Title

Multiomics Networks of Cardiovascular Diseases

Author

Cheng, Jenny

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Multiomics Networks of Cardiovascular Diseases

A thesis submitted in partial satisfaction

of the requirements for the degree Master of Science

in Physiological Science

by

Jenny T. Cheng

2022

ABSTRACT OF THE THESIS

Multiomics Networks of Cardiovascular Diseases

by

Jenny T. Cheng

Master of Science in Physiological Science

University of California, Los Angeles, 2022

Professor Xia Yang, Chair

Spontaneous coronary artery dissection (SCAD) and coronary artery disease (CAD) both represent cardiovascular diseases that may result in myocardial infarction and sudden death. However, the populations predominantly impacted by SCAD and CAD as well as the pathogenesis of both diseases are notably different; while SCAD overwhelmingly affects a young, female population and is non-atherosclerotic, CAD mainly impacts males and develops via atherosclerosis. The genetic architecture and mechanisms leading to SCAD onset are currently poorly understood, and although the genetic and environmental risk factors for CAD are better elucidated, further understanding of the mechanisms in which genetic and environmental factors interact to facilitate pathogenesis is needed. We conducted two studies to examine the networks of SCAD and CAD at various omics layers. In a comprehensive multiomics human SCAD study, we investigated disease-associated biological mechanisms and modeled tissue-specific gene regulatory and protein-protein interaction networks across genetic, transcriptomic, and proteomic levels and predicted potential therapeutics for SCAD treatment. We identified various pathways related to the extracellular matrix, immune function, and blood

clotting activation. We pinpointed key regulatory genes and hub proteins central within our tissue-specific networks, including *HOXB9, CADM1,* and *COL18A1*. Female-specific drugs, such as medroxyprogesterone and progesterone receptor agonist, were prioritized in SCAD drug repositioning. In a single-cell RNA-sequencing mouse study, we examined the cell type-specific transcriptome changes of the aorta in *Ldlr-/-* mice, an established atherosclerosis animal model for CAD, under various diet conditions, namely chow, high-cholesterol, and high-cholesterol with added trimethylamine N-oxide (TMAO). We identified significant differences in cellular composition in the modulated smooth muscle cell (SMC) cluster and macrophage M1 subtype and elucidated the transition of SMCs into a protective fibromyocyte phenotype in a data-driven manner. Additionally, we determined cell type-specific and shared differentially expressed genes and their corresponding biological pathways across dietary conditions, pinpointing extracellular structure organization pathways as significant in various cell types between high-cholesterol and chow diets as well as apoptosis-related pathways as enriched in SMC and modulated SMC with the addition of TMAO to the high-cholesterol diet. We further predicted intercellular communications that influence downstream gene expression in modulated SMC and endothelial cells.

The thesis of Jenny T. Cheng is appproved.

Aldons J. Lusis

Pearl Jennine Quijada

Xia Yang, Committee Chair

University of California, Los Angeles

2022

Acknowledgements

I would like to thank my principal investigator Professor Xia Yang for her constant mentorship and guidance through my academic pursuits. I would like to thank my committee members Professors Jake Lusis and Pearl Quijada for their knowledgeable suggestions and feedback on my research projects and the thesis. I would like to thank Dr. David Adlam, Dr. Thomas Webb, Dr. Nabila Bouatia-Naji, and their teams for their work in curating SCAD data and for their collaboration. I would like to thank Dr. Diana Shih for her animal work in the single-cell study and guidance in data analysis. I would like to thank members of the Yang lab for their guidance and assistance. I would like to thank my family and friends for their continual love and support, particularly my parents, grandparents, siblings Meini, Annie, and Kevin, fiancé Michael, nephew Jacob, and niece Adeline. Most importantly, I would like to thank God for His kindness and grace through all my pursuits. This work was partially funded by NIH/NHLBI R01 HL148110 and NIH/NHLBI R01 HL147883.

Table of Contents

1. Introduction and Background 1

Characterization of Spontaneous Coronary Artery Dissection 1

Characterization of Coronary Artery Disease 1

Introduction to Multiomics Network Projects 2

2. Multiomics Systems Biology of Spontaneous Coronary Artery Dissection 4

Introduction 4

Methods 6

Results 11

Discussion 20

3. Single-Cell Level Understanding of Dietary Risk Factors on an Atherosclerotic Mouse Model 24

Introduction 24

Methods 25

Results 29

Discussion 38

4. Conclusions and Future Directions 41

References 64

List of Tables and Figures

Tables..... 38

Table 1. Clinical characteristics of the French, UK, and Mayo Clinic SCAD study populations. 43

Table 2. Clinical characteristics of the CanSCAD/MGI, DEFINE-SCAD, and VCCRI SCAD study populations..... 44

Table 3. Top significant GWAS canonical pathways and coexpression modules associated with SCAD. 45

Table 4. Top significant TWAS canonical pathways and coexpression modules associated with SCAD. 46

Table 5. Top significant PWAS canonical pathways and coexpression modules associated with SCAD. 47

Table 6. Top significant canonical pathways and coexpression modules shared between GWAS and TWAS..... 47

Table 7. Significant prioritized drugs within cardiovascular tissue for SCAD treatment via network-based repositioning. 48

Table 8. Top significant prioritized drugs within Homo Sapiens cardiovascular tissue for SCAD treatment via overlap-based repositioning. 49

Table 9. HC vs. chow DEGs and biological pathways. 50

Table 10. HC+TMAO vs. HC DEGs and biological pathways. 51

Table 11. HC+TMAO vs. chow DEGs and biological pathways. 52

Figures..... 48

Figure 1. Multiomics systems biology of SCAD: study overview. 53

Figure 2. Top GWAS key drivers within a vascular-specific network. 54

Figure 3. Top transcript key drivers within a vascular-specific network. 55

Figure 5. Cell type identification and specific cell-type markers. 56

Figure 6. Modulated SMC subclustering to reveal three distinct clusters. 57

Figure 7. Macrophage subclustering to reveal distinct M0, M1, and M2 macrophage types. 58

Figure 8. Cell type distribution by diet..... 59

Figure 9. Overlap in DEGs between HC and chow conditions. 60

Figure 11. Overlap in DEGs between HC+TMAO and chow conditions. 61

Figure 12. Intercellular interactions targeting modulated SMC genes in response to TMAO. 62

Figure 13. Intercellular interactions targeting endothelial genes in response to TMAO. 63

1. Introduction and Background

Characterization of Spontaneous Coronary Artery Dissection

Spontaneous coronary artery dissection (SCAD) is a non-atherosclerotic cause of myocardial infarction occurring when an intramural hematoma occludes an epicardial artery[1,2,3]. The true incidence of SCAD remains unknown, largely because most patients classify into the lowest risk categories for conventional atherosclerotic disease, thus leading to underdiagnosis[4]. SCAD overwhelmingly impacts a young, predominantly female population and has demonstrated association with pregnancy and systemic arteriopathies, including fibromuscular dysplasia[2]. Clinical presentation of SCAD has proven variable and may include chest pain, ventricular fibrillation, and sudden death[3]. There are currently no clinical trials to direct optimal medical management following SCAD, and report incidences of recurrent SCAD range from 5% to 19% of cases[5,6,7,8,9,10]. While variants in several genes, including *PHACTR1/EDN1, COL3A1, and PKD1*, have been identified, the genetic architecture of SCAD remains poorly elucidated. Therefore, a more comprehensive understanding of SCAD mechanisms will help develop new preventative and therapeutic strategies targeting key regulatory genes involved in SCAD pathogenesis.

Characterization of Coronary Artery Disease

Coronary artery disease (CAD) is caused by atherosclerosis, or the buildup and hardening of plaque in arteries, and is the most common type of heart disease[11]. CAD primarily impacts men over 45 years of age[12], and CAD mortality rates are higher in men compared to women[13]. However, after menopause the prevalence of CAD in women increases. While lifestyle modifications, medication, or surgery may reduce the risk of CAD development and progression, CAD nonetheless remains one of the leading causes of mortality worldwide[14]. Numerous genetic and environmental factors, including hyperlipidemia[15], lack of exercise[16],

smoking[17], and diets rich in high cholesterol and red meat contribute to CAD onset, and this disease may lead to arrhythmia, myocardial infarction, heart failure, or sudden cardiac death[18]. Diets high in red meat and high cholesterol contain higher levels of choline, phosphatidylcholine, and L-carnitine, and microbial enzymes act on these substrates to generate trimethylamine (TMA), and trimethylamine N-oxide (TMAO), the metabolized form of TMA. Circulating TMAO is associated with promotion of atherosclerosis and increased risk for CAD[19]. Due to the polygenic nature of and environmental contributions towards CAD, investigating the mechanisms by which environmental dietary factors contribute to CAD pathogenesis will help to refine and advance existing treatments and develop novel treatment strategies.

Introduction to Multiomics Network Projects

With the ongoing development and advancements in sequencing technology, data at varying molecular levels, termed “multiomics” data[20], may be integrated to allow for a system-level understanding of multifactorial diseases. Multiomics data encompasses genomics, epigenomics, transcriptomics, proteomics, metabolomics and metagenomics. Specifically, genomics, referring to the genome or DNA sequence, allows for insight into the organization and function of genes as well as how genetic variants may contribute to disease or biological processes. Epigenomics studies non-genetic contributions to changes in gene regulation and encompasses such processes as DNA methylation and chromatin modification. Transcriptomics measures mRNA levels to determine gene expression changes in response to external stimuli, and this omics layer aggregates genetic and epigenetic landscapes to inform on the transcription capability of a specific locus. Further, proteomics measures protein expression levels and can delineate altered protein expression in a disease state and contribute to the identification of disease biomarkers and therapeutic targets. Metabolomics examines endogenous low-molecular-weight structure, including lipids, amino acids, peptides, and organic acids, as small-molecule metabolites and metabolic pathways hold integral roles in biological

systems and in disease development. Finally, metagenomics is the genomic study of uncultured microorganisms, particularly in the microbiome, and the intimate relationship between the microbes and overall homeostasis is increasingly highlighted in various diseases.

We carried out two projects to apply multiomics data to dissect the pathogenic mechanisms of SCAD and CAD by employing systems biology and bioinformatic concepts and approaches. To integrate various levels of omics data that each represent a fragmented layer of a whole system, we utilized tissue-specific network modeling. Networks serve to depict large numbers of molecules associated with disease-related pathways obtained from multiomics data as nodes and regulatory relationships and interactions between such molecules as edges or connections. Varied gene expression and biological function by tissue warrant the implementation of tissue-specific networks to reflect the biological setting most accurately. In networks with scale-free topologies, meaning the fraction of nodes with number of edges *k* follows a power law *k*^-α, peripheral non-hub nodes with few connections surround central hub genes with high connectivity within the network. Central hub nodes may be more pathologically relevant due to their high connectivity and therefore may have a larger impact within pathogenesis. We hypothesize that perturbations to tissue-specific networks underlie pathogenesis, and multiomics data can not only inform these networks but also guide therapeutic predictions for disease treatment. Our holistic systems biology-based approach employing various levels of omics data further advances the understanding of two complex and contrasting cardiovascular diseases.

In the first project, we utilized SCAD human multiomics datasets to identify key biological mechanisms involved in disease onset, model SCAD regulatory gene and protein-protein interaction (PPI) networks, and translate findings for drug repositioning. Although various genetic loci have been previously proposed to contribute to SCAD[1,2,3,21], there is certainly still a need for the identification and interpretation of additional precipitating factors and meaningful disease-associated molecular and biological functions to eventually guide the establishment of

an efficacious SCAD treatment plan. In our study, we integrated SCAD genome-wide association studies (GWAS), transcriptomics, and proteomics with functional genomics to identify disease-associated biologically relevant pathways and functional coexpression modules. We further modeled gene regulatory and PPI networks to pinpoint hub genes and proteins central within the networks that represent key candidates involved in SCAD pathogenesis. Based on the network structure and genes, we performed drug repositioning to propose potential therapeutic treatment options for SCAD as a standardized treatment plan supported by data-driven findings is lacking.

In the second project focusing on CAD, which is on the other end of the cardiovascular ddisease spectrum in contrast to SCAD, we utilized single cell transcriptomics to uncover cell-type specific changes and predict intercellular communication in the mouse aorta in response to varying diet conditions. The lack of a comprehensive understanding of the connection between genetic and environmental factors that together lead to CAD onset requires further insight, particularly at a single-cell resolution to account for cell heterogeneity and differential gene expression. Single-cell RNA-sequencing, a cutting-edge throughput analysis, allows for transcriptional profiling of individual cells. Specifically, we aimed to uncover cell type-specific pathways altered in the aorta of the low-density lipoprotein receptor knockout (*Ldlr-/-*) mouse model of atherosclerosis with differing diet conditions that are known to promote CAD, including chow, high-cholesterol, and high-cholesterol with trimethylamine-*N*-oxide (TMAO). Additionally, we modeled intercellular interactions within the heterogeneous cell population to determine how ligands from one cell type may modulate downstream effectors in another to pinpoint key vascular cell types and regulators of CAD.

2. Multiomics Systems Biology of Spontaneous Coronary Artery Dissection

Introduction

Both recessive and dominant modes of inheritance have been implicated in SCAD as 5 familial cases were identified in the Mayo Clinic SCAD Registry consisting of 412 patients[22], but additional familial clustering of SCAD via pedigree studies are still needed[23]. SCAD pathogenic variants have been identified in several genes, including *PKD1, COL3A1, SMAD3, TGFB2, LOX, MYLK,* and *YY1AP1*, but these pathogenic variants only explained 14/384 cases in a cohort of SCAD survivors[2]. Additionally, *PHACTR1/EDN1* on chromosome 6q24 was pinpointed as another genetic locus associated with SCAD[24] and has been reported to confer risk for CAD and cervicocerebral artery dissection. However, there remains uncertainty in how genetic and molecular factors may interact and ultimately precipitate SCAD, and this gap certainly serves as a detriment in cultivating effective treatment plans for SCAD patients.

Integrating data across multiomics domains to connect risk factors to pathogenic pathway perturbations has proven invaluable in fully leveraging big data to better understand complex diseases such as CAD[25], T2D[25], psoriasis[26], and Alzheimer's[27], and to identify novel targets for effective preventative or therapeutic treatment. Therefore, multiomics integration to dissect the pathogenesis of additional complex diseases such as SCAD is warranted. The establishment of the omnigenic disease model[28], in which genes interconnected in networks lead to complex diseases when perturbed, in conjunction with the multiomics systems biology discipline[29], which models network perturbations in complex diseases, facilitates a comprehensive understanding of the contribution of genetic variants to disease onset. Additionally, the omnigenic disease model prioritizes key hub genes as central regulators that likely contribute more to pathway perturbations compared to other peripheral disease-associated genes. Over the past decade, tissue-specific multiomics systems biology has emerged to resolve the omnigenic disease model and its multi-dimensional complexities by pinpointing key tissues, regulatory genes, pathways, networks, and cross-tissue interactions involved in pathophysiology[29]. We aimed to fully leverage multi-tissue multiomics human SCAD

data to produce tissue-specific networks that unravel the complexities of SCAD pathogenesis at various molecular levels and translate findings to perform drug repositioning.

Drug repositioning, an approach that uses existing FDA-approved drugs for new indications, serves to address our current limitations of preventing SCAD, particularly since SCAD has no standardized treatment plan. Therapeutic drugs exert additional complex effects on surrounding pathways and tissues of the known therapeutic target, and therefore may have beneficial effects on additional diseases for which the drugs are not initially designed. Network-based drug repositioning has the capacity to identify such hidden connections and therapeutic options. Pharmomics[30] is a novel species- and tissue-specific network-based drug repositioning tool which is based on *in vivo* molecular studies of drugs. Pharmomics has been shown to outperform existing repositioning tools, such as CMap[31] and LINCS L1000[32], which depend on *in vitro* cell line data. We aimed to use Pharmomics to couple the multiomics-informed SCAD networks with drug networks to determine therapeutic options that target top SCAD regulatory genes and provide drug therapies for treatment. Beyond network-based drug repositioning, we also employed traditional gene overlap-based repositioning via Pharmomics to assess direct overlap between existing drug gene signatures and our SCAD network genes.

Methods

Overall Study Design

As illustrated in **Figure 1**, we queried curated SCAD human multiomics data (genetic, transcriptomic, and proteomic) to determine significantly enriched pathways and coexpression modules associated with SCAD. Canonical pathways were retrieved from various public repositories[33,34,35] while coexpression modules were constructed using SCAD transcript and protein data. Subsequently, key hub genes or molecules were prioritized within regulatory gene or protein interaction networks, and drug repositioning was performed based on either network structures or direct overlap between drug gene signatures and gene inputs. In this study, GWAS

refers to the association between single nucleotide polymorphisms (SNPs) within the genomes of SCAD patients and healthy controls with disease status. We defined transcriptome-wide association studies (TWAS) as the analysis of differentially expressed genes in fibroblasts between SCAD patients and healthy volunteers. Here, protein-wide association studies (PWAS) were defined as the analysis of differentially expressed proteins in the plasma between SCAD patients and healthy volunteers.

Multiomics Datasets

The GWAS meta-analysis included participants of European ancestry from eight studies: DISCO-3C, SCAD-UK I, SCAD-UK II, Mayo Clinic, DEFINE-SCAD, CanSCAD/MGI, VCCRI I and VCCRI II. SCAD patients presented similar clinical characteristics, homogeneous diagnosis, and inclusion criteria (**Tables 1-2**). The overall sample size for the GWAS meta-analysis was 1,913 SCAD patients and 9,292 controls. All studies were approved by national and/or institutional ethical review boards. The fibroblast transcriptome (n=50 SCAD vs n=29 controls) and plasma proteome (n=50 SCAD vs n=50 controls) data were collected exclusively from the SCAD UK studies.

SNP to Gene Mapping Methods

To map SNPs to potential target genes with changes in the expression levels or alternative splicing, tissue-specific expression quantitative trait loci (eQTL) and splicing quantitative trait loci (sQTL) were used, respectively. Human eQTLs and sQTLs were obtained from the Genotype-Tissue Expression (GTEx) project (version 8) [36] for tissues relevant to SCAD, including aorta, coronary artery, tibial artery, and cultured fibroblasts. SNPs were mapped to genes using tissue-specific eQTLs or sQTLs with a cutoff of 50kb distance. SNPs were also mapped to genes using protein quantitative trait loci (pQTLs) from Emilsson et al.[37]

Canonical Pathways and Data-Driven Coexpression Modules

Canonical pathways were curated from public repositories, including the Kyoto Encyclopedia of Genes and Genomes (KEGG)[33], Reactome[34], and Biocarta[35]. Tissue-specific coexpression networks were constructed from SCAD fibroblast transcriptomic and plasma proteomic data, using two network approaches: Weighted Gene Co-expression Network Analysis (WGCNA)[38] and Multiscale Embedded Gene Co-expression Network Analysis (MEGENA)[39]. WGCNA can identify biologically meaningful and relevant modules, but modules are larger in size, with some modules including thousands of genes; this may increase noise and decrease interpretability in downstream analyses. The second coexpression network construction method, MEGENA, can define smaller and more coherent modules that still remain biologically relevant, thereby overcoming a limitation of WGCNA. Both network methods were based on hierarchical clustering to assign co-regulated genes into the same coexpression module. Agglomerative hierarchical clustering is used in WGCNA, whereas divisive clustering is used in MEGENA. In WGCNA, 1 minus topological overlap matrix (TOM), hence dissTOM=1-TOM, was used as the distance measure. TOM is based on the correlation score (edge weight) between two genes (nodes) but also considers the edge weights of common neighbors of these two nodes in the network. In MEGENA, a shortest path distance (SPD) measure was used. To create compact modules, a nested k-medoids clustering, which defines k-best clusters at each step that minimizes the SPD within each cluster, was used. Nested k-medoids clustering was run until no more compact child clusters could be defined. Subsequently, Gene Ontology (GO) was used to assign biologically meaningful functions to WGCNA or MEGENA coexpression modules.

Marker Set Enrichment Analysis (MSEA)

The MSEA component of the Mergeomics[40,41] pipeline was used to integrate knowledge-driven canonical pathways and data-driven coexpression modules with SCAD GWAS and

functional genomics (eQTLs, sQTLs, pQTLs), transcriptomics, or proteomics to reveal significant disease-related pathways. Before performing MSEA, the GWAS dataset was specifically corrected for linkage disequilibrium (LD) for SNPs, and a LD cutoff of r^2<0.5 was applied to remove redundant SNPs. Pathways and coexpression modules were assessed for enrichment of disease associations using a Chi-square like statistic: $\boxed{\chi = \sum_{i=1}^{n} \frac{O_i - E_i}{\sqrt{E_i + \kappa}}}$, where n denotes the number of quantile, O_i and E_i denote the observed and expected counts of positive findings (i.e. signals above each quantile point), and κ is a stability parameter to reduce artifacts from low expected counts for small marker sets. Pathways and coexpression modules were considered significant if the Benjamini-Hochberg false discovery rate (FDR)< 0.05. Hence, through MSEA, multiomics data and various functional genomic data (QTLs, pathways, coexpression networks) were integrated to capture the tissue-specific causal processes related to the disease.

Meta-MSEA

Meta-MSEA was used to integrate the various levels of omics data for meta-analysis at the pathway and coexpression module level. Meta-MSEA conducts a meta-analysis across datasets to retrieve consistent pathways and coexpression modules. Pathway enrichment Z-scores from each dataset are first estimated with MSEA (described above), and the meta p-value is estimated by integrating individual Z-scores using the Stouffer's method.

Weighted Key Driver Analysis (wKDA)

The second step of Mergeomics involves wKDA, which detects potential key drivers (KDs) of disease-associated pathways or coexpression modules based on the topology of Bayesian Networks (BNs), which include the causal and directed relationships between genes, based on genomic, transcriptomic, and proteomic data. We constructed BNs with an established

method, RIMBANet[42,43,44,45], using genetic and transcriptome data from GTEx V8 aorta, coronary, and tibial arteries, cultured fibroblast, hybrid mouse diversity panel[46,47] (HMDP) aorta, and the SCAD fibroblast transcriptome data. Protein-protein interaction (PPI) networks were curated from STRING[48]; network edges were ranked by their combined score, and the top 3% of edges were used. A KD is defined as a gene connected to a significantly larger number of disease-associated genes compared to the expected number for a randomly selected network gene within a given BN based on a Chi-square like statistic $\chi = \frac{O-E}{\sqrt{E-N}}$, where O and E represent the observed and expected ratios of genes from disease-associated gene sets in a hub subnetwork, and $E = \frac{N_k N_p}{N}$ is estimated using the hub degree N_k, disease gene set size N_p, and the order of the full network N. Statistical significance of the disease-gene enriched KDs was estimated by permuting the network gene labels 10,000 times and estimating the P-value based on the null distribution. wKDA used Benjamini-Hochberg FDR<0.05 as the significance cutoff. Regulatory networks for a subset of KDs were visualized using Cytoscape[49].

Drug Repositioning Using Pharmomics

For Pharmomics[30], network-based drug repositioning was contingent on network similarities between drug signatures and disease genes. Similarities between two gene networks matched by tissue (e.g., a drug network vs. a disease network) were determined by a distance measure derived from the mean of shortest path lengths between a drug gene signature (A) and a disease signature (B) in a given Bayesian gene regulatory network. This distance measure was adapted from a previous study using protein interaction network[50], where $\mathrm{distance}(\mathrm{B},\mathrm{A}) = \frac{1}{\|\mathrm{A}\|}\sum_{\mathrm{a}\in\mathrm{A}} \min \mathrm{b} \in \mathrm{B}\ \mathrm{distance}(\mathrm{b},\mathrm{a})$. To obtain a null distribution for shortest path lengths, we randomly selected genes in the drug or disease network 1000 times and calculated a z-score based on the mean and standard error of the null distribution. We also used a standard gene overlap-based drug repositioning via Pharmomics by assessing the direct

overlap between drug gene signatures and input genes. The overlap Jaccard score was defined as $J(A,B) = \frac{|A \cap B|}{|A \cup B|}$. Predicted drugs were considered significant if $p < 0.05$.

