DISSERTATIONES
MEDICINAE
UNIVERSITATIS
TARTUENSIS
386

KATYAYANI SUKHAVASI

Single-cell RNA sequencing analysis integrated with human gene-regulatory networks provides mechanistic insights of advanced atherosclerosis in men and women

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS
386

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS
386

KATYAYANI SUKHAVASI

Single-cell RNA sequencing analysis integrated with human gene-regulatory networks provides mechanistic insights of advanced atherosclerosis in men and women

UNIVERSITY OF TARTU
Press
1632

Department of Cardiology, Institute of Clinical Medicine, Faculty of Medicine, University of Tartu, Estonia

This dissertation is accepted for the commencement of the degree of Doctor of Philosophy in Medicine on August 27th, 2025, by the Council of the Faculty of Medicine, University of Tartu, Estonia.

Supervisors:	Arno Ruusalepp, MD, PhD, Associate Professor in Cardiac Surgery, Department of Cardiology, Institute of Clinical Medicine, University of Tartu
	Johan. L.M. Björkegren, MD, PhD, Principal investigator and Associate Professor in Molecular Medicine, Department of Medicine, Huddinge, Karolinska Institutet; Adjunct Professor at the Department of Genetics and Genomic Sciences, Mount Sinai Hospital, New York; Visiting Professor of Tartu university and Tokyo College.
Reviewers:	Sander Pajusalu, MD, PhD, Associate Professor in Clinical Genetics, Department of Genetics and Personalized Medicine, Institute of Clinical Medicine, University of Tartu, Estonia.
	Jaak Kals, MD, PhD, Professor in Vasology, Department of Surgery, Institute of Clinical Medicine, University of Tartu, Estonia.
Opponent:	Prof. Dr. rer. nat. Tanja Zeller (PhD), Institute for Cardiogenetics, University of Lübeck, Germany.
Commencement:	November 4th, 2025

The publication of this dissertation is granted by the University of Tartu.

ISSN 1024-395X (print) ISSN 2806-240X (pdf)
ISBN 978-9916-27-991-5 (print) ISBN 978-9916-27-992-2 (pdf)

University of Tartu Press
www.tyk.ee

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Tables:

Figures:

LIST OF ORIGINAL PUBLICATIONS

I. **Sukhavasi, K**., Mocci, G., Ma, L., Hodonsky, C. J., Diez Benevante, E., Muhl, L., Liu, J., Gustafsson, S., Buyandelger, B., Koplev, S., Lendahl, U., Vanlandewijck, M., Singha, P., Örd, T., Beter, M., Selvarajan, I., Laakkonen, J. P., Väli, M., den Ruijter, H. M., ... Björkegren, J. L. M. (2025). Single-cell RNA sequencing reveals sex differences in the subcellular composition and associated gene-regulatory network activity of human carotid plaques. *Nature Cardiovascular Research*, *4*(4), 412–432.

II. Mocci, G.,* **Sukhavasi, K.,*** Örd, T., Bankier, S., Singha, P., Arasu, U. T., Agbabiaje, O. O., Mäkinen, P., Ma, L., Hodonsky, C. J., Aherrahrou, R., Muhl, L., Liu, J., Gustafsson, S., Byandelger, B., Wang, Y., Koplev, S., Lendahl, U., Owens, G. K., ... Björkegren, J. L. M. (2024). Single-Cell Gene-Regulatory Networks of Advanced Symptomatic Atherosclerosis. *Circulation Research*, *134*(11), 1405–1423.

III. Koplev, S., Seldin, M., **Sukhavasi, K**., Ermel, R., Pang, S., Zeng, L., Bankier, S., Di Narzo, A., Cheng, H., Meda, V., Ma, A., Talukdar, H., Cohain, A., Amadori, L., Argmann, C., Houten, S. M., Franzén, O., Mocci, G., Meelu, O. A., ... Björkegren, J. L. M. (2022). A mechanistic framework for cardiometabolic and coronary artery diseases. *Nature Cardiovascular Research*, *1*(1), 85–100.

* Shared first authorship.

Author's contribution to each publication (I–III) is as follows:

Study I: Conceived the study, designed and performed the experiments, analyzed, interpreted the data, and wrote the original manuscript.

Study II: Conceived the study, designed and performed the single cell experiments of human carotid plaques, analyzed the human scRNA data, and wrote the original manuscript with G. Mocci.

Study III: Designed and performed experiments; isolation of RNA from relevant tissues and whole blood, quality and quantity measurement of RNA. Isolation of DNA from whole blood, quality and quantity measurement of DNA. Responsible for the STARNET biobank. Contributed to writing the original manuscript.

LIST OF ABBREVIATIONS

A

ABCA1	ATP-binding cassette transporter
ABCC9	ATP binding cassette subfamily C member 9
ABCG1	ATP-binding cassette sub-family G member 1
Adenine (A)	Adenine
AE study	Athero-Express
AOR	Atherosclerotic aortic root

B

BMI	Body mass index

C

CAD	Coronary artery disease
CCDs	Common complex diseases
CD	Cluster of differentiation
CD3E	T-cell surface glycoprotein CD3 Epsilon chain
CD-RV	Common disease rare variant hypothesis
CEL-seq	Cell Expression by Linear amplification and Sequencing
CHD	Coronary heart disease
CLDN5	Claudin5
CMDs	Cardiometabolic disorders
CNV	Copy-number variants
CTRL	Control
CVDs	Cardiovascular diseases
Cytosine (C)	Cytosine

D

DEGs	Differential gene expression analysis
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate

E

EC	Endothelial cells
ECM	Extracellular matrix
EndICLT	Endothelial to immune cell like transition
EndoMT	Endothelial cells transition to mesenchymal-like cells
eQTLs	Expression quantitative trait loci
eSNPs	Expressed Single Nucleotide Polymorphism
ET-1	Endothelin

F

FACS	Fluorescence-activated cell sorting

FB	Fibroblasts
FCS-A	Forward scatter-area
FDRs	False discovery rates

G

GBD	Global Burden of Disease
GGES	Genetics of Gene Expression Studies
GO	Gene Ontology
GRNs	Gene regulatory networks
GSEA	Gene set enrichment analysis
GTEx	Gene–Tissue Expression
Guanine (G)	Guanine
GWAS	Genome-wide association studies

H

H^2	Broad-sense heritability
h^2	Narrow-sense heritability
HAEC	Human aortic endothelial cells
HCASMCs	Human coronary artery smooth muscle cells
HMDP	Hybrid mouse diversity panel

I

ICAM-1	Intracellular cell adhesion molecule-1
IFCs	Integrated fluidics circuits
IL-8	Interleukin 8
IMT	Intima-media thickness
IVT	In vitro transcription

K

KDRs	Key driver genes
KNN	K-nearest neighbor

L

LD	Linkage disequilibrium
LDL	Low-density lipoprotein
LIV	Liver
LNA	Locked nucleic acid

M

MACS	Magnetic-Activated Cell Sorting
MAF	Minor allele frequency
MAM	Internal mammary artery
MCP-1	Macrophage chemoattractant protein-1
MI	Myocardial infarction
MMLV	Moloney murine leukemia virus
MMP-1	Metalloproteinase-1

MMP-9	Metalloproteinase-9
MOD	Morbidity of death
MP	Macrophages
MR	Mendelian Randomization
mRNA	Messenger ribonucleic acid
MYH11	Myosin11

N

NES	Normalized enrichment scores
NGS	Next generation sequencing
NO	Nitric oxide
NR4A1-3	Nuclear receptor subfamily 4
NSTEMI	non-ST elevation MI

P

PAD	Peripheral artery disease
PCA	Principal Component Analysis
PCs	Pericytes
polyA$^-$	Non polyadenylated
polyA$^+$	Polyadenylated

Q

qPCR	Quantitative PCR

R

RNA-Seq	RNA-sequencing
rRNA	Ribosomal ribonucleic acid
RT	Reverse transcription
RTase	Reverse transcriptase
RVs	Rare variants

S

SCRB-seq	Single-Cell RNA-Barcoding and Sequencing
scRNA-Seq	Single-cell RNA-Seq
SF	Subcutaneous fat
SKAT	Sequence Kernel association test
SKLM	Skeletal muscle
SMART-seq2	Switch Mechanism at the 5' End of RNA Template sequencing
SMC	Smooth (involuntary) muscle cells
SNP	Single nucleotide polymorphism
SNVs	Single-nucleotide variants
SPP1	Osteopontin
SR-B1	Scavenger receptor B1
SSC-A	Side scatter area
STAGE	Stockholm Atherosclerosis Gene Expression

STARNET	Stockholm Tartu Atherosclerosis Reverse Network Engineering Task
STEMI	ST-elevation MI
STRT-seq	Single-cell tagged reverse transcription sequencing

T

Thymine (T)	Thymine
TIA	Transitory ischemic attacks
tRNA	Transfer ribonucleic acid
t-SNE	t-distributed stochastic neighbor embedding
TSO	Template switching oligos

U

UMAP	Uniform manifold approximation and projection
UMIs	Unique molecular identifiers

V

VAF	Visceral abdominal fat
VCAM-1	Vascular cell adhesion molecule-1

W

WES	Whole-exome sequencing
WGCNA	Weighted correlation network analysis
WGS	Whole-genome sequencing
WNT	Wingless-related integration site

1. INTRODUCTION

Cardiovascular diseases (CVDs) are a global health crisis, claiming millions of lives each year, three-quarters of which take place in low- and middle-income countries. As per the World Health Organization (WHO) 2022, 19.8 million died from CVDs representing 32% of all global deaths. Of these deaths, 85% were due to heart attacks and stroke. The Global Burden of Disease (GBD) study 2019 shows that the prevalent total CVD nearly doubled from 271 million in 1990 to 523 million in 2019 (Roth et al., 2020). CVDs encompass a range of disorders affecting the heart or blood vessels. The main CVD involving the arterial wall is atherosclerosis ("hardening of arteries") leading to coronary artery disease (CAD), cerebrovascular disease (CVD) and peripheral artery disease (PAD). While PAD mainly affects the blood flow leading to lower limb ischemia particularly following physical activity (i.e., intermittent claudication), but CAD also cause the formation of blood clots risking a myocardial infarction or if these clots escape the heart travelling to the brain, a plausible stroke.

Atherosclerosis is a common complex degenerative disease that develops over time by altering the cellular composition of the arterial wall mainly in the form of the infiltration of macrophages that become lipid-loaded forming foam cells eventually leading to the formation of the atherosclerotic plaques (Libby, 2021; Lusis et al., 2004). The degree of atherosclerosis development depends on the interaction of environmental exposures (e.g., lifestyle choices) with the genetic makeup in any given individual. This interaction of genetic and environmental factors over time primarily alters gene expression not only in the arterial wall but also in organs affecting the metabolic status of the individual, like in the liver, adipose tissues and skeletal muscle. Changes to gene expression affecting the metabolic homeostasis, like the development of the "metabolic syndrome" with increased body weight, inflammation as well as higher levels of circulating blood glucose and lipids (e.g., low-density lipoproteins (LDL)), accelerates the development of atherosclerosis (Mach et al., 2019; Tegner et al., 2007). Notably, recent studies have also highlighted that atherosclerotic development and risk thereof differ between males and females, where females initially are somewhat protected until menopause, whereafter their risk increases as do their development of atherosclerosis (Maas & Appelman, 2010; Witteman et al., 1989). However, whether these differences are reflected in the actual pathology and cellular composition of the atherosclerotic plaques in males versus females remain unknown.

Common complex diseases (CCDs), like atherosclerosis, are more difficult to study than rare disorders following a strict Mendelian pattern of inheritance (Manolio et al., 2009; Pritchard & Cox, 2002; Visscher et al., 2012). The multifactorial nature (genetic and environmental) of complex diseases makes it challenging to pinpoint any single pathway or for that matter gene that has any major influence on the trait. Instead, their underlying causes rely on hundreds if not thousands of genes with shifting impact over time. Nonetheless in the early

nineties, with the advancements in genetic screening technologies following their successful use for rare disorders often pinpointing causal pathways and key genes (Gusella et al., 1983; Tsui et al., 1985) the genome-wide association studies (GWAS) were introduced. These studies relied on the "common disease, common variant" hypothesizing that a limited number of genetic variants with high frequency (typically above 5%) in the general population would substantially contribute to common disease susceptibility (Sivapalaratnam et al., 2011). CAD has a significant genetic component, with family studies estimating heritability between 40% and 60%. However, although GWAS of CAD and myocardial infarction (MI) since 2007 have yielded significant discoveries, including the identification of the chromosome 9p21 risk locus by multiple research groups (Helgadottir et al., 2007; McPherson et al., 2007; Samani et al., 2007), and over 241 independent variants in 198 risk loci, (Silva et al., 2023), their combined contribution to CAD heritability is limited, with estimates typically ranging from 10–25%. More recent findings suggest that common variants, identified by GWAS, explain a smaller proportion of the total heritability (around 17% or less), as a significant portion is now attributed to ultra-rare variants (Rocheleau et al., 2024). Furthermore, these risk variants have for the most part not been found within the coding areas of given genes but in the non-coding vicinity of many genes. Thus, the "candidate gene approach" which focuses on associations between genetic variants and a limited number of predefined genes that was successfully applied in studies of single-gene disorders, has proven inadequate for complex diseases (Joel N. Hirschhorn & Mark J. Daly, 2005; Manolio, 2010).

This "missing heritability" suggests that other genetic factors, such as rare variants, structural variants, or gene-gene interactions, may play a role in CAD risk (Zeng et al., 2019). Others have suggested that CAD common risk variants identified by GWAS mainly affect the slow or early phase of atherosclerosis development while the "missing variants" are those that affect the later disease stages where environmental exposures such as the state of inflammation is critical, supposedly that are missed by current GWAS (Khan et al., 2023). To address the missing heritability, the importance of genetics-of-gene-expression studies in disease-relevant tissues and cells has increasingly been recognized. This approach aims to better define genes responsible for risk on GWAS loci and those that capture micro-environmental factors affecting atherosclerosis development (Björkegren et al., 2015; Bjorkegren & Lusis, 2022). The transcriptome is the complete set of RNA molecules produced in a cell or organism at a given time point. RNA-sequencing (RNA-Seq) is a powerful technology that has revolutionized transcriptome analysis, surpassing microarray technologies, with a resolution at the single base level. By combining the analysis of RNA-Seq with genotype data, single nucleotide polymorphism (SNPs) that modulate transcript levels referred to as expression quantitative trait loci (eQTLs) or expression single nucleotides polymorphisms (eSNPs) can be identified associating local (*cis*) or distal (*trans*) eSNPs with gene expression levels (Grundberg et al., 2012). Applying machine learning (Huynh-Thu et al., 2010) or Bayesian probabilistic modelling to the combination of genotype and gene expression data, it is possible

to go beyond identifying disease-causal genes one-by-one and instead infer gene regulatory networks (GRNs) containing hundred to thousands of genes representing cellular and organ pathophysiology including CAD-relevant pathways. Using the directionality of gene interactions in these GRNs, genes critical for the overall activity of GRNs named *key driver genes* (KDRs) can be identified and ranked. Top-ranked KDRs with GRNs associated with the extent of atherosclerosis has shown to be suitable targets to prevent atherosclerosis development. Overall, GRNs inferred from functional data (e.g., RNAseq) generated from disease-relevant tissues is rapidly emerging as a valuable mechanistic framework to decipher complex disease etiology (Bjorkegren & Lusis, 2022). The framework of GRNs provides a much needed "top-down", systems view of complex diseases to parallel the well-established "bottom-up" approach of isolated candidate genes and pathways one-by-one.

In our research, the STARNET ("Stockholm Tartu Atherosclerosis Reverse Network Engineering Task") study, one of the largest human genetics-of-gene expression studies combining RNA-seq analysis from blood, metabolic organs, and the arterial walls isolated from both males and females, is central (Franzen et al., 2016). In STARNET, tissues have been isolated CAD patients undergoing coronary artery bypass grafting (CABG) surgery ("cases") as well as patients without CAD (confirmed by pre-operative angiography) undergoing other forms of open thorax surgery (e.g., isolated mitral valve repair) ("controls"). In 2022, we identified 224 GRNs from STARNET by integrating genetic and clinical data from patients with (n = 600) and without (n = 250) CAD together with RNA-seq data from seven disease-relevant tissues including blood, the early atherosclerotic internal mammary artery (MAM), the advanced atherosclerotic aortic root (AOR), subcutaneous fat (SF), visceral abdominal fat (VAF), skeletal muscle (SKLM), and liver (LIV). eQTLs of these 224 GRNs were found to explain >54% of CAD heritability beyond the 21 % explained by GWAS risk loci of CAD, and particularly the 89 cross-tissue GRNs were found associated with CAD severity in STARNET (S. Koplev et al., 2022) forming a first mechanistic framework of CMDs and CAD. Although STARNET through these GRNs give an important first systems view of CAD, the GRN framework is inferred from RNA isolated from whole tissues ("bulk RNA"). Therefore, a portion (often 10 to over 50%) of the variation in bulk RNAseq data does not reflect actual mRNA variation but rather differences in the cellular composition of the tissue samples (Gustafsson et al., 2020).

Single-cell RNA-Seq (scRNA-Seq) overcomes this limitation by analyzing gene expression at the individual cell level. This allows for the identification of rare cell populations, characterization of cell types, and a more comprehensive understanding of cellular diversity within a tissue. Various scRNA-Seq methods exist, each with unique criteria for cell capture and sequencing depth (Svensson et al., 2018). Common steps include cell isolation, single-cell capture, RNA amplification, library preparation, sequencing, and data analysis (Potter, 2018). Major scRNA-Seq technologies include plate-based, microfluidic-based, and droplet-based methods. Switch Mechanism at the 5' End of RNA Template

sequencing (SMART-seq2) is a plate-based method that provides improved sensitivity, accuracy, and full-length coverage across transcripts. The current limitations are the lack of strand specificity and the inability to detect non polyadenylated (polyA$^-$) RNA (Picelli et al., 2014). scRNA-Seq data is not without limitations: First, compared to bulk RNA, the sequencing depth is substantially less. Secondly, smaller cell populations in tissues are easily overseen creating an incomplete picture of the overall gene activity in a tissue.

This thesis investigates bulk RNA-based GRNs, and key drivers identified in STARNET as critical for CAD and understands single-cell gene expression patterns of arterial wall cells such as endothelial, smooth muscle, and macrophages that are involved in atherosclerosis. Using SMART-seq2 technology, we revealed that vascular cells become highly plastic during atherosclerosis progression, undergoing dedifferentiation and trans-differentiation. Our deep single-cell sequencing identified the most gene-rich atherosclerosis subclusters to date. We also found no sex difference in major cell types but observed distinctions in their subcellular clusters. Integrating scRNA insights with STARNET GRNs for clinical and pathophysiological relevance, we validated GRN39 and GRN195 as critical for smooth muscle cell transformation into an osteogenic phenotype and endothelial cell proliferation in males, respectively. These findings reveal both potential and sex-specific therapeutic targets for the disease.

2. REVIEW OF LITERATURE

2.1. Atherosclerotic Cardiovascular Disease (ACVD): a Disease of Arteries

2.1.1. Arteries and their structure

Arteries are blood vessels that carry oxygenated blood under high pressure from the heart to tissues of the body with the pulmonary arteries being the only exception. Large arteries have three distinct layers (tunica): intima, media, and adventitia. Tunica intima is the closest to the arterial lumen and is composed of a single layer of endothelial cells (EC) that are attached to a fine network of the connective tissue basement membrane. The EC layer acts as a selectively permeable layer between blood in the lumen and vessel walls delivering oxygen and nutrients, regulating blood flow, modulating immune cell trafficking, and maintaining tissue homeostasis (Rajendran et al., 2013). Tunica media, the middle layer is thicker in large arteries and is made up of smooth (involuntary) muscle cells (SMC) intermingled with elastic fibers. This layer not only provides strength but also changes the vessel diameter to regulate blood flow and pressure (Webb, 2003). As arteries become smaller, the number of elastic fibers decreases while the number of smooth muscle fibers increases. Tunica adventitia/externa, the outermost layer, is the strongest of all three layers consisting of connective tissue with varying amounts of collagenous and elastic fibers that provide a limiting barrier, protecting the vessel from overexpansion as shown in Fig.1. This layer comprises a heterogeneous population of cells like macrophages (MP), T-cells, B-cells, dendritic cells, and fibroblasts (FB). Another characteristic of this layer is the presence of small blood vessels called the vasa vasorum that supply nutrients to the walls of larger arteries (Xu et al., 2015). In large arteries, a distinct layer of elastic connective tissues separating the tunica intima with tunica media is called the internal elastic membrane, and the tunica media from the outer tunica adventitia is called the external elastic membrane.

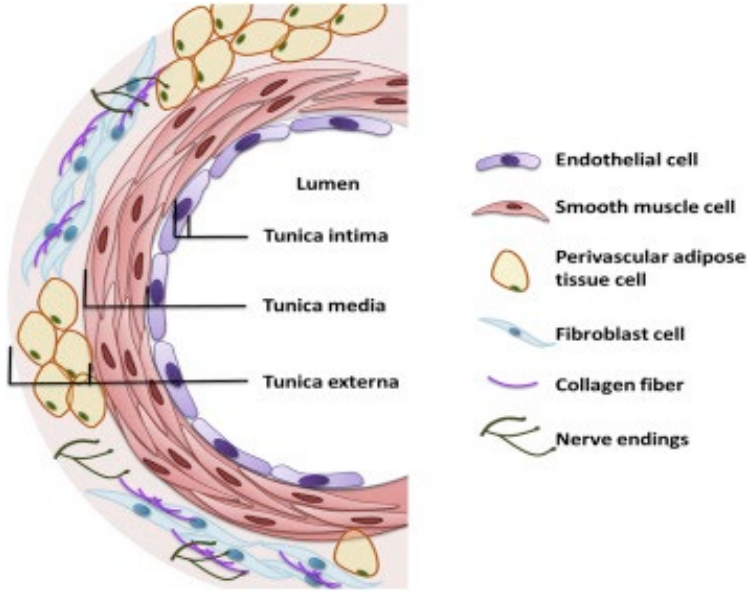

Figure 1: Structure of arterial wall (Yingzi Zhao et al., 2015).

2.1.2. Atherosclerosis pathophysiology

The term "atherosclerosis" was coined by German physician Felix J. Marchand (1846–1928), derived from the Greek words "atherē" meaning gruel, and "sclerosis" meaning hardening. While Rudolf Virchow (1821–1902), a German pathologist, first proposed it as an inflammatory process, the current understanding is more nuanced. Atherosclerosis is now considered a chronic disease of the inner wall of large and medium-sized arteries, including the aorta, coronary arteries, carotid arteries, and lower limb arteries. It has both inflammatory and degenerative components and can lead to devastating conditions like heart disease and stroke. However, the statement that most adults develop atherosclerosis in their early twenties needs clarification (Strong et al., 1999). While fatty streaks, an early stage of plaque buildup, can be present in young adults, the development of clinically significant atherosclerosis typically occurs later in life. The concern lies not in the absolute presence of the disease but in the extent to which it progresses and manifests as complications. These complications can arise from two main mechanisms: reduced blood flow (ischemia) in the affected artery or blood clot (thrombus) formation. Clots can either remain at the site or break off and cause blockage in distant locations (P. Libby & P. Theroux, 2005).

The basic mechanism involving atherosclerosis is determined by multiple factors, among which genetic disposition, inflammation, and oxidation are the most important (Libby, 2021). The most widely used model at present to explain atherosclerosis is the "response to injury hypothesis" proposed by Russel Ross that states atherosclerosis begins with endothelial cell injury leading to endo-

thelial dysfunction caused by modified lipoproteins; mainly LDL, free radicals, toxic substances, high blood pressure, infectious agents or combination of these factors (Minelli et al., 2020). The endothelial dysfunction leads to endothelial activation characterized by loss of endothelial vasodilation e.g. nitric oxide (NO) (Förstermann & Sessa, 2011) and increase of endothelial vasoconstrictors e.g. endothelin (ET-1). Endothelial activation is characterized by a pro-inflammatory, proliferative, and procoagulatory state accompanied by increased expression of adhesion molecules such as vascular cell adhesion molecule-1 (VCAM-1) and intracellular cell adhesion molecule-1 (ICAM-1) which promote the adhesion of leukocytes like monocytes and T-lymphocytes to the arterial wall. These adhered monocytes move to sub-endothelial space by a process called diapedesis under the influence of chemo attractants, in particular, the chemokine macrophage chemoattractant protein-1 (MCP-1) and other mediators like interleukin 8 (IL-8). Once there, they differentiate into macrophages that express several scavenger receptors such as scavenger receptor A, scavenger receptor B1 (SR-B1), a cluster of differentiation (CD)36 and CD68 on their surface that recognize polyanionic macromolecules and may have physiological functions in the recognition and clearance of pathogens and apoptotic cells.

Increased permeability of the dysfunctional endothelium allows LDL from the bloodstream into the subendothelial space, where it becomes trapped and modified by oxidation or enzymatic activity. Macrophage scavenger receptors take up this modified LDL via receptor-mediated endocytosis, which lacks negative feedback regulation. This leads to a massive accumulation of cholesterol within the macrophage cytoplasm, giving them a foamy appearance, hence the term “foam cells” – a hallmark of fatty streaks. Foam cells can efflux cholesterol via ABCA1 and ABCG1 transporters, but when overwhelmed, cholesterol builds up to toxic levels within the macrophages, impairing cell membrane fluidity as depicted in Fig.2. Moreover, intracellular cholesterol can trigger programmed cell death (apoptosis) or unregulated cell death (necrosis). Over time, this results in the deposition of lipids, forming the “necrotic core”, consisting of cell debris and cholesterol, a typical feature of advanced atherosclerotic lesions. Lesion growth is accompanied by the transformation of contractile smooth muscle phenotype to a proliferative state and migration from media to intima. Over time intimal SMCs secrete extracellular matrix proteins such as collagens giving rise to a protective “fibrous cap”. Some initial SMCs engulf LDL, forming foam cells and contributing to the growth and expansion of the lesion. A few SMCs can also differentiate into cells that produce mineralized matrix, undergoing osteoblastic differentiation and resulting in calcification (Demer & Tintut, 2008; Owens et al., 2004; Reynolds et al., 2004)). Ox-LDL can induce metalloproteinase-1 (MMP-1) and metalloproteinase-9 (MMP-9) in the vascular wall matrix causing degradation of extracellular collagen leading to weak and unstable atherosclerotic plaque. Adaptive as well as innate immunity drives chronic inflammation of atherosclerosis, T-cells can activate as well as suppress immune activation as well as help B cells to produce antibodies (Bjorkegren & Lusis, 2022).

