epigenetic modifications to regulate coronary artery disease gene expression. *Genome Med.* **11**, 23 (2019).

217. Liu, X., Li, Y. I. & Pritchard, J. K. Trans effects on gene expression can drive omnigenic inheritance. *Cell* **177**, 1022-1034.e6 (2019).

218. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An expanded view of complex traits: from polygenic to omnigenic. *Cell* **169**, 1177–1186 (2017).

219. Kumra, H. *et al.* Roles of fibronectin isoforms in neonatal vascular development and matrix integrity. *PLoS Biol.* **16**, e2004812 (2018).

220. Greenbaum, J. *et al.* A multi-ethnic whole genome sequencing study to identify novel loci for bone mineral density. *Hum. Mol. Genet.* (2021) doi:10.1093/hmg/ddab305.

221. Cowling, B. S. *et al.* Identification of FHL1 as a regulator of skeletal muscle mass: implications for human myopathy. *J. Cell Biol.* **183**, 1033–1048 (2008).

222. van der Pijl, R. J. *et al.* The titin N2B and N2A regions: biomechanical and metabolic signaling hubs in cross-striated muscles. *Biophys. Rev.* **13**, 653–677 (2021).

223. Sun, X. *et al.* Mechanosensing through Direct Binding of Tensed F-Actin by LIM Domains. *Dev. Cell* **55**, 468-482.e7 (2020).

224. Coghill, I. D. *et al.* FHL3 is an actin-binding protein that regulates alpha-actinin-mediated actin bundling: FHL3 localizes to actin stress fibers and enhances cell spreading and stress fiber disassembly. *J. Biol. Chem.* **278**, 24139–24152 (2003).

225. Samson, T. *et al.* The LIM-only proteins FHL2 and FHL3 interact with alpha- and beta-subunits of the muscle alpha7beta1 integrin receptor. *J. Biol. Chem.* **279**, 28641–28652 (2004).

226. Ottolini, M. *et al.* Mechanisms underlying selective coupling of endothelial Ca2+ signals with eNOS vs. IK/SK channels in systemic and pulmonary arteries. *J Physiol (Lond)* **598**, 3577–3596 (2020).

227. Artamonov, M. V. *et al.* RSK2 contributes to myogenic vasoconstriction of resistance arteries by activating smooth muscle myosin and the Na+/H+ exchanger. *Sci. Signal.* **11**, (2018).

228. Maravall, M., Mainen, Z. F., Sabatini, B. L. & Svoboda, K. Estimating intracellular calcium concentrations and buffering without wavelength ratioing. *Biophys. J.* **78**, 2655–2667 (2000).

229. Daneva, Z. *et al.* Endothelial pannexin 1-TRPV4 channel signaling lowers pulmonary arterial pressure in mice. *eLife* **10**, (2021).

230. Yasuda, R. *et al.* Imaging calcium concentration dynamics in small neuronal compartments. *Sci. STKE* **2004**, pl5 (2004).

231. Woodruff, M. L. *et al.* Measurement of cytoplasmic calcium concentration in the rods of wild-type and transducin knock-out mice. *J Physiol (Lond)* **542**, 843–854 (2002).

232. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

233. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).

234. Xie, Z. *et al.* Gene Set Knowledge Discovery with Enrichr. *Curr. Protoc.* **1**, e90 (2021).

235. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2**, 100141 (2021).

236. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

237. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

238. Janssens, D. CUT&RUN: Targeted in situ genome-wide profiling with high efficiency for low cell numbers. (2019).

239. Yu, F., Sankaran, V. G. & Yuan, G.-C. CUT&RUNTools 2.0: A pipeline for single-cell and bulk-level CUT&RUN and CUT&Tag data analysis. *BioRxiv* (2021) doi:10.1101/2021.01.26.428013.

240. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

241. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

242. Picard Tools - By Broad Institute. https://broadinstitute.github.io/picard/.

243. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

244. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187-91 (2014).

245. Zhu, J. *et al.* Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks. *Nat. Genet.* **40**, 854–861 (2008).

246. Cortes, A., Albers, P. K., Dendrou, C. A., Fugger, L. & McVean, G. Identifying cross-disease components of genetic risk across hospital data in the UK Biobank. *Nat. Genet.* **52**, 126–134 (2020).