Results

Biological Pathways and Coexpression Modules Associated with SCAD GWAS

SCAD GWAS signals were mapped to their corresponding genes via relevant vasculature eQTLs and sQTLs and were subsequently tested for disease enrichment against 1,840 curated canonical pathways and 23,196 coexpression modules constructed from SCAD transcriptome and proteome data. Queried pathways and coexpression modules were considered significantly enriched if FDR<5%, and significant pathways and coexpression modules were merged into 73 nonoverlapping supersets. Significant pathways and coexpression modules from GWAS with biologically relevant annotations are shown in **Table 3**. SCAD GWAS signals were also mapped to their corresponding proteins via pQTLs, but no significant pathways or coexpression modules were uncovered.

Top significant gene sets associated with SCAD based on GWAS summary statistics included the positive control gene set for SCAD and positive control gene set for coronary heart disease, both curated via the online GWAS catalog[51]. SCAD and CAD patients often have contrasting phenotypes, and observational studies in SCAD have accounted for a low frequency of coincidental atherosclerotic CAD[24]. Further, as SCAD is nonatherosclerotic and spontaneous while CAD is characterized by atherosclerosis, it is possible shared genetic factors associated with both diseases may increase risk for one while serving as protective for the other. Therefore, the enrichment of both SCAD and CAD gene control sets supports the hypothesis that the genetic factors of these two cardiovascular diseases may converge in pathogenesis; however, further study into the direction of the genetic signal is warranted. For example, *PHACTR1* represents a top gene shared between these positive control gene sets and has been implicated

to confer risk for both SCAD and CAD[24]. Specifically, rs9349379, a common noncoding variant in the *PHACTR1/EDN1* locus, has been estimated to contribute to an increased SCAD risk among carriers of the common rs9349379-A allele, but this risk allele for SCAD has proven to be the protective allele for CAD and acute myocardial infarction; rs9349379-G, the minor allele, associates with increased risk for CAD and acute myocardial infarction. This genetic variant at the PHACTR1 locus has been reported to lie in a putative enhancer region for the endothelin-1 (ET-1) gene, *EDN1*. ET-1 is a potent vasoconstrictive peptide produced primarily by vascular endothelial cells and has been proposed to mediate various vascular diseases genetically linked to rs9349379 via vasoconstriction and endothelial proliferation[52]. The rs9349379-A allele, associated with increased SCAD risk, has been shown to correlate with decreased expression of plasma ET-1; this decrease in circulating vasoconstrictive ET-1 is consistent with the lack of atherosclerosis in SCAD patients. Estrogen has been proposed to downregulate the vasoconstrictive properties of ET-1[53], however further study into how sex hormones interact with risk factors to produce the sex specificity of SCAD is necessary.

Another top module identified from MSEA based on SCAD GWAS was regulation of endothelial cell proliferation (*ERI1, P3H1, APMAP, SSR4, PRDX4)*. SCAD was initially proposed to occur via an “outside-in” endothelial-intimal disruption in which blood enters the sub-intimal space. However, recently an alternative “inside-out” hypothesis in which the primary event for SCAD is a *de novo* intramural bleed leading to intimal disruption has been supported, as there was no evidence of endothelial/intimal injury while examining SCAD coronary histopathology[54]. Endothelial cells may remain intact, but our results indicate that the proliferation of this cell type is involved in SCAD as evidenced by the enrichment of the endothelial cell regulation coexpression module. It is plausible that the technology to examine the endothelial monolayer lining the lumen is currently lacking as immunohistochemistry to stain for mature endothelial markers (CD31, PECAM1) is insufficient in studying dynamic changes

and proliferation of this cell monolayer as a whole in SCAD. Additional studies utilizing real-time tracking tools to investigate the contrasting hypotheses about SCAD etiology are warranted.

The vascular extracellular matrix (ECM) provides structural and mechanical properties necessary for vessel function as well as signals that stabilize vascular cell phenotypes. Various vascular ECM molecules interact with vascular cell types to regulate gene expression[55]. Extracellular structure organization, assembly of collagen fibrils, integrin signaling, and glycosaminoglycan metabolism all represent ECM-related pathways and modules significantly associated with SCAD GWAS. Estrogen has been implicated in increasing the release of matrix metalloproteinases[56], which can result in degradation of ECM proteins and impaired structural integrity of the vessel wall, and this function of estrogen aligns with the predominantly female population impacted by SCAD. At the 1q21.3 locus, *ECM1, ADAMTSL4, and C1orf54* have been implicated in SCAD susceptibility, and these three genes represent top genes with stronger SCAD GWAS associations in the extracellular structure organization, glycosaminoglycan metabolism, and focal adhesion pathways and modules. *ECM1* encodes a glycoprotein that is secreted to interact with several ECM proteins involved in angiogenesis induced by endothelial cell proliferation[1]. *ADAMTSL4* encodes a secreted glycoprotein localized to the ECM that may contribute to cell-cell or cell-ECM adhesion[1]. *C1orf54* encodes a protein of unknown function but has been previously reported to be upregulated in a rabbit model of carotid artery aneurysms[1]. Also involved in the ECM is collagen, the main structural protein of the ECM that provides tensile strength, regulates cell adhesion, and supports cell migration and tissue development[57]. *COL4A1*, a top collagen gene within the assembly of collagen fibrils pathway, has a rare disruptive variant associated with SCAD[58]. Therefore, defects in ECM integrity and cell-cell adhesion may affect the vascular structure and contribute to SCAD development.

Mechanisms related to vascular injury, such as complement and coagulation cascades, intrinsic prothrombin activation pathway, fibrin clot clotting cascade, and acute myocardial

infarction, were also identified as significantly enriched pathways based on SCAD GWAS. The fibrin clot clotting cascade occurs immediately following vascular injury to minimize blood as enzymatic activations and protein conformational changes convert fibrinogen to fibrin and activate platelets[59]. The intrinsic prothrombin activation pathway is one mechanism in which interaction of exposed collagen in the injured vessel initiates activation of Factor XII that ultimately results in thrombin generation and blood clot[60]. The complement system is described as a proteolytic cascade in blood plasma and mediator of innate immunity, which is related to the coagulation cascade because protease-activated receptors that are activated by thrombin function as a mediator of innate immunity[33]. It is plausible that the coagulation pathway activation or inflammatory cytokine effects may lead to hematoma and subsequent coronary dissection[61], thereby supporting the enrichment of the complement and coagulation cascade in SCAD.

Curated gene sets for both low density lipoprotein (LDL) cholesterol and high-density lipoprotein (HDL) cholesterol also showed significant enrichment for SCAD GWAS association. In a case-control study involving young (≤55 years) female patients, SCAD patients displayed lower levels of LDL cholesterol and of total cholesterol compared to CAD patients[62]. Hypercholesterolemia, in which increased levels of LDL cholesterol circulate in the blood, serves as a risk factor for CAD and promotes atherosclerosis. Considering again the contrasting natures of SCAD and CAD, the level of circulating LDL cholesterol has proven to be another physiological difference between diseases, though further study is warranted to determine whether LDL or HDL cholesterol holds a significant involvement in SCAD. At the 12q13.3 locus, the rs11172113-T allele, localized to the first intron of *LRP1* which was identified as a top gene in the HDL cholesterol gene set, was observed to have higher prevalence among SCAD samples[1,3]. LRP1 is proposed to function in a variety of physiological functions, including focal adhesion disassembly and reorganization, cellular uptake of lipoproteins, reverse cholesterol transport, and vascular wall integrity[63]. Additionally, rs11172113-T and rs11172113-C serve as

risk alleles for SCAD and CAD, respectively, which further supports the dichotomy of opposing risk alleles between the two cardiovascular diseases. Disrupting *Lrp1* in vascular smooth muscle cells in mice resulted in increased susceptibility to atherosclerosis and decreased vascular wall integrity[64].

Biological Pathways and Coexpression Modules Associated with SCAD TWAS

SCAD transcriptomic data was likewise queried against canonical pathways and coexpression modules to determine gene sets enriched in the disease state, as shown in **Table 4**. Significant pathways and coexpression modules at the transcript level were merged into 207 nonoverlapping supersets. Pathways and modules enriched at the transcript level that shared similar annotation terms with GWAS pathways and modules included ECM-related pathways, mRNA metabolism, platelet activation and aggregation, cholesterol biosynthesis, the tricarboxylic acid (TCA) cycle, and the insulin signaling pathway. Various immune pathways and coexpression modules were also implicated at the transcript level, including regulation and activation of nuclear factor kappa B (NFkB), cross presentation and processing of antigens, interleukin and cytokine signaling, and cell apoptosis.

Biological Pathways and Coexpression Modules Associated with SCAD PWAS

We further queried SCAD protein data against canonical and coexpression modules, and those significantly enriched are listed in **Table 5**. Significant pathways and coexpression modules at the protein level included adaptive immune system, amino acid metabolism, complement and coagulation cascades, and regulation of the mitotic cell cycle.

Biological Pathways and Coexpression Modules Associated with SCAD GWAS and TWAS in a Meta-Analysis

We additionally performed meta-MSEA to determine shared significant canonical pathways and coexpression modules between the omics levels. Although PWAS did not share significant pathways and modules with the other data types, various significant pathways and modules were shared between GWAS and TWAS (**Table 6**). The Wnt signaling pathway, an integral pathway involved in cell fate determination, migration, and polarity as well as neural patterning and organogenesis[65], was identified as a shared annotated coexpression module between GWAS and TWAS. The Wnt signaling pathway was uniquely identified by our pathway analysis and has not yet been linked to SCAD previously. Wnt signaling is vital to heart development, patterning, and adult homeostasis and adaptation[66]. Also, a coexpression module derived from SCAD transcriptome data and annotated as neurotransmitter secretion showed significance across GWAS and TWAS. Migraines have been associated with SCAD presentation, and one hypothesis for the etiology of migraine has been the genetic variants for receptors and neurotransmitters associated with adrenergic, GABAergic, and nitroxidergic nerves[67]. Additional significant coexpression modules shared between GWAS and TWAS were annotated as regulation of endothelial cell proliferation and terms relating to the formation of blood clot, including the fibrin clot clotting cascade and the extrinsic pathway.

Key Driver Genes and Proteins Within SCAD Networks

After identifying the significant pathways and coexpression modules associated with SCAD, we next modeled multi-tissue gene regulatory Bayesian networks (BNs) and protein-protein interaction (PPI) networks to pinpoint key regulators, or key drivers (KDs), central within the disease networks. For gene regulatory networks based on significantly enriched SCAD GWAS and TWAS gene sets, BNs informed by GTEx aorta, coronary, and tibial arteries, cultured human fibroblast, mouse aorta, and the SCAD fibroblast transcriptome data were used. We queried these cross-tissue BNs using the 73 nonoverlapping pathways and modules which comprised 5,678 genes based on SCAD GWAS analysis to identify KDs. The KDs were

involved in anatomical structure development and morphogenesis (*CADM1, HOXB9, ATRNL1, AC091814.3, NTF3, ABCC9, COL18A1, IGF1, RP11-357H14.17, GPR4, APCDD1, RCAN2*), assembly of collagen fibrils (*COL1A1, COL6A1, COL6A3, ADAMTS2, COL5A1*), focal adhesion (*COL3A1, COL1A1, COL6A3, ACTN1, MYL9*), glycosaminoglycan metabolism (*ADAMTS15*), Wnt pathway and intercellular signaling (*Col18A1*), and amino acid phosphorylation and intracellular signaling cascade (*HSPA5*). The top five KDs of each pathway or module and their surrounding subnetwork genes are illustrated in **Figure 2**.

We next queried the cross-tissue BNs using the 207 significant nonoverlapping pathways and modules which comprised 12,409 genes based on SCAD TWAS analysis. Top KDs were involved in such pathways as anatomical structure development (*CADM1, HOXB9, RP11-357H14.17, NTF3, IGF1, ABCC9*), translation regulation (*HOXB9*), membrane organization and ligand binding and activation (*EDN1*), and Nerve Growth Factor (NGF) signaling (*SEC31B*). The top five KDs of each pathway or module and their surrounding subnetwork genes are depicted in **Figure 3**. The *PHACTR1/EDN1* locus was previously identified as a genetic locus associated with SCAD[24], and *EDN1* was pinpointed as a KD in our TWAS network modeling. As discussed above, *EDN1* encodes vasoconstrictive peptide ET-1 which has decreased expression in patients with the SCAD *EDN1* risk allele[52].

Shared KDs between SCAD GWAS and TWAS data include genes involved in anatomical structure development, including *CADM1, HOXB9, RP11-357H14.17, NTF3,* and *IGF1*. *HOXB9* has been proposed as a positive regulator of inflammatory molecule expression in aorta endothelial cells and a regulatory of endothelial cell turnover[68]. Low wall shear stress in the aorta was shown to activate a BMP4-HOXB9-TNF signaling pathway to initiate focal arterial inflammation[68], highlighting the contribution of HOXB9 to vascular pathophysiology. *HOXB9* has been shown to be transcriptionally regulated by estrogen, supporting the female-biased nature of SCAD[69]. CADM1 has been shown to be involved in cell-cell adhesion[70,71] and immunity[72]. In humans, *CADM1* SNP variants have been associated with vascular disease and dysfunction.

PPI networks were queried using the seven nonoverlapping modules comprised of 940 genes based on the SCAD PWAS analysis. The hub proteins central to the network, C4A and SERPING1, were both relevant to the complement and coagulation cascade and individually related to formation of tubulin folding intermediates and platelet activation signaling and aggregation, respectively (**Figure 4**). As discussed above, mechanisms leading to thrombus formation contribute to SCAD onset, thereby supporting the role of complement protein C4A and SERPING1 in SCAD at the protein level.

Querying the combined gene BNs encompassing arterial and fibroblast tissue using the consistent significant pathways and coexpression modules across SCAD GWAS and TWAS in the meta-analysis, we identified *COL18A1* as the top significant KD (**Figure 5**). *COL18A1* was prioritized as a highly ranked gene in a rare variant collapsing analysis and proposed to be a credible candidate for SCAD based on function, expression, and mouse phenotype[2]. COL18A1 is a basement membrane proteoglycan that functions in presenting the enzyme lipoprotein lipase, which breaks down triglycerides, to the luminal side of the vascular endothelium[73]. *Col18a1-/-* mice demonstrated hypertriglyceridemia from decreased lipoprotein lipase activity, and humans with homozygous deficiency of COL18A1 were observed to have higher blood triglyceride levels than controls[74]. Additionally, endostatin, a proteolytic fragment of COL18A1, inhibits endothelial proliferation and angiogenesis[75], which may be a more relevant pathway through which *COL18A1* is involved in SCAD onset. Further validation of this candidate gene is warranted as the impact of COL18A1 within SCAD is yet to be elucidated.

Prioritized Therapeutic Options Pinpointed via Network-based Drug Repositioning

Network- and gene overlap-based drug repositioning via Pharmomics using SCAD networks informed by SCAD GWAS and TWAS was performed to reveal FDA-approved drugs that may be translated to SCAD treatment (**Table 7-8**). Top candidate drugs were filtered for function within the human cardiovascular system as curated by the Pharmomics tool[30] to match

the informative and relevant tissuses for SCAD. It is important to note that while these drug repositioning techniques propose therapeutics relating to SCAD, the direction by which these proposed drugs impact SCAD must be studied further through experimental testing.

SCAD GWAS network-based repositioning (**Table 7**) identified medroxyprogesterone, etonogestrel, and contraceptive as top significant candidates for SCAD treatment. Medroxyprogesterone is a progestin hormone capable of treating menstruation problems caused by a hormone imbalance and can prevent pregnancy in an injectable form[76]. Etonogestrel is an artificial active metabolite of the synthetic progestin Desogestrel and prevents the release of luteinizing hormone which prevents ovulation, and etonogestrel subdermal implants are considered a clinically effective and safe contraceptive method[77]. Contraceptive is the general term for mechanisms that prevent pregnancy, including medications, devices, and surgical procedures. The direction by which contraceptives and hormonal treatments affect SCAD must be elucidated as pregnancy hormone modulation, infertility treatments, postmenopausal hormonal therapy, and oral contraceptives can lead to a weakened arterial media[5,6,78,79,80,81,82,83].

SCAD TWAS network-based repositioning (**Table 7**) likewise pinpointed medroxyprogesterone and etonogestrel, highlighting the candidacy of these sex-specific and hormone-based drugs on the female-biased disease. The repositioning analysis further identified rofecoxib, an anti-inflammatory cyclooxygenase-2 (COX-2) inhibitor used to treat conditions like rheumatoid arthritis and migraine, as a significant treatment candidate. COX-2 is predominantly localized in the brain, renal, and endothelial cells and is significantly upregulated in various inflammatory infections, and rofecoxib inhibits the COX-2 enzyme to contain inflammation[84]. Rofecoxib was found to be an effective and overall well-tolerated acute treatment of migraine[85], a condition commonly impacting SCAD patients[86].

Prioritized Therapeutic Options Pinpointed via Gene Overlap-based Drug Repositioning

We next used the traditional gene-overlap based matching between SCAD networks and drug signatures to identify drug candidates. Using SCAD network informed by GWAS, gene overlap-based drug repositioning (**Table 8**) also identified hormone-modulating drugs, such as contraceptive, medroxyprogesterone, progesterone receptor agonist, and etonogestrel as top significant therapeutic candidates for SCAD. Additionally, pain relief (analgesic), anti-inflammatory (COX-2 inhibitor, celecoxib), and proton pump inhibitor (omeprazole, anti-ulcerative) drugs showed significant gene overlap with SCAD GWAS network genes.

Using SCAD network informed by TWAS, gene overlap-based drug repositioning (**Table 8**) uncovered hormone (medroxyprogesterone, progesterone receptor agonist, contraceptive), proton pump inhibitor (omeprazole, anti-ulcerative, antisecretory), and clotting promotor and antihemorrhagic (thrombin, blood coagulation factor, hemostatic) drugs.

Across network and gene-overlap based drug repositioning analysis, hormones and anti-inflammatory medication are consistent findings. The gene-overlap based analysis also revealed proton pump inhibitors, which irreversibly inhibit the hydrogen/potassium pump (H+/K+ ATPase pump) in gastric parietal cells, thereby reducing gastric acid secretion[87]. A significant increase in morbidity due to cardiovascular disease and all-cause mortality was observed in patients using protein pump inhibitors[88]. Long-term use of proton pump inhibitors has been associated with cardiac dysfunction and myocardial infarction as this drug class may impair endothelial function and accelerate endothelial aging[89], so it is plausible proton pump inhibitor use facilitates SCAD precipitation. Additional experimental validation studies are needed to confirm whether these drugs promote or mitigate SCAD.

Discussion

In this multiomics integrative study of SCAD leveraging GWAS, transcriptome, and proteome data and tissue-specific functional genomics and networks, we identified various canonical pathways and coexpression modules enriched in the SCAD disease state. These

include immune function, extracellular matrix organization and integrity, and blood clot mechanisms. Significant pathways or modules were used to query vascular gene regulatory BNs and PPI networks to identify KDs and central proteins, respectively. Separate GWAS and TWAS analyses both identified *CADM1, HOXB9, RP11-357H14.17, NTF3,* and *IGF1* as KDs involved in the anatomical structure development pathway. PWAS analysis highlighted SERPING1 and C4A as KDs in the complement and coagulation cascade. In the GWAS/TWAS meta-analysis, *COL18A1* was pinpointed as the top significant KD across these two omics layers. Lastly, network- and overlap-based drug repositioning revealed various sex-specific treatment options for SCAD, including medroxyprogesterone, etonogestrel, and estradiol.

The numerous significant pathways and modules identified at specific omics layer or shared across omics levels indicate various possible mechanisms of SCAD pathogenesis that may cooperate to ultimately result in disease manifestation. Current hypotheses for SCAD etiology include i) an endothelial-intimal disruption with blood entering the sub-intimal space ("outside-in")[90], ii) a *de novo* intramural bleed in which increased pressure from the hematoma causes intimal disruption ("inside-out")[24], and iii) injury to the vascular endothelium from eosinophilic cytokines[54]. Our data-driven pathway enrichment results identified regulation of endothelial proliferation at the GWAS, TWAS, and GWAS/TWAS meta-analysis levels, emphasizing endothelial proliferation rather than injury or disruption in SCAD onset. Additionally, blood clotting-related pathways and the numerous immune-related pathways and modules identified at all omics levels support the immune and hematoma involvement. Although the traditional "outside-in" and novel "inside-out" proposals provide contrasting etiologies for SCAD, immune function, potentially by mechanism of eosinophils, is certainly involved regardless of the cause and origin of the thrombus.

Immune function is closely involved in SCAD pathogenesis and disease manifestation, and coincidental cases of various autoimmune and inflammatory conditions, such as rheumatoid arthritis, Crohn's disease, and Kawasaki disease, have been presented in SCAD patients. Also,

as myocardial infarction is one manifestation of SCAD, various resident immune cells, such as macrophages, mast cells, and innate lymphoid cells, coordinate to minimize the damaged myocardial infarction region by removing damaged cells; additionally, monocytes and neutrophils are generated and quickly recruited to the injury site. The identification of immune-related pathways and coexpression modules at the GWAS, transcriptome, and proteome levels establishes the integral role of the immune system in SCAD. Specifically, at the GWAS level, the mitogen-activated protein kinase (MAPK) pathway was significantly enriched. MAPKs are involved in all aspects of immune response, ranging from initiation of innate immunity and adaptive immunity to cell death[91], and they regulate the production of immunomodulatory cytokines which subsequently promote Th1 and Th2 responses[92]. At the transcriptome level, various aspects of the immune system were implicated by the significant pathways and coexpression modules; namely, regulation and activation of nuclear factor kappa B (NFkB), cross presentation and processing of antigens, interleukin and cytokine signaling, and cell apoptosis. NFkB is considered an integral mediator of inflammatory responses by inducing the expression of numerous pro-inflammatory genes and by contributing to the regulation of innate immune cells and T cells[93]. Interleukin and cytokine signaling also contributes to inflammatory response, and an inflammation mechanism has previously been proposed for SCAD: inflammatory mediators promote the release of cytotoxic products from eosinophils which cause injury to the vascular endothelium[94]. Apoptosis, or programmed cell death, is critical in immune system homeostasis through the deletion of self-recognizing immune cells and controlled cytotoxic killing[95]. Lastly, at the proteome level, the coexpression module annotated with adaptive immune system, endoplasmic reticulum phagosome pathway, and neurotrophin signaling was significantly enriched for SCAD. While immune dysfunction has been proposed for SCAD pathogenesis, our pathway enrichment analyses certainly revealed novel, specific immune mechanisms reflected in the various omics datasets associated with SCAD.

The sex specificity of SCAD was highlighted at every stage of analysis and across omics datasets. In pathway and coexpression module enrichment, the female hormone estrogen has been shown to modulate ET-1, involved in the *PHACTR1/EDN1* locus (SCAD positive control gene set), and also contributes to the release of matrix metalloproteases which can impair vessel wall integrity (ECM-related terms). *HOXB9,* a KD for the anatomical structure development and morphogenesis subnetwork identified by both GWAS and TWAS network analysis, is transcriptionally regulated by estrogen. Finally, network- and gene overlap-based drug repositioning for GWAS and TWAS prioritized medroxyprogesterone, etonogestrel, and contraceptive; these treatments all correspond to female hormone levels and are currently implemented to treat female conditions, such as menstruation and pregnancy prevention and menopausal side effects. If additional clinical characteristics pertaining to contraceptive usage within the GWAS or TWAS cohorts become available, we can investigate the association of contraceptive usage and type with SCAD manifestation in future studies.

To further dissect the multiomics of SCAD, we plan to query metabolites collected from the UK cohort of SCAD patients and healthy controls. Metabolomics refers to the low molecular weight compounds serving as substrates, intermediates, or products of metabolic reactions that participate in diverse molecular and cellular functions ranging from epigenetic regulation of gene expression to protein modification[96]. Adding this omics layer to our SCAD study will further facilitate the development of a holistic understanding of SCAD as each omics layer solely reflects a partial, fragmented perspective of disease development.