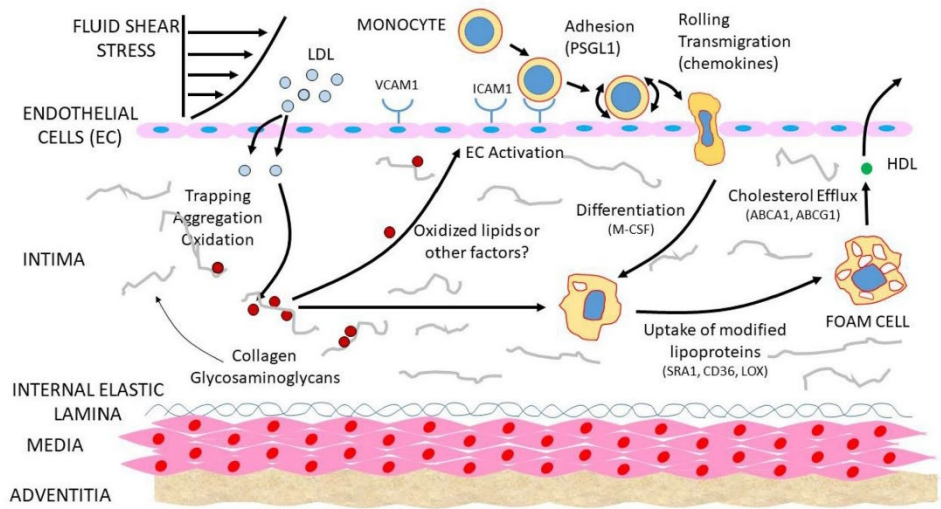

Figure 2: Development of fatty streak lesions (Björkegren et al., 2022).

2.1.3. Clinical presentation of atherosclerosis

When vascular injury occurs, the usually inaccessible matrix components and SMCs are exposed to circulation leading to the activation of platelets, part of the hemostatic system that forms a thrombus to prevent blood loss. In atherothrombotic conditions, there is endothelial injury but no blood loss, resulting in the formation of a thrombus within the vessel and culminating in CAD. Advanced plaques can erode or rupture leading to thrombus formation due to the activation of the clotting cascade (Furie & Furie, 2008) as shown in Fig.3. Thrombus formation can occlude the artery either partially or completely or travel to a distant location from the site of formation, leading to MI or cerebrovascular event (stroke).

Ruptured plaques have thin cap fibroatheroma with a large lipid core and show recent signs of intraplaque hemorrhage (Virmani et al., 2006). In contrast, in eroded plaques, the endothelial layer is eroded, exposing the SMC and extracellular matrix (ECM) to circulation. It is estimated that in ~40% of patients, the culprit lesion responsible for acute coronary syndrome (ACS) is eroded rather than ruptured. Plaque erosion is typically associated with a non-ST-elevation MI (NSTEMI) presentation of ACS, while plaque rupture is more common in ST-elevation MI (STEMI) patients (Baaten et al., 2024).

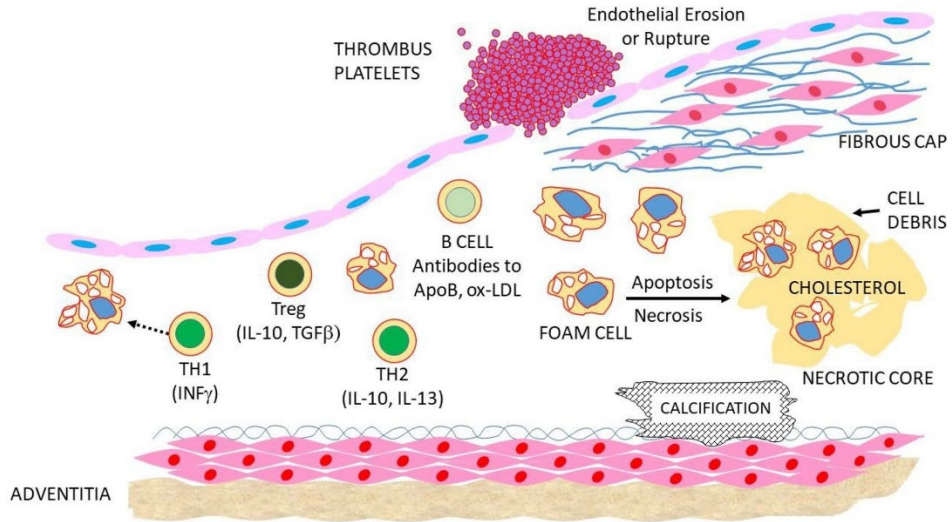

Figure 3: Advanced atherosclerotic lesions (Björkegren et al., 2022).

2.1.4. Atherosclerosis is a multifactorial process

Atherosclerosis development involves the interaction of multiple genetic and environmental factors (Peter Libby & Pierre Theroux, 2005). These risk factors can be either modifiable (environmental) or nonmodifiable (genetic) and can have both direct effects (casual) and indirect effects on disease development. Risk factors can cause either local molecular events at the atherosclerotic plaque site or at distant locations like the liver. Modifiable risk factors include unhealthy diet, airborne pollutants including smoking, alcohol consumption, behavioral habits like degree of stress and exercise (sedentary lifestyle), increased hip: waist circumference, high body mass index (BMI), metabolic syndrome like elevated LDL, triglycerides, hypertension, low HDL, impaired fasting glucose (Diabetes Mellitus). Nonmodifiable risk factors include age, sex, and family history. The net effect of the modifiable risk factors filtered through the individual genetic makeup is reflected in their blood components and flow as shown in Fig.4. Changes in gene expression are reflected in circulation where inflammatory markers synthesized in the organs can be detected. For example, the measurement of high cholesterol levels in plasma synthesized by the liver i.e. hypercholesterolemia (Tegner et al., 2007).

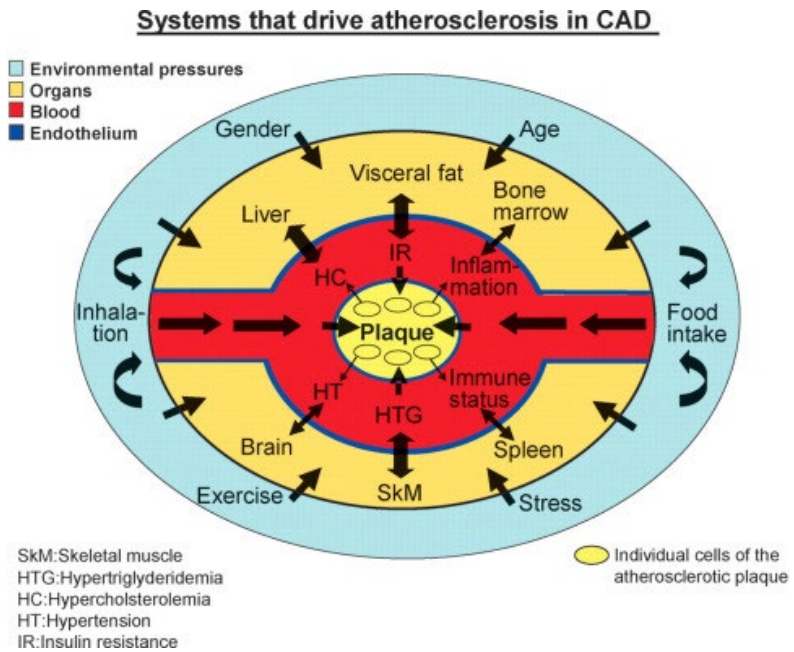

Figure 4: Systems that drive atherosclerosis in CAD (Tegner et al., 2007).

Subclinical atherosclerosis is the presence of atheromatous disease before any signs, symptoms, or events attributable to clinical manifestations, which is widely termed asymptomatic atherosclerosis. People with subclinical atherosclerosis are at higher risk of clinical events like MI or stroke (Kawai et al., 2024). Risk factors for subclinical atherosclerosis do not differ from those of clinical ACVDs. Carotid Intima-Media Thickness (IMT) is a non-invasive measurement used to assess subclinical atherosclerosis, correlating with the burden of coronary artery disease and predicting future risks of MI or stroke (Fox et al., 2003; Lorenz et al., 2007). The intima-media layer, comprised of endothelial cells, connective tissue, and myocytes, is where lipids accumulate to form atherosclerotic plaques (Moghtaderi et al., 2014). The common carotid artery is the preferred site for IMT measurement due to the prevalence of reversible foam cell lesions in this elastic artery. This reversibility helps explain why IMT can decrease with treatments like statins, making it a valuable surrogate endpoint in clinical trials (Dalager et al., 2007). Foam cell lesions are common in elastic arteries like the common carotid and aorta. Notably, the proximal coronary arteries and carotid bifurcation, which are transitional zones between elastic and muscular arteries, also develop early foam cell lesions and lipid core plaques, sometimes even from childhood in coronary arteries (Dalager et al., 2007). Therefore, a thickened carotid IMT is a reliable indicator of generalized atherosclerosis. IMT is considered an intermediate phenotype for early atherosclerosis because it's easily and non-invasively measurable. While IMT reflects overall atherosclerosis, classic cardiovascular risk factors have varying impacts on different arterial territories: choles-

terol primarily affects ischemic heart disease (CAD), hypertension impacts ischemic stroke (CVD), and smoking and diabetes are significant for intermittent claudication (PAD) (Kannel, 1994).

Mice are crucial for atherosclerosis research, allowing us to uncover the molecular and cellular mechanisms of the disease and identify genes influencing susceptibility, due to their genetic manipulability and diverse strains. This enables the creation of mouse models that mimic human atherosclerosis, aiding in understanding the genetic basis of the disease. However, mouse models do not perfectly replicate human atherosclerosis. Key differences include their dominant HDL lipoprotein profile, which necessitates genetic modification to induce human-like LDL-driven hyperlipidemia, and a lower incidence of human-relevant features like plaque rupture and thrombosis. The primary mouse models for studying atherosclerosis are ApoE−/− and LDLR−/− mice. ApoE−/− mice spontaneously develop significant atherosclerosis and hypercholesterolemia, even on a regular diet, due to their inability to clear lipoproteins. LDLR−/− mice, on the other hand, develop atherosclerosis with elevated LDL cholesterol but typically require a high-fat "Western-type" diet to accelerate lesion progression. In our study, we used Ldlr-/-Apob100/100 mice that exhibit a human-like plasma lipoprotein profile, with most of their plasma cholesterol carried by LDL particles (i.e., 200–250 mg/dL) (Skogsberg, Lundstrom, et al., 2008). This model naturally develops severe atherosclerosis even on a standard diet.

2.2. Genetics of rare and common diseases

2.2.1. The Genome and Genetic Variation

All the DNA (deoxyribonucleic acid) contained in one cell is called the genome. DNA is packaged into a thread-like structure called a chromosome. In humans, 23 pairs of chromosomes are present, of which 22 pairs are autosomes and 1 pair is allosome (sex chromosome). Eukaryotes (animals, plants, fungi, and protists) store most of their DNA inside the cell nucleus (20,000 -23,000 protein coding genes in humans) and some in the mitochondria (37 in humans,13 being protein coding genes). DNA is a double helical structure composed of sugar and phosphate molecules linked by nitrogenous bases, together called nucleotides. The building blocks of DNA are four bases: Adenine (A), Guanine (G), Cytosine (C), and Thymine (T). These bases pair specifically, with A always pairing with T and G pairing with C (Watson & Crick, 1953). One copy of the human genome contains 3 billion nucleotides distributed among 23 chromosomes. Humans are diploid; two copies of the human genome, each inherited from one parent. Genes are segments of DNA that are transcribed into messenger RNA (mRNA) and then translated into proteins (central dogma: DNA-mRNA-protein) showing that DNA determines which proteins are produced inside the cells first by transcription followed by translation. The number of protein-coding genes is estimated to be around 20,000, forming about 2% of DNA (coding DNA). The remaining 98% is

non-coding DNA, which is still not fully understood but is believed to be important in controlling gene activity.

Differences among human genomes are called genetic variants. All individuals are 99.9% genetically identical and only 0.1% of genetic variations is responsible for the phenotypic difference. Genetic variation is caused by spontaneous mutations or recombination that remain conserved through evolution. A locus is a specific fixed position on the chromosome where a particular gene or genetic marker is located. There are multiple types of genomic variants; single-nucleotide variants (SNVs) which reflect differences in single nucleotides. SNVs are the most common type of genomic variation. To be considered an SNP, an SNV must be present in at least 1% of the human population. On average, there is one SNP found in every 300 base pairs in the human genome (Kruglyak & Nickerson, 2001). Each SNP has two alleles. Individuals with two identical alleles are called homozygous, and those with two different alleles are called heterozygous. Two SNPs are said to be in linkage disequilibrium (LD) when their alleles are highly correlated in the population, often due to their close physical proximity on a chromosome, which limits recombination between them.

InDels reflect extra (insertion) or missing (deletions) DNA nucleotides in the genome typically fewer than 50 nucleotides. The most common type of insertion/deletion variants are tandem repeats or microsatellites which are short stretches of nucleotides that are repeated multiple times and are highly variable among people (Ellegren, 2004). Though less frequent than SNVs, microsatellites can also have significant health consequences, particularly through changes in their repeat numbers. Structural variants (SVs) are large-scale genomic differences spanning at least 50 nucleotides, up to thousands, involving insertions, deletions, inversions, or translocations. Large tandem repeats (over 50 nucleotides) constitute nearly half of all structural variants. A specific type of structural variant, copy-number variants (CNVs), refers to differences in the total number of nucleotides involved that span over 1 kilobase of DNA sequence (Ku et al., 2010).

2.2.2. Monogenic and Polygenic Diseases

Genetic disorders can be caused by pathogenic variants within a single gene (monogenic) or by the complex interaction of multiple genetic variants across several genes, often combined with environmental factors (polygenic or multifactorial). Monogenic, or Mendelian, diseases are solely caused by genetic factors and are transmitted from parents to offspring following Mendelian inheritance patterns (Klug, 2012). They are typically caused by low-frequency, high-penetrance, rare single-gene mutations with a large effect size. Single gene diseases run in families and can be dominant or recessive, and autosomal or sex-linked. These disease-causing genetic variants can be inherited in an autosomal dominant, autosomal recessive, X-linked, or Y-linked manner, leading to clinical syndromes such as Hypertrophic cardiomyopathy, cystic fibrosis, and Hemophilia A, respectively. Autosomal dominant inheritance requires one copy of the faulty gene, one from each parent. In contrast, autosomal recessive inheritance requires

two copies of the faulty gene, one from each parent, to cause the disease. Sex-linked, primarily X-linked, inheritance requires the faulty gene to be present on a sex chromosome (X), one copy in males and two copies in females (Haines & Pericak-Vance, 2007). These diseases are generally easier to detect compared to polygenic diseases. Pedigree analyses of large families with many affected members are useful for determining single-gene disease inheritance patterns. However, it's important to note that this ease of detection applies primarily to genetic analysis; a clinical diagnosis alone may not always be straightforward.

Polygenic, or non-Mendelian, diseases are multifactorial and are often caused by several high-frequency, low-penetrance gene mutations with small effect sizes. These diseases are also influenced by other factors such as age, sex, nutrition, and exercise. The cumulative effect of genetic and environmental factors plays a definitive role in the manifestation of CCDs like diabetes, coronary artery disease, and arthritis. Unlike Mendelian disorders, their inheritance patterns are not traceable through families in the classic sense, requiring large-scale association studies to identify contributing factors. Recent studies suggest a shared genetic basis between complex traits and Mendelian disorders. Genes implicated in Mendelian disorders are involved in pathways underlying complex diseases. Large-effect coding variants disrupt individual genes, leading to severe Mendelian phenotypes. In contrast, non-coding variants collectively dysregulate the expression of these genes, resulting in more nuanced or tissue-specific complex traits. This suggests a spectrum of genetic architectures, rather than a simple dichotomy between Mendelian and complex traits (Freund et al., 2018).

Genome-wide association studies (GWAS) represent a powerful tool for investigating the architecture of ACVDs by examining the frequencies of SNPs across populations and identifying those that segregate with a given trait. There are hundreds of millions of known SNPs scattered throughout the human genome, with a SNP defined as a variant present in at least 1% of the population (Karczewski et al., 2020). Minor allele frequency (MAF) is a measure of how often the less common version of a gene (allele) appears in a population. Based on MAF, genetic variants can be categorized as common or frequent variants: MAF > 0.05 (5%), low-frequency variants: MAF = 0.01–0.05 (1–5%), and rare variants: MAF < 0.01 (<1%) (Bodmer & Bonilla, 2008). An important initial hypothesis for CCDs was the "common disease, common variant" hypothesis, which assumes that common or frequent genetic variants are the main contributors to the genetic susceptibility of common diseases (Blanco-Gomez et al., 2016). The majority of GWASs for complex traits and diseases have indeed identified significant associations in non-coding genomic regions, often pointing common variants (Freund et al., 2018). SNPs occurring within a gene-coding sequence implicated in disease susceptibility will demonstrate a phenotypic effect. If it appears to be segregating with a disease, or if it is more prevalent in affected versus unaffected subjects, then the SNP is physically close to disease-causing mutation (Botstein & Risch, 2003; J. N. Hirschhorn & M. J. Daly, 2005; Lander & Botstein, 1986). GWAS's success relies on LD, a population-level association between SNPs. LD occurs when alleles of nearby SNPs are inherited

together more often than expected by chance, allowing SNPs to act as markers for disease-causing variants (Joiret et al., 2019).

A central question in CCDs is the extent to which genetic and environmental factors influence phenotypic variation (Visscher et al., 2008). R. A. Fisher (Fisher, 1918) pioneered the partitioning of phenotypic variance into genetic and environmental components, laying the groundwork for heritability estimation (Visscher & Goddard, 2019). Heritability is the proportion of phenotypic variance attributable to genetic factors. Broad-sense heritability (H^2) considers all genetic effects including additive, dominance, and epistatic. Narrow-sense heritability (h^2) considers additive genetic effects relevant for predicting evolutionary responses to selection. It is typically estimated from family studies and can be expected to vary across environments (Manolio et al., 2009). Limitations of GWAS studies are that the majority have been done in European ancestry whereas the highest genetic variation is seen in African and non-European studies. Most common variants individually or in combination confer relatively small increments in risk (1.1–1.5-fold) and explain only a small proportion of heritability (Manolio et al., 2009). GWAS do not necessarily pinpoint causal variants and genes directly responsible for the disease. This non-causality stems primarily from LD, where a genotyped SNP (often termed a 'tag SNP') is not inherited in isolation but rather tends to be co-inherited alongside other physically adjacent variants, which may include the actual causal variant, as an unbroken block on the chromosome. Consequently, an associated SNP serves more as a marker for a genomic region rather than the direct etiological factor itself. Hence, GWAS is fundamentally a discovery phase, adept at identifying promising regions of interest within the genome that then require rigorous, focused investigation to establish true causality.

The Wellcome Trust Case Control Consortium identified a chromosome 9 locus linked to CAD (Samani et al., 2007). More than 200 susceptibility loci associated with CAD have been discovered through GWAS. While these variants are widespread, their individual effects on CAD risk are generally small (Aragam et al., 2022). GWAS studies, while successful in identifying common variants associated with complex traits, have struggled to account for the "missing heritability". This has led to the hypothesis that rare variants may also contribute significantly to the genetic component of these traits, challenging the exclusive focus on common variants. Rare variants, with a MAF, less than 1%, are challenging to detect with GWAS due to their low frequency in the population. This requires larger sample sizes or meta-analysis studies.

Additionally, traditional SNP microarrays often underrepresented rare variants, limiting their detection power. Next-generation sequencing technologies, such as whole-exome and whole-genome sequencing, offer a more powerful approach to identifying rare variants. These techniques have been successfully used to identify causal mutations for rare Mendelian diseases and are increasingly being applied to study complex traits (Blanco-Gomez et al., 2016) as illustrated in Fig. 5.

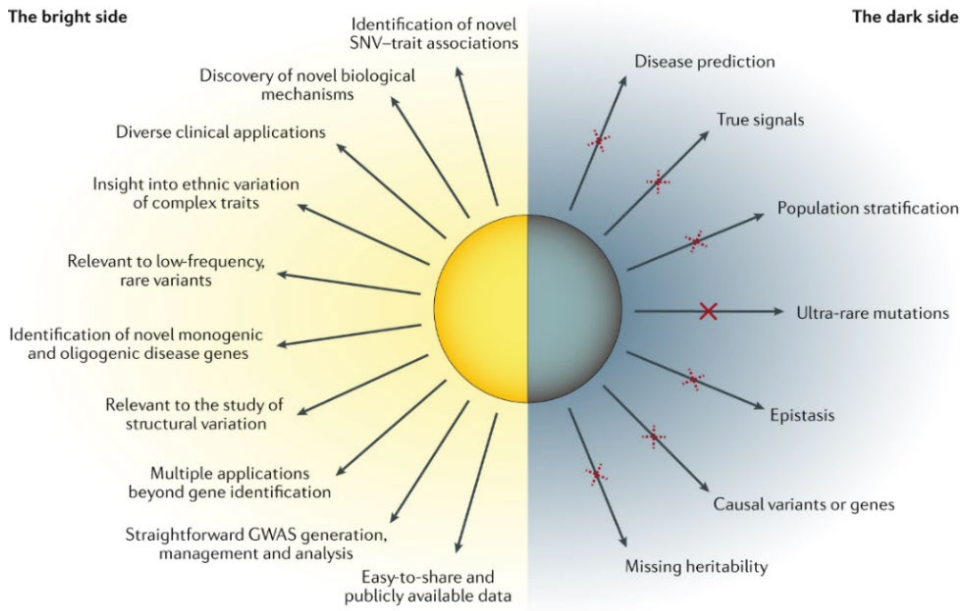

Figure 5: A visual depiction of the current benefits (the bright side) and limitations (the dark side) of GWAS (Vivian Tam et al., 2019).

2.2.3. Sequencing

High-throughput NGS technologies, whole-exome sequencing (WES), and whole-genome sequencing (WGS) are increasingly used to study rare diseases and complex traits. Unlike array-based genotyping in GWAS, WES/WGS directly captures relevant variation, including rare- variants (RVs) (Wu et al., 2011). While WES has limited SV sensitivity due to its exonic focus, WGS offers comprehensive genomic coverage, making it ideal for robust SV detection and characterization**.** The advent of sequencing was considered to have a potential impact on explaining missing heritability through RVs (Manolio et al., 2009) leading to the common disease rare variant hypothesis (CD-RV) that considers that common complex traits may be the result of multiple RVs that impact one or multiple genes that would not be tagged by conventional GWAS SNPs. Adequate power requires extremely large sample sizes, which are often impractical. Additionally, allelic heterogeneity suggests multiple RVs within a gene may affect the same trait. Therefore, RV analysis using next generation sequencing (NGS) data typically employs “aggregative” testing, where variants are tested collectively based on their physical overlap with pre-defined genomic regions. Many aggregative RV analysis methods have been proposed, primarily falling into two categories: burden tests and variance-component (kernel) tests. Among kernel tests, the Sequence Kernel association test (SKAT) and its variations (e.g., SKAT-O, Optimal SKAT) are widely used, though other approaches and modifications exist (Chen

et al., 2022). Gene-burden tests assume homogeneous effects for all RVs, limiting their power when effects are heterogeneous. SKAT, while powerful for variants with opposing effects, can be less powerful than burden tests when effects are predominantly in the same direction. SKAT-O combines burden and SKAT, addressing some limitations but can still be less powerful under specific assumptions. While SKAT and SKAT-O focus on association testing, gene-burden testing can be used in any scenario where a gene needs to be treated as a single unit (Pathan et al., 2024). To accurately detect RVs, WES is used that does not depend on assumptions about genetic architecture. These studies have identified significant RV associations for complex diseases and traits. However, over 98% of genetic variants reside in noncoding regions, and many common variants identified by GWASs are also located in noncoding regions (Li et al., 2022).

Both common and rare genetic variants contribute significantly to trait heritability. However, the interplay between these two types of variants, influenced by LD, varies across traits. While rare variants can explain a substantial portion of the "missing heritability" for certain traits like height and BMI, the limitations of pedigree-based studies and SNP-based analyses hinder the complete recovery of heritability for others. Factors such as sample size, population characteristics, and the exclusion of specific genetic variants contribute to this gap (Pathan et al., 2024).

Limitations of GWAS in identifying central pathological processes for complex diseases can be grouped under:

1. Case/control overlap: In comparing cases with population controls in studies of complex diseases, cases often share underlying processes with healthy individuals. DNA variants associated with these shared processes may not reach genome-wide significance in GWAS but can be associated with disease at the nominal level.
2. Context of shifting environments: DNA variants can be influenced by environmental factors, macroenvironmental factors (lifestyle, diet, stress) and microenvironmental factors (cellular context) can modulate the effect of genetic variants which is not accountable by GWAS.
3. Context of time: DNA variants that influence early disease processes are more likely to be detected by GWAS. Variants affecting late-stage processes may be missed due to their shorter duration and dependence on environmental factors.

GWAS may not fully capture the genetic complexity of complex diseases as shown in Fig. 6. While it can identify significant associations, it may overlook important variants that are influenced by environmental factors or are involved in later stages of disease development.