247. Georges, A. *et al.* Genetic investigation of fibromuscular dysplasia identifies risk loci and shared genetics with common cardiovascular diseases. *Nat. Commun.* **12**, 6031 (2021).

248. Steiner, T. J., Stovner, L. J., Vos, T., Jensen, R. & Katsarava, Z. Migraine is first cause of disability in under 50s: will health politicians now take notice? *J. Headache Pain* **19**, 17 (2018).

249. GBD 2015 Neurological Disorders Collaborator Group. Global, regional, and national burden of neurological disorders during 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Neurol.* **16**, 877–897 (2017).

250. Asghar, M. S. *et al.* Evidence for a vascular factor in migraine. *Ann. Neurol.* **69**, 635–645 (2011).

251. Gormley, P. *et al.* Migraine genetics: from genome-wide association studies to translational insights. *Genome Med.* **8**, 86 (2016).

252. Skinhoj, E. Hemodynamic studies within the brain during migraine. *Arch. Neurol.* **29**, 95–98 (1973).

253. Asghar, M. S. *et al.* Dilation by CGRP of middle meningeal artery and reversal by sumatriptan in normal volunteers. *Neurology* **75**, 1520–1526 (2010).

254. Han, L., Liu, Y., Xiong, H. & Hong, P. CGRP monoclonal antibody for preventive treatment of chronic migraine: An update of meta-analysis. *Brain Behav.* **9**, e01215 (2019).

255. Silberstein, S. D. *et al.* Fremanezumab for the preventive treatment of chronic migraine. *N. Engl. J. Med.* **377**, 2113–2122 (2017).

256. Anderson, L. E. & Seybold, V. S. Calcitonin gene-related peptide regulates gene transcription in primary afferent neurons. *J. Neurochem.* **91**, 1417–1429 (2004).

257. Ma, X. *et al.* Tetralogy of Fallot: aorto-pulmonary collaterals and pulmonary arteries have distinctly different transcriptomes. *Pediatr. Res.* **76**, 341–346 (2014).

258. Patrick, W. L. *et al.* Major aortopulmonary collateral arteries with anatomy other than pulmonary atresia/ventricular septal defect. *Ann. Thorac. Surg.* **104**, 907–916 (2017).

259. Mainwaring, R. D. *et al.* Prevalence and anatomy of retroesophageal major aortopulmonary collateral arteries. *Ann. Thorac. Surg.* **102**, 877–882 (2016).

260. Nørgaard, M. A. *et al.* Major aorto-pulmonary collateral arteries of patients with pulmonary atresia and ventricular septal defect are dilated bronchial arteries. *Eur J Cardiothorac Surg* **29**, 653–658 (2006).

261. Brown, S. C., Eyskens, B., Mertens, L., Dumoulin, M. & Gewillig, M. Percutaneous treatment of stenosed major aortopulmonary collaterals with balloon dilatation and stenting: what can be achieved? *Heart* **79**, 24–28 (1998).

262. Shinkawa, T. *et al.* One-stage unifocalization and palliative right ventricular outflow tract reconstruction. *Ann. Thorac. Surg.* **79**, 1044–1047 (2005).

263. Reddy, V. M. *et al.* Early and intermediate outcomes after repair of pulmonary atresia with ventricular septal defect and major aortopulmonary collateral arteries: experience with 85 patients. *Circulation* **101**, 1826–1832 (2000).

264. Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).

265. Vivian Li, W. & Li, Y. scLink: Inferring Sparse Gene Co-expression Networks from Single-cell Expression Data. *Genomics Proteomics Bioinformatics* **19**, 475–492 (2021).

266. Luo, Q., Yu, Y. & Lan, X. SIGNET: single-cell RNA-seq-based gene regulatory network prediction using multiple-layer perceptron bagging. *Brief. Bioinformatics* **23**, (2022).