In summary, we analyzed and integrated multiomics SCAD datasets to identify disease-associated pathways and modules, such as immune-, ECM-, and clotting-related terms, and modeled vascular gene regulatory and PPI networks to pinpoint hub genes and proteins, respectively. We highlighted the sex specificity of SCAD development in the network key drivers and the predicted drugs from our network- and overlap-based repositioning approaches.

Experimental validation of the key genes, pathways, and drugs identified in this study is warranted.

3. Single-Cell Level Understanding of Dietary Risk Factors on an Atherosclerotic Mouse Model

Introduction

Hypercholesteremia, in which circulating LDL or total cholesterol level is elevated, serves as a major risk factor for CAD. Familial hypercholesterolemia (FH) is an autosomal inheritance primarily attributed to mutations in the low-density lipoprotein receptor (*LDLR*) gene localized to chromosome 19[12]. The LDL receptor, located on the cell surface, mediates removal of LDL particles from plasma via endocytosis. However, in the development of atherosclerotic lesions, a dysfunctional vessel wall endothelium allows for LDL to enter the vessel wall, become oxidized, and initiate monocyte recruitment. At the site of injury, monocytes differentiate into macrophages which uptake lipids and become lipid-laden and “foamy”, smooth muscle cells (SMCs) de-differentiate and proliferate, and eventually the atherosclerotic plaque expands with cholesterol deposits and toxic lipids resulting in cellular death. A fibrous cap comprised of migrated SMCs embedded in dense extracellular matrix overlies the necrotic core and serves to protect from plaque rupture and myocardial infarction. If the plaque ruptures, a thrombus forms and may occlude the vessel, potentially leading to permanent myocardial tissue damage or fatality.

In addition to the genetic basis of hypercholesterolemia through mutation of *LDLR* and other related genes, there are various environmental contributions to atherogenesis, including high intake of cholesterol diets and of dietary nutrients that metabolize to trimethylamine N-oxide (TMAO). Microbial enzymes act on choline, phosphatidylcholine, and L-carnitine from diets high in red meat and full-fat dairy products to generate trimethylamine (TMA), and TMA is metabolized into TMAO by hepatic flavin monooxygenases (FMO3)[97]. TMAO is typically cleared

by the kidney, however circulating TMAO is associated with promotion of atherosclerosis and increased risk for CAD[19]. In mice, dietary supplementation with TMAO, carnitine, or choline reduced reverse cholesterol transport, and chronic dietary L-carnitine supplementation markedly enhanced TMA and TMAO synthesis, increasing atherosclerosis[97]. However, how TMAO modifies atherosclerosis in the vasculature in the presence of high cholesterol is poorly understood.

Therefore, we modeled *in vivo* atherosclerosis in the *Ldlr-/-* mouse model and carried out single-cell RNA-sequencing (scRNAseq) analysis of the aorta to identify cell type-specific changes in gene expression and top representative functional pathways under chow, high-cholesterol (HC), and high-cholesterol with added TMAO (HC+TMAO) dietary conditions. We further predicted intercellular communications across aortic cell populations to identify key cell types and genes mediating cell-cell interactions important for atherogenesis.

Methods

Animals and Dietary Treatment

Female *Ldlr-/-* mice were divided into 3 treatment groups: (1) chow diet group (chow), age of mice: 14 weeks, (2) Cholesterol (0.5%) diet for 3 months (HC), age of mice: 22 weeks, and (3) Cholesterol (0.5%) + 0.125% TMAO diet for 3 months (HC+TMAO), age of mice: 22 weeks.

Aorta Dissection and Single Cell Dissociation

For each treatment group, we had 3 aorta pools with n=2 mice/pool, for a total of 9 independent aorta samples. Aorta samples from the above groups were collected from the lesion prone areas of aorta, including ascending aorta, aortic arch, and thoracic aorta. Aorta cell dissociation protocol was based on the Wirka et al. study[98]. Briefly, the aorta samples were cut into < 1 mm pieces and incubated in 1 ml of Hanks' Balanced Salt solution containing 2 units of liberase and 24 units of elastase at 37 ^{o}C for 1 hour. Fetal bovine serum (10% of total volume)

was then added to inactivate the enzymes. The digested cells were then passed through a 70 μm cell strainer and rinsed with 4 ml PBS. After centrifugation at 300g, 4°C, for 10 min, the supernatant was discarded. The cells were then resuspended in 4 ml PBS/0.04% BSA, spun down again, and resuspended in 100 μl PBS/0.04% BSA.

scRNAseq Library Construction and Sequencing

Approximately 16,000 cells pooled from two aortas were used for each single cell library construction. In total, 9 libraries (3 libraries/diet group) were made using the 10x Genomics Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1. The 9 libraries were then sequenced in one lane of NovaSeq S4 2x100bp at 2.4 billion reads.

scRNAseq Data Pre-processing and Quality Control

Sequencing reads were aligned to the mus musculus genome assembly GRCm38 (mm10), and gene counts were calculated using CellRanger software (v 6.0.1) (10X Genomics). Filtered feature-barcode matrices from each library were loaded into Seurat[99] (v 4.1.1) and combined to create three gene expression matrices corresponding to chow, HC, and HC+TMAO. Single cells were selected based on a threshold of between 200 and 5000 genes expressed and between 500 and 20,000 unique molecular identifiers (UMIs) counted. Single cells were further filtered with a mitochondrial gene expression cutoff of <25% and with a hemoglobin gene expression cutoff of <0.01%. Transcript counts of each gene were normalized by the total counts for that cell, and the values were multiplied by 10,000 and log transformed (LogNormalize).

Doublet Prediction

Doublets, droplets filled with two or more cells in scRNAseq, are technical artifacts that limit cell throughput and confound the analysis by leading to spurious biological conclusions. Therefore, DoubletFinder[100] was implemented to predict doublets according to each real single cell's proximity in gene expression space to artificial doublets created by averaging the transcriptional profile of randomly chosen cell pairs. DoubletFinder first simulates artificial doublets from existing scRNAseq data, and then merge and preprocess real and artificial data using the Seurat pipeline across the original and merged datasets. Subsequently, DoubletFinder performs dimensionality reduction using principal component analysis (PCA) and detects the *k* nearest neighbors for every cell in the principal component (PC) space. With this calculation, each cell's proportion of artificial nearest neighbors (pANN) is computed, and DoubletFinder relies upon the assumption that real and artificial doublets co-localize in PC space to predict real doublets as cells with the top *n* pANN values (*n* = total number of expected doublets).

Cell Clustering and Cell Type Annotation and Distribution

Using Seurat, cells were projected onto two dimensions using T-distributed Stochastic Neighbor Embedding (t-SNE) or Uniform Manifold Approximation and Projection (UMAP) and assigned into clusters using Louvain clustering. Seurat's FindAllMarkers, which employs a non-parametric Wilcoxon rank sum test that is run within each sample and calculates a meta p-value across all samples to assess the significance of each gene's specific membership to a cluster, was used to determine the cluster-specific marker genes consistent across samples. Genes tested were required to be expressed in at least 10% of the cells in one cluster. These markers were manually evaluated for convergence on known cell type marker genes and used to identify cell types. Cell clusters labeled as "unknown" displayed a combination of markers from known cell types. The cell type proportions for all cell types within a diet treatment group was calculated and significant differences between the same cell type across dietary conditions were evaluated by one-way ANOVA and Tukey's Honest Significant Difference (HSD) post-hoc test.

Differential Gene Expression Analysis and Pathway Enrichment

The FindMarkers function from Seurat, which similarly to the FindAllMarkers function performs a non-parametric Wilcoxon rank sum test, was used to compare gene expression between dietary treatment groups to identify differentially expressed genes (DEGs). Genes tested were expressed in at least 10% of the single cells in one cell cluster. Pathway enrichment analysis, utilizing pathways from KEGG[33], BioCarta[35], and Reactome[34], and Hallmark[101] gene sets from the Molecular Signatures Database (MSigDB), was performed using DEGs with adjusted p-value<0.05 and log fold-change (logFC) >0.1. However, for clusters that had insufficient number of DEGs at adjusted p-value<0.05 to identify meaningful significant pathways, a DEGs with a threshold of unadjusted p-value<0.01 were inputted to predict suggestive pathways for these cell types[102]. Significant enrichment of pathways was based on a hypergeometric test followed by multiple testing correction with the Benjamini-Hochberg method. logFC of each pathway was calculated by summing all the individual pathway member gene logFC values.

Cell Lineage Trajectory Analysis

Slingshot[103] was used to perform trajectory analysis on smooth muscle cell (SMC) subpopulations and on macrophage subtypes separately in the UMAP space. To construct the global lineage structure for trajectory analysis, Slingshot uses a cluster-based minimum spanning tree and fits smooth branching curves to the predicted lineages based on simultaneous principal curves. The designated starting cell cluster guided the trajectory prediction via a semi-supervised approach.

Intercellular Interaction Analysis

Nichenet[104] predicts ligand-target connections between interacting cell types by combining cell expression data with prior knowledge on signaling and gene regulatory networks. In addition to predicting intercellular communications, Nichenet also predicts how ligands influence target gene expression in receiver cell types. Curated biological knowledge about ligand-target signaling paths, signal transduction, and gene regulatory interactions were integrated into weighted networks which were optimized to weight each data source to maximize the accuracy of ligand-target predictions. Regulatory potential scores between all pairs of ligands and target genes were calculated, and a pair was given a high regulatory potential score if regulators of the target gene were downstream of the signaling network of the ligand. The Pearson correlation coefficient between prior regulatory potential scores and target genes was used to indicate the ability of each ligand to predict target genes; better predictive ligands were ranked higher. In our study, the modulated SMC and endothelial cell types were designated as the "receiver/target" cell populations, and all the known cell types, including modulated SMC and endothelial, served as the "sender" populations to infer which cell types and genes regulate SMCs and endothelial cells.

Results

Cell Type Identification in the Ldlr-/- Aorta

A single-cell digital gene expression matrix across the 9 aorta samples was generated using Seurat and projected onto two dimensions using UMAP to define cell clusters. We detected ten clusters containing cells sharing similar gene expression patterns (**Figure 5a**). Known aortic cell types recovered from the *Ldlr-/-* mouse models included SMC, modulated SMC, fibroblast, macrophage, pericyte, and endothelial cells. Top cell type-specific markers are illustrated in **Figure 5b-e**. Previously known cell markers, such as *Acta2* for SMC[105], *Spp1* for modulated SMC[98], *Notch3* for pericyte[98], *Egfl7* for endothelial[106], *Cd14* for macrophage[107], and *Pi16* for fibroblast[108], demonstrated cluster-specific expression patterns, highlighting the ability

of our data-driven method to retrieve known cell types. In addition to these known cell types, we also found four cell clusters which expressed a combination of known aortic cell type markers and were designated “unknown 1-4”. We excluded the possibility of these unknown clusters being doublets based on the doubletFinder analysis. Therefore, these unknown clusters were included in downstream analyses. Because the designated unknown clusters expressed combinations of known cell type markers from SMC, fibroblast, and endothelial, their top markers proved to be less specific.

Identification of Subtypes of Modulated SMC in the Ldlr-/- Aorta

Three types of modulated SMC have been proposed as a result of SMC phenotypic modulation during atherosclerosis[98]: (1) pro-inflammatory, dysfunctional macrophage-like cells that are characterized by upregulation of the known macrophage marker Lgals3[109], (2) protective ECM-producing SMCs that may contribute to the fibrous cap that prevents plaque rupture[110,111], and (3) a mesenchymal stem cell-like population of unknown significance[109]. We subclustered the modulated SMC cell type and produced three distinct subclusters, entitled Modulated SMC 0, 1, and 2 (**Figure 6a-b**). SMC phenotypic switching is associated with decreased expression of SMC markers, and therefore the majority of SMCs within atherosclerotic lesions cannot be identified using traditional SMC markers, but may rather express markers of macrophages (LGALS3, CD11b, CD68), myofibroblasts (ACTA2), and mesenchymal stem cells (SCA1)[109,110]. Our inspection of gene expression patterns of potential modulated SMC subtype markers for macrophage, myofibroblast, and mesenchymal stem cell did not produce clear patterns in our modulated SMC subpopulations. Unfortunately, a comprehensive resource detailing the markers specific to various types of phenotypically switched SMCs is lacking and the precise lineage of lesion cells remains unknown[111].

Wirka et al.[98] hypothesized that SMC phenotypic modulation occurs along a single trajectory rather than separating into distinct lineages. They defined SMCs undergoing

phenotypic modulation as 'fibromyocyte,' a fibroblast-like cell derived from contractile SMCs with decreased SMC gene expression and increased protective properties at the fibrous cap; though similar to the myofibroblast characterization, the fibromyocyte term highlights the cell origin as smooth muscle myocytes rather than as fibroblast. Although fibromyocytes express the macrophage marker *Lgals3*, they otherwise lack expression of all other top macrophage markers and instead can be identified by the *Lum* gene, a fibroblast marker. In our three modulated SMC subclusters, we investigated the expression of fibromyocyte markers identified by Wirka et al. across diet conditions (**Figure 6c**). The expression patterns of the fibromyocyte markers in our modulated SMC clusters 1 and 2 closely aligned with those of the Wirka et al. fibromyocyte characterization. Namely, *Cnn1* and *Tagln* showed decreased expression in modulated SMC clusters 1 and 2 across all diet conditions, whereas expression of *Lgals3, Lum, Fn1, Tnfrsf11b, Col1a1* and *Dcn* was upregulated in modulated SMC clusters 1 and 2 in the HC and HC+TMAO diet conditions. We further performed pseudotime trajectory analysis to predict the lineage of SMC and modulated SMC single cells **(Figure 6d**). Originating from the SMC cluster, the trajectory extended to modulated SMC cluster 0 and subsequently bifurcated into modulated SMC clusters 1 and 2. As modulated SMC cluster 0 had relatively higher average expression of *Cnn1* and *Tagln* (**Figure 6c**), it is possible that this cluster represents an earlier timepoint in the SMC phenotypic transition compared to clusters 1 and 2. Together, these data support the Wirka et al. proposal that SMC phenotypically transition along a single trajectory into fibromyocytes.

We further illustrated the expression of top modulated SMC subcluster-specific markers in **Figure 6e**. Namely, *Tcap* served as a subcluster-specific marker for modulated SMC cluster 0; *Tcap* was upregulated in the transcriptome of a high fiber diet from the Human Microbiome Project[112] and is considered to serve a preventative role in various heart diseases, including hypertrophic cardiomyopathy and dilated cardiomyopathy[113,114]. *Gas7*, associated with metabolic traits like obesity in humans and mice[115,116], was subcluster-specific in expression for modulated

SMC cluster 1. Specifically, male mice transgenic for the *GAS7* human gene demonstrated significantly reduced body weight, triglycerides, and unesterified cholesterol as well as significantly increased HDL[116]. For modulated SMC cluster 2, *Col11a2* was highly and specifically expressed. *Col11a2* is an ECM gene involved in collagen biosynthesis and cross-linking[105] and may contribute to the protective fibrous cap of the atherosclerotic plaque.

Identification of Subtypes of Macrophages in the Ldlr-/- Aorta

We likewise subclustered the macrophage cell type as undifferentiated macrophages (M0) undergo polarization to distinguish into M1-like and M2-like macrophages (**Figure7a-b**). "Classically activated" M1 macrophages within innate immunity occur in inflammatory environments and in response to tissue injury whereas "alternatively activated" M2 macrophages are considered anti-inflammatory and involved in wound healing[117]. The expression of previously known M1 and M2 macrophages[118,119] are illustrated in **Figure 7c-d**, confirming the reliability of our data-driven approach in distinguishing the differentiated M1 and M2 macrophage subtypes. Top macrophage subcluster markers for M0, M1, and M2 included *Fermt2, Gngt2,* and *Cd163*, respectively (**Figure 7e**). FERMT2 is a kindlin molecule regulating integrin function and activation and has been shown to be downregulated in atherosclerotic plaques and associated with SMC-rich plaque markers[120]. *Gngt2* was found to be increased more than two-fold change in M1 and decreased more than two-fold change in M2 macrophages[118], thereby establishing *Gngt2* as an M1 marker. *Cd163* is a well-established M2 macrophage marker, and CD163 is necessary for effective hemoglobin clearance after plaque hemorrhage[121].

Cell Proportion Changes between Dietary Conditions

We further observed the cell type proportions at each diet condition to determine significant differences for particular cell types between chow, HC, and HC+TMAO (**Figure 8a**).

Given the cellular composition of the aorta, we expected SMCs (53.6% of all cells tested) and fibroblasts (18.6% of all cells tested) to compose the majority of the single cell population. Since modulated SMC are involved in atherosclerotic SMC phenotypic modulation, we anticipated the number of phenotypically modulated SMCs to be proportionally higher in the HC and HC+TMAO groups compared to the chow group, but there was only a significant difference in the proportion of modulated SMCs between HC (5.1% of HC single cells) and chow (0.03% of chow single cells).

As we also subclustered the modulated SMC and macrophage cell types, we further examined differences in the proportional number of subcluster cells across diets (**Figure 8b-c**). There were no statistically significant differences in the proportions of modulated SMC subclusters across diet conditions (**Figure 8b**). However, modulated SMC subcluster 2 consisted of 2.78% of the chow modulated SMC population, whereas this subcluster comprised 23.8% and 13.7% of the modulated SMCs in HC and HC+TMAO, respectively, suggesting that modulated SMC subcluster 2 is more atherogenic.

Because M1 macrophages function in an inflammatory environment, we expected the number of M1 macrophages to be proportionally high in the HC and HC+TMAO diets compared to chow. Indeed, the proportions of M1 macrophages in the HC (35.7%) and HC+TMAO (33.0%) conditions were significantly higher than that in chow (2.6%) (**Figure 8c**).

Cell Type-specific DEGs and Perturbed Pathways Between Dietary Conditions

To determine cell type-specific genes and pathways that may contribute to atherogenesis and disease development, we identified differentially expressed genes (DEGs) between HC and chow, HC+TMAO and HC, and HC+TMAO and chow at FDR<0.05. We annotated DEGs with biological pathways curated from KEGG[33], Reactome[34], Biocarta[35], and Gene Ontology[122] databases to propose cell type-specific mechanisms affected by CAD risk diets.

Between HC and chow, we identified diverse pathways involved in atherogenesis, including smooth muscle contraction (SMC, Unknown 3), extracellular structure and collagen fibril organization (SMC, modulated SMC, fibroblast, Unknown 1, Unknown 4), complement cascade (fibroblast), regulation of cell proliferation and endothelial migration (endothelial), immune and inflammatory response (macrophage), respiratory electron transport (pericyte), and protein translation (Unknown 2, Unknown 3, Unknown 4) (**Table 9**).

Comparing HC+TMAO and HC diets, regulation of cell proliferation (SMC), regulation of apoptosis and the apoptotic process (SMC, modulated SMC), vasculature development (endothelial), metabolism of mRNA and proteins (modulated SMC, endothelial, macrophage), focal adhesion and extracellular structure organization (Unknown 1), and protein translation (fibroblast, Unknown 2, Unknown 3, Unknown 4) were suggested as pathways affected by TMAO treatment (**Table 10**).

Lastly, between HC+TMAO and chow diets, we identified extracellular structure organization (SMC, fibroblast, Unknown 3, Unknown 4), integrin/ECM binding (modulated SMC, Unknown 1), collagen fibril organization (modulated SMC), platelet degranulation (fibroblast), hemostasis (endothelial), cell proliferation (endothelial), immune activation and macrophage migration (macrophage), and protein translation (Unknown 2) (**Table 11**).

The DEGs and pathways identified from the above dietary comparisons represent functionally relevant mechanisms in atherosclerosis. Between every comparison, extracellular structure or fiber organization was implicated in the fibroblast cell type. Fibroblasts contribute to remodeling of the ECM, and within atherosclerosis, they function to regulate inflammation, ECM organization, collagen production, and plaque structural integrity[123]. The consistent enrichment of ECM-related pathways in fibroblast suggests continual contribution of this cell type to the remodeling of the vascular wall and ECM as well as to the formation of the fibrous cap when HC or HC+TMAO is added to the diet. When comparing the two experimental groups to the baseline chow diet, we found extracellular organization and binding pathways involved in the

SMC and modulated SMC cell types, whereas between HC and HC+TMAO groups, regulation of apoptosis and the apoptotic process were enriched in SMC and modulated SMC, respectively. Our results indicate that the *Ldlr* genetic mutation and HC diet together precipitate SMC and modulated SMC transition towards an ECM-synthesizing, protective phenotype.

Few DEGs were identified in the macrophage cell type between HC and HC+TMAO likely due to a similar inflammatory plaque state between the two dietary conditions, but immune and macrophage activation were identified as significant DEG pathways when comparing each diet to chow. The six DEGs in macrophage when comparing HC and HC+TMAO include *mt-Nd3* (down), *AY036118* (down), *Nmt1* (down), *Rpl10* (up), *Tuba1c* (down), and *Gm42418* (down).

The number of cell type-specific and shared DEGs between cell types when comparing two dietary conditions is depicted in **Figures 9, 10,** and **11**. *Spp1* represents the DEG shared between the greatest number of cell types when comparing HC v. chow (up in all cell types) and HC+TMAO v. chow (up in all SMC, modulated SMC, fibroblast, endothelial, Unknown 1-4) (**Figure 9, 11**). Spp1, a pro-inflammatory cytokine, has been proposed to contribute to both plaque formation and instability[124]. Between HC+TMAO and HC conditions, *mt-Nd3* (down), a subunit of the respiratory chain complex I involved in NADH dehydrogenation and electron transfer, was the DEG shared between all cell clusters except Unknown 2 (**Figure 10**). *mt-Nd3* gene expression has been associated with a higher histological severity of hepatic steatosis, and variants in MT-ND3 yielded a significant association with body-mass index in a gene-based meta-analysis burden test[125]. Mitochondrial genetic variation may contribute to cardiovascular disease susceptibility via alterations in oxidative phosphorylation impacting vascular function[126]. Furthermore, various ribosomal (*Rpl-, Rps-*) genes represented shared DEGs (up) across cell types between HC+TMAO and HC, and dysfunction of ribosomal proteins has been linked to cardiovascular disease progression since ribosomal proteins function in a variety of biological processes, including cell proliferation, differentiation, apoptosis, and DNA repair[127].

Intercellular Communication Between Dietary Conditions

We aimed to investigate how other cell types may interact with modulated SMC and affect downstream signaling in modulated SMC single cells between the HC and HC+TMAO dietary conditions. Based on the Pearson correlation coefficient that represents the ability of each ligand to predict modulated SMC target genes, the top prioritized ligands included Jag1, Plau, Edn1, Pdgfd, Cd320, Icam2, F11r, Lrpap1, Nrtn, Tnfsf9, and Gmfb (**Figure 12a**). We further identified the main cell type in which each ligand is expressed; namely, Jag1, Edn1, Ican2, and F11r were more highly expressed in the endothelial cell type; Plau, Tnfsf9, and Gmfb were more highly expressed in the macrophage cell type; Lrpap1 was more highly expressed in the fibroblast cell type; Pdgfd, Cd320, and Nrtn were more highly expressed in the SMC cell type (**Figure 12b**). **Figure 12c** illustrates the differential expression of genes that encode for the prioritized ligands in a particular cell type between HC and HC+TMAO conditions. Specifically, upregulated genes by the addition of TMAO to the diet include *Nrtn* in SMC, *Jag1* in fibroblast, *Edn1, Icam2,* and *Gmfb* in endothelial, and *Plau, Lrpap1,* and *Tnfsf9* in macrophage. Downregulated genes include *Jag1, Edn1, and Pdgfd* in SMC, *Nrtn* in fibroblast, *Jag1, Pdgfd, F11r, and Nrtn* in endothelial, and *F11r* and *Gmfb* in macrophage. The putative downstream target genes, expressed by modulated SMC, of the prioritized ligands include *Sh3pxd2a, Rpl18, Kdm6b, Sod3,* and *Csrnp1* (**Figure 12d-e**). Endothelial JAG1 signaling promotes inflammatory leucocyte recruitment and atherosclerosis[128] and was predicted to affect downstream signaling of *Sh3pxd2a*, a previously identified genome-wide significant CAD locus[129], in modulated SMC. Edn1 from endothelial cells was predicted to target *Sod3*, extracellular superoxide dismutase 3, in modulated SMC. This antioxidant-related gene in the aorta demonstrated protection against endothelial dysfunction[130]. Further, decreased SOD3 binding to endothelial cells was associated with increased incidence of hypertension and increased risk of ischemic heart disease[131,132]. Higher levels of EDN1, a vasoconstrictive peptide previously linked to atherosclerosis and hypertension, was found in patients carrying the CAD risk allele in the *PHACTR1/EDN1*

locus[133]. The predicted ligand-target relationship between endothelial EDN1 and modulated SMC *Sod3* supports the protective properties of modulated SMC in the disease state of our study; with higher levels of endothelial EDN1 associated with atherosclerosis affecting downstream modulated SMC *Sod3* expression, modulated SMCs may function protectively via *Sod3* to impede additional endothelial damage, but additional studies are warranted to support this hypothesis. Additionally, Icam2, intercellular adhesion molecule 2, was a prioritized endothelial ligand predicted to target *Kdm6b* in modulated SMC. Hypomethylation in the untranslated region of *ICAM2* in young CAD patients within an Indian population implied increased adhesion and aggregation processes occurring during atherosclerosis which contribute to endothelial dysfunction[134]. *Kdm6b* is the predicted modulated SMC target gene for numerous prioritized ligands. Although *Kdm6b* has not yet been studied in SMCs, macrophage Kdm6b regulates the pro-fibrotic signature of peritoneal foam cells[135], and a myeloid Kdm6b deficiency was shown to result in more advanced atherosclerosis[136].