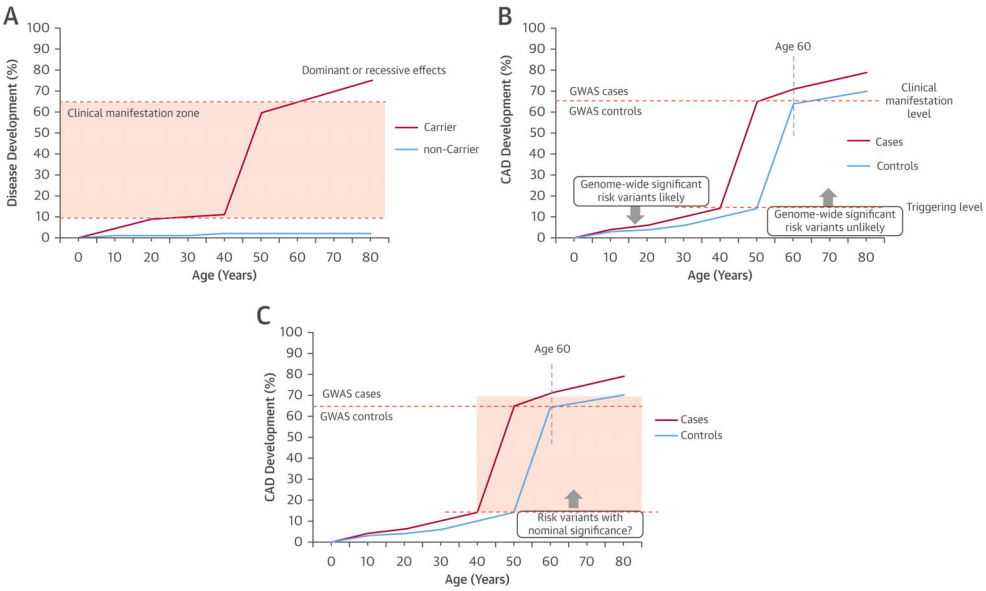

Figure 6: Development of rare vs complex diseases (source: Johan LM Björkegren et al., 2015). A. Rare diseases are caused by specific DNA variations, and their severity can vary. The sigmoidal curve is a common biological pattern but may not apply to all rare diseases. B. Complex diseases like CAD develop slowly, often over decades. GWAS studies primarily focus on early-stage disease, making identifying genetic variants associated with these initial stages easier. C. The influence of environmental factors, and the shorter duration of these stages, genetic variants associated with late-stage disease often have weaker statistical significance in traditional genetic studies.

To solve this, Systems genetics, a new approach to complex diseases like CAD analyzes various cellular activities (RNA, proteins, etc.) to understand the intricate networks of genes driving the disease. This information is then combined with genetic data to explain disease heritability (Lusis & Weiss, 2010). The underlying rationale for this approach is that the functioning of a complex system cannot be predicted by its parts, even when these are fully understood. Applying a holistic approach that integrates data from multiple sources can improve understanding of the underlying mechanisms that determine the biological system's characteristics as illustrated in Fig.7 (van den Beucken, 2019).

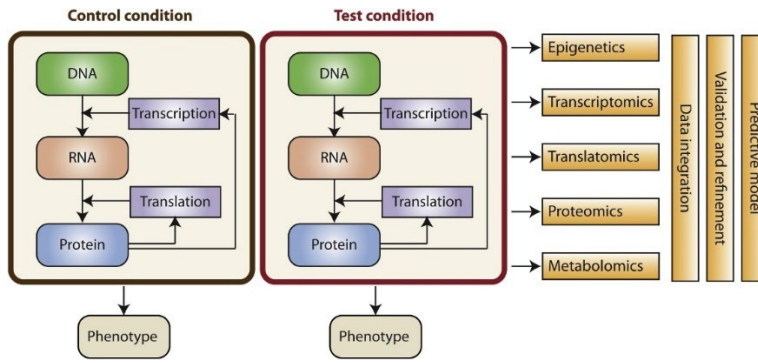

Figure 7: Systems biology for studying cellular functions in complex diseases (Twan van den Beucken et al., 2019).

2.2.4. Genetics of Gene Expression Studies (GGES)

Currently, disease networks are primarily inferred using genetic and gene expression data. This involves applying network inference algorithms to identify relationships between genes (Barabási et al., 2011). While traditional methods often rely on microarray data, newer approaches can analyze more complex data types like RNA sequencing, offering higher resolution and a broader dynamic range (Ozsolak & Milos, 2011; Wang et al., 2009). The STAGE (Stockholm Atherosclerosis Gene Expression) and STARNET (Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task) datasets are examples of GGES. Both studies recruited patients undergoing heart surgery, those with CAD were cases, while those without CAD were controls. STAGE is a smaller pilot study, while STARNET has a larger sample size. Parallel sampling of up to 9 CAD-relevant tissues from each patient is a key aspect of the STAGE and STARNET studies (Franzen et al., 2016; Hagg et al., 2009). The 9 RNA samples were then converted into microarray data (STAGE, custom-made HuRSTA-2a520709 Affymetrix arrays [Affymetrix, Santa Clara, California]) and RNA sequence data (STARNET). RNA expression datasets are used to infer causal regulatory disease-driving molecular processes, as reflected in gene networks operating both within and across tissues to cause CAD; and to identify DNA variants that modulate these networks (Schadt, 2009). These gene expression data can be used to identify eQTLs, which are DNA variants (frequently an SNP) that influence gene expression levels. By comparing genetic information with gene expression data, researchers can identify SNPs associated with changes in gene activity. The gene expression data was used to identify groups of genes that work together in networks based on similar co-expression across the 9 tissues. These networks were then linked to specific patient characteristics, such as the severity of coronary artery disease. By using Bayesian network analysis, key genes that drive these networks and contribute to the development of CAD are identified.

Systems genetics for CAD can be expanded beyond DNA and RNA to include other molecular data like proteins, metabolites, and lipids. This integrated approach,

along with clinical information, can provide a more comprehensive understanding of CAD disease mechanisms as shown in Fig.8 (Joshi et al., 2021). While analyzing these molecules (proteins, metabolites, and lipids) at a genome-wide level in individual tissues is challenging, studying them in blood plasma can be informative. Combining different molecular data types can enhance the predictive power of disease models. The focus on RNA in STAGE and STARNET is due to its stability and the availability of advanced technologies for RNA analysis. While protein analysis is more challenging due to its rapid turnover, studying plasma proteins can still provide valuable insights. Integrating multiple data types, including DNA, RNA, proteins, and metabolites, can improve the predictive power of CAD disease networks (Björkegren et al., 2015).

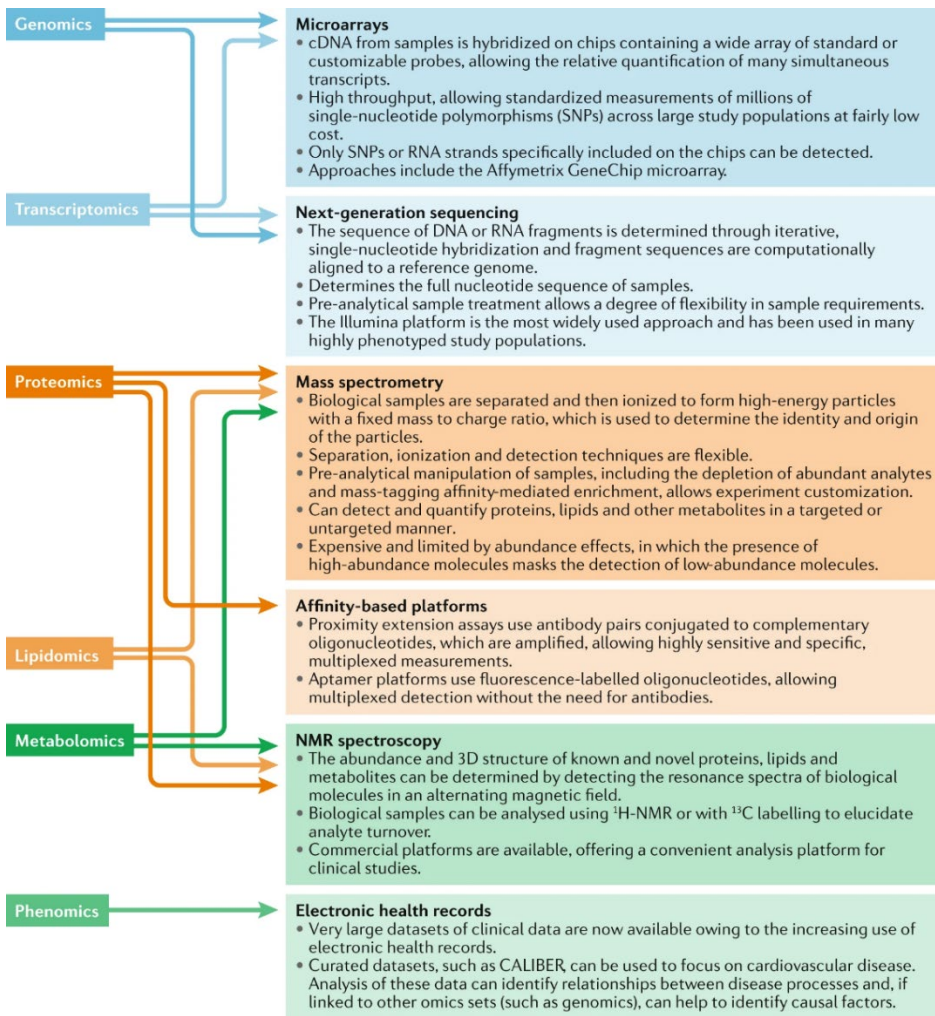

Figure 8: Omics platforms used in systems biology. Multiple analytical platforms are used to interrogate samples and systems at a molecular level (Joshi et al., 2021).

Systems genetics offers a powerful, integrated approach to understand how genetic variation interact with environmental risk factors to impact complex disease development as reflected in the RNAseq or other forms of functional omics data. Instead of focusing on individual genes like in GWAS (Björkegren et al., 2015), networks of genes are in focus that are generated from diverse 'omics datasets (e.g., genomics, transcriptomics, proteomics and metabolomics) to provide a holistic view of how genetic perturbations interact with both macro- and micro-environments to drive complex disease development (Tegner et al., 2007). Central to this approach are GRNs, which map the regulatory gene interactions across tissues controlling gene expression. Within these GRNs, key drivers (highly connected and high hierarchical network nodes) exert significant control over the overall network activity and downstream phenotypes, making their dysregulation profoundly impactful (Simon Koplev et al., 2022). Key drivers have been validated through experimental perturbations via targeted gene knockout/knockdown studies or by overexpression in cell culture or in mouse models (Hartman et al., 2021).

2.3. Transcriptomics

2.3.1. Single-cell RNA sequencing (scRNA-seq)

Bulk RNA sequencing and microarray of atherosclerotic plaques from human patients or animals highlight networks and pathways accelerating lesion progression. However, the tissue's heterogeneous nature complicates the interpretation of the transcriptomic signals such as pinpointing responsible cells and identifying the effects of rare but crucial cells due to signal noise (Slenders et al., 2022). Traditional bulk tissue assays provide only a weighted average of the population. Cell-to-cell variation is a universal property of multicellular organisms, even within a seemingly homogeneous tissue, diverse cell populations exist with varying functions, morphologies, and gene expression profiles. Single-cell transcriptomics, however, can resolve these heterogeneous cell populations at an unprecedented scale and address questions related to the stochastic nature of gene expression through the advent of transformative single-cell technologies, such as fluorescence-activated cell sorting (FACS), microfluidic devices, and mass cytometry. Single-cell studies have revealed unexpected cell plasticity, challenging traditional views of cell fate determination during development and differentiation. Designing optimal single-cell transcriptomic experiments is complex due to the need for tailored approaches. Factors like sample type, cell number, preparation method, scRNA-seq technique, sequencing parameters, and computational analysis strategy significantly impact the interpretability of results as depicted in Fig.9 (Lafzi et al., 2018).

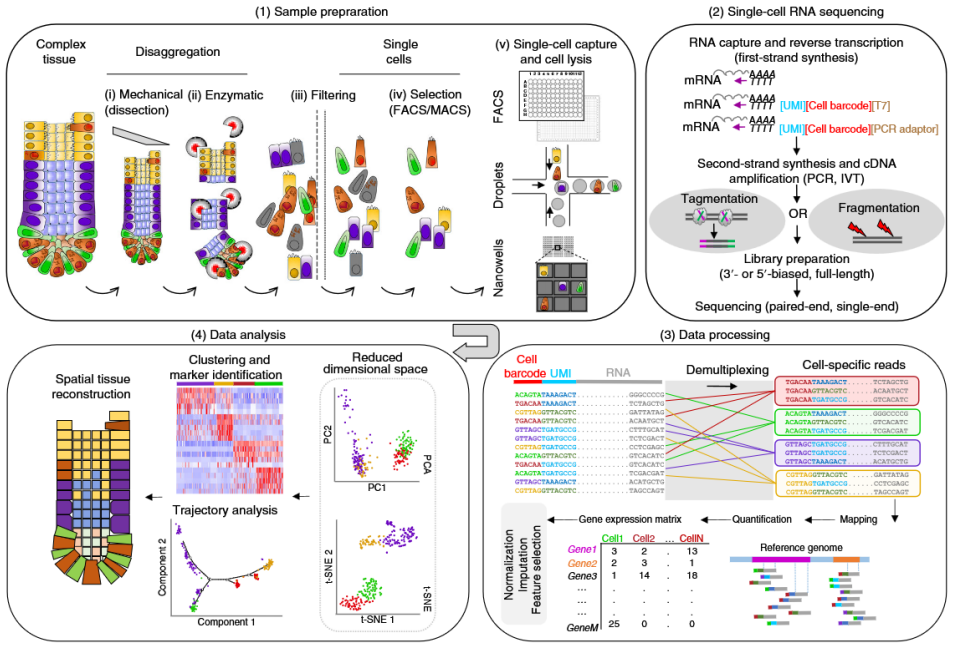

Figure 9: The design of single-cell transcriptomics experiments (Lafzi et al., 2018).

ScRNA-seq workflow begins with sample preparation, a crucial step influencing data quality and interpretation (Rostom et al., 2017; Ziegenhain et al., 2017). Optimal results stem from fresh, viable single-cell suspensions to minimize transcriptional artifacts (Ding et al., 2020). However, practical considerations often necessitate alternatives like cryopreserved cells or nuclear RNA from frozen tissue; these approaches, while enabling sample banking, can introduce biases in gene expression and cell recovery (Regan & Preall, 2022). To ensure high-quality data, rapid processing (within 30 minutes) is vital to mitigate cell degradation and aggregation. Essential to this stage is the effective removal of cellular debris, dead cells, and reverse transcription (RT) inhibitors through optimized wash steps, solution compositions, centrifugation, and filtration. For solid tissues, dissociation involves a combination of mechanical disruption and enzymatic digestion, with enzyme choice and digestion time being tissue specific (Saliba et al., 2014). DNase I treatment is routinely used to prevent cell clumping by degrading extracellular DNA, followed by filtration to remove remaining debris. Balancing efficient dissociation with minimizing cellular stress is paramount, as prolonged or harsh conditions can activate stress-related genes, bias results and damaging sensitive cell types (van den Brink et al., 2017). Conversely, insufficient digestion can lead to incomplete cell separation and the exclusion of important cell populations.

Next, single cell capture physically isolates individual cells for transcriptome profiling (Hwang et al., 2021). Low-throughput experiments may employ micro-

manipulation or pipetting, while high-throughput methods primarily utilize FACS or microfluidics (Lafzi et al., 2018). FACS enables precise cell sorting, allowing for the exclusion of unwanted cells and enrichment of specific populations into microtiter plates. Microfluidic systems, encompassing IFCs, droplets, and nanowells, offer advantages such as reduced reagent usage and efficient cell capture. Pre-processing with FACS or MACS can further enhance cell suspension quality by removing debris and damaged cells. Achieving an unbiased view of cellular composition necessitates capturing a diverse range of cell types, including cells of various sizes that might otherwise be excluded by certain isolation methods. The required number of cells per experiment depends on sample heterogeneity and subpopulation frequency; more cells are generally needed to identify rare cell types or resolve fine-grained subpopulations (Hedlund & Deng, 2018). Pilot experiments are therefore recommended to determine optimal cell numbers and the necessity of subpopulation enrichment (Mereu et al., 2020).

Single-cell RNA sequencing itself typically targets poly(A)-tailed RNA. This RNA is reverse-transcribed into stable cDNA, during which most methods incorporate single-cell-specific barcodes within poly(T) oligonucleotides. These barcodes enable sample multiplexing and serve as UMIs to track individual molecules and correct for amplification bias. Sequencing can yield either full-length transcript data or only the 3'/5'-end of transcripts. Full-length scRNA-seq generally incurs higher library preparation costs as it often lacks early UMI and cell barcoding, but it offers improved detection of lowly expressed transcripts and insights into splicing. Plate-based approaches like Smart-seq (Ramskold et al., 2012, 2020) and Smart-seq2 (Picelli et al., 2013) utilize template-switching reactions, with Smart-seq2 efficiently capturing full-length transcripts. STRT-seq (Islam et al., 2012) and STRT-seq-2i (Hochgerner et al., 2017) (the latter in nanowells) are plate- or nanowell-based 5'-end sequencing methods incorporating UMIs. Other plate-based methods include SCRB-seq, its optimized version mcSCRB-seq (Bagnoli et al., 2018), as well as Quartz-seq (Sasagawa et al., 2013), Quartz-seq2 (Sasagawa et al., 2018), CEL-seq (Hashimshony et al., 2012), CEL-seq2 (Hashimshony et al., 2016), and MARS-seq (Jaitin et al., 2014). Microfluidic systems-based approaches have enhanced efficiency, throughput, and cost-effectiveness compared to plate-based methods. The three primary microfluidic capture strategies are IFCs (microfluidic channels with defined capture sites), droplet-based approaches (cells encapsulated with reagents for RT and barcoding), and nanowell-based systems (cells captured in nanowells for individual processing), such as Seq-Well (Gierahn et al., 2017). Droplet-based systems, including inDrops (Klein et al., 2015) and Drop-seq (Macosko et al., 2015), provide greater scalability and flexibility by encapsulating cells in nanoliter droplets (as shown in Table 1), enabling high-throughput processing and precise cell capture control. However, their random barcode assignment can hinder the direct linking of cell identities to visual information and complicates doublet detection.

Table 1: Comparison of different scRNA technology approaches.

Method	Capture format	Cell loading	Single-cell indexing	Molecule identifier	cDNA amplification	Fragmentation	Transcript coverage	Sequencing
Smart-seq	Plate	FACS	Tagmentation	NA	PCR	Tagmentation	Full length	Paired end
Smart-seq2	Plate	FACS	Tagmentation	NA	PCR	Tagmentation	Full length	Paired end
STRT-seq	Plate	FACS	TSO	UMI	PCR	DNase I	5′ end	Single end
STRT-seq-2i	Nanowell	FACS/Poisson	TSO	UMI	PCR	Tagmentation	5′ end	Single end
SCRB-seq	Plate	FACS	Oligo(T) primer	UMI	PCR	Tagmentation	3′ end	Paired end
mcSCRB-seq	Plate	FACS	Oligo(T) primer	UMI	PCR	Tagmentation	3′ end	Paired end
Quartz-seq	Plate	FACS	Oligo(T) primer	NA	PCR	Ultrasound	Full length	Paired end
Quartz-seq2	Plate	FACS	Oligo(T) primer	UMI	PCR	Ultrasound	3′ end	Paired end
CEL-seq	Plate	FACS	Oligo(T) primer	NA	IVT	KOAc, MgOAc	3′ end	Paired end
CEL-seq2	Plate	FACS	Oligo(T) primer	UMI	IVT	Random priming	3′ end	Paired end
MARS-seq	Plate	FACS	Oligo(T) primer	UMI	IVT	Zinc	3′ end	Paired end
Seq-Well	Nanowell	Poisson	Oligo(T) beads	UMI	PCR	Tagmentation	3′ end	Paired end
inDrops	Droplets	Poisson	Oligo(T) beads	UMI	IVT	KOAc, MgOAc	3′ end	Paired end
Drop-seq	Droplets	Double Poisson	Oligo(T) beads	UMI	PCR	Tagmentation	3′ end	Paired end

Library preparation and sequencing for short-read approaches involve fragmenting the amplified cDNA (PCR) or RNA (IVT) before adding sequencing adaptors. Fragmentation can be achieved enzymatically (e.g., with tagmentase or DNase), chemically (e.g., with zinc, KOAc, or MgOAc), or mechanically (e.g., ultrasound). 3′- or 5′-based libraries are then amplified with primers specific to the transcript end or start, respectively. At this stage, a pool-specific index can be introduced for multiplexed sequencing of multiple experiments. In contrast, full-length methods introduce cell-specific barcodes only after fragmentation, slowing initial cell pooling. Microfluidic approaches typically sequence to lower depths (<100,000 reads per cell), whereas plate-based approaches often require higher read numbers (~500,000 reads per cell).

Finally, data processing transforms raw sequencing reads into gene expression matrices. After generating FASTQ reads and performing quality checks, the crucial next step is de-multiplexing reads using cell barcodes. These de-multiplexed reads are then mapped to reference genomes using alignment tools. scRNA-seq data is inherently noisy and variable due to technical factors like dropout events, biased amplification, and incomplete sequencing. Biological variations, such as cell size, cell cycle stage, and gender differences, further complicate analysis. Normalization techniques are vital to address these challenges and enable meaningful data interpretation. Beyond noise, scRNA-seq data also exhibits significant sparsity, characterized by numerous zero values in the expression matrix due to non-expressed genes and technical dropouts, which can hinder a comprehensive understanding of each cell's transcriptome.

Data analysis aims to identify co-expression patterns and cluster cells by similarity. Cell clusters can be interpreted by identifying defining “marker genes” (Trapnell, 2015). While supervised clustering uses known markers, unsupervised clustering is often preferred for hypothesis-free exploration of data, employing methods like hierarchical clustering (agglomerative or divisive) or k-means clustering (Kiselev et al., 2019). To visualize cellular subpopulations in two or three dimensions, dimensionality reduction techniques are commonly used. These include PCA, t-SNE (Kobak & Berens, 2019), diffusion maps (Haghverdi et al., 2015), and UMAP (Becht et al., 2018).

2.3.2. Switch Mechanism at the 5' End of RNA Template (Smart-seq2)

The Smart-seq protocol (Ramskold et al., 2012), revolutionized single-cell RNA sequencing by generating and amplifying full-length cDNA from individual cells. This method utilizes a Moloney murine leukemia virus (MMLV) reverse transcriptase (RTase) that leverages template-switching and terminal transferase activity. This unique enzymatic property enriches for intact 5′ ends and obviates the need for a separate second-strand synthesis step. An advancement in the field, Smart-seq2 (Picelli et al., 2013), built upon its predecessor by offering improved RT efficiency, enhanced template switching, and optimized preamplification.

These refinements collectively lead to a higher cDNA yield, increased sensitivity, and a reduction in technical bias. The initial steps of the Smart-seq2 protocol involve careful cell preparation: cells are gently disaggregated into a single-cell suspension under near-physiological conditions to preserve cellular integrity and prevent RNA degradation. Subsequently, individual cells undergo lysis in a buffer containing a ribonuclease inhibitor to stabilize RNA molecules. A mild hypotonic lysis buffer, supplemented with free deoxynucleotide triphosphates (dNTPs) and tailed oligo-dT primers (comprising a 30-nucleotide poly-dT stretch and a 25-nucleotide universal 5′ anchor sequence), is employed to prime the RT reaction on polyadenylated RNA sequences. This buffer composition is specifically designed to minimize interference with the subsequent RT reaction.

The RT reaction in Smart-seq2 is meticulously controlled. It is initiated at 42 °C for 90 minutes, followed by a brief increase to 50 °C for 2 minutes to facilitate the unfolding of any RNA secondary structures. The temperature is then lowered to 42 °C for an additional 2 minutes to ensure the complete synthesis of cDNA. This cycle is repeated a total of ten times to maximize cDNA yield. Betaine, a methyl group donor, plays a crucial role in optimizing this reaction by stabilizing proteins and destabilizing DNA helices, particularly G-C rich base pairs. This action significantly enhances the efficiency of template switching and the yield of full-length cDNA. To further optimize the RT reaction, the concentration of magnesium chloride is increased to 9–12 mM in the presence of 1 M betaine, which amplifies the DNA-destabilizing effect of betaine, leading to even greater cDNA yields. Smart-seq2 incorporates a modified template switching oligo (TSO) featuring two riboguanosines and a locked nucleic acid (LNA) modification at the 3' end. These LNAs improve the thermal stability of the TSO and enhance its annealing to the untemplated 3' extension of the cDNA, thereby contributing to more efficient template switching and improved cDNA yield. Furthermore, the use of KAPA HiFi HotStart ReadyMix eliminates the need for bead-based purification of first-strand cDNA, which improves sensitivity and facilitates automation of liquid handling.

Following cDNA amplification, tagmentation is employed for rapid and efficient construction of sequencing libraries. This process utilizes a hyperactive Tn5 transposase that simultaneously fragments the DNA and ligates adapters, thereby eliminating the need for separate fragmentation and size selection steps. The resulting tagmented DNA fragments, typically ranging from 200 to 600 base pairs, are then ready for PCR amplification. Illumina's dual-indexing strategy, incorporating both i7 and i5 indices, allows for the pooling of up to 96 samples for efficient sequencing on a single lane of an Illumina sequencer (as depicted in Fig. 10). Typically, each cell yields between 1 and 20 million sequence reads, depending on the desired sequencing depth and multiplexing level. When sequencing 50-base pair single-end reads, approximately 60% of the reads uniquely map to the reference genome, while 20% exhibit multimapping, and 20% do not map at all. While Smart-seq2 enables the detection of a broader range of genes, including those with higher GC content, and provides full-length cDNA sequences, it is selective for polyA+ RNA, and the generated reads do not retain messenger RNA strand information.