267. Van de Sande, B. *et al.* A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* **15**, 2247–2276 (2020).

268. Liao, M. *et al.* Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).

269. Lambrechts, D. *et al.* Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat. Med.* **24**, 1277–1289 (2018).

270. Ramachandran, P. *et al.* Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* **575**, 512–518 (2019).

271. Zheng, Q. & Zhao, Y. The diverse biofunctions of LIM domain proteins: determined by subcellular localization and protein-protein interaction. *Biol. Cell* **99**, 489–502 (2007).

272. Sheikh, F. *et al.* An FHL1-containing complex within the cardiomyocyte sarcomere mediates hypertrophic biomechanical stress responses in mice. *J. Clin. Invest.* **118**, 3870–3880 (2008).

273. McGrath, M. J. *et al.* Four and a half LIM protein 1 binds myosin-binding protein C and regulates myosin filament formation and sarcomere assembly. *J. Biol. Chem.* **281**, 7666–7683 (2006).

274. Emmanuele, V. *et al.* Fhl1 W122S causes loss of protein function and late-onset mild myopathy. *Hum. Mol. Genet.* **24**, 714–726 (2015).

275. Cisse, I. I. *et al.* Real-time dynamics of RNA polymerase II clustering in live human cells. *Science* **341**, 664–667 (2013).

276. Mahmood, S. R. *et al.* β-actin dependent chromatin remodeling mediates compartment level changes in 3D genome architecture. *Nat. Commun.* **12**, 5240 (2021).

277. Wei, M. *et al.* Nuclear actin regulates inducible transcription by enhancing RNA polymerase II clustering. *Sci. Adv.* **6**, eaay6515 (2020).

278. Oughtred, R. *et al.* The BioGRID interaction database: 2019 update. *Nucleic Acids Res.* **47**, D529–D541 (2019).

279. Stark, C. *et al.* BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **34**, D535-9 (2006).

280. Perisic Matic, L. *et al.* Phenotypic modulation of smooth muscle cells in atherosclerosis is associated with downregulation of LMOD1, SYNPO2, PDLIM7, PLN, and SYNM. *Arterioscler. Thromb. Vasc. Biol.* **36**, 1947–1961 (2016).

281. Liu, X. *et al.* An AP-MS- and BioID-compatible MAC-tag enables comprehensive mapping of protein interactions and subcellular localizations. *Nat. Commun.* **9**, 1188 (2018).

282. Liu, X., Salokas, K., Weldatsadik, R. G., Gawriyski, L. & Varjosalo, M. Combined proximity labeling and affinity purification-mass spectrometry workflow for mapping and visualizing protein interaction networks. *Nat. Protoc.* **15**, 3182–3211 (2020).

283. Cho, K. F. *et al.* Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nat. Protoc.* **15**, 3971–3999 (2020).

284. Gonzalez-Teran, B. *et al.* Transcription factor protein interactomes reveal genetic determinants in heart disease. *Cell* **185**, 794-814.e30 (2022).

285. Bodea, C. A. *et al.* Pathway specific polygenic risk scores identify pathways and patient clusters associated with inflammatory bowel disease risk, severity and treatment response. *medRxiv* (2021) doi:10.1101/2021.11.19.21266549.

286. Barbu, M. C. *et al.* Pathway-based polygenic risk scores for schizophrenia and associations with clinical and neuroimaging phenotypes in UK Biobank. *medRxiv* (2022) doi:10.1101/2022.07.12.22277553.

287. O'Reilly, P. *et al.* The power of pathway-based polygenic risk scores. *Res. Sq.* (2021) doi:10.21203/rs.3.rs-643696/v1.

288. Örd, T. *et al.* Dissecting the polygenic basis of atherosclerosis using disease associated cell state signatures. *Res. Sq.* (2022) doi:10.21203/rs.3.rs-1775130/v1.

289. Xia, B. *et al.* Widespread transcriptional scanning in the testis modulates gene evolution rates. *Cell* **180**, 248-262.e21 (2020).

290. Cheng, P. *et al. smad3* regulates smooth muscle cell fate and governs adverse remodeling and calcification of atherosclerotic plaque. *BioRxiv* (2020) doi:10.1101/2020.09.15.299131.

291. Davis, M. J. & Hill, M. A. Signaling mechanisms underlying the vascular myogenic response. *Physiol. Rev.* **79**, 387–423 (1999).