We also examined how known aorta cell types may interact with the endothelial cell type and affect downstream signaling in endothelial single cells between the HC and HC+TMAO dietary conditions. The top prioritized ligands included Anxa1, Ntn1, Itgam, Vegfc, Col2a1, Il1rn, Dll4, Hgf, Il1a, Edn1, Vegfa, Pdgfb, Nampt, Dll1, Psen1, Il1b, Apoe, Jag1, Bdnf, and Fgf2 (**Figure 13a**). We identified the main cell type in which each ligand is expressed; namely, Jag1, Edn1, Icam2, and F11r were more highly expressed in the endothelial cell type; Anxa1, Col2s1, Hgf, Bdnf, and Fgf2 were more highly expressed in the modulated SMC cell type; Ntn1, Edn1, Vegfa, Dll1, and Jag1 were more highly expressed in the endothelial cell type; Itgam, Il1rn, Il1a, Pdgfb, Nampt, Il1b, and Apoe were more highly expressed in the macrophage cell type (**Figure 13b**). **Figure 13c** illustrates the differential expression of genes that encode for the prioritized ligands in a particular cell type between HC and HC+TMAO conditions. The putative downstream target genes, expressed by endothelial cells, of the prioritized ligands include *Hes1, Ptgs2,* and *Igfbp5* (**Figure 13d-e**). *Ptgs2*, prostaglandin-endoperoxide synthase 2 or

cyclooxygenase-2 (COX2), mediates prostaglandin production, which can induce an inflammatory response[137], and has been shown to promote metalloproteinase-induced plaque rupture in symptomatic lesions[138]. Numerous prioritized ligands were predicted to target this endothelial target gene, and *Ptgs2* was previously identified as a hub gene in a network of inflammatory genes associated with CAD[137]. The connections between endothelial *Ptgs2* and ligands from modulated SMC, endothelial, fibroblast, and macrophage highlight endothelial inflammation as a process involved when TMAO is added to the HC diet. *HES1,* an endothelial target gene functioning within the Notch pathway, has been proposed to play a protective role with regards to inflammation and oxidative stress, and dysregulation of the Notch pathway has been associated with atherosclerosis and chronic obstructive pulmonary disease (COPD). Particularly in COPD, which is characterized by chronic lung inflammation, *HES1* reduction has been shown in the endothelium. Il1b, an inflammatory M1-associated marker, was predicted to affect *Ptgs2* and *Igfbp5* in endothelial cells. IGFBP5, insulin-like growth factor-binding protein 5, was observed to be downregulated in advanced atherosclerotic plaques, and may repress endothelial inflammation[139]. In our study, average expression of athero-protective *Hes1* was significantly higher in the HC diet compared to the HC+TMAO diet (data not shown), implying the addition of TMAO may exacerbate the atherosclerotic inflammatory state. Interestingly, average expression of *Ptgs2* and *Igfbp5* was also significantly higher in the HC diet compared to HC+TMAO, suggesting that while the atherosclerotic inflammatory environments persist in both HC and HC+TMAO dietary conditions, pro- and anti-inflammatory gene expression may vary between states.

Discussion

High-throughput single cell sequencing analysis of the *Ldlr-/-* mouse aorta after various dietary treatments allowed us to identify aorta cell types and subclusters to study the transcriptomic architecture of lesion-prone areas of the aorta under control physiological and

atherosclerotic conditions. Our analyses of cell proportion changes, cell type-specific and shared DEGs across cell types, and perturbed biological pathways in response to dietary intervention collectively highlight the varying degrees of susceptibility of the aortic cell types to atherogenesis. HC and HC+TMAO dietary conditions in the *Ldlr-/-* mouse were both capable of establishing atherogenic environments as shown by their cellular compositions (increased M1 macrophages and modulated SMCs), DEGs when compared to chow *(Spp1* (up), *Vcam1* (up), *Lgals3* (up*),* and their perturbed pathways (extracellular structure organization, cell proliferation, immune and inflammatory response). The contribution of TMAO, associated with promoting atherosclerosis, to the disease state transcriptome was demonstrated in DEG analyses comparing HC+TMAO and HC conditions, including identification of enriched biological pathways (apoptosis regulation, extracellular structure organization, mRNA and protein metabolism) and intercellular communications targeting such DEGs, such as endothelial Edn1 targeting modulated SMC DEG *Sod3*.

Our data-driven single cell analysis maintained the Wirka et al. fibromyocyte model[98] for modulated SMC in which SMCs follow a single trajectory to phenotypically transition into a fibroblast-like cell. Because Wirka et al. and our study both relied on single cell transcriptomic data of *ApoE-/-* and *Ldlr-/-* mouse atherosclerosis models, our findings on the modulated SMC phenotype may be more reliable than previous knowledge-based conclusions. While our data-driven modulated SMC observations refuted the canonical belief that SMCs phenotypically transition into multiple lineages, our macrophage subclustering supported the known macrophage polarization from M0 into M1 and M2 types. Particularly, the proportion of the M1 inflammatory macrophage subtype was significantly greater in HC and HC+TMAO cellular compositions compared to chow, thereby highlighting the inflammatory environment of atherosclerotic plaques.

With the addition of TMAO to the HC diet, genes pertaining to apoptotic pathways in SMC and modulated SMC were identified. *Lgals1,* a DEG upregulated in SMCs by TMAO and

functioning in the regulation of apoptosis pathway, encodes galectin-1 (Gal-1) which can promote resolution of acute and chronic inflammation by influencing innate and adaptive immune responses[140]. Gal-1 has been shown to be upregulated during SMC growth and proliferation[141] and to inhibit cell apoptosis[142]. *Ptgis* encodes prostacyclin synthase and was likewise upregulated between HC+TMAO and HC in the regulation of apoptosis pathway; the exact function of this gene in SMC apoptosis remains unclear, but *Ptgis* has been studied in the context of atherosclerosis. In a previous study utilizing the *Ldlr-/-* mouse model, prostacyclin synthase, believed to have an athero-protective effect, was significantly increased after 8 weeks on a high-fat diet but significant reduced after 16 weeks when compared to the control[143]. The *Ldlr-/-* mice in our study were treated for 3 months via dietary intervention, and our results align with the prostacyclin synthase expression profile at the earlier 8-week aortic atherosclerotic lesion[143] and serve as an additional timepoint regarding the timeline of prostacyclin synthase expression changes in atherosclerosis. Interestingly, though *Ptgis* may be athero-protective, high activity of this gene has been previously observed in the failing hearts of rats and humans[144]. Additionally, *Btg1* was downregulated between HC+TMAO and HC relating to the regulation of apoptosis pathway, and BTG1 has been shown to induce apoptosis within a cancer context[145,146,]. From the enrichment of such genes as *Lgals1, Ptgis,* and *Btg1* in the apoptosis regulation pathway for SMC and understanding their putative functions, it is possible that programmed cell death is hindered or dysfunctional in SMCs with the addition of TMAO to the HC diet, though additional validation is certainly warranted to substantiate this hypothesis.

We acknowledge the limitations of a mouse study for atherosclerosis because mice tend to develop atherosclerotic lesions in the aorta and carotids while humans mostly develop lesions in coronary arteries[147], and myocardial infarction is rare in mice. Despite these limitations of the mouse model, the mouse genome still shares 95% of the protein-coding genes with that of humans, and this animal model can nonetheless provide insight into disease mechanisms and physiological perturbations that can be translated to human study.

In conclusion, our *Ldlr-/-* mouse atherosclerosis models treated with varying dietary conditions highlighted the genetic and dietary impacts on the mouse transcriptome. Our cell type-specific findings contribute to the understanding of CAD pathogenesis and can direct future mechanistic studies by providing insight on cell type-specific target genes and pathways in response to atherogenic stimuli.

4. Conclusions and Future Directions

In our multiomics integrative SCAD human study, we identified numerous canonical pathways and coexpression modules enriched in SCAD pathology, including extracellular matrix organization, clotting mechanisms, and various immune functions. Significant pathways and coexpression modules informed arterial and fibroblast gene regulatory BNs, and we pinpointed KDs *CADM1, HOXB9, RP11-357H14.17, NTF3,* and *IGF1* as shared between our separate SCAD GWAS and TWAS analyses. PWAS KDs included SERPING1 and C4A in the complement and coagulation cascade. *COL18A1* was the top significant KD in our GWAS/TWAS meta-analysis. Further, network- and overlap-based drug repositioning based on GWAS and TWAS network modeling revealed various sex-specific treatment options for SCAD, including medroxyprogesterone, etonogestrel, and estradiol.

In our single cell dietary study of the *Ldlr-/-* mouse aorta, we examined the transcriptomic architecture of lesion-prone areas of the aorta under varying dietary conditions. HC and HC+TMAO dietary conditions in the *Ldlr-/-* mouse demonstrated atherogenic properties as shown by their cellular compositions (increased M1 macrophages and modulated SMCs), DEGs when compared to chow *(Spp1* (up), *Vcam1* (up), *Lgals3* (up*),* and their perturbed pathways (extracellular structure organization, cell proliferation, immune and inflammatory response). The addition of TMAO to the HC diet highlighted apoptosis-related pathways as relevant to SMC and modulated SMC cell types, and mt-Nd3 (down) was the DEG shared

between the greatest number of cell types. Additionally, intercellular communications targeting modulated SMC and endothelial DEGs revealed endothelial Edn1 to target *Ptgs2* in endothelial and *Sod3* in modulated SMC.

In both cardiovascular disease networks studies, EDN1 was highlighted as a key driver gene in SCAD TWAS network modeling and as a top prioritized endothelial ligand influencing modulated SMC and endothelial downstream target genes. EDN1 was highlighted at the transcriptomic level in both SCAD and CAD studies, and while EDN1 has been implicated in the pathogenesis of these two cardiovascular diseases, we propose EDN1 as an integral molecule in the continued development of both cardiovascular diseases. SCAD and CAD demonstrate opposite risk alleles at the *PHACTR1/EDN1* locus[24], but the role of EDN1 within the atherosclerotic plaque remains unknown. While we determined EDN1 to impact downstream antioxidant gene expression in modulated SMCs and inflammatory signals in endothelial cells, further study into the atherogenic effects of EDN1 is warranted to determine the potential of this peptide as a therapeutic target. Several pathways identified in our SCAD and CAD studies converged, including those related to ECM organization, immune function, and cell proliferation. Dysfunctions in such pathways certainly compromise vascular wall integrity and can contribute to SCAD onset or CAD plaque instability.

For our SCAD study, we plan to integrate SCAD metabolomics to further clarify the holistic view of the disease, identifying metabolic pathways and molecules centrally involved in relevant tissues that may contribute to disease when perturbed. We also plan to examine the effect that various predicted drugs exert on disease susceptibility, particularly broadly used medications such as contraceptives. Also, for both cardiovascular disease studies, we aim to experimentally validate genes prioritized in our analysis that we hypothesize may be central in disease development and progression and observe the direction of effect for each disease. Overall, we aim to continue dissecting how genetic and external environmental factors contribute to pathogenesis in both cardiovascular diseases.

Tables

Table 1. Clinical characteristics of the French, UK, and Mayo Clinic SCAD study populations.

Q1: 25% quantile, Q3: 75% quantile, FMD: fibromuscular dysplasia, BMI: body-mass index, HTN: hypertension, T2D: type 2 diabetes.

Study	French Study		SCAD-UK Study I		SCAD-UK Study II		Mayo Clinic Study	
	Cases (DISCO)	Controls (3C-Study)	Cases	Controls	Cases	Controls	Cases	Controls
Type	Clinical based	Population based	Clinical based	Population based	Clinical based	Population based	Clinical based	Healthy volunteers
Inclusion criteria	age> 18, 1) retrospective with a diagnostic of SCAD made from 2010, or 2) prospective at the time of hospitalisation during which the diagnosis of SCAD was made.	Geographic sampling	SCAD confirmed on invasive angiography		SCAD confirmed on invasive angiography		SCAD confirmed by angiogram	No reported SCAD
Total (n)	313	1487	383	1925	139	815	506	1549
Women (n,%)	285 (91)	876 (58.9)	361 (94.2)	1815 (94.3)	115 (82.7)	665 (81.6)	484	1477
Age at study inclusion (Median, Q1,Q3)	51, 44, 59	74.36 ± 5.5 [65 - 94]		56,49,62		56, 48, 61	46.6 ± 9.2	64 ± 14.5
Age at SCAD (Median, Q1/Q3)	52.2, 44.55, 60	NR	47, 41, 52	NR	49.0, 43, 54	NR	46.6, 39, 53	NR
SCAD Type (1, 2, 3)	49, 237, 32	NR		NR		NR	Single vessel-397; Multi-vessel-87	NR
FMD (Yes, No, NA)	140, 152, 21	NR	104,108,171		20,71,48		175, 140, 169	unknown
BMI (kg/m2)	**58**	NA	median 25 (23,29) 6 missing values		27 (23,31) 1 missing value)	NA	26.0 ± 5.9	unknown
HTN (n,%)	96 (30.7)	1171 (78.7)	94 (24.6; 1 missing value)		25 (18.0)	NA	157 (32.4)	unknown
T2D (n,%)	12 (3.2)		7 (1.8)		2 (1.4%)	NA	14 (2.9)	unknown
Migraine (n,%)	76 (28.4)		176 (48,9; 23 missing values)		67 (48.2; 2 missing values)	NA	175 (36.2)	unknown
Smoking (Non/Current/Ex)	Ever 220/93		266/14/103		94/9/36 (67, 6, 26)	NA	343/ 12 /129 (71 /2.5/26.7)	unknown

Table 2. Clinical characteristics of the CanSCAD/MGI, DEFINE-SCAD, and VCCRI SCAD study populations.

Q1: 25% quantile, Q3: 75% quantile, FMD: fibromuscular dysplasia, BMI: body-mass index, HTN: hypertension, T2D: type 2 diabetes.

Study	CanSCAD/MGI Study		DEFINE-SCAD Study		VCCRI Study I		VCCRI Study II	
	Cases	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**
Type	Clinical based	Population based	Clinical based	Clinical based	Clinical based	Population based	Clinical based	Clinical Based
Inclusion criteria	SCAD diagnosis was confirmed on coronary angiography by the UBC core laboratory research team, and categorized according to previously established Saw classification	Age, Sex, PC (PC1-PC3) matched controls	SCAD confirmed on invasive angio	Vascular disease excluded on history and physical exam. Also matched to SCAD cases by age, BMI, sex.	SCAD confirmed by angiogram	No reported SCAD	SCAD confirmed by angiogram	No Reported SCAD
Total (n)	357	2125	42	153	88	1127	85	111
Women (n,%)	315 (88.2%)	1873 (88.1%)	41 (97.6%)	153 (100%)	80, 90.9%	672, 59.6%	83, 97.6%	46, 41.4%
Age at study inclusion (Median, Q1,Q3)	53, 46, 60	53, 46, 61	49, 41.5, 53.75	50 (43-58)	50, 44, 59	all >70 years old	52, 48, 60	61, 52, 67
Age at SCAD (Median, Q1/Q3)		NR	45.5, 36, 50.25 6 missing values	NA	44, 39, 52	NA	49, 43, 56	NA
SCAD Type (1, 2, 3)	117,193,36	NR	4; 32; 1; NA: 5	NR	32, 50, 4, (2 NR)	NA	29, 48, 6	NA
FMD (Yes, No, NA)	149,123,85	NR	31, 10, 1	NR	14, 32, 42	NR	10, 22, 53	NR
BMI (kg/m2)	median 25.5 (22,30;9 missing values)	NA	24.55 (22, 29.15)	23.85 (20.98, 26.6)	26.32	27.5	27.04	NR
HTN (n,%)	108 (30%); 33 missing values	NA	18 (42.86%)	4 (2.61%)	17, 19%	NR	19, 22.3%	NR
T2D (n,%)	9 (2.5%); 33 missing values	NA	0 (0%)	1 (0.65%)	3, 3.4%	NR	6, 7%	NR
Migraine (n,%)	100 (28%); 33 missing values	NA	9 (21.4%)	NA	41, 46.6%	NR	48, 56.5%	NR
Smoking (Non/Current/Ex)	211/25/88	NA	34/1/7	NA	54/ 3/ 28	NR	51/ 3/ 31	NR

Table 3. Top significant GWAS canonical pathways and coexpression modules associated with SCAD.

Annotated pathways and modules with biological relevance and FDR<0.05 are shown. The pathway/module size refers to the number of genes within the pathway or module. Top 5 genes involved in each pathway or module are listed.

Source	Pathway Description	Pathway/ Module Size	FDR	Top Genes
GWAS catalog	Positive control gene set for SCAD	14	4.32E-03	*C1orf54; ECM1; ADAMTSL4; MRPS21; PHACTR1*
SCAD fibroblast transcriptome module	Regulation of endothelial cell proliferation	10	1.23E-02	*ERI1; P3H1; APMAP; SSR4; PRDX4*
GWAS catalog	Positive control gene set for coronary heart disease	88	1.27E-02	*PHACTR1; MRPS6; COL4A2; COL4A1; HLA-B*
SCAD fibroblast transcriptome module	Parkin pathway, neurotransmitter secretion	118	1.27E-02	*C1orf54; ECM1; ADAMTSL4; MRPS21; IGF2R*
SCAD fibroblast transcriptome module	Regulation of developmental process	49	1.39E-02	*ADK; MLKL; CIDECP; LMF2; CHMP7*
KEGG	Complement and coagulation cascades	45	1.43E-02	*C4A; C4B; CYP21A1P; CYP21A2; NOTCH4*
SCAD fibroblast transcriptome module	Extracellular structure organization	11	1.47E-02	*C1orf54; ECM1; ADAMTSL4; MRPS21; SERPIND1*
KEGG	Protein translation	11	1.47E-02	*PROCR; TRPC4AP; EDEM2; EIF6; GGT7*
KEGG	Acute myocardial infarction	14	1.49E-02	*COL4A1; COL4A2; DDB2; ACP2; NR1H3*
SCAD fibroblast transcriptome module	Formation of fibrin clot clotting cascade	56	1.63E-02	*C1orf54; ECM1; ADAMTSL4; MRPS21; PRCP*
SCAD fibroblast transcriptome module	MAPK pathway	12	1.92E-02	*VHL; BRK1; CRELD1; EMC3-AS1; MAP3K2*
SCAD fibroblast transcriptome module	Wnt pathway; intercellular signaling	136	2.10E-02	*C1orf54; ECM1; ADAMTSL4; MRPS21; IGF2R*
KEGG	VEGF, hypoxia, and angiogenesis	21	2.18E-02	*RP11-54A4.2; CERS2; GOLPH3L; HORMAD1; ARNT*
SCAD fibroblast transcriptome module	Protein metabolism	19	2.21E-02	*SMG1P6; TBX6; YPEL3; BOLA2; GDPD3*
SCAD fibroblast transcriptome module	mRNA metabolism	17	2.26E-02	*AM86B3P; C14orf79; PHYKPL; ARFIP1; RNGTT*
Reactome	Regulation of hypoxia-inducible factor by oxygen	20	2.27E-02	*RP11-54A4.2; CERS2; GOLPH3L; HORMAD1; ARNT*
KEGG	Renal cell carcinoma	51	2.28E-02	*RP11-54A4.2; CERS2; GOLPH3L; HORMAD1; ARNT*
SCAD fibroblast transcriptome module	Regulation of angiogenesis	41	2.31E-02	*LXN; PROCR; TRPC4AP; EDEM2; EIF6*
KEGG	Intrinsic prothrombin activation pathway	16	2.35E-02	*COL4A2; COL4A1; DDB2; ACP2; NR1H3*
Reactome	Signaling by PDGF	133	2.36E-02	*STAT6; COL4A1; COL4A2; SMG1P6, TBX6*
KEGG	Insulin signaling pathway	107	2.39E-02	*PPP1R18; FLOT1; SMG1P6; TBX6; YPEL3*
GWAS catalog	Positive control gene set for low density lipoprotein cholesterol	41	2.41E-02	*FRK; PPP1R3B; ABO; DNAH11; MYLIP*
KEGG	AKT/mTOR pathway	31	2.43E-02	*GSK3B; EIF4E; PDK2; EIF2S1; PDPK1*
SCAD fibroblast transcriptome module	Glycosaminoglycan metabolism	28	2.77E-02	*C1orf54; ECM1; ADAMTSL4; MRPS21; PEAR1*
GWAS catalog	Positive control gene set for high density lipoprotein cholesterol	63	2.86E-02	*LRP1; MACF1; PABPC4; PPP1R3B; DDB2*
SCAD fibroblast transcriptome module	Extracellular matrix receptor interaction; focal adhesion	72	2.86E-02	*C1orf54; ECM1; ADAMTSL4; MRPS21; COL4A1*
SCAD fibroblast transcriptome module	Extrinsic pathway	13	2.89E-02	*C1orf54; ECM1; ADAMTSL4; MRPS21; ZFYVE28*
KEGG	Integrin signaling pathway	32	3.29E-02	*ITGA1; SMG1P6; TBX6; YPEL3; BOLA2*
SCAD fibroblast transcriptome module	TCA cycle and respiratory electron transport	55	3.30E-02	*RNASEH2C; KAT5; BANF1; SIPA1; MAP3K11*
SCAD fibroblast transcriptome module	DNA base excision repair	19	3.41E-02	*MCL1; ARL5B; AQP11; RP11-91P24.1; RP11-91P24.5*
SCAD fibroblast transcriptome module	Glucose transport; glycerolipid metabolism	133	3.71E-02	*PSMD4; PABPC4; TNFRSF10B; RSRC1; TMEM120B*
SCAD fibroblast transcriptome module	Mitotic cell cycle	29	4.59E-02	*ANP32E; CENPQ; NUP155; ZNF714; ARSJ*
Reactome	Assembly of collagen fibrils	40	4.77E-02	*COL4A2; COL4A1; RP11-54A4.2; CERS2; GOLPH3L*

Table 4. Top significant TWAS canonical pathways and coexpression modules associated with SCAD.

Annotated pathways and modules with FDR<0.05 are shown. The pathway/module size refers to the number of genes within the pathway or module. Top 5 genes involved in each pathway or module are listed.