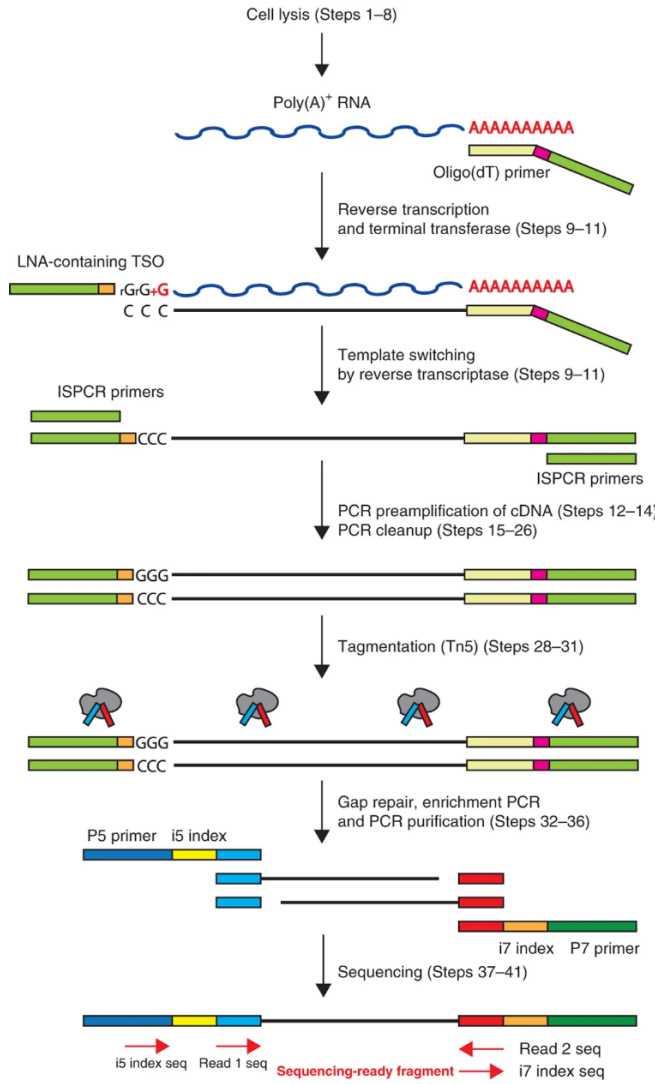

Figure 10: Flow chart for Smart-seq2 library preparation (Picelli et al., 2014).

2.4. Summary of literature review

CVD is a very broad term that refers to any disease affecting the heart or blood vessels. ACVD is a specific type of CVD that is caused by atherosclerosis – the buildup of plaque in the arteries. ACVD is a pervasive and multifaceted inflammatory condition that represents a leading cause of morbidity and mortality worldwide. This chronic disease, characterized by the accumulation of plaque within the walls of large and medium-sized arteries, ultimately culminates in severe clinical events such as myocardial infarction (heart attack) and stroke

(brain attack). For many years, the prevailing understanding of ACVD progression focused primarily through the lens of lipid accumulation. While dyslipidemia remains a critical component, it's now widely recognized that ACVD involves a far more intricate interplay of diverse risk factors, encompassing not only traditional culprits like high cholesterol and blood pressure but also non-traditional contributors such as chronic inflammation and genetic predispositions. The advent of GWAS has fundamentally revolutionized our comprehension of ACVD's genetic underpinnings. These powerful, unbiased approaches have successfully identified over 200 genetic loci associated with CAD alone, marking a significant advancement compared to the limited success of earlier genetic linkage studies. GWAS have pinpointed regions of the genome where common genetic variants (SNPs) are statistically associated with an increased or decreased risk of developing ACVD.

However, despite these remarkable successes, a significant challenge persists: the common variants identified by GWAS individually exert only a small effect on disease risk. Collectively, they explain only a fraction of ACVD's estimated heritability, a phenomenon widely referred to as "missing heritability". This gap highlights that the exact biological mechanisms underlying many GWAS-identified associations often remain unclear. Furthermore, precisely pinpointing the specific causal gene or variant responsible for an observed effect within the often-broad genomic regions identified by GWAS proves challenging. To compound this complexity, over 98% of identified genetic variants, including many common variants implicated by GWAS, reside in noncoding genomic regions. Deciphering the functional implications of these noncoding variants, which may regulate gene expression or impact chromatin structure, presents a formidable interpretational hurdle. This unresolved "missing heritability" and the challenges in interpreting noncoding variants have fueled alternative hypotheses, notably the "common disease, rare variant" hypothesis. This theory suggests that common complex traits like ACVD may also result from the combined effects of multiple RVs. RVs, by definition, occur at low frequencies in the population (typically less than 1%) and might exert stronger individual effects than common variants. However, traditional GWAS methods are often statistically underpowered to reliably detect these RVs due to their infrequent occurrence, leaving a significant portion of genetic susceptibility unexplained.

To begin unraveling these profound complexities, systems genetics has emerged as a crucial scientific approach. It systematically examines how genetic variations within a population influence complex traits and diseases by studying their effects on intermediate molecular phenotypes, such as RNA expression, protein levels, and metabolite concentrations. By analyzing these multi-omics changes and their intricate relationships to observable disease traits, systems genetics helps map the complex pathways that connect genes to disease. Our research group has been at the forefront of applying this strategy through the STARNET project. This crucial RNA-seq resource, derived from seven atherosclerosis-relevant tissues (atherosclerotic aortic wall, mammary artery, liver, skeletal muscle, as well as abdominal and subcutaneous fat) collected from living

donors undergoing open thoracic surgery in Tartu, Estonia, exemplifies this approach. By meticulously integrating genetic and clinical data from 600 CAD patients and 250 controls with multi-tissue RNA-seq data, STARNET successfully identified 224 GRNs. These GRNs captured the delicate molecular interplay underlying CAD, collectively explaining over 54% of its heritability (Study 3; http://starnet.mssm.edu). While STARNET provided a powerful bulk tissue-level understanding of GRNs and offered invaluable insights into broad genetic and molecular drivers, it simultaneously underscored the inherent complexity that necessitates an even higher resolution of analysis.

The limitation of bulk RNA sequencing, as employed in STARNET, is its provision of an average global gene expression profile, obscuring crucial cell-specific insights. To overcome this, scRNA-seq (Smart-seq2 method) was employed in atherosclerotic plaques to address this limitation for capturing cell-to-cell variation. This method is profoundly critical, especially given the known plasticity of cellular phenotypes during atherosclerosis progression, allowing us to gain unprecedented insights into specific cell types within plaque composition, disease progression dynamics, and even sex differences in disease presentation. For instance, integrating scRNA-seq data from both human and mouse carotid atherosclerotic plaques with established STARNET GRNs successfully pointed specific SMC and macrophage subtypes, along with their associated GRNs (e.g., GRN39 for SMCs), as critical drivers of advanced, symptomatic atherosclerosis (Study 2). Similarly, scRNA-seq was used to meticulously uncover crucial sex-specific differences in human carotid artery plaques, revealing distinct subcellular compositions and GRNs between males (e.g., GRN195 for ECs) and females (Study 1). These studies not only captured cellular plasticity but also demonstrated the immense potential of integrating single-cell resolution with population-level GRNs to uncover specific cellular transformations underlying disease.

This thesis synthesized and expanded on foundational findings regarding atherosclerotic plaque buildup in the coronary and carotid arteries. Using a systems-based approach, we integrated genetic, RNA-seq, and clinical data from the STARNET study that led to the discovery and validation of 224 GRNs that promote CAD development and showed that the genetic regulation of these GRNs explains 54–60% of CAD heritability, a significant increase from previous findings (22%). Our single-cell analysis (Smart-seq2) mapped cellular and molecular effects within specific cell types, revealing key insights into cell plasticity and sex-specific differences in plaque progression. By integrating this single-cell data with our population-level GRNs, we successfully pinpointed the causal cellular mechanisms and functional pathways of ACVD. While this work has advanced the potential for more precise treatments, a complete understanding of the disease remains an ongoing challenge.

3. AIMS AND OBJECTIVES

The overall aim of this thesis was to study the process of atherosclerosis development primarily in humans using a top-down systems network approach including the roles of metabolic organs and sex on the trans-differentiation of key atherosclerosis cell types as depicted in Fig. 11.

1) To study how cell patterns and associated GRNs of clinically significant atherosclerosis differ between men and women (Study 1).
2) To uncover the cellular and clinical mechanisms of atherosclerosis progression by integrating gene regulatory networks with single-cell gene expression data from atherosclerotic plaques of a human-like mouse model and carotid stenosis patients with and without clinical symptoms. (Study 2).
3) To elucidate the interplay between genetic and non-genetic primarily metabolic risk factors in the development of coronary artery disease (CAD) by mapping a mechanistic framework of gene regulatory networks (GRNs) that span major metabolic organs, blood, and the arterial wall (Study 3).

4. MATERIALS AND METHODS

Detailed description of the materials and methods presented in this thesis are found in the original publications. This section provides a summary of the methods used in the studies.

4.1. Overall experimental plan

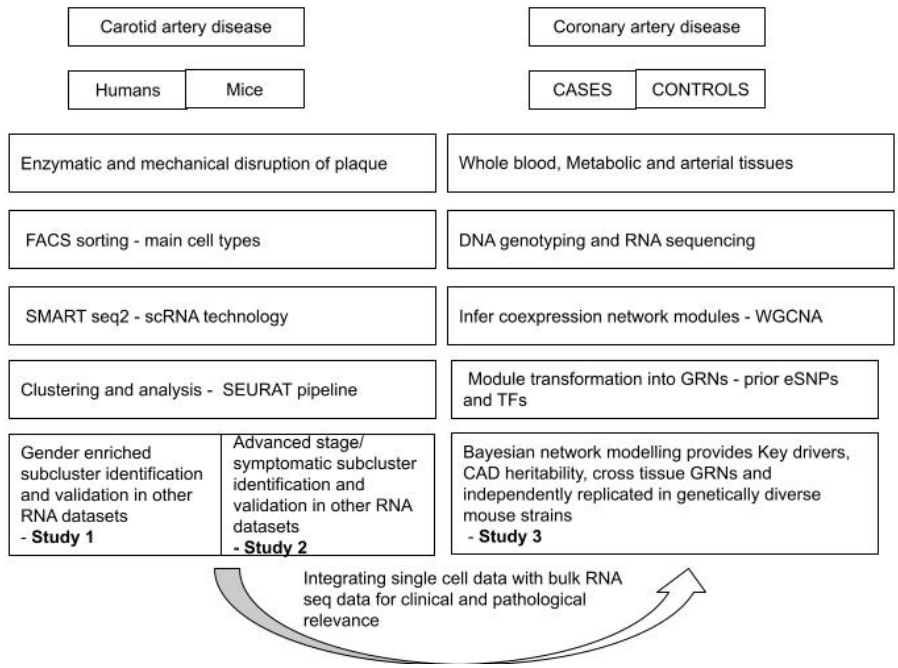

Figure 11: Flow chart of the overall study design.

4.2. Ethical statement

The study protocols used in this research were approved by the Institutional Ethics Committee on Human Research of the University of Tartu (IRB nos: 289/T-12, 2771T,17). Informed consent was obtained from all subjects (Ethical Approvals Dnr 154/7 and 188/M-12). All mouse experiments were carried out in accordance with Swedish and Finnish legislation, local guidelines, and regulations for animal welfare and were approved by the Animal Research Ethics Committee in Linköping; Sweden (dnr729), and the National Experimental Animal Board of Finland.

4.3. Patient selection

Patients eligible for carotid endarterectomy surgery because of significant stenosis (>70%) in the proximal part of their internal carotid arteries at the Department of Surgery, Tartu University Hospital were included. Patients were characterized as asymptomatic (no clinical signs of brain ischemia during last 6 months) or symptomatic (clinical signs of brain ischemia such as minor stroke, transitory ischemic attacks (TIA), retinal infarction or *amaurosis fugax* (brief visual disturbance during last 6 months). Non-diseased or minimal carotid plaque samples were not obtained as surgical operation of carotid endarterectomy aims to remove the entirety of the plaque; hence unaffected ("normal") arterial wall areas cannot be obtained (Study 1 and 2).

Individuals with CAD eligible for coronary artery bypass surgery and those eligible for open-thorax surgery like valve replacement following aortic stenosis and a preoperative angiogram ruling out obstructive CAD (hence the need for CABG) were enrolled as cases and controls respectively at the Department of Cardiac Surgery, Tartu University Hospital (Study 3).

Patient exclusion criteria

Patients with severe systemic disease e.g., active systemic inflammatory disease or cancer were excluded. Each enrolled participant completed a questionnaire to assess disease history, current drug regimens, and lifestyle (e.g., daily activity, alcohol consumption, and smoking). All the recruited patients signed an informed written consent form before participating in the study.

4.4. Collection of clinical samples (Study 1–3)

4.4.1. Human biopsies

In total 15 carotid plaques including 7 females and 8 males (Study 1) and 7 asymptomatic and 8 symptomatic patients (Study 2) who underwent carotid endarterectomy surgery were collected. Whole blood and seven disease relevant tissue biopsies from 600 individuals with CAD and 250 controls who underwent open-thorax surgery were collected (Study 3). Whole blood was collected in EDTA tubes and the biopsies like AOR, MAM (only in individuals with CAD), VAF, SF, SKLM and a tiny bit of left lateral LIV lobe were collected in eppendorf tubes with RNAlater in it and stored at -80 freezer till RNA was extracted.

4.4.2. Ldlr-/-Apob100/100 mouse model

The Ldlr-/-Apob100/100 mice exhibit a human-like plasma lipoprotein profile with most of its plasma cholesterol carried by LDL particles mirroring human-like plasma cholesterol concentrations (i.e., 200–250 mg/dL). This model naturally develops severe atherosclerosis even on a standard diet. Mice were housed

in groups of up to 6 in standard ventilated cages with a 12-hour light/dark cycle, ad libitum access to water and fed a chow diet. We used 46 independent Ldlr-/-Apob100/100 mice bred on 100% C57Bl6/J genetic background divided into four baseline (non-atherosclerotic, 10 weeks old), nine early atherosclerotic (20–30 weeks old), and ten advanced (45–60 weeks old) atherosclerotic pairs of aortic arches isolated from 2 independent mice. Number of mice used at baseline: 8 mice (2 female, 6 male); early stage: 18 mice (6 female, 12 male) and advanced stage: 20 mice (10 female, 10 male).

4.5. STARNET STUDY (Study 3)

4.5.1. STARNET normalization of RNA seq data and eSNPs inference

RNA from seven tissues (AOR, blood, LIV, MAM, SKLM, SF, VAF; n=534–560/tissue) was isolated (Qiagen RNeasy Mini) and single-end sequenced (Illumina Hiseq/Hiseq 4000) using mainly polyA and Ribo-Zero protocols. Initial sequencing (50–100bp, 20–30M reads) was performed for matching CAD cases (n=600) and controls based on univariate distribution of age, sex, BMI using Kolmogorov-Smirnoff test. Only individuals with ≥3 tissues were included. The resulting 1,496 samples were randomized and sequenced (100bp, 20–30M reads). Samples with <1M aligned reads were excluded. Reads were mapped to GRCh37 (STAR v.2.3.0e). Gene counts (≥6 in ≥10% of samples/tissue) from seven tissues were normalized (L2 penalized regression) adjusting for technical (sequencing lab, read length, RNA extraction) and biological (age, sex) covariates, plus four surrogate variable analysis (SVAs) and flow cell information. SVA used all STARNET phenotype data (excluding age, sex, IDs) with imputed missing values. Cis-eQTLs (±1Mbp Transcription Start Site) were identified (Matrix eQTL, model LINEAR, cisDist < $1x10^{-6}$, FDR-controlled). Data are in dbGAP (phs001203.v1. p1).

4.5.2. Phenotype associations and multi-tissue RNA seq differential expression

Gene expression was associated with CAD phenotypes (SYNTAX score, Duke score, diseased vessels, lesions, BMI, CRP, HbA1c, waist/hip ratio, cholesterol, triglycerides) using Pearson correlation, Student's t-test, and DESeq2 (adjusted for age and sex, FDR < 0.01). Co-expression module meta-analysis (Fisher's method) was performed, aggregating transcript p-values per module. Module associations with CAD measurements were corrected for multiple hypotheses (Benjamini-Hochberg) considering number of modules (224) and number of features (three). Tissue-specific and cross-tissue modules were clustered hierarchically (Euclidean distance, complete linkage) using phenotype association -log(p-values) (min. 10^{-16}) and scaled tissue fractions -log ($10^{-16} * 7^{-1}$).

For multi-tissue case-control differential expression (adjusted for metabolic phenotypes, drugs, and genotype), RNA-seq reads were aligned to GRCh38 v89 (STAR 2.5.0). Gene expression (median counts ≥5) was normalized (DESeq size factors) and standardized. Differential expression (DESeq2, adjusted for sex and age) was determined (Benjamini-Hochberg correction, FDR < 0.01, fold change > ±30%). Covariate influence on differential expressions was assessed by comparing p-values before and after adjusting for CMD traits (BMI, lipids, glucose), therapies, and multidimensional scaling components. Subsequent analyses used differential expression statistics adjusted for sex, age, and BMI.

4.5.3. Co-expression network modules and replication using Genotype- Tissue Expression (GTEx) project

Normalized STARNET gene expression data (151,774 transcripts, 19.4% missing data) across seven tissues per patient was used to infer co-expression modules with scale-free properties using Weighted Gene Co-expression Network Analysis (WGCNA). The WGCNA β parameter was optimized (ranging 1.0–10.0) by evaluating scale-free R^2 for connectivity vs. log(pk) within and between all tissue pairs. For each tissue combination, the lowest β with $R^2 > 0.85$, or if none existed, a β within 95% of the optimal R^2 and closest to 3 (cross-tissue) or 6 (tissue-specific) was selected as cross-tissue correlations tend to be weaker than tissue-specific correlations. Three correlation adjustment schemes were compared: a single average β value for any tissue combination ("single"), all specific β values per combination ("complete"), and average tissue-specific and cross-tissue β values ("dual"). Using the dual approach (β=5.2 tissue-specific, β=2.7 cross-tissue), blockwise WGCNA (five blocks via projective k-means) was performed, and modules were identified via hierarchical clustering and dynamic tree cut. Modules with >5% of transcripts from a secondary tissue were classified as "cross-tissue"; otherwise, they were "tissue-specific". Each gene's expression in a tissue was assigned to only one module. To assess STARNET co-expression network reproducibility and WGCNA assumptions, an independent, TPM-normalized GTEx v7 RNA-seq data (six overlapping tissues, excluding mammary) was used. For each non-mammary STARNET module (224 in total), correlation concordance with GTEx using a t-statistic (n-2 df) and Benjamini-Hochberg correction were tested. STARNET-GTEx correlation R^2 values was also calculated. Additionally, NetRep permutation tests to avoid normality assumptions for correlations and weights was used. Gene set enrichment analyses were calculated by Fisher's exact test with hypergeometric distribution. Co-expression modules were annotated using GO term enrichment (biological process, cellular component, molecular function) via WGCNA, displaying the top 100 terms per category found in <50 modules. Cross-tissue module-associated terms were identified by comparing log enrichment P-values with tissue-specific modules (16,581 GO terms) using Mann-Whitney tests and Benjamini-Hochberg correction (FDR < 0.01).

4.5.4. Gene Regulatory Networks

GRNs were inferred within each co-expression module (fewer than 3,000 transcripts) using GENIE3, constrained by eQTL and transcription factor edges. Standardized gene expression data with imputed missing values were used. Key driver analysis (Mergeomics) prioritized influential genes within each subnetwork. While alternative key driver identification strategies exist (e.g., based on global tissue causal structures), this study ranked key drivers within each co-expression module according to their statistical significance. GWAS gene enrichment for cardiometabolic traits (blood lipids, glucose metabolism, obesity) was performed using Fisher's exact test on co-expression modules, using trait-associated genes from the NHGRI-EBI GWAS Catalog and CAD GWAS studies. Using individual-level genotype data from nine CAD GWAS datasets, heritability contributions of eSNPs within STARNET GRNs was calculated. Genome relationship matrices for each eSNP list were generated using LDAK and adjusted for linkage disequilibrium and minor allele frequency (5%). Further adjustments were made for population stratification and study batch effects. CAD variance explained (liability model, 5% prevalence, 40% heritability) was calculated using REML with GCTA. Lead CAD SNPs and SNPs in LD were excluded. Mediated expression score regression was also used to assess the combined heritability contribution of the 224 GRNs.

Causal GRNs were inferred for 119 tissue-specific modules (95% nodes from one tissue, <500 genes) using Mendelian Randomization (MR) via the Findr Python package. Findr tested for trans-eQTL associations between key driver cis-eQTLs and target genes, controlling for horizontal pleiotropy (key driver and target gene were not independently associated with the *cis* eQTL). Local false discovery rates (FDRs) were calculated and thresholds set per tissue to achieve a global FDR of 10%. 3,601 unique key drivers were identified across seven tissue groups. These key drivers were then used in one- and two-sample MR analyses to assess causal relationships with 11 STARNET traits (8 cardiometabolic disease; BMI, waist/hip ratio, CRP, total cholesterol, LDL and HDL cholesterol, triglycerides and HbA1c) and 3 coronary artery disease (number of lesions and SYNTAX and Duke scores). One sample MR used the 'ivreg' R package. Two-sample MR used an inverse-variance weighted method in the Mendelian Randomization R package, selecting lead cis-eQTLs ($p \leq 1 \times 10^{-3}$) with available GWAS data. Directed interactions between eigengenes of co-expression GRN modules were inferred using a Bayesian network approach (fast greedy equivalence search) via the rcausal R package (Tetrad suite). Network structure likelihood was evaluated using Bayesian information criteria, with a maximum node out-degree of 100. The supernetwork layout was generated using the Fruchterman-Reingold algorithm (igraph R package). Supernetwork edge reproducibility was assessed via bootstrap sampling (1,000 replicates), with P-values calculated from edge frequencies. This analysis was performed globally and for 32 CAD-associated modules. Eigengene network influence was quantified using betweenness centrality, calculated from the bootstrap P-value matrix.

4.5.5. Endocrine factor identification

Endocrine signaling candidates were identified from STARNET multi-tissue gene expression data by analyzing 2,444 UniProt-annotated secreted proteins. Candidates were selected based on: (1) presence in a cross-tissue co-expression module, and (2) significant correlation (Pearson's) between their mRNA expression and an eigengene from a tissue-specific module *other* than the candidate's origin. Only genes from modules with <5,000 transcripts were considered. Multiple testing correction (Benjamini-Hochberg, FDR < 0.2) was applied. This approach integrates cross-tissue co-expression, super network interactions, and endocrine signaling.

4.5.6. GRN78 Endocrine factor validation

To validate adipose-to-liver endocrine axis identified in the STARNET super-network linking cross-tissue GRNs with liver-specific GRNs in study 3, eigengene values were computed for each of the liver-specific GRNs that were targeted by one or more cross-tissue GRNs. Next, Pearson correlations were calculated between the mRNA levels of the endocrine candidate present in the cross-tissue GRNs (and believed to mediated the cross-tissue connection) with the eigengene value of the target liver GRN. Next, for endocrine candidates showing this association in GRN78, we sought to validate using independent human data from GTEx (Consortium et al., 2015) and the morbid obesity data (MOD) (Greenawalt et al., 2011) as well as data from the Hybrid mouse diversity panel (HMDP) mice (fed chow, high-fat, and apoE-Leiden diets) (Ghazalpour et al., 2012). For HMDP data, gene symbols of the STARNET endocrine candidates validated in GTEx and MOD were mapped to their corresponding mouse homologs. Next, HMDP adipose and liver microarray data (n=96–108 strains, 2–5 mice/strain) were used, except for apoE-Leiden adipose (n=63, RNA-seq, STAR 2.6.0c alignment to GRCm38 v95). Unlike in the HMDP mice, the human GTEx and MODs, subcutaneous and VAF tissues were distinct.

For 10 endocrine candidates successfully validated in independent data, recombinant proteins (CMV-overexpressed, TEV-cleaved, 9xHis-tagged) were synthesized in HEK293 cells, purified via nickel-cobalt column, and dialyzed. Purity (>80%) was confirmed by Coomassie gel. Eight-week-old male C57BL/6J mice (Jackson Laboratory) were housed at 23°C on a standard chow diet with a 12-h day/night cycle. All procedures were approved by the UCLA IACUC (protocol 92–169) and conducted according to UCLA animal care guidelines and ARRIVE guidelines. Four to six mice per protein were injected intraperitoneally daily for 4 days (0.1 μg/g body weight). Experimenters were blinded to injections until metabolite data analysis. Mice were fasted for 3 hours before isoflurane euthanasia. Tissues were snap frozen. For RNA-seq, livers were pulverized, and RNA was extracted (Qiagen RNeasy Mini kit). Samples with RIN > 6.8 were used. Five samples were excluded due to low RIN. Libraries were prepared from 800 ng RNA (KAPA HyperPrep kit) and sequenced (PE, 50 bp) on a HiSeq 4000.

Reads were aligned to GRCm38 v95 (STAR 2.6.0c). Differential expressions were assessed using DESeq2. Select proteins (FCN2, LBP, FSTL3, EPDR1) were cross validated using purchased proteins.

4.6. Single-cell isolation and capture (Study 1 and 2)

4.6.1. Tissue preparation for single cell isolation

Harvested carotid plaques were enzymatically digested in HBSS solution with 10 mg/ml of Collagenase type I, II and Elastase (Sigma) for one hour at room temperature. Gentle mechanical disruption was then applied to obtain a single-cell suspension. The cell suspension was filtered through a 70-micron cell strainer and subsequently centrifuged. Finally, the cell pellets were resuspended in a FACS buffer (PBS supplemented with 0.5% BSA, 2 mM EDTA, 25 mM HEPES).