292. Uchida, E. & Bohr, D. F. Myogenic tone in isolated perfused vessels. Occurrence among vascular beds and along vascular trees. *Circ. Res.* **25**, 549–555 (1969).

293. Jackson, W. F. Myogenic tone in peripheral resistance arteries and arterioles: the pressure is on! *Front. Physiol.* **12**, 699517 (2021).

294. Gunst, S. J. & Zhang, W. Actin cytoskeletal dynamics in smooth muscle: a new paradigm for the regulation of smooth muscle contraction. *Am J Physiol, Cell Physiol* **295**, C576-87 (2008).

295. Liu, C. *et al.* Modeling hypercholesterolemia and vascular lipid accumulation in LDL receptor mutant zebrafish. *J. Lipid Res.* **59**, 391–399 (2018).

296. Luo, H. *et al.* Chronological in vivo imaging reveals endothelial inflammation prior to neutrophils accumulation and lipid deposition in HCD-fed zebrafish. *Atherosclerosis* **290**, 125–135 (2019).

297. Roman, B. L. *et al.* Disruption of acvrl1 increases endothelial cell number in zebrafish cranial vessels. *Development* **129**, 3009–3019 (2002).

298. Corti, P. *et al.* Interaction between alk1 and blood flow in the development of arteriovenous malformations. *Development* **138**, 1573–1582 (2011).

299. Zhu, X.-Y. *et al.* Ponatinib-induced ischemic stroke in larval zebrafish for drug screening. *Eur. J. Pharmacol.* **889**, 173292 (2020).

300. Singh, A. P. *et al.* αKlotho Regulates Age-Associated Vascular Calcification and Lifespan in Zebrafish. *Cell Rep.* **28**, 2767-2776.e5 (2019).

301. Ando, K. *et al.* Clarification of mural cell coverage of vascular endothelial cells by live imaging of zebrafish. *Development* **143**, 1328–1339 (2016).

302. Santoro, M. M., Pesce, G. & Stainier, D. Y. Characterization of vascular mural cells during zebrafish development. *Mech. Dev.* **126**, 638–649 (2009).

303. Stratman, A. N. *et al.* Interactions between mural cells and endothelial cells stabilize the developing zebrafish dorsal aorta. *Development* **144**, 115–127 (2017).

304. Chen, X., Gays, D., Milia, C. & Santoro, M. M. Cilia control vascular mural cell recruitment in vertebrates. *Cell Rep.* **18**, 1033–1047 (2017).

305. Rödel, C. J. *et al.* Blood flow suppresses vascular anomalies in a zebrafish model of cerebral cavernous malformations. *Circ. Res.* **125**, e43–e54 (2019).

306. Machikhin, A. S., Volkov, M. V., Burlakov, A. B., Khokhlov, D. D. & Potemkin, A. V. Blood Vessel Imaging at Pre-Larval Stages of Zebrafish Embryonic Development. *Diagnostics (Basel)* **10**, (2020).

307. Ridker, P. M. *et al.* Modulation of the interleukin-6 signalling pathway and incidence rates of atherosclerotic events and all-cause mortality: analyses from the Canakinumab Anti-Inflammatory Thrombosis Outcomes Study (CANTOS). *Eur. Heart J.* **39**, 3499–3507 (2018).

308. Ridker, P. M., Rifai, N., Stampfer, M. J. & Hennekens, C. H. Plasma concentration of interleukin-6 and the risk of future myocardial infarction among apparently healthy men. *Circulation* **101**, 1767–1772 (2000).

309. Kithcart, A. & MacRae, C. A. Using Zebrafish for High-Throughput Screening of Novel Cardiovascular Drugs. *JACC Basic Transl. Sci.* **2**, 1–12 (2017).

310. Stoletov, K. *et al.* Vascular lipid accumulation, lipoprotein oxidation, and macrophage lipid uptake in hypercholesterolemic zebrafish. *Circ. Res.* **104**, 952–960 (2009).

311. Han, J. *et al.* Zebrafish model for screening antiatherosclerosis drugs. *Oxid. Med. Cell. Longev.* **2021**, 9995401 (2021).