Source	Pathway Description	Pathway/ Module Size	FDR	Top Genes
SCAD fibroblast transcriptome module	Cytosolic sulfonation of small molecules	47	1.39E-13	*COL11A2; MAPK8IP3; HYPK; SPACA6P; RAPGEFL1*
SCAD fibroblast transcriptome module	Meiotic synapsis; packaging of telomere ends	100	1.053E-10	*OBFC1; AKIP1; OSTF1; TAC1; ZNF3*
SCAD fibroblast transcriptome module	Regulation of NFkB cascade	68	1.22E-10	*TPPP3; GHDC; ANO10; PELI1; KLHL29*
SCAD fibroblast transcriptome module	TCA cycle intermediate metabolic process	65	1.55E-10	*MRPL40; SOD1; OBFC1; AIP; FRS2*
SCAD fibroblast transcriptome module	Molecule transport	22	2.32E-10	*HYPK; MYSM1; SRRM2; SNAPIN; NPIPB11*
SCAD fibroblast transcriptome module	Cross presentation of antigens	35	5.35E-09	*AIP; CYP20A1; CCDC167; PSMD4; MED11*
SCAD fibroblast transcriptome module	Mitotic cell cycle	163	9.72E-09	*PL7P46; ZNF587; MAPK8IP3; SNAPC4; NXF1*
SCAD fibroblast transcriptome module	Oxidative phosphorylation; electron transport	58	1.30E-08	*PSMB6; RFX7; ADSL; SNRPD2; COX6B1*
SCAD fibroblast transcriptome module	Platelet activation, signaling, and aggregation	100	5.15D-07	*FXYD5; CNPY3; TECR; TXLNG; CALM1*
SCAD fibroblast transcriptome module	Cardiomyopathy	32	2.16E-06	*BNC2; UCP2; ST3GAL6; PRICKLE1; SESTD1*
SCAD fibroblast transcriptome module	Interleukin signaling	37	1.23E-05	*QRICH2; NAV3; EIF4EBP2; SLC30A1; AC144831.1*
SCAD fibroblast transcriptome module	Insulin signaling pathway	60	2.18E-05	*SDE2; FRS2; KLHL15; CCNT1; YTHDF3*
SCAD fibroblast transcriptome module	Glucose transport; glycerolipid metabolism	133	2.73E-05	*PSMD4; PABPC4; TNFRSF10B; RSRC1; TMEM120B*
Reactome	mRNA metabolism	214	2.77E-05	*RPL26L1; PSMB6; PSMA7; PSMB4; PSMC1*
SCAD fibroblast transcriptome module	Cholesterol biosynthesis	318	1.03E-04	*LSMEM1; ZNF445; NDUFS8; TCF20; AC093162.5*
SCAD fibroblast transcriptome module	ECM receptor interaction; focal adhesion integrin cell surface interactions	57	1.42E-04	*SRRM2; PKD1P1; NPIPB11; ATAT1; PSMC1*
SCAD fibroblast transcriptome module	DNA base excision repair	19	2.25E-04	*MCL1; ARL5B; AQP11; RP11-91P24.1; RP11-91P24.5*
SCAD fibroblast transcriptome module	Cell apoptosis	19	3.15E-04	*FRS2; ROR2; HIAT1; FAM53C; FASTKD5*
SCAD fibroblast transcriptome module	Cytokine signaling in immune system	77	3.30E-04	*SCG2; MLPH; TRPV2; CCND2; PDLIM3*
SCAD fibroblast transcriptome module	Wnt pathway; intercellular signaling	136	3.30E-04	*C1orf54; ECM1; ADAMTSL4; MRPS21; IGF2R*
SCAD fibroblast transcriptome module	Parkin pathway, neurotransmitter secretion	118	3.74E-04	*C1orf54; ECM1; ADAMTSL4; MRPS21; IGF2R*
SCAD fibroblast transcriptome module	Antigen processing	47	3.75E-04	*GSTM4; ADSL; CCDC62; HSD17B10; MEA1*
SCAD fibroblast transcriptome module	Membrane organization	42	6.16E-04	*CTTNBP2; NRCAM; MIF4GD; UBE2QL1; KL*
SCAD fibroblast transcriptome module	Anatomical structure development	377	2.06E-03	*HSPA2; SCG2; LINC00517; TOM1L1; TRPV2*
SCAD fibroblast transcriptome module	Glycoprotein metabolism	32	2.34E-03	*MARCH4; MIF4GD; BASP1; EBF1; SOX4*
SCAD fibroblast transcriptome module	TCA cycle and respiratory electron transport	55	3.97E-03	*RNASEH2C; KAT5; BANF1; SIPA1; MAP3K11*
Reactome	Activation of NFkB in B cells	61	9.05E-03	*PSMB6; PSMA7; PSMB4; PSMC1; PSMB7*
SCAD fibroblast transcriptome module	Regulation of endothelial cell proliferation	12	1.88E-02	*LMAN2; MYDGF; SSR4; PRDX4; ERP29*
KEGG	IL-10 anti-inflammatory signaling pathway	15	3.14E-02	*IL1A; BLVRB; BLVRA; STAT3; IL10RB*
SCAD fibroblast transcriptome module	Extrinsic pathway	13	3.50E-02	*C1orf54; ECM1; ADAMTSL4; MRPS21; ZFYVE28*
SCAD fibroblast transcriptome module	Extracellular matrix organization	63	3.84E-02	*RPL26L1; ZNF275; FGF5; CAPN2; DAB2IP*
SCAD fibroblast transcriptome module	Formation of fibrin clot clotting cascade	69	4.72E-02	*TFB1M; CNIH2; MRPL9; TMEM62; STAU2*
SCAD fibroblast transcriptome module	Actin cytoskeleton organization and biogenesis	13	4.97E-02	*GSTT2; GSTT2B; FAM124A; DYRK1A; AP000350.5*

Table 5. Top significant PWAS canonical pathways and coexpression modules associated with SCAD.

Annotated pathways and modules with FDR<0.05 are shown. The pathway/module size refers to the number of genes within the pathway or module. Top 5 genes involved in each pathway or module are listed.

Source	Pathway Description	Pathway/ Module Size	FDR	Top Genes
SCAD plasma protein module	Adaptive immune system; phagosome pathway, neurotrophin signaling	473	1.75 E-08	*ADH5; ICAM4; ORC4; DCP1A; IGBP1*
SCAD plasma protein module	Arginine and proline metabolism	56	5.09E-07	*DCP1A; IGBP1; PNKD; TPRKB; RAB22A*
SCAD plasma protein module	Mitotic G1 S phases	66	5.76E-06	*ORC4; ASH2L; CDIPT; AFAP1; YTHDF1*
SCAD plasma protein module	Complement and coagulation cascades, platelet activation and signaling pathway	23	7.48E-05	*RHOG; PF4V1; ALB; SERPING1; SERPINF2*
SCAD plasma protein module	Regulation of mitotic cell cycle, antigen processing ubiquitination proteasome degradation	59	6.62E-03	*ORC4; CDIPT; GIMAP8; DIRAS1; SPR*

Table 6. Top significant canonical pathways and coexpression modules shared between GWAS and TWAS.

Annotated pathways and modules significant across GWAS (FDR<5%) and TWAS (FDR<5%) with a significant meta-FDR<5% are listed.

Source	Pathway Description	GWAS FDR	TWAS FDR	Meta-FDR
SCAD fibroblast transcriptome module	Parkin pathway; neurotransmitter secretion	1.27E-02	3.74E-04	1.19E-06
SCAD fibroblast transcriptome module	Glucose transport; glycerolipid metabolism	3.71E-02	2.73E-05	1.99E-06
SCAD fibroblast transcriptome module	Wnt pathway; intercellular signaling	2.10E-02	3.30E-04	3.09E-06
SCAD fibroblast transcriptome module	DNA base excision repair	3.41E-02	2.25E-04	9.34E-06
SCAD fibroblast transcriptome module	Regulation of endothelial cell proliferation	1.23E-02	1.88E-02	5.26E-05
SCAD fibroblast transcriptome module	TCA cycle and respiratory electron transport	3.30E-02	3.97E-03	1.21E-04
SCAD fibroblast transcriptome module	Formation of fibrin clot clotting cascade	1.63E-02	4.72E-02	2.89E-04
SCAD fibroblast transcriptome module	Extrinsic pathway	2.89E-02	3.50E-02	7.01E-04

Table 7. Significant prioritized drugs within cardiovascular tissue for SCAD treatment via network-based repositioning.

Drugs functioning within the cardiovascular system were considered statistically significant if $p<0.05$. Top overlap genes refer to the genes shared between the input disease and drug networks.

SCAD Dataset	Drug Name	Drug Category	Z score	Rank	P value	Top Overlap Genes
SCAD GWAS	Medroxyprogesterone	Hormone	-4.07	1	2.35E-05	*CRISPLD2, MGP, GPX3, ACTG2, EFHD1*
SCAD transcript	Medroxyprogesterone	Hormone	-3.03	0.99	1.21E-03	*H19, MGP, GPX3, CFD, IGFBP2*
SCAD transcript	Etonogestrel	Hormone	-2.58	0.98	4.89E-03	*H19, MGP, BMP6, SOX4, IGFBP2*
SCAD GWAS	Etonogestrel	Hormone	-2.44	0.98	7.42E-03	*CRISPLD2, MGP, DSTN, CPE, HTRA1*
SCAD GWAS	Contraceptive	Hormone	-2.30	0.98	1.06E-02	*CRISPLD2, MGP, AXIN2, PAWR, GAS6*
SCAD transcript	Rofecoxib	Anti-inflammatory	-1.93	0.94	2.68E-02	*TGFB2, KLF13, EDN1, JUNB, ZNF331*

Table 8. Top significant prioritized drugs within Homo Sapiens cardiovascular tissue for SCAD treatment via overlap-based repositioning.

The top 10 significant drugs ($p<0.05$) functioning within the homo sapiens cardiovascular system from each of the GWAS and TWAS overlap-based repositioning are shown. Top overlap genes refer to the genes shared between drug signature and input disease genes.

SCAD Dataset	Drug Name	Drug Category	Jaccard Score	Odds Ratio	P value	Top Overlap Genes
SCAD GWAS	Contraceptive	Hormone	0.051	12.2	3.88E-27	*CRISPLD2, MGP, AXIN2, PAWR, GAS6*
SCAD GWAS	Medroxyprogesterone	Hormone	0.051	19.1	4.64E-24	*CRISPLD2, MGP, GPX3, ACTG2, EFHD1*
SCAD GWAS	Progesterone receptor agonist	Hormone	0.034	6.94	1.32E-15	*CRISPLD2, MGP, GPX3, ACTG2, EFHD1*
SCAD transcript	Medroxyprogesterone	Hormone	0.036	13.8	9.68E-14	*H19, MGP, GPX3, CFD, IGFBP2*
SCAD transcript	Progesterone receptor agonist	Hormone	0.031	7.16	1.16E-13	*BMP6, H19, MGP, IGFBP2, CTHRC1*
SCAD transcript	Contraceptive	Hormone	0.032	8.27	1.62E-13	*H19, MGP, BMP6, IGFBP2, AXIN2*
SCAD transcript	Omeprazole	Proton pump inhibitor	0.035	13.9	4.37E-13	*GPX3, NR4A2, SELP, CYTL1, NR4A3*
SCAD transcript	Anti-ulcerative	Anti-inflammatory; proton pump inhibitor	0.035	13.9	4.37E-13	*GPX3, NR4A2, SELP, CYTL1, NR4A3*
SCAD transcript	Proton pump inhibitor	Proton pump inhibitor	0.035	13.9	4.37E-13	*GPX3, NR4A2, SELP, CYTL1, NR4A3*
SCAD transcript	Antisecretory	Proton pump inhibitor; antimuscarinic	0.035	13.9	4.37E-13	*GPX3, NR4A2, SELP, CYTL1, NR4A3*
SCAD GWAS	Etonogestrel	Hormone	0.028	13.7	8.13E-11	*CRISPLD2, MGP, DSTN, CPE, HTRA1*
SCAD transcript	Thrombin	Clotting promoter	0.028	7.60	2.11E-10	*NR4A2, CTGF, RGS2, ALDH3A1, EGR3*
SCAD transcript	Blood coagulation factor	Clotting promoter	0.028	7.60	2.11E-10	*NR4A2, CTGF, RGS2, ALDH3A1, EGR3*
SCAD transcript	Hemostatic	Antihemorrhagic	0.028	7.60	2.11E-10	*NR4A2, CTGF, RGS2, ALDH3A1, EGR3*
SCAD GWAS	Analgesic	Pain relief	0.026	6.13	2.03E-09	*KLF13, EPS8L2, C3orf52, ELN, GDF10*
SCAD GWAS	COX-2 inhibitor	Anti-inflammatory	0.026	6.13	2.03E-09	*KLF13, EPS8L2, C3orf52, ELN, GDF10*
SCAD GWAS	Celecoxib	Anti-inflammatory	0.024	11.6	1.07E-08	*KLF13, C3orf52, ESM1, EPS8L2, MYLK*
SCAD GWAS	Omeprazole	Proton pump inhibitor	0.023	8.07	1.04E-07	*GPX3, CLMP, COL4A6, LYVE1, TMEM37*
SCAD GWAS	Anti-ulcerative	Anti-inflammatory; proton pump inhibitor	0.023	8.07	1.04E-07	*GPX3, CLMP, COL4A6, LYVE1, TMEM37*
SCAD GWAS	Proton pump inhibitor	Proton pump inhibitor	0.023	8.07	1.04E-07	*GPX3, CLMP, COL4A6, LYVE1, TMEM37*

Table 9. HC vs. chow DEGs and biological pathways.
"Up" refers to DEGs with increased expression in HC compared to Chow. "Down" refers to DEGs with decreased expression in HC compared to Chow. Cell types denoted in parentheses did not have a sufficient number of DEGs at FDR<5% to reveal significant pathways, so DEGs reaching a threshold of $p < 0.01$ (number of DEGs for this threshold in parentheses) were used to predict suggestive pathways for these cell types.

Cell Type	No. DEGs	Top DEG pathways	Top 5 representative DEGs
SMC	94	Smooth muscle contraction	Down: *Tpm1, Myl6, Myl9, Tpm2, Acta2*
		Extracellular structure organization	Up: *Spp1, Lum;* Down: *Sparc, Cst3, Itgb1*
Modulated SMC	101	Extracellular structure organization	Up: *Lum, Vcam1, Fn1, Spp1, Bgn*
		Collagen fibril organization	Up: *Lum, Col3a1, Aebp1, Col1a1, Col5a2*
Fibroblast	109	Triggering of complement pathway	Up: *C1s, C3, C2, Cfb, C4b*
		Extracellular structure organization	Up: *Spp1, Lum;* Down: *Ccn2, Myh11, Eln*
Endothelial	37	Regulation of cell differentiation	Up: *Spp1, Igfbp5, Nfkbia, Gdf10, Hes1*
		Migration of endothelial cells	Up: *Ptgs2, Nr4a1, Pdgfb, Dcn, Cdh13*
Macrophage	275	Immune response	Up: *Fabp5, Lgals3, Syngr1, Ctss, Itgax*
		Inflammatory response	Up: *Spp1, Lpl, Adam8;* Down: *Pf4, Cd163*
(Pericyte)	4 (197)	Respiratory electron transport	Up: *mt-Nd3, mt-Nd2, mt-Atp8, mt-Nd4, mt-Co3*
		Cell growth	Up: *Igfpb5, Caprin2, Vegfa, Rgs4, Ntrk3*
Unknown1	90	Extracellular structure organization	Up: *Spp1, Vcam1, Lum, Crispld2;* Down: *Cst3*
(Unknown2)	5 (237)	Protein translation	Down: *Rps8, Rpl10, Rpl26, Rps12, Rps13*
(Unknown3)	16 (229)	Protein translation	Down: *Rps9, Rps14, Rpl13, Rpl23a, Rps15*
		Smooth muscle contraction	Down: *Myl6, Tpm2, Tpm1,Mylk, Myl9*
Unknown 4	196	Protein translation	Down: *Rpl21, Rpl8, Rps5, Rpl28, Rpl10*
		ECM receptor interaction	Up: *Spp1, Col2a1, Fn1, Col1a2, Col3a1*

Table 10. HC+TMAO vs. HC DEGs and biological pathways.
"Up" refers to DEGs with increased expression in HC+TMAO compared to HC. "Down" refers to DEGs with decreased expression in HC+TMAO compared to HC. Cell types denoted in parentheses did not have a sufficient number of DEGs at FDR<5% to reveal significant pathways, so DEGs reaching a threshold of $p < 0.01$ (number of DEGs for this threshold in parentheses) were used to predict suggestive pathways for these cell types.

Cell Type	No. DEGs	Top DEG pathways	Top 5 representative DEGs
SMC	68	Regulation of cell proliferation	Up: *Tpm1, Itgb1;* Down: *Fbln5, Zfp36l1, Btg1*
		Regulation of apoptosis	Up: *Ptgis, Lgals1, Rpl10;* Down: *Zfp36l1, Btg1*
Modulated SMC	32	Metabolism of mRNA and proteins	Up: *Rpl9, Rps5, Rpl18, Rpl17, Rpl13*
		Apoptotic process	Up: *Rpl10, Rps3a1, Rps3, Rps7;* Down: *Hk2*
Fibroblast	89	Protein translation	Up: *Rpl10, Rps15, Rpl17, Rps14, Rpl12*
		Fiber organization	Up: *Tpm2, Myh11, Tmsb4x;* Down: *Tmsb10, Col1a1*
(Endothelial)	9 (263)	Metabolism of mRNA	Down: *Rps29, Rpl41, Rps28, Rpl38, Rpl6*
		Vasculature development	Down: *Ptgs2, Hes1, Tnfaip2, Smoc2, Nr4a1*
(Macrophage)	6 (100)	Metabolism of proteins	Up: *Rpl10, Rps9, Rps7, Rpl26;* Down: *Tuba1c*
(Pericyte)	0 (197)	Cell growth	Down: *Igfbp5, Caprin2, Vegfa, Rgs4, Ntrk3*
(Unknown1)	3 (170)	Focal adhesion	Up: *Tln1;* Down: *Col3a1, Tnc, Pik3r1, Spp1*
		Extracellular structure organization	Up: *Cst3;* Down: *Col3a1, Fbln5, Eln, Tnc*
Unknown2	38	Protein translation	Up: *Rps8, Rpl10, Rpl26, Rps12, Rps13*
(Unknown3)	7 (229)	Protein translation	Up: *Rps9, Rps14, Rpl13, Rpl23a, Rps15*
		Collagen fibril organization	Down: *Dpt, Col3a1, Col1a1, Col1a2, P4ha1*
Unknown4	101	Protein translation	Up: *Rpl9, Rpl10, Rpl19, Rpl18, Rps14*
		Muscle contraction	Up: *Myl6, Tpm2, Tpm1, Myl9, Tcap*

Table 11. HC+TMAO vs. chow DEGs and biological pathways.
"Up" refers to DEGs with increased expression in HC+TMAO compared to Chow. "Down" refers to DEGs with decreased expression in HC+TMAO compared to Chow. Cell types denoted in parentheses did not have a sufficient number of DEGs at FDR<5% to reveal significant pathways, so DEGs reaching a threshold of $p < 0.01$ (number of DEGs for this threshold in parentheses) were used to predict suggestive pathways for these cell types.

Cell Type	No. DEGs	Top DEG pathways	Top 5 representative DEGs
SMC	44	Extracellular structure organization	Up: *Spp1, Lum;* Down: *Sparc, Ccn2, Ccn1*
Modulated SMC	90	Integrin binding	Up: *Vcam1, Fn1, Spp1, Col3a1;* Down: *Adam23,*
		Collagen fibril organization	Up: *Lum, Aebp1, Col3a1, Sfrp2, Col5a2*
Fibroblast	47	Extracellular structure organization	Down: *Sparc, Eln, Col1a1, Mfap4, Serpinh1*
		Platelet degranulation	Up: *Serpina3n, Fn1;* Down: *Sparc, Manf, Igf1,*
(Endothelial)	11 (303)	Cell proliferation	Up: *Lrg1, Lgmn, Vcam1, Mdk, Sox4*
		Hemostasis	Up: *Clu, Selp, Timp1, C1qbp, Pde1a*
Macrophage	151	Immune activation	Up: *Fabp5, Lgals3, Syngr1, Fth1, Trem2*
		Macrophage migration	Up: *Lgals3, Trem2, Mmp14, P2rx4;* Down: *Cd81*
(Pericyte)	0 (130)	Respiratory electron transport	Down: *Ndufa2, Ndufa3, Ndufs6, Cox6c, Idh3a*
(Unknown1)	15 (169)	Cell adhesion	Up: *Spp1, Fn1, Itga5, Lamb2, Ctnnb1*
		Extracellular matrix binding	Up: *Spp1, Lgals3, Sparc1l, Ecm1;* Down: *Lgals1*
Unknown2	54	Protein translation	Up: *Rpl23a, Rps12, Rps25, Rps9, Rps14*
(Unknown3)	7 (83)	Extracellular structure organization	Up: *Spp1, Vcam1;* Down: *Sparc, Mfap5, Ccn1*
Unknown4	65	Extracellular structure organization	Up: *Spp1, Vcam1, Lum, Fn1, Timp1*

Figures

Figure 1. Multiomics systems biology of SCAD: study overview.

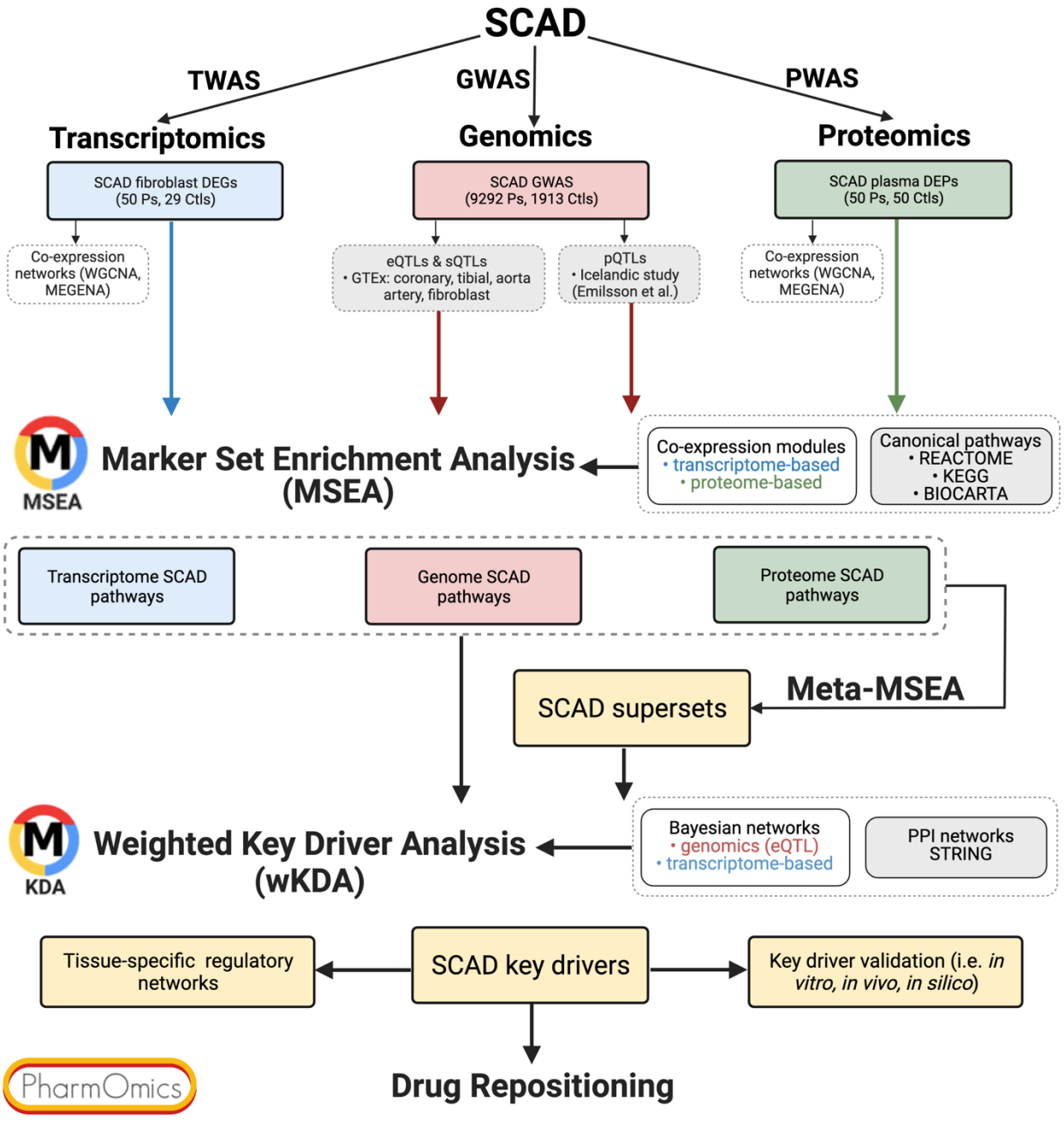


Figure 2. Top GWAS key drivers within a vascular-specific network.
Key drivers are denoted by their larger size. Annotated pathways or coexpression modules of KDs are labeled.