The Ldlr-/-Apob100/100 mice after being euthanized by cervical dislocation were perfused with PBS, to isolate the aortic arch from 23 age- and sex-matched pairs at three stages: baseline (n=8), early (n=18), and advanced (n= 20). The aortic arch was dissected from the aortic root to the third rib, cleaned from the adventitia and immediately placed into ambient PBS solution (DPBS, Thermo-Fisher Scientific). To generate sufficient cells two aortic arches were used in a two-step process. First to isolate ECs initial incubation was done in HBSS (Thermo-Fisher Scientific) with 2% BSA (Sigma), supplemented with 10 mg/ml of Collagenase type I, II and IV and 2 mg/ml of Elastase (all from Sigma) at 37 °C with horizontal shaking at 500–800 rpm for 10 minutes. In the second incubation, the remaining aortic arch tissues were cut into smaller pieces and incubated in HBSS solution with 10 mg/ml of Collagenase type I, II and Elastase (Sigma) in three cycles followed by pipetting in 20 min intervals. The cell suspensions were then sequentially filtered through 70 micron and 40-micron cell strainers. To maximize cell retention, the 70 μm cell strainer was additionally washed with 5 ml of DMEM (ThermoFisher Scientific). After centrifugation at 300 × g at 20°c for 5 min, cell pellets were resuspended in the FACS buffer.

4.6.2. Fluorescence activated cell sorting analysis (FACS)

To isolate specific cell populations SMCs, ECs, and immune cells for flow cytometry analysis, the human single-cell suspension was enriched using fluorophore-conjugated antibodies against PDGFRβ (SMCs), CD31/CD144 (ECs), and CD45 (immune cells). Calcein green AM was used to identify viable cells. Similarly, the mouse single-cell suspension was enriched using fluorophore-conjugated antibodies against PDGFRβ (SMCs), CD31 (ECs), and CD45 (immune cells), along with Calcein green AM. The labeled cell suspensions were diluted and placed on ice prior to FACS analysis.

To capture specific cell populations, the Calcein green and antibody-stained cell suspensions were sorted using a FACS Melody cell sorter equipped with a 100-micron nozzle and deposited into 384-well plates containing 2.3 μl lysis

buffer (0.2% Triton X-100, 2 U/ml RNase inhibitor, 2 mM dNTPs, 1 μM Smart-dT30VN, ERCC 1:4 × 104 dilution). To maximize yield, initial gating for forward scatter-area/ side scatter-area (FCS-A/SSC-A, linear scale) was used to exclude cell debris and red blood cells. Next, doublet discrimination was implemented using FCS-A/FSC-height and SSC-A/SCC-height including a generous threshold for the distance of events from the diagonal line to prevent possible bias toward round shaped cells. Last, cells passing the first two criteria were selected for Calcein-green and then sorted for the cell-specific antibody into individual wells of 384-well plates. To ensure correct gating, samples stained with single, or no antibodies were used as negative controls. During sorting, plates were always kept at 4°C and thereafter immediately stored at −80°C.

4.6.3. Single-cell library preparation and sequencing

Single-cell cDNA libraries were prepared using the Smart-Seq2 protocol. In brief, single cell mRNA was transcribed into cDNA using oligo(dT) primers and SuperScript II reverse transcriptase (ThermoFisher Scientific) and amplified by PCR for 24–25 cycles. Next, cDNA samples with sufficient quality according to DNA High Sensitivity chip analysis in a 2100 Bioanalyzer (Agilent Biotechnologies) were fragmented, tagged (tagmented) and indexed using Tn5 transposase and the Illumina Nextera XT index kits (Set A–D). Indexed cDNA libraries were pooled and sequenced using a HiSeq3000 (Illumina) with a dual indexing and single-end, 50 base-pair read protocol.

4.7. Single-cell RNA seq data analysis (Study 1 and 2)

4.7.1. scRNA seq Data pre-processing

Raw sequences were demultiplexed and converted into fastq files using Illumina bcl2fastq pipeline with default settings. The reference mouse and human genomes, GRCm38 (mm10) and GRChg38 applying the TopHat2 with Bowtie2 option were used for mapping. Alignments in the resulting BAM files were used for sorting. The samtools software was used to remove duplicate reads and FeatureCounts from the Subread package to calculate raw read counts. The 92 ERCC RNA included in the lysis buffer was used to ensure accurate mapping and as technical controls for raw read counts. ENSEMBLE annotation was performed using the org.mM.eg.db and org.Hs.eg.db packages (version 3.14.0) provided in the R-software.

4.7.2. scRNA-seq data clustering and analysis

Combined cell raw read counts in gene expression matrixes were analyzed using the Seurat pipeline. Libraries with abnormally high (≥750,000 suggesting possible cell doublets) or low (≤50,000) read counts or high relative content of mitochondrial or ERCC genes (≥10%) were removed. Genes expressed in < 3

cells, or with a cumulative count < 300 reads were considered absent. The 2000 most variable genes were used to calculate principal components of variation. Dimension reduction plots were visualized using the UMAP function (UMAP: uniform manifold approximation and projection). The graph-based clustering approach with the K-nearest neighbor (KNN) graph from the Seurat pipeline was used for clustering. Based on this clustering, possible batch effects related to RNAseq runs of individual 384-plates were examined and excluded. Within each cell type (SMCs, MPs and ECs), subcellular cluster genes were defined as those that were differentially expressed in a particular subcluster in relation to all other subclusters (i.e., background) of a given cell type. The Chi-square test is used to determine if the proportion of male and female cells within each major cell type significantly deviates from what would be expected by chance. Chi-square analysis is applied to categorical variables. It compares the size of any discrepancies between the expected results and the actual results, given the size of the sample and the number of variables in the relationship. For these tests, degrees of freedom are used to determine if a certain null hypothesis can be rejected based on the total number of variables and samples within the experiment. The Monocle3 pipeline was used for trajectory and pseudo time analyses. Venn diagrams were drawn with the VennDiagram R-software package (https://CRAN.R-project.org/package=VennDiagram). Biological pathway enrichment was assessed using Enrichr and GO platforms. GRNs were reproduced from the STARNET database (http://starnet.mssm.edu/) using the network package in the R-software (https://CRAN.R-project.org/ package=network). Statistical significance was assessed using the Wilcoxon rank sum test.

4.8. Independent data and experimental GRN validation for scRNA seq data

4.8.1. GRN122, GRN33 and GRN195 validation (Study 1)

Ischemic and non-ischemic human coronary artery tissues were obtained from consented diseased heart transplant donors and normal human coronary artery biospecimens were obtained from non-diseased donor hearts rejected for orthotopic heart transplantation at Stanford University. Tissues were transferred in frozen conditions and processed at the University of Virginia. RNA was extracted; libraries prepared using Illumina kits and sequenced on a NovaSeq S4. Reads were processed (FASTQC, Trimgalore, STAR, Picard, WASP) and quantified (RNA-SeQC, RSEM). Tissues were de-identified and clinical and histopathology information was used to classify ischemic, non-ischemic hearts and lesion and non-lesion containing arteries.

i) The NetRep R package was used to assess network preservation.

1. GRN33 (Female): Four external independent bulk RNAseq datasets from female patients. a) Coronary arteries obtained from freshly explanted hearts from orthotopic heart transplant recipients at Stanford

University (n =47). b) Post-mortem coronary artery samples downloaded from GTEx (v6) (n =47). c) The human cell type specific RNA-seq dataset primary blood MP from STARNET (n =134). d) Primary blood MP isolated in the multi-ethnic NGS-PREDICT study (n = 44)

2. GRN122 (Female): a) Post-mortem aorta samples downloaded from GTEx (n =70). b) Post-mortem coronary artery samples downloaded from GTEx (v6) (n =47). c) The human cell type specific RNA-seq dataset primary blood MP from STARNET (n =134). d) Primary blood MP isolated in the multi-ethnic NGS-PREDICT study (n = 44)
3. GRN195 (Male): Three external independent bulk RNAseq datasets comprising only males; coronary arteries a) Coronary arteries obtained from freshly explanted hearts from orthotopic heart transplant recipients at Stanford University (n = 92). b) Post-mortem coronary artery samples downloaded from GTEx (v6) (n = 70). c) Human aortic endothelial cells (HAEC) from distinct HAEC cultures generated from the human endothelial cells of male donors (n = 115).

ii) To validate the independent phenotypic association of GRN 33, 122, and 195: Differential gene expression analysis (DEGs) was performed on a subset of the Stanford dataset. Normal (n=23) vs. ischemic (n=38) donors, and lesion-containing (n=28) vs. non-lesion containing (n=26) coronary arteries were compared. DEseq2 (v3.1) normalized raw read counts, and DEGs were identified (FDR < 0.05) after adjusting for age, sex, RIN score, and ancestry, using the Benjamini & Hochberg method for multiple testing correction.

iii) To assess cell-specific expression of GRN 122, 33, and 195: snATAC-seq data from 41 Stanford coronary artery samples (10x Genomics platform) was utilized. Data, processed using the Cell Ranger ATAC pipeline (v1.2.0) and mapped to hg38, was further analyzed with ArchR (v1.0.1) to retain high-quality nuclei. Cell types were annotated by integrating snATAC-seq data with existing coronary artery scRNA-seq data from 4 individuals using ArchR, assigning scATAC-seq cells the gene expression profiles of their closest scRNA-seq matches. Finally, ArchR's "getPeak2GeneLinks" function correlated chromatin accessibility peaks with integrated RNA expression levels.

iv) The Athero-Express (AE) study includes over 3500 patients undergoing endarterectomy for atherosclerotic disease at the University Medical Center Utrecht. The biobank includes pathological and clinical data, 3-year follow-up, whole transcriptome (n=700), whole genome methylation (n=700), SNP data (n=2000), and single-cell sequencing data (n=45). DEGs in the single-cell RNAseq carotid plaque AE data between 20 females and 26 males was used to validate the enriched GRN in each cell type.

v) Functional validation of GRN 195:

vi) Cell culture and lentiviral transduction: TeloHAEC (ATCC CRL-4052) and primary HAECs (ATCC-PCS-100-011) were transduced with lentiviral vectors overexpressing PLVAP, FAM110D, or a non-targeting control (CTL). Cells were cultured in Vascular Cell Basal Medium (ATCC PCS-100-030) with

VEGF growth kit (ATCC PCS-100-041) and antibiotics. 75,000 cells/ml were seeded and transduced the following day (MOI 2-5). Blasticidin selection (6 μg/mL) was performed for 6–7 days. RNA was extracted (Monarch® Total RNA Miniprep Kit), cDNA synthesized (RevertAid First Strand cDNA Synthesis Kit), and qPCR performed (KiCqStart® SYBR® Green ReadyMix™). PLVAP and FAM110D expression was quantified using the ΔΔCt method, with RPLP0 as the housekeeping gene. RNA-Seq libraries were prepared using the Lexogen QuantSeq 3′ mRNA-Seq V2 kit with 100–200 ng of RNA and 12 nt dual indexes, then sequenced on an Illumina NextSeq 2000 (100 cycles, v3 kit). Reads were processed using the nf-core RNA-Seq pipeline (v3.12.0) with STAR alignment and Salmon quantification (hg38, Ensembl release 98). Lowly expressed genes (minimum 5 counts in any sample, total count 15) were filtered using edgeR (v3.32.1). Differential expressions were analyzed with DESeq2 (v1.30.1). Gene set enrichment analysis (GSEA) was performed using fgsea (v1.25.1) with default parameters and the DESeq2 DE statistic. Human GO Biological Process gene sets (MSigDB release 2023.2. Hs, 10–1000 genes) were used.

Proliferation assay: 7,500 Telo-HAEC cells/well were seeded in 96-well plates. Cell proliferation was monitored using an Incucyte® S3 Live-Cell Analysis System with 10x magnification, phase channel, and standard scan type imaging every two hours for 2–4 days. Data was normalized to time zero and figures generated using GraphPad Prism.

Cell painting: 0.6x10^4 HAEC cells/well were seeded in 96-well Revvity COL PhenoPlates. Following lentiviral transduction (MOI 5) and selection, cells were incubated for 6 days. Mitochondrial staining (PhenoVue JUMP kit) was performed for 30 min at 37°C before 4% PFA-PBS fixation. Cells were washed, then stained for 30 min with PhenoVue Hoechst 33342, Fluor 512, and Fluor 641. After washing, plates were imaged on an Opera Phenix Plus system (40x water immersion objective). Harmony software was used for image analysis (maximum intensity projections). Cell segmentation enabled measurement of morphological, SER texture, and Haralick features from nucleus, ring, cytoplasm, membrane, and cell regions. Data was exported and ratio-normalized to control to assess phenotypic changes.

4.8.2. GRN 39 validation (Study 2)

Immunofluorescence: Mouse aortas were harvested, fixed in 4% formaldehyde for 4–12 h at 4°C, and then in 20–30% sucrose/PBS for at least 24 h at 4°C. Tissues were embedded in cryo-medium, sectioned (14μm) using a CryoStat NX70, and stored at -80°C. Sections were dried, washed in PBS, and blocked (DAKO Serum-free protein blocking solution with 0.2% Triton X-100) for at least 60 min. Primary antibodies (anti-alpha-SMA-Cy3, anti-CD31, anti-CD68) were incubated overnight at 4°C, followed by secondary antibodies (goat anti-rabbit IgG Alexa Fluor-488 and rabbit anti-goat IgG Alexa Fluor-647) for 1h at room temperature. Sections were mounted with ProLong® Gold/DAPI and imaged using a Leica TCS SP8 confocal microscope (LAS X v3.5.7.23225).

Images were processed using ImageJ/FIJI and are presented as maximum intensity projections unless otherwise stated.

i) GRN39 conservation was assessed using the NetRep R package with three independent bulk RNA-seq datasets: (1) Stanford (n=139 explanted hearts), (2) GTEx coronary arteries (n=118), and (3) GTEx aortic wall (n=197). Differential gene expression analysis (DEGs) was performed on a Stanford subset (normal vs. ischemic donors, and lesion-containing vs. non-lesion containing arteries). DEseq2 (v3.1) normalized raw read counts, and DEGs were identified (FDR < 0.05) after adjusting for age, sex, RIN score, and ancestry, using Benjamini & Hochberg correction. To assess cell-specific expression of GRN 39, snATAC-seq data from 41 Stanford coronary artery samples (10x Genomics platform) was utilized.

ii) Primary vascular SMCs were isolated from ascending aortas of 136 healthy heart transplant donors (UCLA) and vascular SMCs isolated from the ascending aortas were purchased from Lonza/PromoCell (n=15). All cells were maintained in Smooth Muscle Cell Basal Medium with supplements and used for up to 5 passages. SMCs were RNA sequenced and characterized for migration, proliferation, and calcification. Quiescence was induced with serum-free media, and calcification assessed with high inorganic phosphate/osteogenic stimuli. Proliferation was measured with PDGF-BB, TGF-β1, or IL-1β. Migration was assessed using a modified Boyden chamber assay with a PDGF-BB gradient. RNA-seq libraries (Illumina TruSeq) were prepared from quiescent and proliferative (5% FBS) SMCs and sequenced (~100M reads, 300bp paired-end, STAR alignment to hg38). The number of expressed genes (> 6 reads/million in at least 80% of the samples) ranged from 18,078 in proliferative to 18,635 in quiescent SMC conditions.

iii) GRN39 ALCAM siRNA: HCASMCs were cultured in M231 medium with SMGS. For siRNA transfection, 1.5 x 10^5 cells/well were seeded in 12-well plates and grown to 70–80% confluency. Cells were washed and Opti-MEM was added. Control siRNA and ALCAM siRNA (80 nM and 180 nM, respectively) were complexed with Oligofectamine, incubated for 5 min, and transfected into cells for 4 hours. Complete M231 medium was added for 12 hours, followed by serum starvation (0.5% FBS) for 8 hours. Cells were then treated with CTL, TGFb (10 ng/ml), or Cholesterol (70 μg/ml) in starvation media for 48 hours. RNA was extracted using the Monarch Total RNA Miniprep Kit.

iv) GRN39 FRZB overexpression: Stable HCASMC pools expressing FRZB or empty vector were generated using lentiviral transduction (VectorBuilder pLV[Exp]-EGFP: T2A:Bsd-hPGK>hFRZB, VB230301-1111hbu, MOI 2). Transduced cells were selected with blasticidin (6 μg/ml) for 5–6 days. Prior to TGFb (10 ng/ml) and Cholesterol (70 μg/ml) treatment, cells were serum-starved (0.5% FBS) for 12 hours. After 48 hours of treatment, cells were collected for RNA extraction.

v) HCASMC calcification assay: 1.5 x 10^5 control and FRZB cells were seeded in 24-well plates. At 80% confluence, cells were treated with calcifying medium (3.7mM inorganic phosphate) changed every 72 hours for 14 days. Calcein was added 24 hours before imaging. Medium was replaced with normal

medium before imaging to reduce background. Calcification was visualized and quantified using IncuCyte and Fiji/ImageJ. Statistical analysis was performed using GraphPad Prism (v10.0.2).

vi) HCASMC qPCR and RNA seq: Knockdown and overexpression was validated by quantitative PCR (qPCR) for RPLP0, FRZB, and ALCAM using specific primers. The ΔΔCt method was used for relative gene expression. RNA-seq libraries were prepared using the Lexogen QuantSeq 3′ mRNA-Seq kit with dual indexing and sequenced on an Illumina NextSeq 2000. Raw data was deposited in NCBI GEO (GSE246269). RNA-seq reads were processed using the nf-core RNA-Seq pipeline (v3.12.0) with STAR alignment and Salmon quantification (hg38, Ensembl release 98). Lowly expressed genes (minimum 5 counts in any sample, total count 15) were filtered using edgeR (v3.32.1), resulting in 15,012 genes. Differential expressions were analyzed with DESeq2 (v1.30.1). GSEA was performed using fgsea (v1.25.1) with default parameters and the DESeq2 DE statistic. Human GO Biological Process gene sets (MSigDB release 2023.2. Hs, 10–1000 genes) were used.

5. STATISTICAL ANALYSIS

Carotid endarterectomy patients: No statistical method was used to predetermine sample size. No data were excluded from the analysis. The experiments were neither blinded nor randomized. Ldlr-/-Apob 100/100 mice: At least four biological replicates (one biological replicate: one pair of aortic arches) per group were used to minimize possible biological variability and technical background from the SMARTseq2 technique. As the animals were not subjected to any type of treatment, blinding and randomization were not necessary. Gene annotations of single-cell RNA data from mice and human were performed using the org.mM.eg.db and org.Hs.eg.db packages (version 3.14.0) provided in R software. The male/female proportion of cells within major cell types was statistically analyzed using the chi-square ($\chi 2$) test, which measures how a model compares to actual observed data.

A volcano plot shows differentially expressed subcluster genes (red, upregulated; blue, downregulated) using all other SMC subclusters as background. Cell type-specific gene enrichment was calculated using the Wilcoxon rank-sum test (log2-fold change > 0.3, Bonferroni-adjusted $P<0.005$). Dot plots show top-ranked biological processes according to GO. GO enrichment −log10 P values were calculated with Fisher exact test. Box plots show sex-specific expression of top-ranked key driver isolated from female (n = 7) and male (n = 8) carotid plaques and Ldlr$^{-/-}$*apob*$^{100/100}$ mice, female (n = 18) or male (n = 28) where the center line denotes the median value (50th percentile), and the box contains the 25th to 75th percentiles of the dataset. The whiskers mark the 5th and 95th percentiles. Radar plots show statistical significance of key cardiometabolic phenotypes associations with cell type specific enriched GRNs. The significance of GRN-phenotype associations (−log10; P=0–100) was calculated by aggregating GRN gene-level phenotype associations (Pearson correlation two tailed t-test) corrected for the total number of STARNET GRNs (n=135) and the number of genes in each GRN using the Benjamini-Hochberg procedure.

To validate the independent phenotypic association of GRN 33, 39, 122, and GRN 195 external independent bulk RNAseq datasets namely GTEx and a subset of Stanford dataset was used in which the raw read counts were normalized in DEseq2 (v3.1) and DEGs were called at FDR <0.05 significance cut-off after correcting for age, sex, RIN Score, and ancestry. P-values were adjusted for multiple testing using the Benjamini & Hochberg method. Average pairwise correlation of GRN195 coding genes across different conditions, comparing Mock, oePLVAP, and oeFAM110D groups with four replicates each. In the box plot, the center line denotes the median value (50th percentile), and the box contains the 25th to 75th percentiles of the dataset. The statistical test was conducted using t-test (two-sided). No adjustment was made and there was no multiple comparison. The normalized enrichment scores (NES) indicate the direction and strength of gene set enrichment in RNAseq data of HAECs overexpressing PLVAP and FAM110D. The statistical test used for gene set

enrichment analysis (GSEA), employs a one-sided test to assess whether a predefined gene set is significantly enriched at either the top or bottom of the ranked gene list. P-values were calculated using permutation testing, and multiple comparisons were adjusted using the FDR correction.

Statistical analysis conducted for relative proliferation in the PLVAP/FAM110D group compared to the control (CTRL) group by using repeated measures ANOVA (two-sided) Dunnett's multiple comparison test, $P < 0.0001$. Data represented as Mean ± SD. Figures made using GraphPad Prism9. Representative images of cell painting showing nuclear (blue), mitochondria (red), nucleic acid (yellow) stains at d6 in LV- overexpressing HAEC cells. Scale bar = 50 μm. 4 wells/group including 10 regions/well were imaged (n=40 images/group) with two biological replicates. Heatmap illustrating the effect of LV-FAM110D and LV-PLVAP overexpression on cell phenotypic features in HAECs at d6 detected by cell painting. 4 wells/group including 10 regions/well were imaged (n=40 images/group) with Opera Phenix Plus High Content Screening System and cell morphological signature was analyzed by Harmony analysis software. Two independent experiments were performed. Heatmap values showing fold changes over the LV-EGFP control vector. Relative gene expression of PLVAP and FAM110D in Telo-HAECs and HAECs compared to control lentivirus transduced cells. The qPCR was performed in four replicates. Data represented as Mean ± SEM.

Gene expression of PLVAP, FAM110D, ALCAM and FRZB was quantified using the ΔΔCt method with RPLP0 as the housekeeping gene. HCASMC quantification of calcified area was done using Fiji/ImageJ. Statistical analysis of data was done using GraphPad prism (version 10.0.2). HCASMC RNA-Seq data were processed using the nf-core RNA-Seq pipeline (version 3.12.0) with default settings (STAR aligner and salmon quantification) using the hg38 genome with Ensembl release 98 transcript definitions. Lowly expressed genes were filtered out using the edgeR (version 3.32.161) function filterByExpr. Differential gene expressions were assessed using DESeq2 (version 1.30.162) with default parameters. GSEA was performed using the fgsea package (version 1.25.163). Statistical analysis was performed using R (version 4.3.2). All graphs were edited for appearance using Adobe Illustrator (v25.2.3).

6. RESULTS

6.1. Single-cell RNA sequencing studies of sex differences in the subcellular composition of carotid plaques (Study 1) and atherosclerotic progression in human-like *Ldlr$^{-/-}$Apob$^{100/100}$* mice in relation to symptomatic plaques in humans (Study 2)

6.1.1. Marked sex-differences in the subcellular composition of major carotid plaque cell types (Study 1)

Carotid plaque samples from 15 patients (7 female, 8 male) were analyzed using single-cell RNA sequencing (Smart-seq2). Established cell-type markers for initial clustering (Butler et al., 2018) of the scRNA seq data revealed 3 major cell-types corresponding to ECs, MPs, and SMCs, as well as two smaller cell-type clusters represented by T-cells and pericytes (Fig. 12a–b). While overall clinical characteristics were similar between sexes (except for age, expected higher median age for female), clustering of 7690 cells (ECs, MPs, SMCs, T-cells, pericytes) revealed similar numbers of major cell types between sexes (Fig. 12c), but sex-biased distribution within subcellular clusters of these major cell types, suggesting diverse pathobiological functions.

Sub-clustering of 3263 SMCs identified 9 subtypes (SMC1-9) with distinct sex biases showing a trajectory from contractile to immune-activated, chondrogenic, and osteogenic phenotypes (Fig.12d). SMC3 and SMC8 were predominantly female, SMC3 was characterized by myofibroblast-like markers and functions related to collagen fibril organization and ossification, while SMC8 showed osteogenic-like markers and bone mineralization functions. Male-dominated SMC subtypes included SMC1 (contractile markers, muscle organ development), SMC4 and SMC5 (chondrocyte-like markers, collagen fibril organization and chondrocyte development), and SMC9 (endothelial-like markers, endothelial cell proliferation and angiogenesis). SMC6, exclusively male and from a single patient, was excluded from further analysis.

Sub-clustering of 2796 MPs revealed 8 subclusters (MP1-8) with varying sex distribution and functional profiles (Fig. 12e). A trajectory from MP1 to MP8 showed evolving markers for resident-like, dendritic-like, smooth muscle cell-like, and lipid-associated/inflammatory-like phenotypes. Female-dominant MP subclusters (MP2, MP3, MP5, MP8) were enriched in immune response-related processes like antigen presentation, immunoglobulin-mediated immune response, and response to interferon-gamma. Male-dominant MP subclusters included MP1 (resident-like markers, regulation of autophagy and toll-like receptor signaling) and MP7 (hemoglobin and tissue remodeling markers, lysosome organization, blood vessel endothelial cell migration, and bone resorption regulation).

Sub-clustering of 892 ECs identified 5 subclusters: EC1-EC3 (vasa vasorum) and EC4/EC5 (carotid lumen). EC1 showed angiogenesis markers, EC2 has SMC

and immune cell markers, and EC3 contains inflammation markers. EC4 and EC5 showed endothelial progenitor cell and ECM markers, respectively. While EC numbers were similar between sexes (Fig. 12c), EC1 was male-dominant, showing EC proliferation, migration, and RAP1 signaling, with functions in angiogenesis and T cell-mediated cytotoxicity. Female-dominant EC subclusters were EC2 (EndoMT and EndICLT markers, MHC class II antigen presentation, and immunoglobulin-mediated immune response) and EC4 (tissue-resident vascular endothelial progenitor markers, response to oxidative stress, and autophagy) (Fig. 12f).