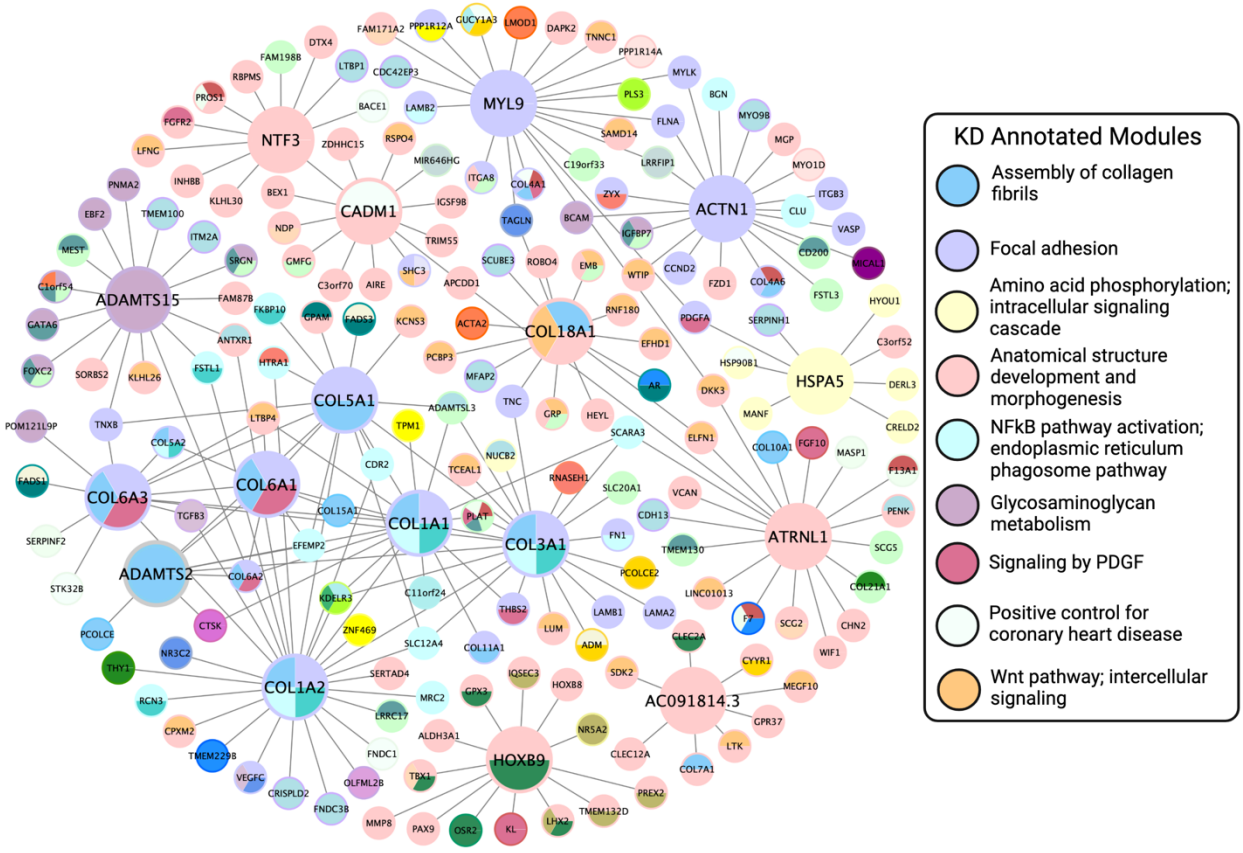


Figure 3. Top transcript key drivers within a vascular-specific network.
Key drivers are denoted by their larger size. Annotated pathways or coexpression modules of KDs are labeled.

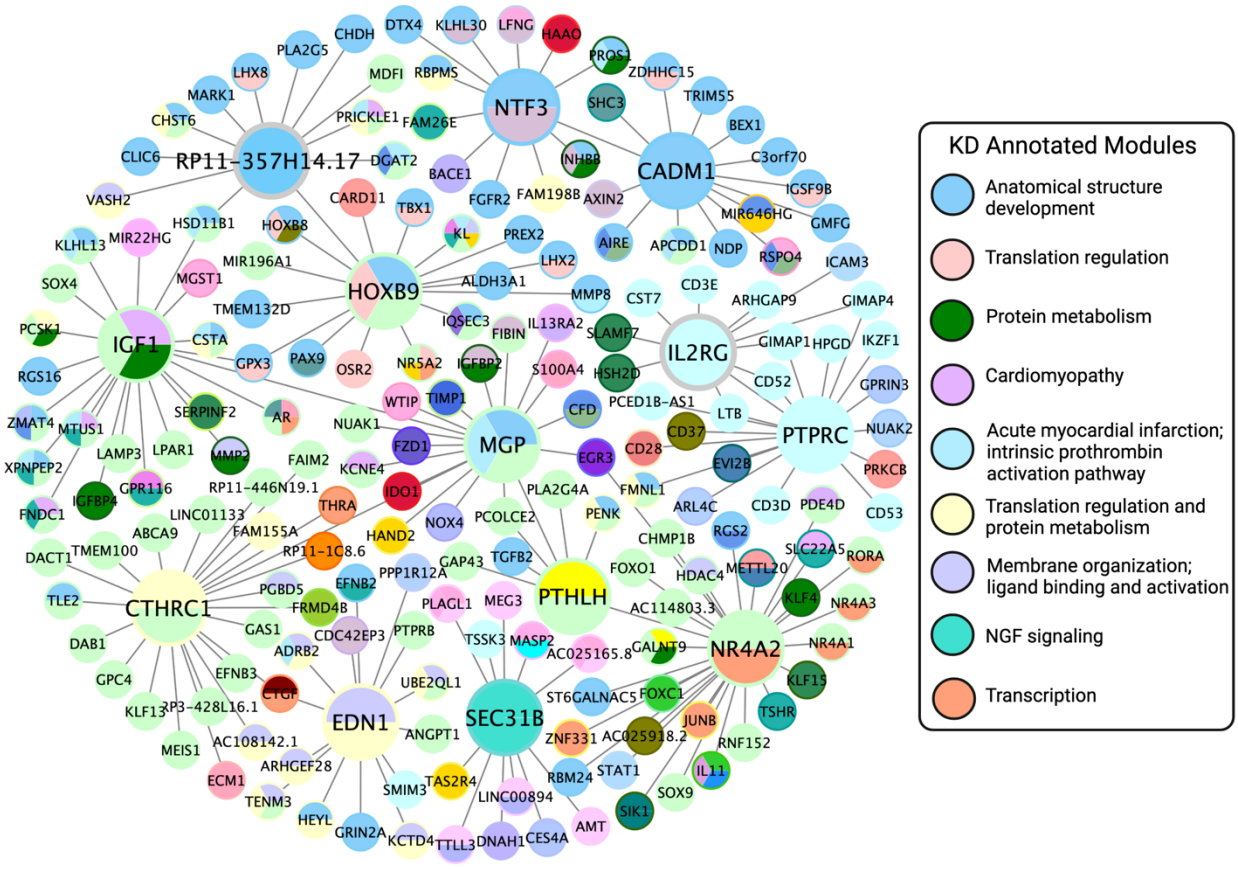


Figure 4. Top protein key drivers within a PPI network.
Hub proteins central to the network are denoted by their larger size. All annotated pathways or coexpression modules involved in the network are labeled.

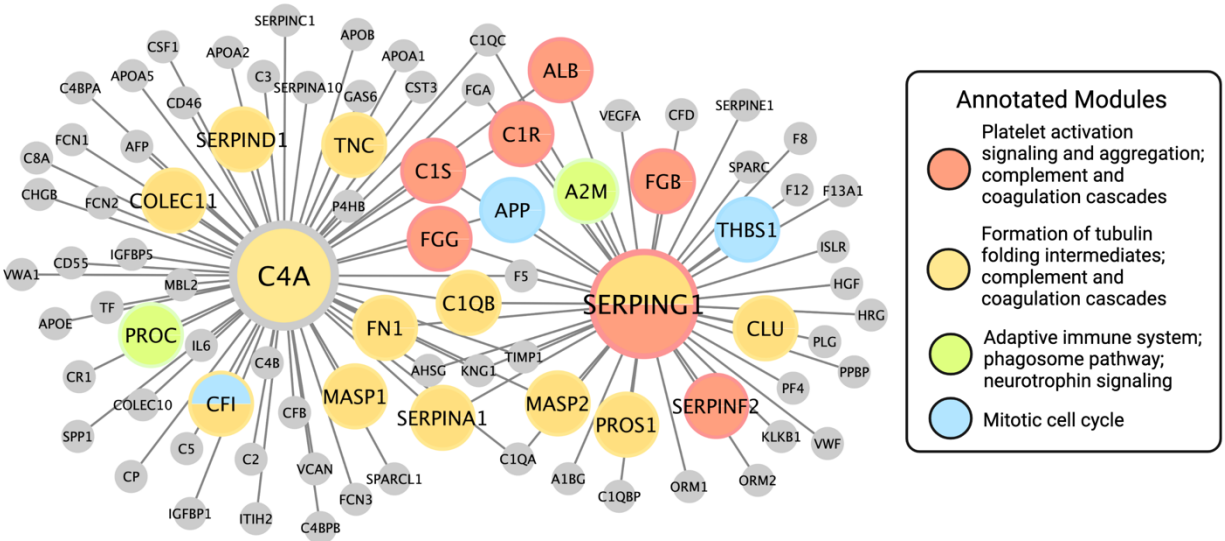


Figure 5. Cell type identification and specific cell-type markers.
(a) Labeled UMAP plot of single cells in the two-dimensional space by diet condition. **(b)** Top cell type-specific markers: *Mylk4* for SMC, *Spp1* for modulated SMC, *Cygb* for fibroblast, *Pecam1* for endothelial, *Lyz2* for macrophage, *Timp4* for pericyte. The lack of specificity for Unknown clusters 1, 2, and 4 demonstrate how a combination of markers from known cell types constitute these clusters. **(c-e)** Additional canonical markers of endothelial (*Egfl7*), macrophage (*Cd14*), and fibroblast (*Pi16*) within the UMAP space and highlighted in their respective clusters.

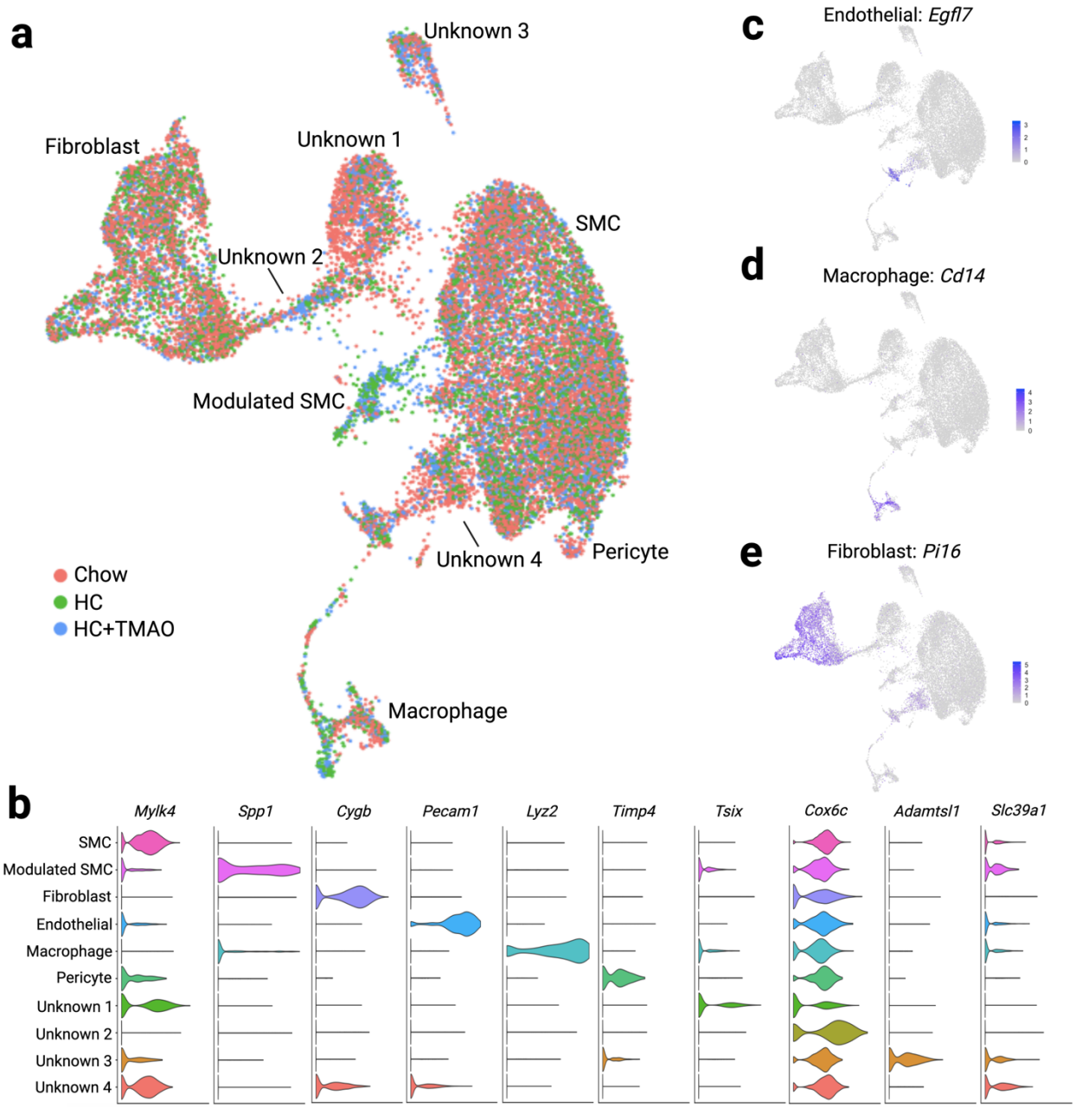


Figure 6. Modulated SMC subclustering to reveal three distinct clusters.
(a) UMAP plot of modulated SMC single cells in the two-dimensional space by diet condition. **(b)** UMAP plot of modulated SMC single cells from HC and HC+TMAO conditions. **(c)** Average expression of fibromyocyte markers by modulated SMC subcluster and by diet condition. mSMC=modulated SMC. **(d)** Trajectory analysis for SMC and modulated SMC subclusters. SMC (dark green) was assigned as the origin cluster for the trajectory analysis. **(e)** Top subcluster-specific markers for subcluster 0 (*Tcap*), 1 (*Gas7*), and 2 (*Col11a2*).

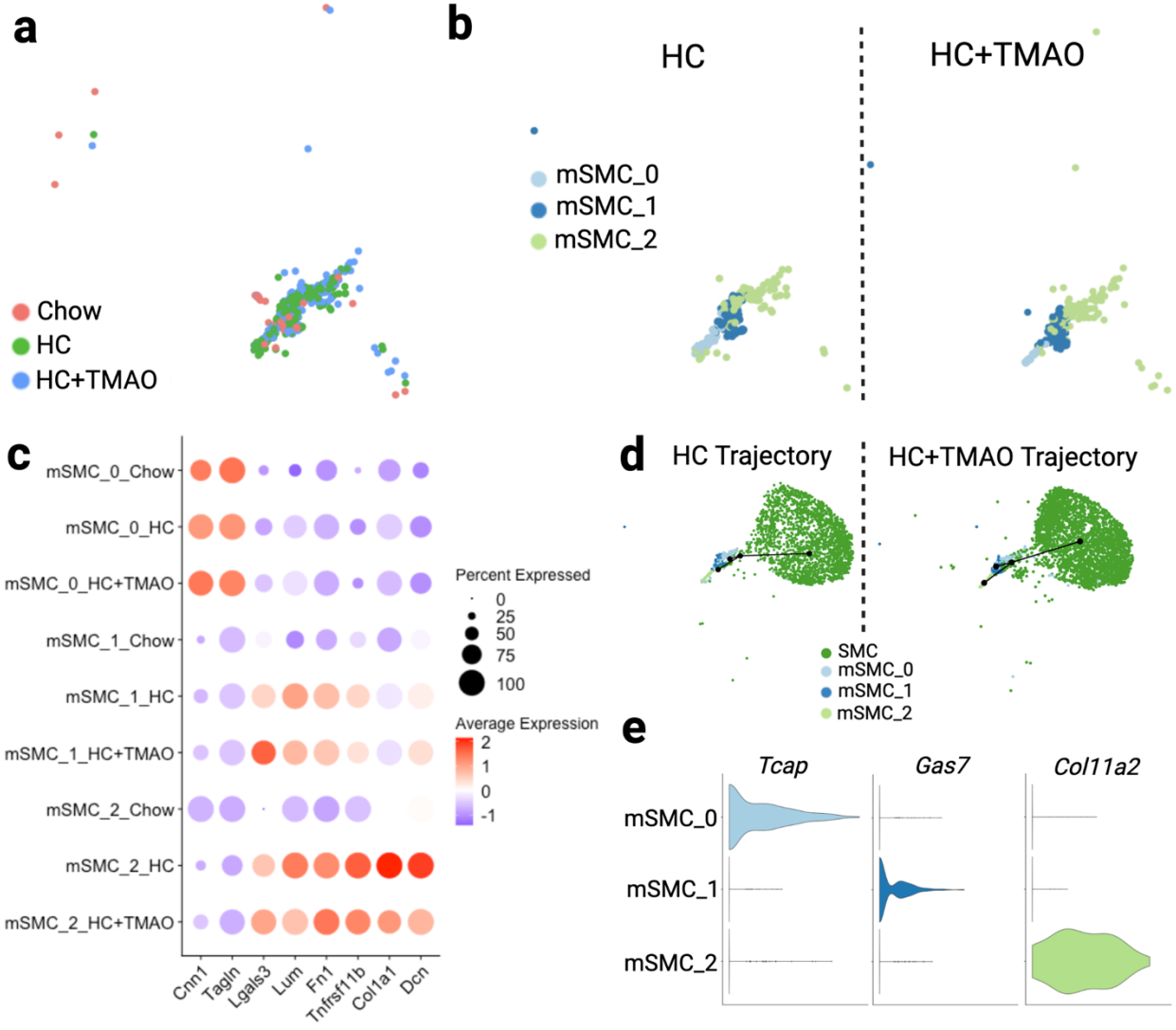


Figure 7. Macrophage subclustering to reveal distinct M0, M1, and M2 macrophage types. **(a)** UMAP plot of macrophage single cells in the two-dimensional space by diet condition. **(b)** UMAP plot of macrophage single cells from chow, HC, HC+TMAO conditions. **(c)** Scaled expression of previously known M1 markers by macrophage subtype and diet condition. **(d)** Scaled expression of previously known M2 markers by macrophage subtype and diet condition. **(e)** Top subcluster-specific markers for subcluster M0 (*Fermt2*), M1 (*Gngt2*), and M2 (*Cd163*). **(f)** Trajectory analysis for M0, M1, and M2. M0 (green) was assigned as the origin cluster for the trajectory analysis.

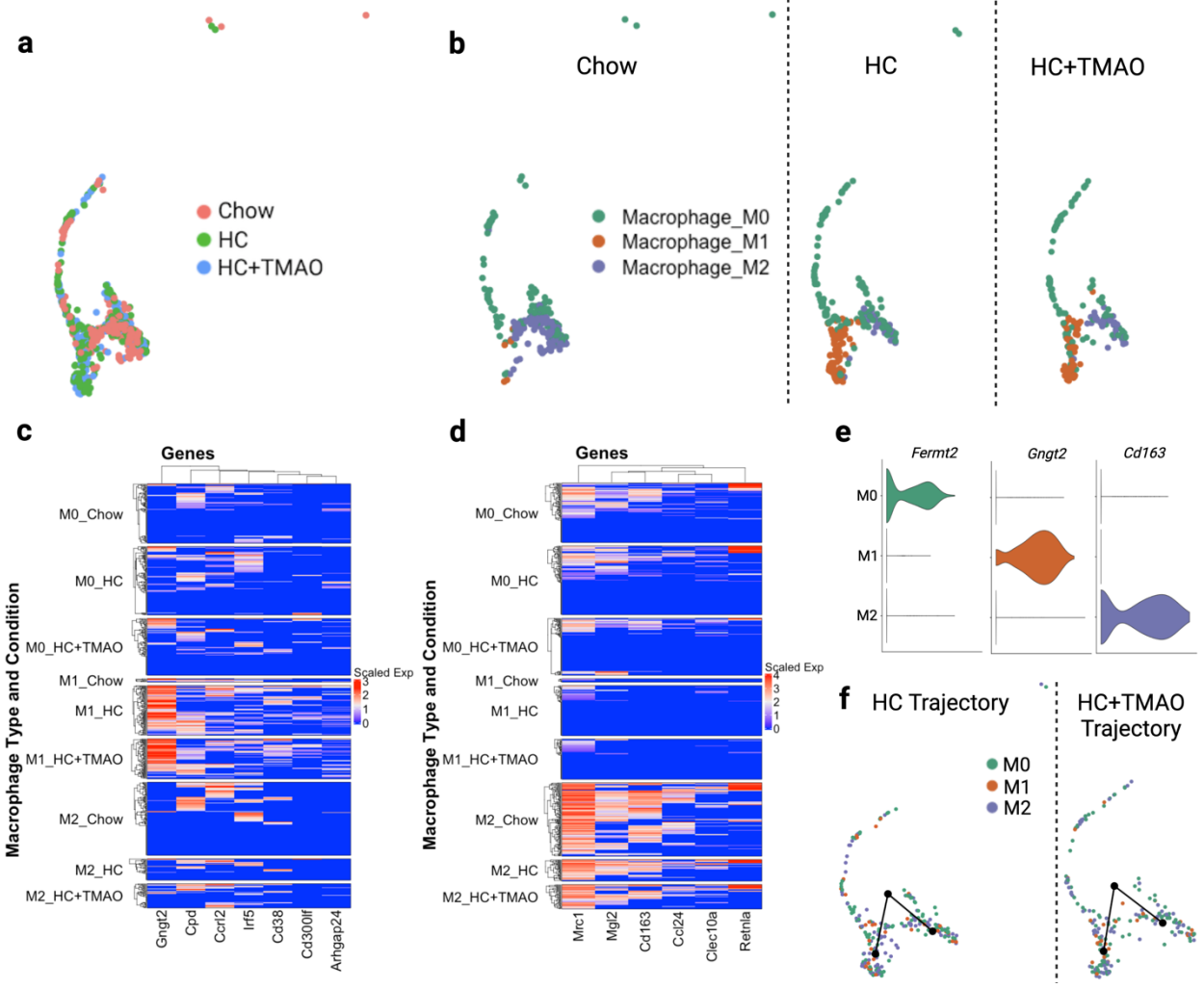


Figure 8. Cell type distribution by diet.
The proportions of each cell type or subcluster within each diet are illustrated. Significance between two diet conditions for a single cell type or subcluster was considered $p < 0.05$. **(a)** The proportionate numbers of each cell type by diet. **(b)** The proportionate numbers of each modulated SMC subcluster by diet. mSMC=modulated SMC. **(c)** The proportionate numbers of each macrophage subcluster (M0, M1, M2) by diet.

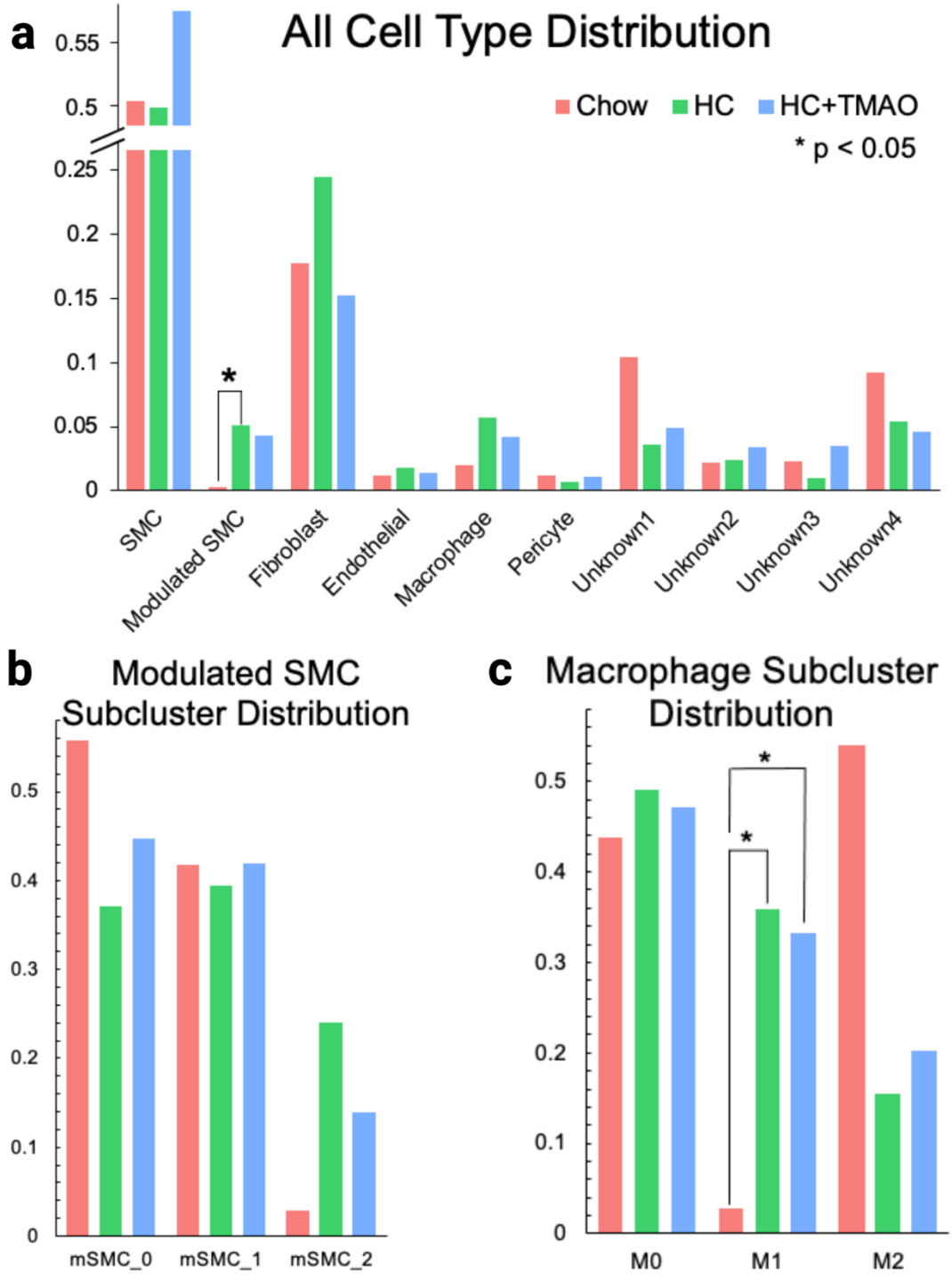


Figure 9. Overlap in differentially expressed genes between high-cholesterol and chow conditions.
Significant (FDR<5%) cell-type specific DEGs (red) and shared DEGs between cell types (black) between the Chow and HC dietary conditions. Shared DEGs between cell types are labeled if the number of genes ≤ 3. The set size indicates the total number of DEGs per cell type (sum of cell type-specific and shared).

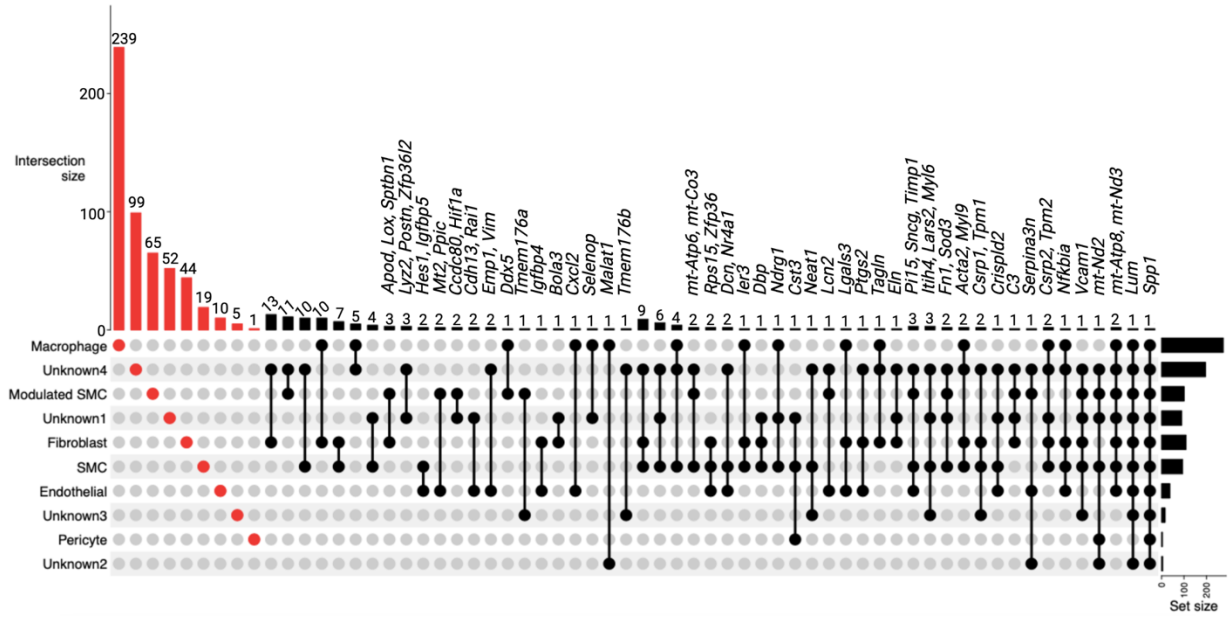


Figure 10. Overlap in differentially expressed genes between high-cholesterol+TMAO and high-cholesterol conditions.
Significant (FDR<5%) cell-type specific DEGs (red) and shared DEGs between cell types (black) between the HC and HC+TMAO dietary conditions. Shared DEGs between cell types are labeled if the number of genes ≤ 3. The set size indicates the total number of DEGs per cell type (sum of cell type-specific and shared).

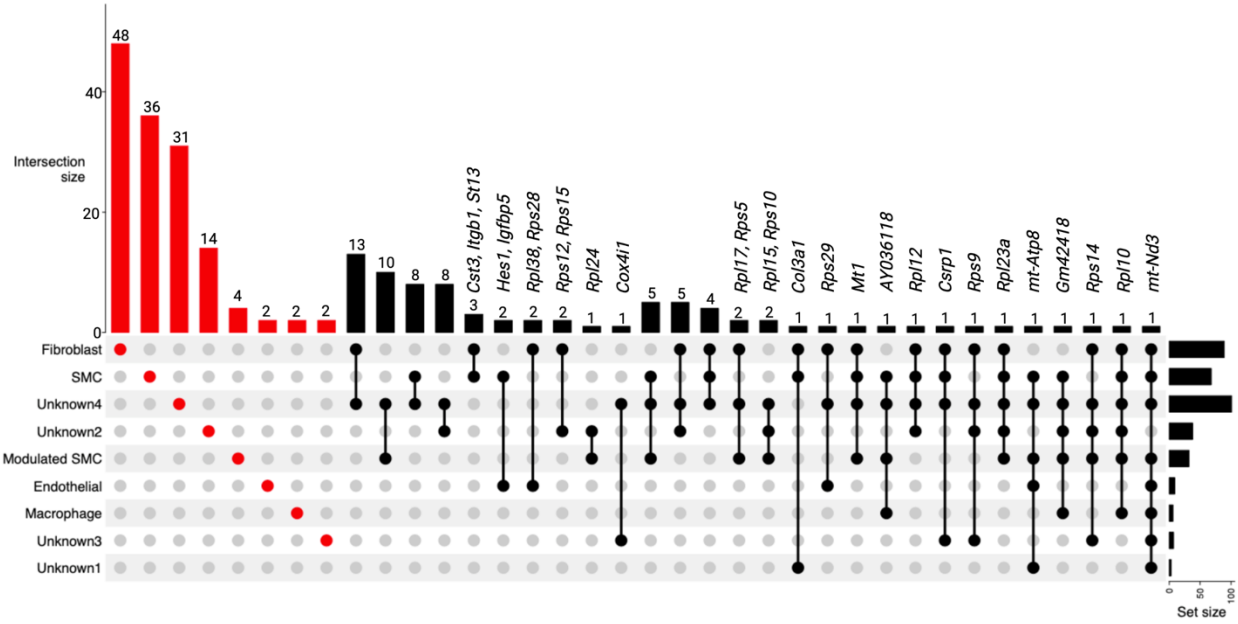


Figure 11. Overlap in differentially expressed genes between high-cholesterol+TMAO and chow conditions.
Significant (FDR<5%) cell-type specific DEGs (red) and shared DEGs between cell types (black) between the chow and HC+TMAO dietary conditions. Shared DEGs between cell types are labeled if the number of genes ≤ 3. The set size indicates the total number of DEGs per cell type (sum of cell type-specific and shared).

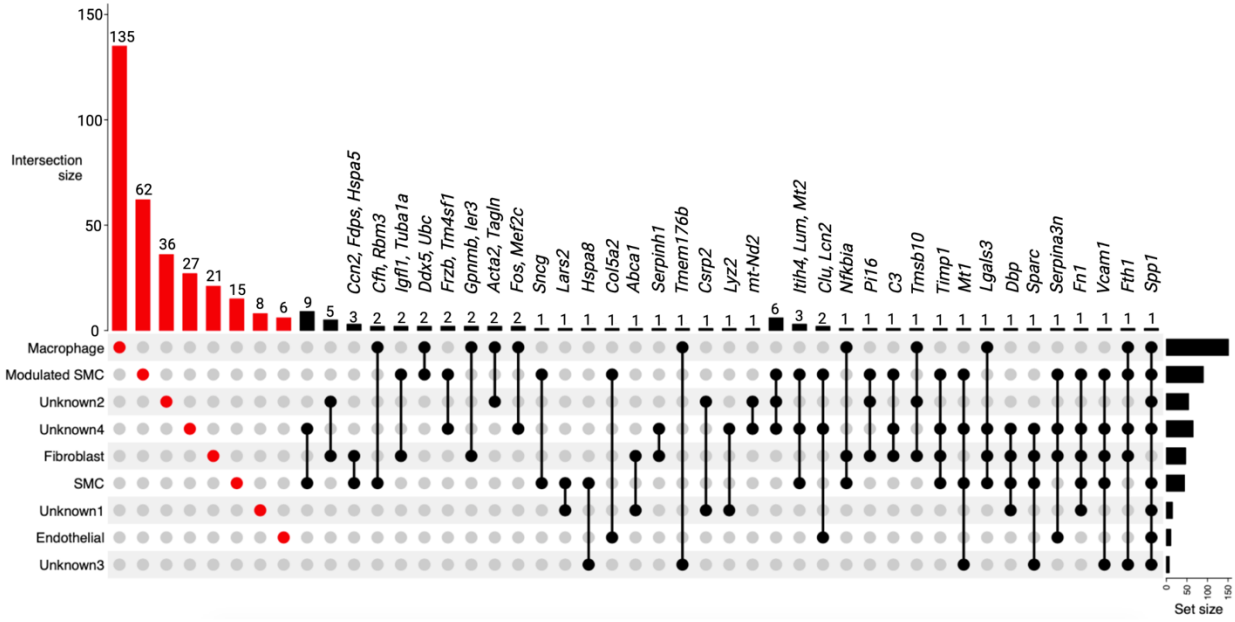


Figure 12. Intercellular interactions targeting modulated SMC genes in response to TMAO.

(a) Top predicted ligands targeting modulated SMC genes as ranked by Pearson correlation coefficient between prior regulatory potential scores and target gene set assignments. **(b)** Percent expression and average expression of each top predicted ligand in the known aortic cell types. mSMC=modulated SMC. **(c)** Log fold-change (LFC) of each top predicted ligand in the known aortic cell types between HC+TMAO and HC conditions. **(d)** Modulated SMC active target genes of the top predicted ligands. **(e)** Summary of top prioritized ligands from each cell type and their corresponding target genes in modulated SMC. Weight of the edge represents the regulatory potential of the ligand on the downstream target gene.

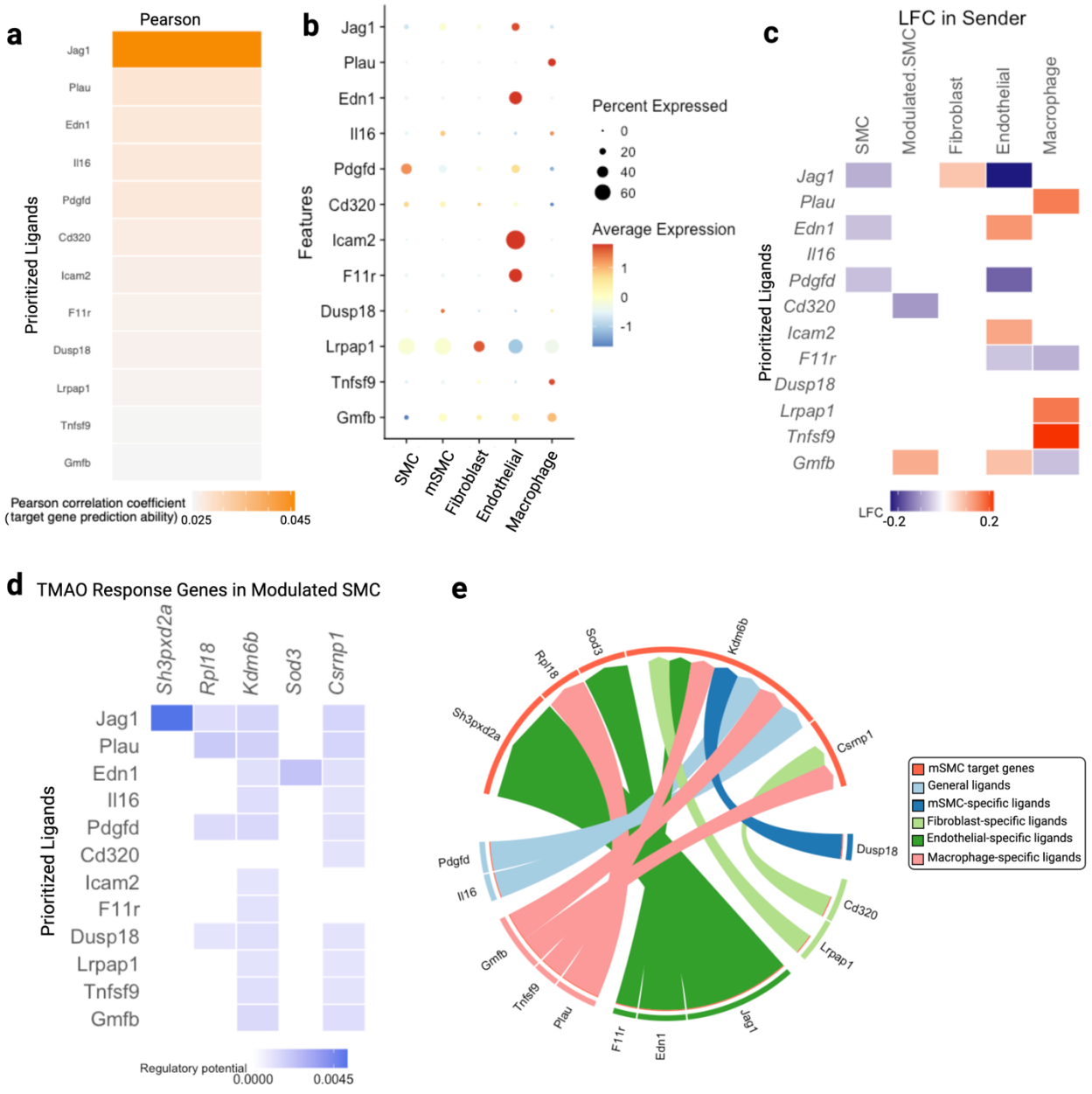
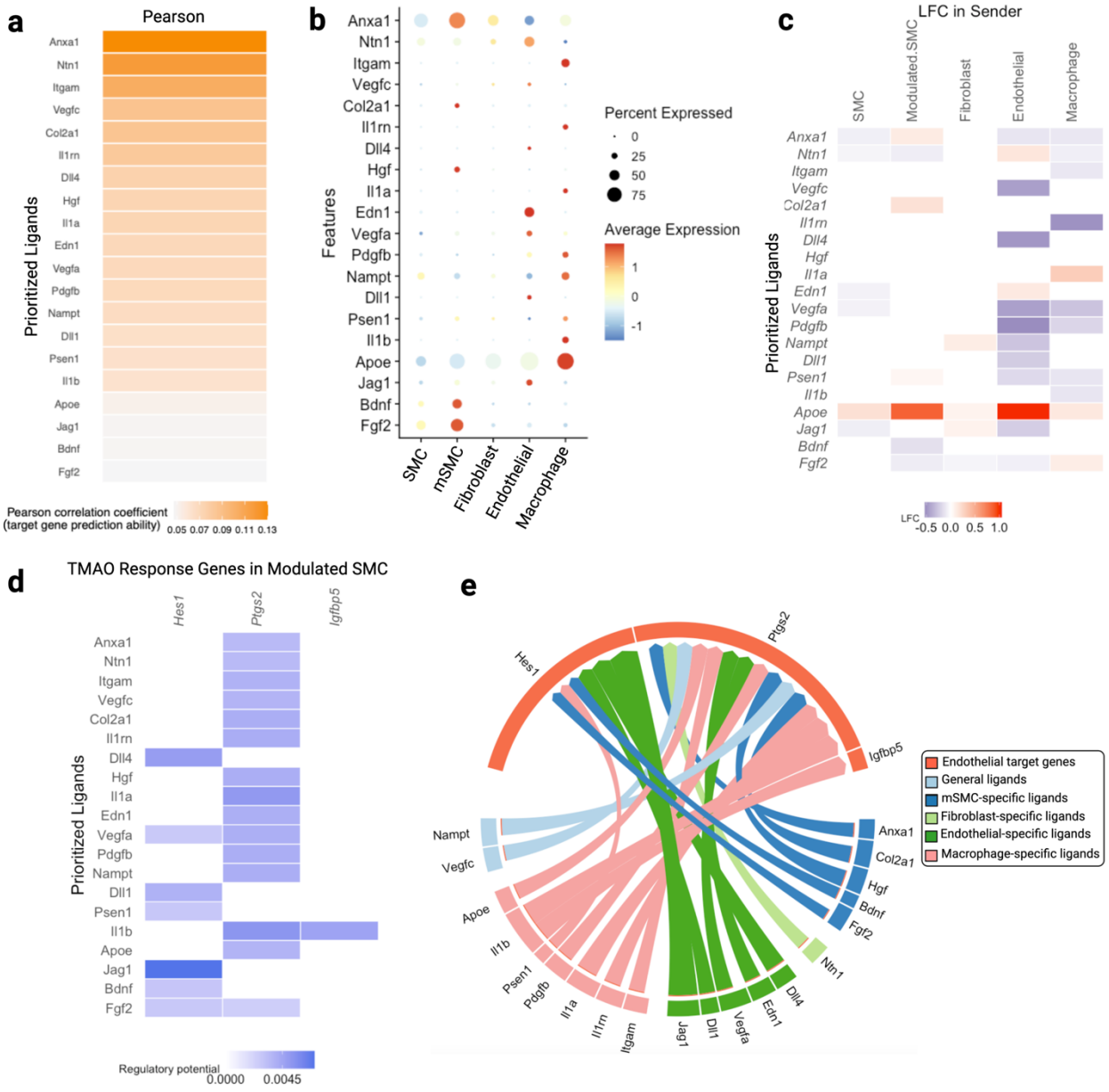


Figure 13. Intercellular interactions targeting endothelial genes in response to TMAO. **(a)** Top predicted ligands targeting endothelial cell genes as ranked by Pearson correlation coefficient between prior regulatory potential scores and target gene set assignments. **(b)** Percent expression and average expression of each top predicted ligand in the known aortic cell types. **(c)** Log fold-change (LFC) of each top predicted ligand in the known aortic cell types between HC+TMAO and HC conditions. **(d)** Endothelial cell active target genes of the top predicted ligands. **(e)** Summary of top prioritized ligands from each cell type and their corresponding target genes in the endothelial cell type. Weight of the edge represents the regulatory potential of the ligand on the downstream target gene.



References

1. Turley, T. N. *et al.* Identification of Susceptibility Loci for Spontaneous Coronary Artery Dissection. *JAMA Cardiol.* **5**, 929–938 (2020).

2. Carss, K. J. *et al.* Spontaneous Coronary Artery Dissection. *Circ. Genomic Precis. Med.* **13**, e003030 (2020).

3. Saw, J. *et al.* Chromosome 1q21.2 and additional loci influence risk of spontaneous coronary artery dissection and myocardial infarction. *Nat. Commun.* **11**, 4432 (2020).

4. Al-Hussaini, A. & Adlam, D. Spontaneous coronary artery dissection. *Heart* **103**, 1043–1051 (2017).

5. Tweet, M. S. *et al.* Clinical Features, Management, and Prognosis of Spontaneous Coronary Artery Dissection. *Circulation* **126**, 579–588 (2012).

6. Saw, J. *et al.* Spontaneous Coronary Artery Dissection. *Circ. Cardiovasc. Interv.* **7**, 645–655 (2014).

7. Lettieri, C. *et al.* Management and Long-Term Prognosis of Spontaneous Coronary Artery Dissection. *Am. J. Cardiol.* **116**, 66–73 (2015).

8. Rogowski, S. *et al.* Spontaneous Coronary Artery Dissection. *Catheter. Cardiovasc. Interv.* **89**, 59–68 (2017).

9. Nakashima, T. *et al.* Prognostic impact of spontaneous coronary artery dissection in young female patients with acute myocardial infarction: A report from the Angina Pectoris–Myocardial Infarction Multicenter Investigators in Japan. *Int. J. Cardiol.* **207**, 341–348 (2016).

10. Eleid, M. F. *et al.* Coronary Artery Tortuosity in Spontaneous Coronary Artery Dissection. *Circ. Cardiovasc. Interv.* **7**, 656–662 (2014).

11. Shao, C., Wang, J., Tian, J. & Tang, Y. Coronary Artery Disease: From Mechanism to Clinical Practice. in *Coronary Artery Disease: Therapeutics and Drug Discovery* (ed. Wang, M.) 1–36 (Springer, 2020). doi:10.1007/978-981-15-2517-9_1.

12. Austin, M. A., Hutter, C. M., Zimmern, R. L. & Humphries, S. E. Familial Hypercholesterolemia and Coronary Heart Disease: A HuGE Association Review. *Am. J. Epidemiol.* **160**, 421–429 (2004).

13. Nowbar, A. N., Gitto, M., Howard, J. P., Francis, D. P. & Al-Lamee, R. Mortality From Ischemic Heart Disease. *Circ. Cardiovasc. Qual. Outcomes* **12**, e005375 (2019).