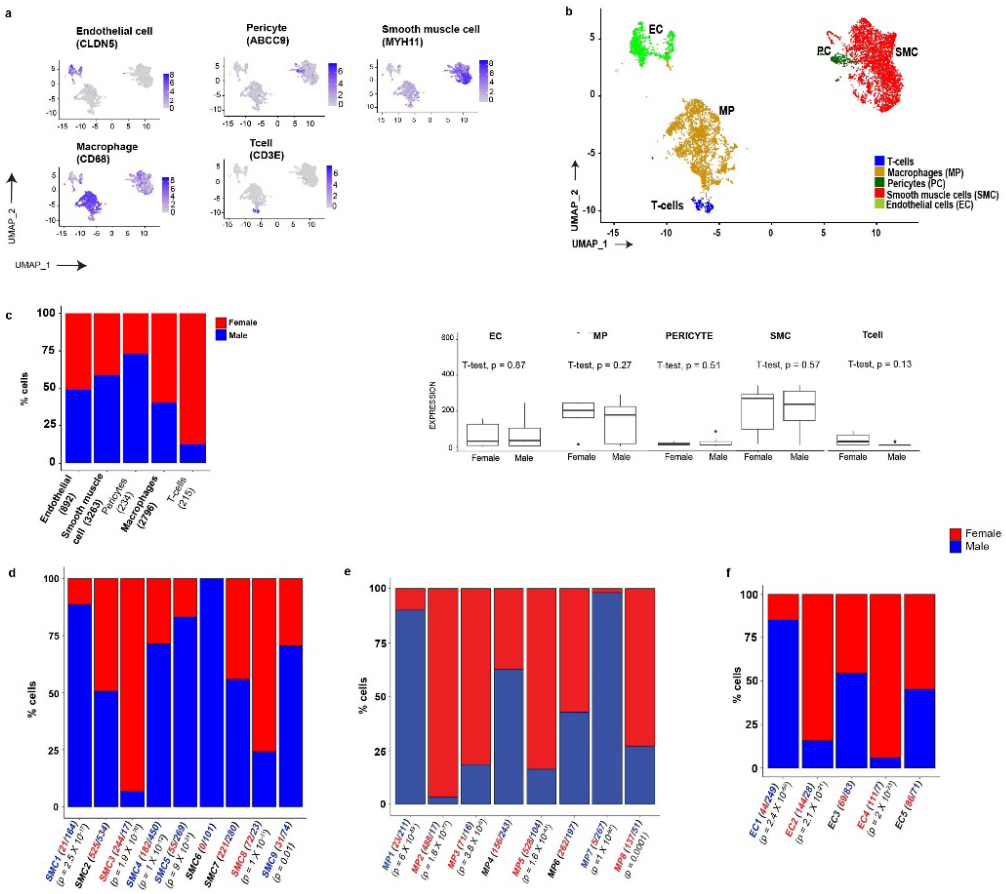

Figure 12: Single-cell RNA sequencing data of carotid endarterectomy patients.

a. Uniform Manifold Approximation and Projection (UMAP) featuring single cell gene expression levels of established cell-type markers of endothelial cells (ECs, Claudin5 (CLDN5)), macrophages (MPs, (CD68)), pericytes (PCs, ATP binding cassette subfamily C member 9 (ABCC9), smooth muscle cells (SMCs, Myosin11 (MYH11)) and T-cells (T-cell surface glycoprotein CD3 Epsilon chain (CD3E)).

b. UMAP of major cell types.

c. Bar plots showing the relative contribution of major cell types isolated from female (n=7) and male (n=8) carotid plaques. (), indicates the total number of cells in each cluster.

d. Bar plots showing the relative sex-specificity of cells in each SMC subcluster. (), shows the total number of cells collected in each category. *P*-values were calculated using chi-square statistics assessing the statistical significance between observed and expected values.
e. Bar plots showing the relative sex-specificity of cells in each MP subcluster. (), shows the total number of cells collected in each category. *P*-values were calculated using chi-square statistics assessing the statistical significance between observed and expected values.
f. Bar plots showing the relative sex-specificity of cells in each EC subcluster. (), shows the total number of cells collected in each category. *P*-values were calculated using chi-square statistics assessing the statistical significance between observed and expected values.

6.1.2. Single-cell RNA sequencing of atherosclerosis progression in human-like *Ldlr$^{-/-}$Apob$^{100/100}$* mice (Study 2)

Single-cell RNA sequencing (Smart-Seq2) was performed on 8898 cells from healthy (10 weeks, C57BI6/J mice) and atherosclerotic (20, 30, 45, and 60 weeks, *Ldlr$^{-/-}$Apob$^{100/100}$* mice (C57BI6/J)) aortic arches of mice to study vascular cell transformation during atherosclerosis progression. The *Ldlr−/−Apob100/100* mouse model carries most of its plasma cholesterol in circulating human-like LDL particles developing advanced atherosclerosis despite being on a regular chow diet (Bjorkegren et al., 2014; Skogsberg, Dicker, et al., 2008; Skogsberg, Lundstrom, et al., 2008). SMCs, ECs, and inflammatory cells (Pdgfrβ+, Cd31+, Cd45+) were isolated from lesions of the aortic arch using FACS. Mice with *early atherosclerosis* (week 20–30) showed signs of infiltrating macrophages and had more myoFB, EC, and SMC clusters than mice at baseline (Fig. 13a). Mice with advanced atherosclerosis (week 45–60) with increased lesion size showed more distinct SMC and MP subtype patterns. Notably, B and T cell clusters remained relatively stable compared to early lesions. GO analysis of differentially expressed genes revealed distinct functional changes in subcellular clusters for SMCs, ECs, and macrophages during atherosclerosis progression.

FACS-sorted 7690 cells of main cell types, SMCs, ECs, and CD45+ from carotid plaques of 15 patients (8 symptomatic, 7 asymptomatic) were clustered revealing SMC, EC, macrophage, T-cell, and pericyte populations. While EC contribution was slightly higher in asymptomatic patients, other cell types were similar between symptomatic and asymptomatic patients (Fig. 13b). Differential gene expression analysis showed changes primarily in macrophages, with some overlap between SMCs and ECs. GO analysis revealed biological processes highly relevant to atherosclerosis and its main cell types including ECM disassembly (MPs), ECM organization (SMCs), cytokine signaling (ECs), and mitochondrial ATP synthesis/coupled electron transport (ECs/SMCs).

SMCs were the most abundant atherosclerosis cell type in both in the mice (n=3778) and the human carotid plaques (n=3235). Mouse SMCs clustered into 7 subtypes (mSMC1-7), with a trajectory showed progression from mSMC1/2 to mSMC3/4, then branching to mSMC5 and mSMC6/7. mSMC6/7 showed high

expression of inflammation and osteogenesis markers. Human SMCs (hSMC) also had subtypes with hSMC6 and hSMC7 showing low contractility and high collagen assembly markers, respectively. hSMC7, along with the smaller hSMC8, were most prevalent in symptomatic plaques. Comparing mSMC6/7 and hSMC7 revealed a 35% gene overlap between at least two clusters, with hSMC7 sharing 43% of its genes with mSMC6/7, suggesting shared functions, particularly ECM organization (Fig.13c). These clusters not only captured already established advanced-stage atherosclerosis SMC genes (Wirka et al., 2019) but due to the sequencing depth of SMARTseq2 compared to earlier single cell atherosclerosis studies mostly using 10x (Alencar et al., 2020), expanded the number of advanced atherosclerosis genes related to osteogenic and inflammatory pathways by three-fold, among other revealing new genes involved in the signaling and complement activation pathways.

MPs were the second most abundant cell type in both the mice (n=1218) and the human carotid plaques (n=2888). Mouse macrophages are clustered into six subtypes (mMP1-6), with mMP1-4 being monocyte-like and mMP5/6 exhibiting pro-inflammatory (M1) and Trem2-high lipid-associated macrophage phenotypes, respectively. Pseudotime analysis showed mMP1 evolving into mMP2, mMP5, and mMP6. Human macrophages were clustered into eight subtypes (hMP1–8). hMP2-4 were monocyte-like, while hMP5-7 expressed osteopontin, a marker of advanced plaques. hMP7, originating from hMP5, was enriched in cells from symptomatic plaques. mMP5/6 and hMP7 contained 80 genes that overlapped (12.5% of mouse genes and 23% of the human genes) (Fig.13d). This overlap was enriched for several pathways including cytokine signaling and ECM disassembly, which was consistent with a combined M1/Trem2-high phenotype. mMP5/6 and hMP7 in capturing most of already established advanced-stage macrophage genes (Cochain et al., 2018; Winkels et al., 2018), also contained three times the established genes in novel advanced stage macrophage genes including new pathways like protein ubiquitination, mitochondrial dysfunction, and RHOGDI/EIF2 signaling.

scRNA-seq analysis of the mouse ECs revealed four clusters (mEC1-4), with markers indicated that mEC4 originated from luminal ECs and mEC1-3 from the vasa vasorum capillaries. The human ECs clustered into six subtypes (hEC1-6), again with hEC4-6 being luminal and primarily from asymptomatic plaques and hEC1-3 being from the vasa vasorum (Fig.13e). Both mouse and human luminal ECs were associated with ECM organization. T-cell and FB numbers increased with atherosclerosis progression, but no T-cell or FB subcluster were identified. T-cell gene expressions in advanced atherosclerotic lesions were associated with macrophage activation in mice and with general immune responses in humans' carotid plaques. FB processes shifted from ECM organization (early) to complement regulation (advanced) in both mice and humans. Pericyte and B-cell numbers remained stable throughout the progression of atherosclerosis in mice. In human carotid plaques, pericytes and T-cells lacked subclusters and were equally distributed between symptomatic and asymptomatic plaques.

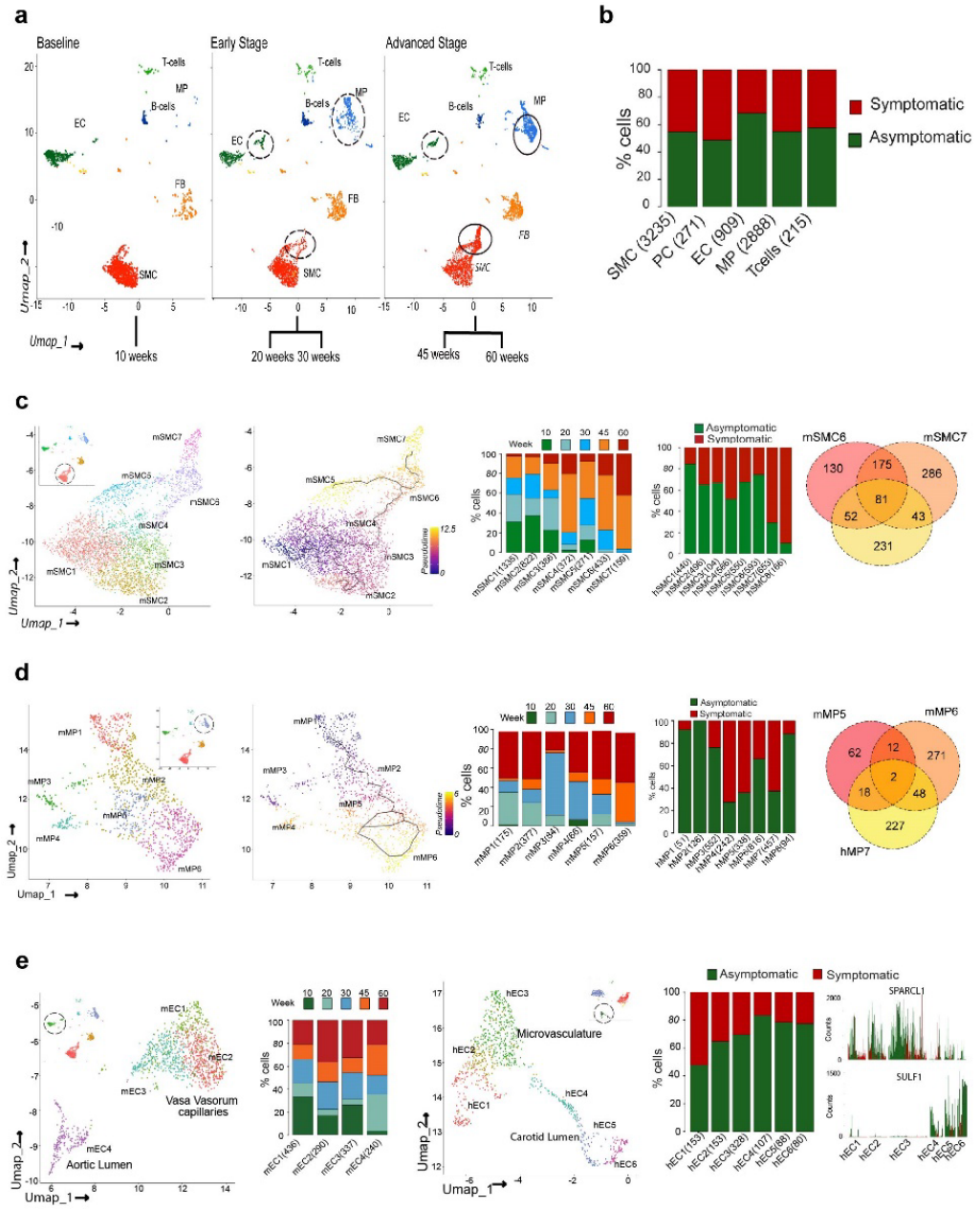

Figure 13: Single-cell RNA sequencing of vascular cells during atherosclerosis progression in *Ldlr−/−Apob100/100* mice.

a. Clusters of individual B-cells, endothelial cells (ECs), macrophages (MPs), myofibroblasts (FBs), smooth muscle cells (SMCs), and T-cells from the atherosclerosis-free baseline, the early atherosclerosis stage, and the advanced atherosclerosis stage visualized by Uniform Manifold Approximation and Projection (UMAP). Circles indicate major changes to the cell clusters during the progression of atherosclerosis.

b. Bar plots showing the relative contribution of symptomatic and asymptomatic cells of the main cell types in the carotid plaques. Values in parenthesis indicate the total number of collected cells for each cell type.
c. Subclusters and trajectory (color coded by pseudotime) visualized with Uniform Manifold Approximation and Projection (UMAP) of SMCs during atherosclerosis progression in *Ldlr*$^{-/-}$*Apob*$^{100/100}$ mice. Bar plot showing the relative cell contributions of the seven SMC clusters (*x*-axis) in relation to progression of atherosclerosis in *Ldlr*$^{-/-}$*Apob*$^{100/100}$ mice (weeks 10 to 60). Values in parenthesis indicate the total number of cells obtained in each of the seven SMC subclusters. Bar plot showing the relative contribution of carotid plaque cells in the 8 SMC subclusters isolated from asymptomatic (n=7) or symptomatic (n=6) carotid plaques. Values in parenthesis indicate the total number of cells obtained in each SMC subcluster. Venn diagram of subcluster genes in mSMC6/mSMC7 and hSMC7.
d. Subclusters and trajectory (color coded by pseudotime) visualized with Uniform Manifold Approximation and Projection (UMAP) of MPs during atherosclerosis progression in *Ldlr*$^{-/-}$*Apob*$^{100/100}$ mice. Bar plot showing the relative cell contributions of the six MP clusters (*x*-axis) in relation to progression of atherosclerosis in *Ldlr*$^{-/-}$*Apob*$^{100/100}$ mice (weeks 10 to 60). Values in parenthesis indicate the total number of cells obtained in each of the seven MP subclusters. Bar plot showing the relative contribution of carotid plaque cells in the 8 MP subcluster isolated from asymptomatic (n=7) or symptomatic (n=6) carotid plaques. Values in parenthesis indicate the total number of cells obtained in each SMC subcluster. Venn diagram of subcluster genes in mMP5/mMP6 and hMP7.
e. EC subclusters visualized by UMAP of the vasa-vasorum capillary, and the aortic arch lumen isolated during atherosclerosis progression. Bar plot showing the relative contribution of cells in the 4 mouse EC subcluster in relation to progression of atherosclerosis (weeks 10-60). (), indicate the total number of cells in each of the 4 EC subclusters. EC subclusters visualized by UMAP of the vasa-vasorum capillary, and the carotid lumen isolated from human carotid plaques. Bar plot showing the relative contribution of symptomatic and asymptomatic cells in the six EC subclusters isolated from the human carotid plaques. Bar plots showing the cellular expression of the human EC microvasculature marker, SPARCL1 (top panel) and the EC lumen marker, SULF1 (lower panel).

6.1.3. Cross-tissue and tissue-specific co-expression GRNs capture a large portion of variation associated with CMDs linking them to CAD (Study 3)

STARNET is a large genetics-of-gene expression (i.e., RNAseq) study of blood, four major metabolic organs and the advanced and early atherosclerotic arterial wall isolated from 600 patients with CAD during CABG and from up to 250 CAD-free controls during other forms of open breast surgery (mostly aortic valve replacements). We first explored correlations of CMD trait with individual genes across the seven tissues. A small number of genes were found to correlate with the complex trait of CAD severity (i.e., SYNTAX and Duke scores) whereas for the CMD traits, the number of individual genes that correlated was large. Secondly, we examined differential expressions of individual genes between the CAD cases and CAD-free controls across the seven tissues.

The numbers of DEGs across the STARNET tissues were large despite applying strict statistical criteria (>30% fold change in expression levels and FDR <1%). A significant number of the DEGs was represented by noncoding mRNA. Notably, in the advanced AOR, the numbers of DEGs were equally upregulated and downregulated (n = 1,727 and 1,652, respectively). However, differentially expressed noncoding genes in AOR were predominantly upregulated in the CAD patients (n = 9,047 and 560). This contrasted the patterns in the liver where both coding (n = 436 and 1,555) and noncoding (n = 111 and 4,480) genes were predominantly downregulated. In the VAF, SF, and SKLM the up-and down-regulation of coding genes were balanced whereas noncoding genes were predominantly downregulated in CAD. Adjustments for metabolic factors, drugs, and genotypes did not substantially alter these patterns of DEGs. While the roles of non-coding in CMDs and CAD are yet to understood, the GO analysis of the differentially expressed protein coding genes revealed expected biological processes and pathways. In the case of upregulated genes, AOR showed increased lipoprotein remodeling and coagulation, LIV increased cholesterol biosynthesis, and VAF leukocyte activation. For downregulated genes, AOR showed decreased glucose transport, LIV, decreased cell adhesion and for VAF, decreased regulation of secretory pathways. Other tissue-specific changes associated with CAD, were increased antigen presentation in SF and mitochondrial respiration in SKLM. We concluded that analyzing genes one-by-one or pre-defined pathways is insufficient to embrace the full complexity of CMDs and CAD.

To provide broader functional contexts, systems analysis of the entire RNAseq datasets at once by identifying GRNs has proven useful to capture the full complexity of complex traits, like CMDs and CAD (Björkegren et al., 2015; Bjorkegren & Lusis, 2022; Tegner & Bjorkegren, 2007). For that purpose, we first applied weighted gene co-expression network analysis (WGCNA) across the RNA datasets of the seven STARNET tissues. We identified 224 co-expression modules (135 tissue-specific, 89 cross-tissue) that were largely replicated in independent RNAseq datasets of corresponding tissues (except for MAM) from the GTEx NIH datasets (Consortium et al., 2015). Overall, the co-expression modules showed significant associations with CMD phenotypes, CAD severity, and enrichments for differentially expressed genes in CAD. In sum, the 224 co-expression modules provided a broader functional context for studying CMDs and CAD.

To add regulatory directions to the undirected gene-gene interactions of the co-expression modules, GENIE3 was applied to transform the 224 co-expression modules to GRNs. Transcriptional factors and network genes regulated by SNPs (eSNPs) were used as regulatory priors – meaning that these genes were deemed as being regulators rather than regulated in the GRNs. Based on the directional edges of the GRNs, key drivers (high-hierarchical regulatory GRN nodes with many down-stream regulated genes) were identified (Talukdar et al., 2016). In total there were 41,586 eSNP in the 224 GRNs that combined was found to significantly contribute to broad sense heritability (H2) of CAD, with a combined contribution of 59.8% according to the Restricted Maximum Likelihood (REML)

method (Nelson et al., 2017) as well as with an alternative method mediated expression score regression (Yao et al., 2020), the latter leading to 54.3% beyond the contribution of still GWAS lead SNPs (~22%). Although GRN size (reflected by the number of genes and eSNPs) largely determined its H^2 contribution, the individual H^2 contributions of individual GRNs and eSNPs varied. In addition, a significant portion of candidate genes (932/1037) previously identified for CAD by GWAS were present in at least one GRN. Several GRNs were in fact statistically enriched for CMD/CAD GWAS candidate genes. In summary, the 224 GRNs captured other aspects of CAD heritability than that of GWAS hits perhaps involving SNPs with relative GWAS hits, lower CAD heritability contributions that within the context of several SNPs within GRNs become significant. Next, MR was used to confirm the causal roles of GRNs for the developments of CMDs and CAD by examining the overlap of key drivers in tissue-specific GRNs with those of causal networks inferred by a MR analysis (Yavorska & Burgess, 2017) applied to each tissue. ~50% of MAM- and blood-specific GRNs and ~67% of the AOR, LIV, SKLM, SF and VAF-specific GRNs were enriched with causal key drivers. 218 GRNs with key drivers were found causal for CMDs or CAD ($P \leq 0.001$) either in the one-sample (STARNET data only), or the two-sample (STARNET with GWAS) MR analyses. 28 GRNs were significantly enriched with several CMD/CAD-causal key drivers (FDR < 10%).

The 224 GRNs were organized based on their intra- and inter-organ interactions using a “supernetwork” approach by applying Bayesian network modeling to eigengene values of the WGCNA-inferred co-expression modules. Bootstrapping confirmed the reliability of the supernetwork’s connections by randomly redrawing individual STARNET samples to create 1,000 eigengene datasets from which supernetwork connections were recalculated. 11% were inter-organ interactions (113 out of 1021) involving cross-tissue modules. Of the 113 inter-organ interactions, 101 were represented by 89 cross-tissue modules and only 12 were active in two tissue-specific modules. In their inter-organ functions, cross-tissue GRNs were also observed to be over-represented among GRNs strongly associated with CAD severity and more frequently enriched in genes differentially expressed between the CAD cases and controls. In addition, the cross-tissue GRNs contributed more to CAD heritability (especially per key driver), were more enriched with GWAS candidate genes than the tissue-specific GRNs, highlighting a critical role in mediating interactions between gene activity of metabolic organs with that of the arterial wall driving CAD development.

The notion that cross-tissue GRNs are critical for inter-organ communications was also reflected in the GO enrichment results. Compared to tissue-specific GRNs, cross-tissue GRNs showed greater enrichment in communicative biological processes and molecular functions like “response to stimuli” and “secretion”. Furthermore, secretory proteins within corresponding genes active within cross-tissue GRNs were more likely to correlate with eigengene values of neighboring tissue-specific GRNs than secretory proteins present in tissue-specific GRNs, further supporting their role in organ-to-organ signaling.

In total, there were 374 unique endocrine candidates identified in 89 cross-tissue GRNs, with at least one association with a neighboring tissue-specific GRN, of which 152 (>40%) operated along an adipose (VAF/SF) to liver axis. 42 of the 152 adipose candidates were validated in independent global adipose and liver gene expression data from inbred strains of mice: HMDP showing significant associations with liver network modules. A particularly reliable part of the adipose-to-liver axis between the cross-tissue GRN78 and the liver-specific GRN98 involved 13 adipose endocrine candidates. GO enrichment analysis showed that GRN78 is enriched in genes for "carboxylic/fatty acid processing" and GRN98 genes involved in "sterol/cholesterol metabolism". The thirteen adipose endocrine factors of cross-tissue GRN78 correlated strongly with both eigengene values of GRN98 and their host GRN, GRN78. The correlations of the 13 adipose endocrine factors with GRN98 were confirmed in independent adipose and liver datasets of the HMDP, GTEx, morbid obesity cohorts. In addition, the 13 adipose endocrine candidates correlated with BMI, CAD severity, and lipid/glucose levels in both STARNET and HMDP suggesting a mechanism where the activity of GRN78 regulates the secretion of these 13 factors from adipose tissue, which then travels to the liver to impact liver lipid and glucose homeostasis primarily by modulating activity of genes in GRN98. To experimentally validate this endocrine effect of GRN78, recombinant forms of 10 of the 13 adipose endocrine factors were generated and subsequently injected into C57BL/6N wild-type mice. After 72 hours, 7 of the injected proteins significantly altered hepatic or plasma lipid or blood glucose levels. EPDR1, FCN2, FSTL3, and LBP showed the broadest effects on lipid and glucose levels in mice, mirroring human and HMDP data. Sequencing RNA isolated from livers of the C57BL/6N mice after intraperitoneal injections of these four endocrine factors markedly affected the expression of four key drivers in GRN98. One of these drivers, Dhcr7, strongly correlated with *Pcsk9* expression, a target for cholesterol lowering drugs. Hepatic cholesterol was negatively associated with Hmgcs1 and Rdh11 expressions. While individual endocrine candidate effects varied, their up/downregulation often aligned with known mechanisms. EPDR1 reduced glucose, free fatty acids, and liver triglycerides, consistent with its role in thermogenesis. The experimental validation by injecting these 10 endocrine factors in mice reinforces the relevance of the identified GRNs to modulate CMDs and provides strong evidence for a significant, previously under characterized adipose-to-liver endocrine signaling axis impacting lipid and glucose levels as illustrated in Fig.14.