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

15. Malakar, A. Kr. *et al.* A review on coronary artery disease, its risk factors, and therapeutics. *J. Cell. Physiol.* **234**, 16812–16823 (2019).

16. Powell-Wiley, T. M. *et al.* Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association. *Circulation* **143**, e984–e1010 (2021).

17. Bouabdallaoui, N. *et al.* Impact of smoking on cardiovascular outcomes in patients with stable coronary artery disease. *Eur. J. Prev. Cardiol.* **28**, 1460–1466 (2021).

18. Momiyama, Y., Adachi, H., Fairweather, D., Ishizaka, N. & Saita, E. Inflammation, Atherosclerosis and Coronary Artery Disease. *Clin. Med. Insights Cardiol.* **8s3**, CMC.S39423 (2014).

19. Roncal, C. *et al.* Trimethylamine-N-Oxide (TMAO) Predicts Cardiovascular Mortality in Peripheral Artery Disease. *Sci. Rep.* **9**, 15580 (2019).

20. Subramanian, I., Verma, S., Kumar, S., Jere, A. & Anamika, K. Multi-omics Data Integration, Interpretation, and Its Application. *Bioinforma. Biol. Insights* **14**, 1177932219899051 (2020).

21. Turley, T. N. *et al.* Susceptibility Locus for Pregnancy-Associated Spontaneous Coronary Artery Dissection. *Circ. Genomic Precis. Med.* **14**, e003398 (2021).

22. Goel, K. *et al.* Familial Spontaneous Coronary Artery Dissection: Evidence for Genetic Susceptibility. *JAMA Intern. Med.* **175**, 821–826 (2015).

23. Murad, A. M. *et al.* Spontaneous coronary artery dissection is infrequent in individuals with heritable thoracic aortic disease despite partially shared genetic susceptibility. *Am. J. Med. Genet. A.* **188**, 1448–1456 (2022).

24. Adlam, D. *et al.* Association of the PHACTR1/EDN1 Genetic Locus With Spontaneous Coronary Artery Dissection. *J. Am. Coll. Cardiol.* **73**, 58–66 (2019).

25. Shu, L. *et al.* Shared genetic regulatory networks for cardiovascular disease and type 2 diabetes in multiple populations of diverse ethnicities in the United States. *PLOS Genet.* **13**, e1007040 (2017).

26. Zhao, Y. *et al.* Multi-omics integration reveals molecular networks and regulators of psoriasis. *BMC Syst. Biol.* **13**, 8 (2019).

27. Nativio, R. *et al.* An integrated multi-omics approach identifies epigenetic alterations associated with Alzheimer's disease. *Nat. Genet.* **52**, 1024–1035 (2020).

28. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**, 1177–1186 (2017).

29. Yang, X. Multitissue Multiomics Systems Biology to Dissect Complex Diseases. *Trends Mol. Med.* **26**, 718–728 (2020).

30. Chen, Y.-W. *et al.* PharmOmics: A species- and tissue-specific drug signature database and gene-network-based drug repositioning tool. *iScience* **25**, 104052 (2022).

31. Subramanian, A. *et al.* A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **171**, 1437-1452.e17 (2017).

32. Wang, Z., Clark, N. R. & Ma'ayan, A. Drug-induced adverse events prediction with the LINCS L1000 data. *Bioinformatics* **32**, 2338–2345 (2016).

33. Kanehisa, M., Goto, S., Kawashima, S. & Nakaya, A. The KEGG databases at GenomeNet. *Nucleic Acids Res.* **30**, 42–46 (2002).

34. Joshi-Tope, G. *et al.* Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res.* **33**, D428–D432 (2005).

35. Nishimura, D. BioCarta. *Biotech Softw. Internet Rep.* **2**, 117–120 (2001).

36. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).

37. Emilsson, V. *et al.* Coding and regulatory variants are associated with serum protein levels and disease. *Nat. Commun.* **13**, 481 (2022).

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

39. Song, W.-M. & Zhang, B. Multiscale Embedded Gene Co-expression Network Analysis. *PLOS Comput. Biol.* **11**, e1004574 (2015).

40. Shu, L. *et al.* Mergeomics: multidimensional data integration to identify pathogenic perturbations to biological systems. *BMC Genomics* **17**, 874 (2016).

41. Ding, J. *et al.* Mergeomics 2.0: a web server for multi-omics data integration to elucidate disease networks and predict therapeutics. *Nucleic Acids Res.* **49**, W375–W387 (2021).

42. Zhu, J. *et al.* Stitching together Multiple Data Dimensions Reveals Interacting Metabolomic and Transcriptomic Networks That Modulate Cell Regulation. *PLOS Biol.* **10**, e1001301 (2012).

43. Zhu, J. *et al.* Increasing the Power to Detect Causal Associations by Combining Genotypic and Expression Data in Segregating Populations. *PLOS Comput. Biol.* **3**, e69 (2007).

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

45. Zhu, J. *et al.* An integrative genomics approach to the reconstruction of gene networks in segregating populations. *Cytogenet. Genome Res.* **105**, 363–374 (2004).

46. Lusis, A. J. *et al.* The Hybrid Mouse Diversity Panel: a resource for systems genetics analyses of metabolic and cardiovascular traits. *J. Lipid Res.* **57**, 925–942 (2016).

47. Bennett, B. J. *et al.* Genetic Architecture of Atherosclerosis in Mice: A Systems Genetics Analysis of Common Inbred Strains. *PLOS Genet.* **11**, e1005711 (2015).

48. Timalsina, P., Charles, K. & Mondal, A. M. STRING PPI Score to Characterize Protein Subnetwork Biomarkers for Human Diseases and Pathways. in *2014 IEEE International Conference on Bioinformatics and Bioengineering* 251–256 (2014). doi:10.1109/BIBE.2014.46.

49. Shannon, P. *et al.* Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **13**, 2498–2504 (2003).

50. Cheng, F. *et al.* Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat. Commun.* **9**, 2691 (2018).

51. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).

52. Jankowich, M. & Choudhary, G. Endothelin-1 levels and cardiovascular events. *Trends Cardiovasc. Med.* **30**, 1–8 (2020).

53. Tweet, M. S., Miller, V. M. & Hayes, S. N. The Evidence on Estrogen, Progesterone, and Spontaneous Coronary Artery Dissection. *JAMA Cardiol.* **4**, 403–404 (2019).

54. Margaritis, M. *et al.* Vascular histopathology and connective tissue ultrastructure in spontaneous coronary artery dissection: pathophysiological and clinical implications. *Cardiovasc Res* (2021).

55. Wagenseil, J. E. & Mecham, R. P. Vascular Extracellular Matrix and Arterial Mechanics. *Physiol. Rev.* **89**, 957–989 (2009).

56. Turley, T. N. *et al.* Rare Missense Variants in TLN1 Are Associated With Familial and Sporadic Spontaneous Coronary Artery Dissection. *Circ. Genomic Precis. Med.* **12**, e002437 (2019).

57. Kielty, C. M. & Grant, M. E. The Collagen Family: Structure, Assembly, and Organization in the Extracellular Matrix. in *Connective Tissue and Its Heritable Disorders* 159–221 (John Wiley & Sons, Ltd, 2002). doi:10.1002/0471221929.ch2.

58. Zekavat, S. M. *et al.* Fibrillar Collagen Variants in Spontaneous Coronary Artery Dissection. *JAMA Cardiol.* **7**, 396–406 (2022).

59. Smith, S. A., Travers, R. J. & Morrissey, J. H. How it all starts: Initiation of the clotting cascade. *Crit. Rev. Biochem. Mol. Biol.* **50**, 326–336 (2015).

60. Okwusidi, J. I., Anvari, N. & Ofosu, F. A. Modulation of intrinsic prothrombin activation by fibrinogen and fibrin I. *J. Lab. Clin. Med.* **121**, 64–70 (1993).

61. Krittanawong, C. *et al.* Clinical features and prognosis of patients with spontaneous coronary artery dissection. *Int. J. Cardiol.* **312**, 33–36 (2020).

62. Meng, P. N. *et al.* [Characteristics of acute myocardial infarction caused by spontaneous coronary artery dissection in young female patients]. *Zhonghua Xin Xue Guan Bing Za Zhi* **46**, 536–542 (2018).

63. Xian, X. *et al.* LRP1 integrates murine macrophage cholesterol homeostasis and inflammatory responses in atherosclerosis. *eLife* **6**, e29292 (2017).

64. Boucher, P., Gotthardt, M., Li, W.-P., Anderson, R. G. W. & Herz, J. LRP: Role in Vascular Wall Integrity and Protection from Atherosclerosis. *Science* **300**, 329–332 (2003).

65. Komiya, Y. & Habas, R. Wnt signal transduction pathways. *Organogenesis* **4**, 68–75 (2008).

66. Matthijs Blankesteijn, W. & Hermans, K. C. M. Wnt signaling in atherosclerosis. *Eur. J. Pharmacol.* **763**, 122–130 (2015).

67. Shufelt, C. L., Pacheco, C., Tweet, M. S. & Miller, V. M. Sex-Specific Physiology and Cardiovascular Disease. in *Sex-Specific Analysis of Cardiovascular Function* (eds. Kerkhof, P. L. M. & Miller, V. M.) 433–454 (Springer International Publishing, 2018). doi:10.1007/978-3-319-77932-4_27.

68. Souilhol, C. *et al.* Homeobox B9 integrates bone morphogenic protein 4 with inflammation at atheroprone sites. *Cardiovasc. Res.* **116**, 1300–1310 (2020).

69. Ansari, K. I., Shrestha, B., Hussain, I., Kasiri, S. & Mandal, S. S. Histone Methylases MLL1 and MLL3 Coordinate with Estrogen Receptors in Estrogen-Mediated HOXB9 Expression. *Biochemistry* **50**, 3517–3527 (2011).

70. Jin, J. *et al.* The Implicated Roles of Cell Adhesion Molecule 1 (CADM1) Gene and Altered Prefrontal Neuronal Activity in Attention-Deficit/Hyperactivity Disorder: A “Gene–Brain–Behavior Relationship”? *Front. Genet.* **10**, (2019).

71. Zhang, J. *et al.* GW29-e0967 MicroRNA-6321 contributes the myocardial fibrosis through regulating cell adhesion and proliferation molecules in heart failure. *J. Am. Coll. Cardiol.* **72**, C31–C32 (2018).

72. Takeuchi, A. *et al.* CRTAM confers late-stage activation of CD8+ T cells to regulate retention within lymph node. *J. Immunol. Baltim. Md 1950* **183**, 4220–4228 (2009).

73. Peloso, G. M. *et al.* Association of Low-Frequency and Rare Coding-Sequence Variants with Blood Lipids and Coronary Heart Disease in 56,000 Whites and Blacks. *Am. J. Hum. Genet.* **94**, 223–232 (2014).

74. Bishop, J. R. *et al.* Deletion of the Basement Membrane Heparan Sulfate Proteoglycan Type XVIII Collagen Causes Hypertriglyceridemia in Mice and Humans. *PLOS ONE* **5**, e13919 (2010).

75. Ambade, A. S., Hassoun, P. M. & Damico, R. L. Basement Membrane Extracellular Matrix Proteins in Pulmonary Vascular and Right Ventricular Remodeling in Pulmonary Hypertension. *Am. J. Respir. Cell Mol. Biol.* **65**, 245–258 (2021).

76. Dianat, S. *et al.* Side Effects and Health Benefits of Depot Medroxyprogesterone Acetate: A Systematic Review. *Obstet. Gynecol.* **133**, 332–341 (2019).

77. Moray, K. V., Chaurasia, H., Sachin, O. & Joshi, B. A systematic review on clinical effectiveness, side-effect profile and meta-analysis on continuation rate of etonogestrel contraceptive implant. *Reprod. Health* **18**, 4 (2021).

78. Elkayam, U. *et al.* Pregnancy-Associated Acute Myocardial Infarction. *Circulation* **129**, 1695–1702 (2014).

79. Sheikh, A. S. & O'Sullivan, M. Pregnancy-related Spontaneous Coronary Artery Dissection: Two Case Reports and a Comprehensive Review of Literature. *Heart Views Off. J. Gulf Heart Assoc.* **13**, 53–65 (2012).

80. Yip, A. & Saw, J. Spontaneous coronary artery dissection—A review. *Cardiovasc. Diagn. Ther.* **5**, 37–48 (2015).

81. Saw, J. Pregnancy-Associated Spontaneous Coronary Artery Dissection Represents an Exceptionally High-Risk Spontaneous Coronary Artery Dissection Cohort. *Circ. Cardiovasc. Interv.* **10**, e005119 (2017).

82. Kamel, H., Roman, M. J., Pitcher, A. & Devereux, R. B. Pregnancy and the Risk of Aortic Dissection or Rupture. *Circulation* **134**, 527–533 (2016).

83. Gilhofer, T. S. & Saw, J. Spontaneous coronary artery dissection: update 2019. *Curr. Opin. Cardiol.* **34**, 594–602 (2019).

84. Leão, R. P. *et al.* Identification of New Rofecoxib-Based Cyclooxygenase-2 Inhibitors: A Bioinformatics Approach. *Pharmaceuticals* **13**, 209 (2020).

85. Saper, J. *et al.* Rofecoxib in the Acute Treatment of Migraine: A Randomized Controlled Clinical Trial. *Headache J. Head Face Pain* **46**, 264–275 (2006).

86. Kok, S. N. *et al.* Prevalence and Clinical Factors of Migraine in Patients With Spontaneous Coronary Artery Dissection. *J. Am. Heart Assoc.* **7**, e010140 (2018).

87. Bruno, G. *et al.* Proton pump inhibitors and dysbiosis: Current knowledge and aspects to be clarified. *World J. Gastroenterol.* **25**, 2706–2719 (2019).

88. Shiraev, T. P. & Bullen, A. Proton Pump Inhibitors and Cardiovascular Events: A Systematic Review. *Heart Lung Circ.* **27**, 443–450 (2018).

89. Ariel, H. & Cooke, J. P. Cardiovascular Risk of Proton Pump Inhibitors. *Methodist DeBakey Cardiovasc. J.* **15**, 214–219 (2019).

90. Tan, N. Y. & Tweet, M. S. Spontaneous coronary artery dissection: etiology and recurrence. *Expert Rev. Cardiovasc. Ther.* **17**, 497–510 (2019).

91. Dong, C., Davis, R. J. & Flavell, R. A. MAP Kinases in the Immune Response. *Annu. Rev. Immunol.* **20**, 55–72 (2002).

92. Soares-Silva, M., Diniz, F. F., Gomes, G. N. & Bahia, D. The Mitogen-Activated Protein Kinase (MAPK) Pathway: Role in Immune Evasion by Trypanosomatids. *Front. Microbiol.* **7**, (2016).

93. Yamamoto, Y. & Gaynor, R. B. Role of the NF-kB Pathway in the Pathogenesis of Human Disease States. *Curr. Mol. Med.* **1**, 287–296 (2001).

94. Pitliya, A. *et al.* Eosinophilic inflammation in spontaneous coronary artery dissection: A potential therapeutic target? *Med. Hypotheses* **121**, 91–94 (2018).

95. Ekert, P. G. & Vaux, D. L. Apoptosis and the immune system. *Br. Med. Bull.* **53**, 591–603 (1997).

96. Blencowe, M., Karunanayake, T., Wier, J., Hsu, N. & Yang, X. Network Modeling Approaches and Applications to Unravelling Non-Alcoholic Fatty Liver Disease. *Genes* **10**, 966 (2019).

97. Koeth, R. A. *et al.* Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat. Med.* **19**, 576–585 (2013).

98. 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).

99. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).

100. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* **8**, 329-337.e4 (2019).

101. Liberzon, A. *et al.* The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst.* **1**, 417–425 (2015).

102. Arneson, D. *et al.* Single cell molecular alterations reveal target cells and pathways of concussive brain injury. *Nat. Commun.* **9**, 3894 (2018).

103. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).

104. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).

105. Hartmann, F. *et al.* SMC-Derived Hyaluronan Modulates Vascular SMC Phenotype in Murine Atherosclerosis. *Circ. Res.* **129**, 992–1005 (2021).

106. Fitch, M. J., Campagnolo, L., Kuhnert, F. & Stuhlmann, H. Egfl7, a novel epidermal growth factor-domain gene expressed in endothelial cells. *Dev. Dyn.* **230**, 316–324 (2004).

107. Wu, Z., Zhang, Z., Lei, Z. & Lei, P. CD14: Biology and role in the pathogenesis of disease. *Cytokine Growth Factor Rev.* **48**, 24–31 (2019).

108. Buechler, M. B. *et al.* Cross-tissue organization of the fibroblast lineage. *Nature* **593**, 575–579 (2021).

109. 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).

110. Bennett, M. R., Sinha, S. & Owens, G. K. Vascular Smooth Muscle Cells in Atherosclerosis. *Circ. Res.* **118**, 692–702 (2016).

111. Gomez, D. & Owens, G. K. Smooth muscle cell phenotypic switching in atherosclerosis. *Cardiovasc. Res.* **95**, 156–164 (2012).

112. Brandon Brown, Bandar E Almansouri, Diane E Heck, & Hong Duck Kim. Surveillance utilizes multi-omics in cardiovascular disease: Diet and its potentiality in Preventive index. *GSC Biol. Pharm. Sci.* **17**, 001–009 (2021).

113. Hayashi, T. *et al.* Tcap gene mutations in hypertrophic cardiomyopathy and dilated cardiomyopathy. *J. Am. Coll. Cardiol.* **44**, 2192–2201 (2004).

114. Woulfe, K. C. *et al.* Fibrosis and Fibrotic Gene Expression in Pediatric and Adult Patients With Idiopathic Dilated Cardiomyopathy. *J. Card. Fail.* **23**, 314–324 (2017).

115. Chistiakov, D. A., Bobryshev, Y. V. & Orekhov, A. N. Changes in transcriptome of macrophages in atherosclerosis. *J. Cell. Mol. Med.* **19**, 1163–1173 (2015).

116. Yang, X. *et al.* Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks. *Nat. Genet.* **41**, 415–423 (2009).

117. Atri, C., Guerfali, F. Z. & Laouini, D. Role of Human Macrophage Polarization in Inflammation during Infectious Diseases. *Int. J. Mol. Sci.* **19**, 1801 (2018).

118. Jablonski, K. A. *et al.* Novel Markers to Delineate Murine M1 and M2 Macrophages. *PLOS ONE* **10**, e0145342 (2015).

119. Orecchioni, M., Ghosheh, Y., Pramod, A. B. & Ley, K. Macrophage Polarization: Different Gene Signatures in M1(LPS+) vs. Classically and M2(LPS–) vs. Alternatively Activated Macrophages. *Front. Immunol.* **10**, (2019).

120. Oksala, N. *et al.* Kindlin 3 (FERMT3) is associated with unstable atherosclerotic plaques, anti-inflammatory type II macrophages and upregulation of beta-2 integrins in all major arterial beds. *Atherosclerosis* **242**, 145–154 (2015).

121. Chinetti-Gbaguidi, G., Colin, S. & Staels, B. Macrophage subsets in atherosclerosis. *Nat. Rev. Cardiol.* **12**, 10–17 (2015).

122. Consortium, T. G. O. Creating the Gene Ontology Resource: Design and Implementation. *Genome Res.* **11**, 1425–1433 (2001).

123. Singh, S. & Torzewski, M. Fibroblasts and Their Pathological Functions in the Fibrosis of Aortic Valve Sclerosis and Atherosclerosis. *Biomolecules* **9**, 472 (2019).

124. Chen, P. *et al.* Identification and validation of four hub genes involved in the plaque deterioration of atherosclerosis. *Aging* **11**, 6469–6489 (2019).

125. Kraja, A. T. *et al.* Associations of Mitochondrial and Nuclear Mitochondrial Variants and Genes with Seven Metabolic Traits. *Am. J. Hum. Genet.* **104**, 112–138 (2019).

126. Fetterman, J. L. *et al.* Relations of mitochondrial genetic variants to measures of vascular function. *Mitochondrion* **40**, 51–57 (2018).

127. Wang, W. *et al.* Ribosomal Proteins and Human Diseases: Pathogenesis, Molecular Mechanisms, and Therapeutic Implications. *Med. Res. Rev.* **35**, 225–285 (2015).

128. Nus, M. *et al.* Endothelial Jag1-RBPJ signalling promotes inflammatory leucocyte recruitment and atherosclerosis. *Cardiovasc. Res.* **112**, 568–580 (2016).

129. 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).

130. Lu, Z. *et al.* Extracellular Superoxide Dismutase Deficiency Exacerbates Pressure Overload–Induced Left Ventricular Hypertrophy and Dysfunction. *Hypertension* **51**, 19–25 (2008).

131. Juul, K. *et al.* Genetically Reduced Antioxidative Protection and Increased Ischemic Heart Disease Risk. *Circulation* **109**, 59–65 (2004).

132. Yamada, H. *et al.* Protective Role of Extracellular Superoxide Dismutase in Hemodialysis Patients. *Nephron* **84**, 218–223 (2000).

133. Gupta, R. M. Causal Gene Confusion: The Complicated EDN1/PHACTR1 Locus for Coronary Artery Disease. *Arterioscler. Thromb. Vasc. Biol.* **42**, 610–612 (2022).

134. Ghose, S. *et al.* Investigating Coronary Artery Disease methylome through targeted bisulfite sequencing. *Gene* **721**, 144107 (2019).

135. Neele, A. E. *et al.* Macrophage Kdm6b controls the pro-fibrotic transcriptome signature of foam cells. *Epigenomics* **9**, 383–391 (2017).

136. Neele, A. E. *et al.* Myeloid Kdm6b deficiency results in advanced atherosclerosis. *Atherosclerosis* **275**, 156–165 (2018).

137. Nair, J., Ghatge, M., Kakkar, V. V. & Shanker, J. Network Analysis of Inflammatory Genes and Their Transcriptional Regulators in Coronary Artery Disease. *PLOS ONE* **9**, e94328 (2014).

138. Cipollone, F. *et al.* Overexpression of Functionally Coupled Cyclooxygenase-2 and Prostaglandin E Synthase in Symptomatic Atherosclerotic Plaques as a Basis of Prostaglandin E2-Dependent Plaque Instability. *Circulation* **104**, 921–927 (2001).

139. Xu, S. *et al.* Flow-dependent epigenetic regulation of IGFBP5 expression by H3K27me3 contributes to endothelial anti-inflammatory effects. *Theranostics* **8**, 3007–3021 (2018).

140. Roldán-Montero, R. *et al.* Galectin-1 prevents pathological vascular remodeling in atherosclerosis and abdominal aortic aneurysm. *Sci. Adv.* **8**, eabm7322 (2022).

141. Moiseeva, E. P., Javed, Q., Spring, E. L. & de Bono, D. P. Galectin 1 is involved in vascular smooth muscle cell proliferation. *Cardiovasc. Res.* **45**, 493–502 (2000).

142. Chetry, M. *et al.* Effects of Galectin-1 on Biological Behavior in Cervical Cancer. *J. Cancer* **11**, 1584–1595 (2020).

143. Cyrus, T., Ding, T. & Praticò, D. Expression of thromboxane synthase, prostacyclin synthase and thromboxane receptor in atherosclerotic lesions: Correlation with plaque composition. *Atherosclerosis* **208**, 376–381 (2010).

144. Rodius, S. *et al.* Transcriptional response to cardiac injury in the zebrafish: systematic identification of genes with highly concordant activity across in vivo models. *BMC Genomics* **15**, 852 (2014).

145. Sun, G. G. *et al.* Expression of BTG1 in hepatocellular carcinoma and its correlation with cell cycles, cell apoptosis, and cell metastasis. *Tumor Biol.* **35**, 11771–11779 (2014).

146. Liu, C. *et al.* BTG1 potentiates apoptosis and suppresses proliferation in renal cell carcinoma by interacting with PRMT1. *Oncol. Lett.* **10**, 619–624 (2015).

147. Ma, Y. *et al.* Hyperlipidemia and Atherosclerotic Lesion Development in Ldlr-Deficient Mice on a Long-Term High-Fat Diet. *PLOS ONE* **7**, e35835 (2012).