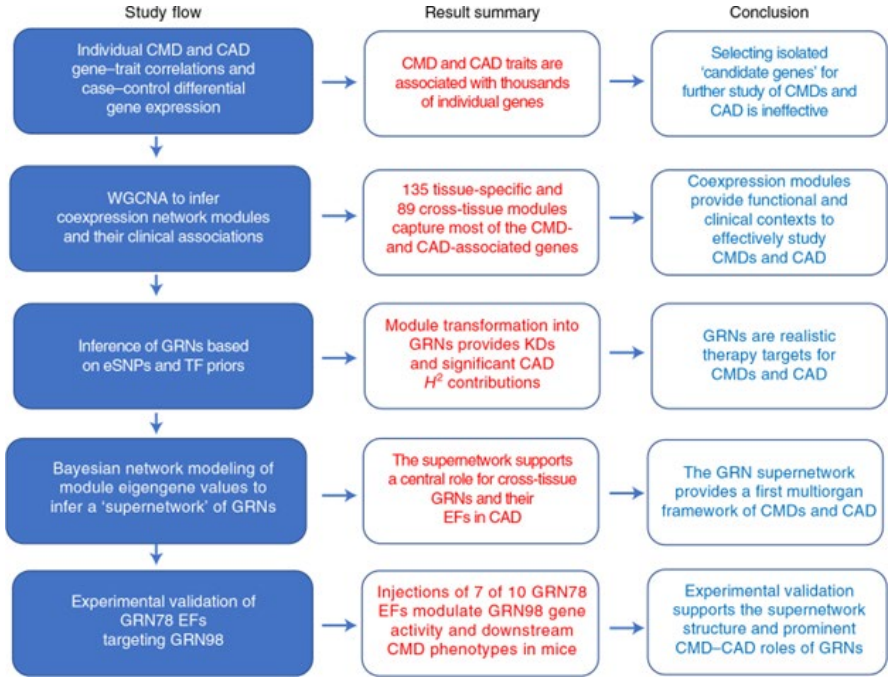

Figure 14: Schematic overview of STARNET study with main data-analysis steps with results and conclusions. TFs, transcriptions factors; GRN78, cross-tissue GRN78 identified in the STARNET study; GRN98, liver-specific GRN98 identified in the STARNET study; EF, endocrine factor; KDs, key disease drivers; H^2, broad-sense heritability contribution.

6.2. Integration of single-cell subcellular clusters with STARNET GRNs

6.2.1. Sex biased subclusters integration with STARNET GRNs (Study 1)

scRNA-seq identifies subclusters of cells in healthy and diseased tissues, including the atherosclerotic arterial wall. Reasonably, only some of the genes within these subcellular clusters are responsible for driving disease. Compared to these subcellular clusters, at least some GRNs offer a better understanding of the clinical implications of the identified subcellular clusters by offering maps of gene interactions underlying disease mechanisms as well as key driver genes as putative targets. To plausibly identify atherosclerosis mechanisms and key driver genes segregating male and female carotid plaque subclusters, we examined the enrichment of sex-biased subcellular cluster genes in the one hundred thirty-five tissue-specific GRNs from the STARNET study. Notably, six arterial wall GRNs were enriched in genes from the sex-biased SMC subclusters as shown in Fig15a. Specifically, arterial wall (AOR and MAM) GRNs 82, 49, and 39 were enriched

in genes from subcellular clusters identified as both males and female. GRN177, however, strongly associated with both SYNTAX and DUKE scores, was strongly enriched in the male-dominant SMC1 subcluster with a SMC contractile phenotype. GRN177 harbors 43 key drivers and 19 CAD GWAS candidate genes jointly contributing to 2.16% of CAD heritability. GRN177 involved the biological processes of actin filaments and cytoskeletal rearrangement and was enriched in genes associated with BMI and LDL/HDL cholesterol levels. *ILK* was identified as a potential target to modulate GRN177 activity specifically in males, both based on the expression patterns in male carotid plaques and during atherosclerosis progression on the $ldlr^{-/-}apoB^{100/100}$ mouse model.

Out of the 6 arterial wall GRNs enriched in the sex-biased genes of the MP subclusters, GRN33/GRN3 and GRN174, were strongly enriched in genes in the female-dominant lipid-laden inflammatory MP8 subcluster and the immune response related MP3 subcluster, respectively as shown in Fig.15b. The MAM-specific early atherosclerotic GRN3 was found to be an early version of the AOR advanced atherosclerotic GRN33 (n=82/115 overlap). GRN33 contained 33 key drivers, contributed 1.14% to CAD heritability, and harbored 10 CAD GWAS candidate genes. GRN33 was also found to be strongly enriched in genes associated with CAD severity, BMI and LDL/HDL cholesterol levels. It is involved in the cellular responses to lipid and cytokine stimuli processes according to GO analysis. CEBPD was identified as a potential female-specific key driver of GRN33, showing upregulation in female carotid stenosis patient with symptoms and during atherosclerosis progression in the of Ldlr−/−Apob 100/100 mice. GRN174, another GRN enriched in genes from the female-dominant MP subclusters, predominantly from the female MP3 subcluster, involved immune-regulatory functions. GRN174 contains 27 keydrivers, contributes to 1.58% of CAD heritability, and harbored 12 CAD GWAS candidate genes. GRN174 was also strongly enriched in genes associated with blood levels of LDL cholesterol and CRP. LYZ was identified as a potential female-specific key driver of GRN174 activity, displaying increased expression during atherosclerosis progression in female mice and upregulation in female carotid plaques.

Three arterial wall GRNs (GRN113, GRN122 and GRN195) were strongly enriched in genes from two sex-biased carotid plaque vasa vasorum EC1 and EC2 subclusters as shown in Fig.15c. GRN113/GRN195 were enriched in by male EC1 subcluster (angiogenesis, T cell cytotoxicity) and GRN122 in the female genes from the EC2 subcluster involved in EndoMT, and EndICLT atherosclerosis mechanisms. GRN195 and GRN122 were also associated with SYNTAX and Duke scores and carotid plaque clinical symptoms. GRN122 contributes to 4.10% to CAD heritability and contains 16 CAD GWAS candidate genes. Although associated with the EC2 subcluster, the genes within this GRN are predominately of macrophage origin and involved in immune response regulation likely suggesting that the EC2 cells are undergoing EndoMT or/and EndICLT. GRN122 was associated with Duke score, BMI, and cholesterol levels. CTSB was identified as a potential female-specific key driver of GRN122. The MAM-specific GRN113 was identified as an early arterial wall version of the advanced AOR

GRN195 (n=73/141 overlapped) GRN195 contributed 0.6% to CAD heritability and contained 7 CAD GWAS candidate genes. GO analysis of GRN195 genes revealed the biological processes of vasculature development and angiogenesis. PLVAP was identified as a potential male-specific key driver of GRN195 according to independent data.

6.2.2. Advanced stage/symptomatic subclusters integration with STARNET GRNs (Study 2)

To understand the pathophysiological relevance and possible clinical implications of the 6 advanced-stage and symptomatic-associated SMC and macrophage clusters, their individual enrichments in 135 tissues specific GRNs inferred from the STARNET study was examined as depicted in Fig.15d. Twelve GRNs were significantly enriched in genes from these SMC clusters. Eight of these 12 GRNs originated from the atherosclerotic arterial wall. Among these, GRN82, GRN49, and GRN39 were most significantly enriched. GRN39 genes were predominantly expressed in SMCs whereas GRN49 and GRN82 genes were expressed across multiple cell types. GRN39 consists of 182 genes, with 41 identified as key drivers, contributed to 1.79% of CAD H^2 and contained 8 CAD GWAS genes. GRN195 was enriched for the biological process of ECM organization. GRN39 was also associated with SYNTAX score, BMI, HbA1c, LDL, and HDL cholesterol levels. The mouse orthologs of the top key drivers, FRZB and ALCAM, showed increased SMC expression during atherosclerosis progression suggesting they may be suitable targets for modulating GRN39 activity in the arterial wall preventing atherosclerosis.

Ten GRNs whereof 6 from the arterial wall were significantly enriched in genes from the MP subclusters. The top 4 enriched GRNs also originated from the atherosclerotic arterial wall. Among these, GRN33 and GRN122 were strongly associated with Duke score, indicating a direct link to the extent of coronary atherosclerosis. GRN33 consisted of 182 genes, with 41 genes identified as key drivers. GRN33 contributed 1.14% of CAD heritability Enriched biological processes was responses to lipids/cytokines and GRN33 also contained 9 CAD GWAS candidate genes and was highly enriched in the AOR and MAM genes differentially expressed between CAD cases and controls and associated with Duke score and levels of LDL cholesterol. The macrophage expression levels of the top key drivers, ATF3 and FOS, show increased expression during atherosclerosis progression in the Ldlr−/−Apob 100/100 mice suggesting these genes may be suitable early targets for modifying GRN33 activity. GRN122 consisted of 766 genes, with 30 key drivers contributing to a significant 4.1% to CAD heritability. GRN22 was strongly associated with Duke score, BMI, and plasma LDL cholesterol levels representing the biological process of immune response regulation. The macrophage expression levels of the key drivers CTSB and CAPG increased with atherosclerosis progression in the Ldlr−/−Apob 100/100 mouse model suggesting these genes may be suitable early targets for modifying GRN122 activity and atherosclerosis development.

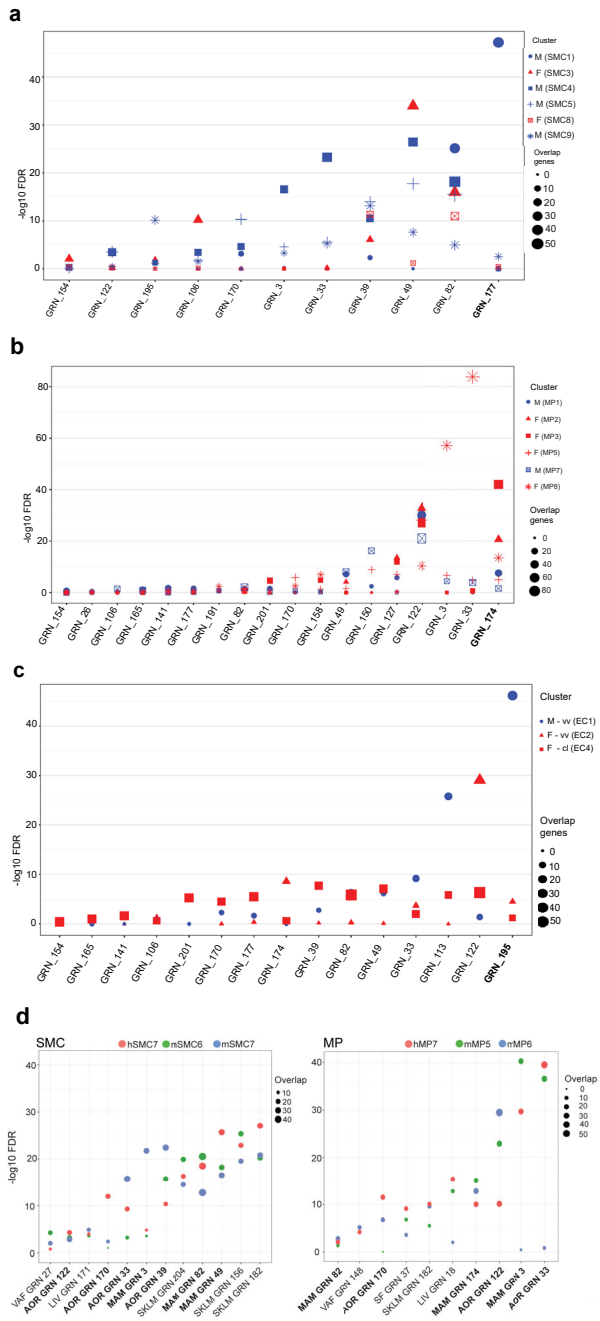

Figure 15: Sex biased, advanced stage and symptomatic subclusters integration with STARNET GRNs.

a. Dot plot showing eleven top-ranked arterial wall GRNs (x-axis) according to their enrichments in genes of sex-specified SMC subclusters in the carotid plaques. y-axis shows -log-10% False discovery rate (FDR, log-10 FDR 10% highlighted). Dot size indicates the number of genes overlapping between the SMC subclusters and GRNs.

b. Dot plot showing eighteen top-ranked arterial wall GRNs (x-axis) according to their enrichments in genes of sex-specified MP subclusters in the carotid plaques. y-axis shows -log-10% False discovery rate (FDR, log-10 FDR 10% highlighted). Dot size indicates the number of genes overlapping between MP subclusters and GRNs.
c. Dot plot showing fifteen top-ranked arterial wall GRNs (x-axis) according to their enrichments in genes of sex-specified EC subclusters in the carotid plaques. y-axis shows -log-10% False discovery rate (FDR, log-10 FDR 10% highlighted). Dot size indicates the number of genes overlapping between EC subclusters and GRNs.
d. Dot plots showing GRNs (*x*-axis) with gene enrichment (*y*-axis, −log10 false discovery rates [FDRs]) for genes in the advanced stage and symptomatic SMC (left, n=12 GRNs) and MP (right, n=10 GRNs) subclusters.

6.3. Reproducibility and clinical validation of sex-biased and clinically relevant GRNs in independent human arterial wall datasets (Study 1 and 2)

GRNs enriched in sex-biased carotid plaque subcellular clusters, GRN33, GRN122, and GRN195 were replicated using independent datasets. GRN33 (enriched in female-biased MP8 subcluster genes) and GRN195 (enriched in male-biased EC1 subcluster genes) were involved in early atherosclerosis development. GRN122 (female MP) was found linked to EndoMT and EndICLT, potentially crucial for plaque erosion in females. Cellular origin of the GRNs were confirmed using independent coronary lesion snATAC/scRNA-seq data as shown in Fig.16a. Reproducibility was assessed using the NetRep application in independent coronary artery bulk RNAseq and cell-specific RNAseq datasets as shown in Fig.16b. The cellular origin of GRN122 was confirmed to MPs using independent coronary lesion snATAC/scRNA-seq data. GRN33 and GRN122 robustly replicated in female-specific coronary bulk, and MP-specific RNAseq data. Similarly, GRN195 robustly replicated in male-specific coronary bulk, and arterial wall EC-specific RNAseq datasets. All three GRNs showed enrichment in DEGs related to atherosclerosis and CAD examined in human coronary arteries between lesion and non-lesion samples, ischemic vs. non-ischemic CAD patients, and a combination of both as shown in Fig.16c.

GRNs enriched in advanced-stage and symptomatic-associated SMC and macrophage clusters; GRN39, GRN33, and GRN122, GRN39 had the highest fraction of single-cell annotated genes (>95%). GRN39 replicated in coronary artery and aortic wall bulk RNAseq datasets. GRN39 genes were also enriched in ischemic status and combined lesion/ischemic status. The SMC origin of GRN39 was reconfirmed using snATAC/scRNA-seq. Gene-gene interactions were further reconfirmed in RNAseq data from proliferative cultured SMCs. Most GRN39 genes were associated with vascular SMC atherosclerosis-related phenotypes confirming GRN39 plays a critical role in transforming SMCs from a contractile to an osteogenic phenotype, promoting advanced and symptomatic atherosclerosis.

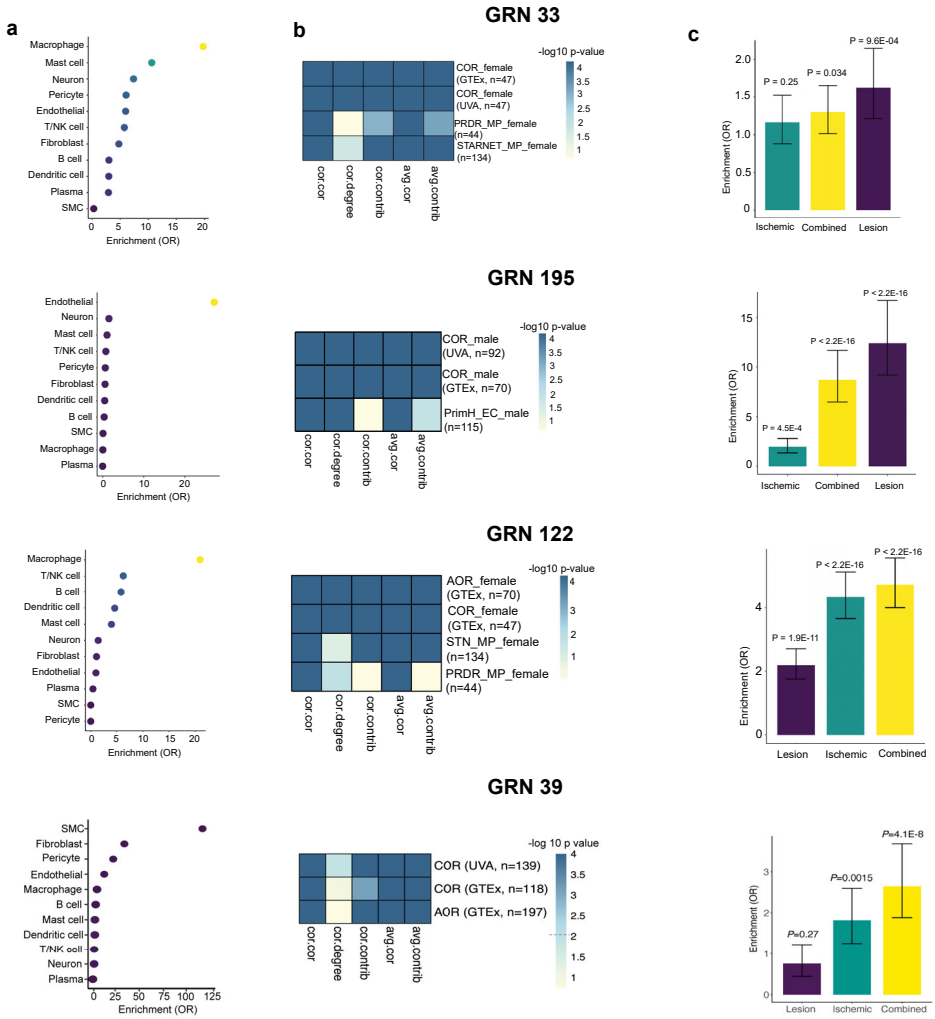

Figure 16: Reproducibility of GRNs in independent human arterial wall datasets.
a. GRN33, GRN195, GRN122 and GRN39 cell-type specificity according to the enrichment of cell-type determined genes in scATAC-seq data obtained from 41 human coronary artery tissue samples. Dot plots show odds ratio of enrichment of GRN33, GRN195, GRN122 and GRN39 module genes for marker genes determined from integrated snATAC-seq and scRNA-seq profiles obtained from atherosclerotic coronary artery tissues.
b. GRN33, GRN195, GRN122 and GRN122 39 reproducibility according to NetRep.
c. Bar plots showing the odds ratio (OR) for GRN33, GRN195, GRN122 and GRN39 enrichments in coronary artery differential expressed genes. The Fisher test was performed on overlapping gene sets, the p-values shown represent the Fischer test (two-tailed) p-values and error bars represent 95% confidence intervals. Data represented as Mean ± SEM.

6.4. Experimental validation of GRN195 (Study 1) and GRN39 (Study 2)

Experimental validation of GRN195 was done by overexpressing its top driver genes, PLVAP and FAM110D in HAECs as shown in Fig.17c. RNA-Seq and cell painting assays were performed. Overexpression of both PLVAP and FAM110D significantly reduced GRN195 connectivity and resulted in negative NES, indicating decreased GRN195 activity as depicted in Fig.17a. Functional assays confirmed a significant decrease in EC proliferation in *FAM110D*-overexpressing cells compared with controls as depicted in Fig.17b. FAM11OD overexpression increased nuclear intensity and reduced nuclear area, suggesting chromatin condensation. Reduced cell and membrane area, consistent with downregulation of cell adhesion genes. Upregulation of genes involved ribonucleoprotein complex biogenesis and cell cycle progression and a significant decrease in endothelial cell proliferation. PLVAP overexpression decreased mitochondrial staining intensity. and generated a trend toward downregulation of oxidative stress response genes including an upregulation of cell cycle-related genes paralleled with an increase in endothelial cell proliferation. Overexpression of FAM110D and PLVAP specifically impacted GRN195 gene activity associated with alterations in atherosclerosis-relevant EC phenotypes, including chromatin condensation, adhesion, mitochondrial activity, and proliferation.

Experimental validation of GRN39, on the other hand, was done by overexpressing and silencing its top key driver genes, FRZB and ALCAM, respectively, in HCASMCs. The HCASMCs were also subjected to different media conditions: untreated, TGFβ-treated (mimicking contractile SMC phenotype), and cholesterol-treated (mimicking synthetic SMC phenotype). After these manipulations a marked upregulation of FRZB and down regulation of ALCAM mRNA levels in the range of 60% to 70% were observed. Perturbing FRZB and ALCAM in HCASMCs resulted in GRN39 genes being significantly more connected than random gene groups for all conditions. Specifically, FRZB overexpression increased GRN39 connectivity while ALCAM knockdown decreased GRN39 connectivity as shown in Fig.17f. FRZB and ALCAM perturbations showed strong associations with GRN39 genes, especially under cholesterol-loading conditions as shown in Fig.17d. FRZB overexpression significantly affected SMC proliferation, chromatin remodeling, and ossification inducing calcification as depicted in Fig.17e. FRZB overexpression also affected the levels of transcription factors ZEB1 and KLF3 as well as chromatin-associated protein HMGB1. ALCAM knockdown increased expression of inflammation-related pathways and induced INHBA, IL1B, and MMP3 expressions. Taken together, these results confirm that GRN39 plays a critical role in several atherosclerosis-relevant functions.

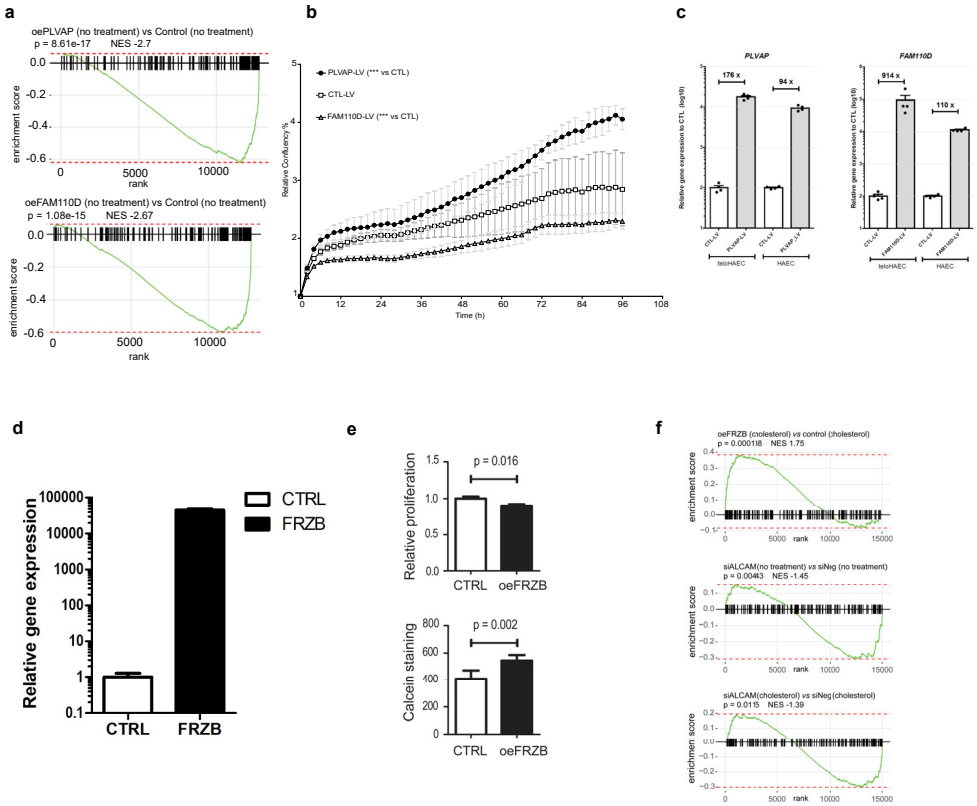

Figure 17: Experimental validations of enriched GRN195 and 39 by targeting its top key drivers.

a. Gene set enrichment analysis (GSEA) of RNAseq data of HAECs overexpressing *PLVAP* and *FAM110D* (GRN195). The normalized enrichment scores (NES) indicate the direction and strength of gene set enrichment. The statistical test used was GSEA, which employs a one-sided test to assess whether a predefined gene set is significantly enriched at either the top or bottom of the ranked gene list. P-values were calculated using permutation testing, and multiple comparisons were adjusted using the False Discovery Rate (FDR) correction.

b. Relative proliferation in the PLVAP group compared to the control (CTRL) group shows a significant increase, while the FAM110D group shows a significant decrease relative to CTRL. Statistical analysis was conducted using repeated measures ANOVA (two-sided) Dunnett's multiple comparison test, $P < 0.0001$. Data represented as Mean ± SD.

c. Relative gene expressions of PLVAP and FAM110D (GRN195) in Telo-HAECs and HAECs compared to control lentivirus transduced cells. The qPCR was performed in four replicates. Data represented as Mean ± SEM.

d. Relative gene expression of FRZB (GRN39) compared to control lentivirus transduced cells. The qPCR was performed in four replicates.

e. Relative proliferation in the FRZB, compared with the control (CTRL) group, was significantly reduced using a paired 2-tailed *t* test (*P*=0.016) and borderline significant using a Wilcoxon 2-tailed matched-pairs signed rank test (*P*=0.13; n=4, mean and SD are shown. Fluorescence intensity indicative of calcification deposit levels in cells transduced with CTRL or FRZB lentivirus following a 14-day inorganic phosphate-induced calcification. Statistical analysis was performed using the Mann-Whitney *U* test. Both mean and SD are shown (n=6).
f. GSEA of RNAseq data of HCASMCs overexpressing *FRZB* and downregulating *ALCAM*.

7. DISCUSSION

7.1. Significance

The broad aim of this thesis was to provide a deeper understanding of the biological diversity of cells that transform healthy tissues into diseased ones, and to explore sex-diverse subcellular cluster in atherosclerosis, the main underlying cause of stroke and myocardial infarction. We did this by generating and analyzing a combination of bulk and deep single-cell RNA sequencing data, the latter using the Smart-seq2 technology. RNA-seq offers a more comprehensive genetic view than DNA sequencing by detecting gene fusions, splicing variations, mutations/indels, and differential gene expression. Bulk transcriptomics treats tissue as uniform providing an important holistic understanding of gene transcription within and across organs. However, in bulk RNA, gene expression variability across individual cells is masked, a limitation that can be overcome by generating scRNA-seq data resolving the cellular and transcriptomic heterogeneity by identifying subpopulations of cells. Elucidating the dynamics of cellular transitions at the single-cell level is believed to be critical for a better and deeper understanding of diseases, particularly complex diseases (Islam et al., 2014; Shapiro et al., 2013). ScRNA-seq lacks the data depth (including lowly expressed genes) and genetic diversity needed for network inference. Traditional genetic studies of atherosclerosis focus on individual genes and pathways (driven by rare disease successes and the current focus on resolving target genes in GWAS risk loci). This "one-by-one" approach is likely insufficient. Current genetic understanding suggests that the overall number of risk alleles, rather than specific isolated pathways, drives disease (Pang et al., 2023). Therefore, atherosclerosis should at least in parallel be viewed as a complex system involving thousands of interacting genes reflected in GRNs not just because of individual pathways. We have shown that GRNs are critical in complex diseases by pinpointing their pathophysiological and molecular dysfunctions (Bjorkegren & Lusis, 2022). Furthermore, in study 3, we show that physiological GRNs typically are confined to specific tissues, while the more disease-relevant GRNs often extend beyond tissue boundaries (S. Koplev et al., 2022). This makes a lot of sense to us not at least from what we already know about complex diseases like CMDs where liver, SKLM and fat, particularly VAF, play crucial roles in regulating levels of plasma lipids, glucose as well as the general inflammatory status of the individual. Then, CMDs drive atherosclerosis in the arterial wall not solely through elevated plasma lipid and glucose levels but as we now have shown, through cross-tissue GRNs.

While single-cell co-expression network inference methods are being developed (Algabri et al., 2022), the limited genetic diversity of current scRNA-seq datasets including those reported in this thesis, hinders the inference of directional networks at the cellular level (i.e., cellular GRNs). . To circumvent this temporary shortcoming, we integrated subcellular clusters identified in study 1 and 2 as critical for advanced atherosclerosis with the GRNs identified in study 3 from the STARNET dataset. In this fashion we were able to pinpoint key drivers, clinical

contexts, heritability contributions, and pathophysiological mechanisms for the male and female subcellular clusters critical for driving the formation of the advanced atherosclerotic plaque.

The generated single cell data of atherosclerosis in humans and mice is a central piece of this thesis. By applying Smart-Seq2, we provide the deepest multi-species scRNA-seq data set of atherosclerosis to date, which allowed us to reidentify and substantially extend our current understanding of the cellular transformation taking place during atherosclerosis progression as well as in the late-stages of atherosclerosis risking adverse clinical symptoms in humans. In addition, given a near equal number of females and male carotid stenosis patients, we for the first time reveal that although absent at the major cell type level (i.e., ECs, SMCs, MPs), the subcellular forms of the main cell types are clearly sex biased. While female carotid stenosis patients undergo subcellular changes consistent with EndoMT or EndICLT risking plaque erosion followed by possible thrombosis, male carotid stenosis patients at risk of adverse clinical events appear to have EC subclusters promoting angiogenesis and T-cell activation.

7.2. Main findings

7.2.1. scRNA-seq provides sex-biased subcellular differences in major cell types of carotid plaques (Study 1)

Variation in plaque size, composition, and morphology have been observed between sexes in patients with carotid stenosis for at least a decade (Maas & Appelman, 2010; Witteman et al., 1989). More recently, the role of sex has also been shown in the pathogenesis of atherosclerosis; where in women, lipid-poor carotid plaques destabilize due to erosion causing thrombosis and in men lipid-rich carotid plaques rupture (Vrijenhoek et al., 2014). The observed sex differences in plaque characteristics are further complicated by the premenopausal period in women, which slows atherosclerosis development likely due to lower inflammatory status. However, in the postmenopausal period, female plaque growth accelerates with the clinical event rate exceeding that of men (Man et al., 2020). Notably, both stable and unstable plaques pose significant cardiovascular risks.

By generating deep scRNA sequencing data (i.e. Smart-seq2 data) from carotid plaques isolated from men and women, the dominant main finding in study 1 of this thesis, was that cell composition differences between sexes, although absent at the level of main cell types, were pronounced between the subcellular clusters of these main cells. In fact, 6 of 9 SMC, 6 of 8 MP and 3 of 5 EC subcellular clusters were significantly sex biased. In SMCs, a transformation from healthy arterial wall contractile forms (SMC1-2) to endothelial-like forms (SMC9) of SMCs in the advanced stages of atherosclerosis occurred in both sexes. However, in women SMCs also transformed to become osteogenic-like SMCs whereas in men, a similar transformation resulted in chondrogenic-like

SMCs. MPs also exhibited sex-diverse transformation with pro-inflammatory MP subclusters being more prevalent in male carotid plaques, while dendritic-like, lipid-laden, and other pro-inflammatory MPs were more common in females. Furthermore, vasa vasorum ECs displayed functions related to angiogenesis, immune responses in male carotid plaques whereas in female plaques, EndoMT/EndICLT prevailed (female-biased).

A second main finding of study 1 was that beyond identifying the sex-biased subcellular clusters, we were able to pinpoint and validate (see further below) some clinical and pathophysiological relevance of these subclusters by examining their enrichments of genes from the 15 sex-biased subclusters in 135 tissue-specific GRNs identified in STARNET (S. Koplev et al., 2022). Reassuringly, among the 9 most enriched GRNs, all were from the arterial wall (AOR or MAM): GRN49, GRN82, and GRN177 were enriched in genes from SMC subclusters; GRN3, GRN33, and GRN174 in MP subclusters; and GRN113, GRN122, and GRN195, in EC subclusters. Among these GRNs, GRN82, GRN49 and GRN39, were enriched in genes from SMC subclusters representing both sexes. GRN49 and GRN82 were identified as being of "cross-cell type" nature meaning that genes in these networks, according to established cell-specific markers in the scRNA-seq data, originate from all three cell-types (i.e., SMCs, MPs and ECs) suggesting that these GRNs may be involved in dedifferentiation and transdifferentiation taking place in all main cell types. In female carotid plaques, GRN33 was found enriched in lipid/inflammatory genes of MP8 subcluster, supposedly involved in driving the transformation to TREM1+ macrophages, a key component of atherogenic inflammation and cerebrovascular events. A key driver of GRN33, CEBPD, is known for promoting a pro-inflammatory PLIN2hi/TREM1hi macrophage state, potentially via a TREM2 transition. GRN3, showing a similar enrichment patterns as GRN33, was in fact noted to likely represent an early MAM-specific version of the AOR-advanced atherosclerosis-specific GRN33. The MAM-specific GRN174 was predominately enriched in genes from the female-dominant MP3, MP5, and MP8 subclusters containing genes critical to immune cell regulation.

Particularly prevalent in women, EndoMT is a process where endothelial cells transition to mesenchymal-like cells believed to play a critical role in atherosclerosis. Depending on the extent of EC adhesion loss, EndoMT can lead to either chain or individual cell migration, with the latter potentially causing plaque erosion and thrombosis. Just like SMC phenotypic switching (i.e., "transformation"), EndoMT is not a one-way transition but based on stimuli can result in either athero-protective (ACTA2+) or atherosclerosis-prone (FN1+) subcellular phenotypes (Gole et al., 2022). ACTA2+ ECs contribute to plaque stability by compensating for SMC loss in the fibrous cap preventing plaque rupture. Notably, in our study, female-dominated EC subclusters were both ACTA2+ (EC2, 84%) and FN1+ (EC4, 94%). GRN122, a network with antigen-presenting genes associated with immune regulation and cell activation, was enriched in genes from both the ACTA2+ EC2 and the FN1+ EC4 subclusters. As such GRN122

may be important to regulate not only EndoMT but also inflammation predominantly in female carotid plaques.

In male carotid plaques, GRN177 was found enriched in genes from the male-dominant SMC1 subcluster consistent with a regulatory role of actin filament-, and cytoskeletal-based biological processes. In addition, 85% of the genes in the vasa vasorum EC1 subcluster were male-derived and strongly enriched in GRN113 and GRN195. Like the GRN3/GRN33 situation, the MAM-specific GRN113 appeared to represent an early version of the AOR-specific GRN195 with 50% of GRN113 genes also being present in GRN195. GRN195 involved an angiogenesis-related biological processes, driven by key drivers like PLVAP and FAM110D. In vasa vasorum, GRN113 and GRN195 as part of EC1 appear to play a role in hypoxia-driven, unstabilizing neovascularization of carotid plaques particularly happening in men, which would increase the risk of intraplaque hemorrhage and plaque rupture.

To validate the cellular origins, reproducibility, and clinical relevance of the seemingly most sex-biased GRNs, GRN33, GRN195, and GRN122, we sought their replication in 6 independent atherosclerosis-related omics datasets (i.e., in female and male snATAC seq, bulk RNAseq, and EC/MP-specific RNAseq datasets). At large, the reproducibility (using the NetRep software, see Methods) and cellular origins (using the EC/MP-specific and snATAC seq datasets) of these 3 GRNs could be reconfirmed. GRN195, standing out as being the most sex-dominant subclusters, was further investigated using experimental validation by overexpressing its key drivers, PLVAP and FAM110D in cultured ECs, HAECs. RNA-sequencing and cell painting revealed that these drivers indeed modulate GRN195 gene activity and impact EC phenotypes relevant to atherosclerosis, including chromatin condensation, adhesion, mitochondrial activity, and proliferation. In sum, GRN195, particularly PLVAP and FAM110D, emerge as potential therapeutic targets to regulate neovascularization to prevent plaque destabilization and plaque rupture particularly in male ACVD patients.

This study, while utilizing deep sequencing of carotid stenosis samples, has limitations including a relatively small cohort (15 patients, slightly older female group), questioning its population representation. Furthermore, rare cell types may be underrepresented due to the FACS-sorting, and the use of GRNs inferred in data from CAD patients may not fully reflect the molecular nuances of the atherosclerotic plaques of carotid stenosis patients. While CAD and carotid stenosis share core atherosclerotic mechanisms, the sex-specific cellular clusters found in carotid plaques may exhibit unique clinical and pathophysiological features not represented by GRNs derived from studies of CAD. Nonetheless, taken together, the findings from study 1 emphasize the need for additional sex-stratified research of carotid stenosis patients with the longer-term goal to better sex-stratify diagnostics and therapies of patients with life-threatening ACVDs.

7.2.2. scRNA-seq elucidates the vascular cell transformations that occur during atherosclerosis progression (Study 2)

Atherosclerosis involves the pathophysiological activation of multiple cell types, including SMCs, ECs, and immune cells. Single-cell genomic technologies have revealed the heterogeneous cell composition of lesions in both murine and human atherosclerosis (Pan et al., 2020). However, the regulation, dynamics, and relationship with CVD risk of these cell types remain unclear. In study 2, we sought to better understand how arterial wall cell plasticity during atherosclerosis progression is essential for determining how stable plaques destabilize, leading to clinical events. The main finding of study 2 was that atherosclerosis cell states/types in both mice, during disease progression and in humans with clinical symptoms undergo significant phenotypic switching through comprehensive scRNA-seq using Smart-Seq2 including the data used in Study 1. In study 2, we analyzed cells isolated from: (1) the healthy aortic arch of C57BI6/J mice at baseline (10 weeks); (2) the atherosclerotic aortic arch of the Ldlr−/−Apob 100/100 mice (C57BI6/J) at early (20 and 30 weeks) and advanced (45 and 60 weeks) stages of atherosclerosis; and (3) carotid plaques from symptomatic (n=8) and asymptomatic (n=7) human patients (data overlapping with Study 1). The Ldlr−/−Apob 100/100 mouse model was selected because, like humans, it accumulates most of its plasma cholesterol in Apob100/100-containing LDL, leading to the rapid development of human-like advanced atherosclerotic lesions, even on a regular chow diet.

Unlike terminally differentiated skeletal and cardiac muscle cells, mature SMCs retain remarkable plasticity, driven by environmental cues and extracellular signals (Gomez & Owens, 2012). In study 2, during murine atherosclerosis progression, seven SMC subclusters (mSMC1-mSMC7) were identified, each characterized by varying expressions of markers related to contraction, immune response, and inflammation/osteogenesis. Pseudotime trajectory analysis showed baseline mSMC1 and mSMC2 transitioning to mSMC3 and mSMC4, followed by a divergence into two paths: one leading to mSMC5, and the other to mSMC6 and mSMC7. Notably, mSMC6 and mSMC7 exhibited high expression of inflammation and osteogenesis-related genes. In human carotid plaques, eight SMC subclusters (hSMC1-hSMC8) were found, with hSMC7 showing the most evident switch from a contractile to an osteogenic phenotype.

During murine atherosclerosis macrophages, the second most abundant cell type, formed six subclusters, transitioning from monocyte-like to M1-type proinflammatory (mMP5) and Trem2-high lipid-associated (mMP6) phenotypes. In patients with carotid plaques, macrophages formed eight subtype clusters (hMP1-8). Excluding the sparse hMP1, hMP2-4 exhibited high expression of the monocyte-like macrophage marker F13A1. Conversely, hMP5-7 showed high osteopontin (SPP1) expression, a marker of advanced atherosclerotic plaques. Notably, hMP7, and its trajectory origin, hMP5, were enriched in cells from symptomatic carotid plaques.

The scRNA-seq data also uncovered four endothelial cell (EC) clusters emerging during murine atherosclerosis progression. Luminal ECs (mEC4) expressed Cytl1, while vasa vasorum capillary ECs (mEC1-mEC3) expressed Kdr. Luminal mEC4, consistently present across all disease stages, primarily reflected extracellular matrix organization. In human carotid plaques, six EC clusters were identified. Consistent with murine data, hEC1-hEC3 originated from microvasculature, whereas luminal hEC4-hEC6, predominantly from asymptomatic lesions, showed a strong association with extracellular matrix organization. T-cell (n=46, 121, 190) and myofibroblast (FB; n=131, 532, 708) numbers increased from baseline to advanced atherosclerosis, but neither formed distinct subtype clusters. FB biological processes shifted from extracellular matrix organization (GO:0030198) in early atherosclerosis to regulation of complement activation (GO:0030449) in advanced stages. Pericytes (n=23, 28, 27) and B-cells (n=93, 177, 88) were also present in the arterial wall during atherosclerosis progression but showed minimal changes in cell numbers and function. In human carotid plaques, pericytes (n=271) and T-cells (n=215) lacked distinct subclusters and were evenly distributed between symptomatic and asymptomatic plaques.

In sum, by generating the most comprehensive multi-species scRNA-seq dataset of atherosclerosis to date using Smart-Seq2. This dataset, and thus the main findings of Study 2 reidentified and significantly expanded upon previously known inflammatory/osteogenic SMC subtypes and M1-type proinflammatory/Trem2-high lipid-associated macrophage subtypes, both crucial in advanced and symptomatic atherosclerosis. The mSMC6, mSMC7, and hSMC7 clusters captured nearly all SMC genes from prior advanced-stage atherosclerosis scRNA-seq studies. These clusters tripled the gene sets associated with the advanced-stage osteogenic and inflammatory phenotype (n=564/176), revealing novel signaling and complement activation pathways. Likewise, the mMP5, mMP6, and hMP7 clusters encompassed nearly all macrophage genes from previous studies, tripling the gene sets for M1-type proinflammatory and Trem2-high lipid-associated phenotypes (n=906/325) and identifying new pathways such as protein ubiquitination, mitochondrial dysfunction, oxidative phosphorylation, and the RHOGDI and EIF2 signaling pathways.

Like study 1, we next sought to provide pathophysiological and clinical context to these six cell clusters through integrative analysis with 135 tissue-specific STARNET GRNs. This analysis identified 3 enriched GRNs all from the arterial wall (AOR and MAM): the macrophage-specific GRN33 and GRN122, and the SMC-specific GRN39, which were successfully validated using independent data and GRN39 by experimental assays. These 3 GRNs represent the second part of our main findings of Study 2.

GRN33: KLF4 (Kruppel-like factor 4), the top-ranked key driver of macrophage-dominated GRN33, has previously been implicated in atherogenesis, influencing SMC switching to a macrophage-like osteogenic phenotype, contributing to plaque calcification and destabilization. KLF4 is also a known regulator within macrophages (Depuydt et al., 2020). Other key drivers in GRN33 included

KLF2 and KLF6, along with compelling factors such as ATF3 (a negative regulator of inflammation-promoting cholesterol metabolism in macrophages (Wang et al., 2022), CEBPD (a promoter of lipid accumulation in M1-type macrophages in atherosclerotic lesions (Lai et al., 2017), EGR1-3 (early growth response transcription factors regulating IL1B and CXCL2), FOS/FOSB/FOSL1&2 and JUN/JUNB/JUND (involved in foam cell formation and AP-1 transcription factor complex regulation), MYC/MXD1 (regulating cellular transformation via the MAX transcription factor), NFIL3 (activating ATFs), and NR4A1-3 (nuclear receptor subfamily 4). Collectively, these key drivers suggest that in advanced atherosclerosis, GRN33 promotes M1-type macrophage formation, contributing to symptomatic plaques through lipid accumulation and increased inflammation.

GRN122: While GRN33 has been partially implicated in atherogenesis (Alencar et al., 2020) macrophage-dominated GRN122, a significant contributor to CAD heritability (H2=4.1%), has not. Although GRN122 contains 16 CAD-associated genes identified by GWAS, their individual contributions (<0.1% on average) do not account for the 4.1% H2. Instead, the genetic regulation of GRN122 genes (eQTLs) carries substantial CAD heritability. GRN122 is notable for its limited number of key drivers (30 out of 766 genes), with only seven top ranked. Consistent with its dominant GO representation, GRN122's significant contribution to CAD heritability and its strong association with coronary atherosclerosis severity suggest an immune-regulatory role crucial for symptomatic plaque progression. Notably, unlike GRN33, enriched only in mMP5 genes (lipid-poor proinflammatory macrophages), GRN122 was uniquely enriched in mMP6 genes (TREM2-high, lipid-rich, potentially athero-protective macrophages) and partially enriched in mMP5 genes. Thus, GRN122 may determine whether lesion macrophages differentiate into proatherogenic or athero-protective phenotypes through its immune-regulatory function.

GRN39: WNT (wingless-related integration site) signaling operates through three main pathways: canonical (transcriptional), noncanonical planar cell polarity (cytoskeletal/cell morphology), and noncanonical Wnt/calcium (intracellular calcium regulation). While the extracellular matrix (ECM) modulates WNT signaling (Astudillo & Larraín, 2014), WNT itself influences cell adhesion (Brembeck et al., 2006), inflammation (George, 2008), calcification (Albanese et al., 2017), and atherosclerosis (Boucher et al., 2020). GRN39 comprises 30 ECM-related genes, 7 proliferation/differentiation genes, 34 intracellular calcium regulation genes, and 28 cell adhesion genes. GRN39's enrichment in ECM organization and intracellular calcium regulation genes suggests it modulates Wnt signaling via the noncanonical Wnt/calcium pathway, promoting SMC proliferation and cell adhesion, thereby driving atherosclerosis. This is supported by the presence of Wnt signaling modifiers DKK3, FRZB, and FZD1 among GRN39's top key drivers, and by evidence that DKK3 ablation attenuates atherosclerosis in mice.

The study limitations of Study 2 largely are the same as in Study 1 with some notable distinctions given the results. First, the number of isolated T-cells in atherosclerotic lesions was low, despite their likely significant role in advanced

and symptomatic atherosclerosis (Fernandez et al., 2019) that is also evident from macrophage-dominated GRN122 that clearly regulates cellular immunity. Second, our single-cell isolation protocol yielded fewer luminal ECs than microvasculature ECs as endarterectomy-isolated carotid plaques contain few luminal ECs, which may unjustly underestimate the role of luminal EC in carotid stenosis. However, carotid neovascularization has indeed been implicated in plaque rupture (Williams et al., 2020). Notably also among our four murine (three vasa vasorum) and six human (three microvasculature) EC subtypes, none of these 10 subclusters were enriched in cells from advanced murine atherosclerosis or patients with symptomatic carotid plaques. Third and like in Study 1, while using both human and murine atherosclerosis tissue sources, rather than a single model system, our data from the mouse aorta and human carotid plaques may not fully represent CAD and vice versa. However, our combined bulk and single-cell RNA sequencing analyses suggest overlapping atherosclerosis mechanisms across these arterial wall locations and ACVDs. Fourth, again like Study 1, Smart-Seq2, a FACS based technology, may result in cell loss due to antibody-based cell-type selection. Fifth related to our independent data validation in Study 2, while serum-free medium-cultured human primary vascular SMCs have been used to model contractile SMCs, the clinical relevance of in vitro SMC phenotyping remains questionable. Finally, like previous scRNA-seq studies (Alencar et al., 2020; Cochain et al., 2018; Depuydt et al., 2020; Fernandez et al., 2019; Winkels et al., 2018) we relied on a relatively small, albeit larger than earlier studies, number of carotid plaques and murine aortic arches.

7.2.3. STARNET facilitated a comprehensive, systems-based analysis of CAD (Study 3)

Traditional CAD research focuses on isolated risk factors while systemic approach advocates also using RNA-seq data to understand the complex, multi-organ processes driving CAD. Unlike DNA-only studies like GWAS, systems studies integrate DNA and RNA sequencing that capture the interplay of genetic and environmental factors, overcoming the limitations of analyzing isolated genetic variations. In study 3, we used a comprehensive approach that integrates DNA genotype, high-quality RNA-seq, and clinical data from seven tissues of well characterized CAD patients (n = 600) and controls (n = 250) who underwent open-heart surgery as part of the STARNET study. In study 3, we created gene coexpression modules and gene regulatory networks, that can identify key "driver genes" which control these networks and influence disease progression, offering a more comprehensive understanding and potential therapeutic targets. These modules formed interacting GRNs within a larger, interconnected CMD–CAD supernetwork. This supernetwork operated across multiple biological domains, including metabolic organs, blood, and the arterial wall, contributing to CAD progression. In study 3, we further integrated with GWAS datasets that revealed these GRNs account for 54–60% of CAD heritability, substantially surpassing the previously recognized 22%. In study 3, we used the Mendelian randomization

(MR) analysis to further demonstrate that KDRs within 218 of the 224 GRNs exhibit a causal relationship with both CMDs and CAD.

Gene regulatory network is defined as directional coexpression network as it is based on correlations in gene expression, with inferred directionality and addresses how genetic variation affects disease traits. This is different from traditional GRN that implies direct, often causal and regulatory interactions. GRNs anchored in cis eQTL associations directly address how genetic variants affect gene expression as they capture the cascading and integrative effects of multiple genetic loci on gene expression. As GRNs are functional (based on expression correlations), it's logical to define both tissue-specific and cross-tissue GRNs.

Cross-tissue GRNs (networks linking gene expression across multiple tissues) are more important for CAD than tissue-specific GRNs (networks confined to a single tissue). To understand how GRNs interact, eigengene values were calculated to link the 224 GRNs into a supernetwork structure. In study 3, bootstrapping was used to validate the reliability of these interactions. Main findings of Study 3 is that these supernetwork show how all the GRNs work together, the GRN78 → GRN98 interaction was identified as the most reliable edge in the supernetwork that was confirmed by experimental validation. Cross-tissue GRNs, crucial for CAD, are mediated by endocrine signaling, revealing 374 novel factors that connect metabolic tissues and the arterial wall. These factors, potential therapeutic targets, were validated in mice, showing that adipose-liver endocrine signaling impacts metabolic markers. Hence, the supernetwork provides a mechanistic framework for understanding molecular interactions in CMDs and CAD that allows for the study of individual candidate genes within a broader molecular context.

Systems studies inherently generate more questions than answers. The limitations of Study 3 are that we observed a striking upregulation of noncoding genes in atherosclerotic arterial walls and a reciprocal downregulation in metabolic tissues of CAD patients. Notably, the number of altered noncoding genes significantly exceeded that of coding genes in these tissues, underscoring their crucial but underexplored role in CMD–CAD pathobiology. Another unexpected finding was the general suppression of hepatic transcription, indicating potential suboptimal liver function in CAD. Furthermore, while we accounted for potential confounders, including medication effects from tissue samples taken during severe CAD surgery, refining GRNs will require additional molecular, anatomical, and functional data. This should encompass a spectrum of human ethnicities, with temporal resolution at tissue, cell type, and single-cell levels, across diverse CMD–CAD model systems and ethnicities.

7.3. Future Directions

In summary this thesis has established a foundational understanding of sex-biased atherosclerosis by identifying distinct, sex-specific subcellular changes and clusters that drive atherosclerosis. It also provides a multi-organ framework for exploring the molecular interplay between CMDs and CAD. Using Smart-seq2 technology, we generated the deepest multi-species scRNA-seq dataset of atherosclerosis to date. This work revealed significant biological diversity within main cell types, showing how they transition from healthy to diseased states. By integrating our findings with human GRNs, we provided critical clinical context and a framework for the pathophysiological mechanisms underlying these sex-biased cell subtypes and those that are critical for advanced and symptomatic atherosclerosis, paving the way for targeted research.

1) **Integrating Proteomics for a Multi-Omics Perspective**: Our future work will extend beyond transcriptomics to include a proteomics layer. By investigating the plasma proteome and tissue-specific protein quantitative trait loci (pQTLs) within the STARNET cohort, we will bridge the gap between gene expression and the functional proteins that drive the disease. This multi-omics integration will provide a more comprehensive view of the cellular and molecular networks at play
2) **Incorporating Spatial Transcriptomics**: While scRNA-seq provides unparalleled resolution of individual cell types, it lacks the critical spatial context of how these cells are organized and interact within the tissue. To address this, we will incorporate spatial transcriptomics. This technology will allow us to create a "molecular map" of the tissue, showing the precise location of gene expression and revealing how cells function and communicate in their native environment. This spatial information is vital for validating our scRNA-seq findings and for discovering novel cellular interactions that are key to atherosclerotic progression.
3) **Focusing on EndMT and Plaque Erosion in Women**: Among our numerous findings, a particularly compelling discovery was the sex-biased distribution of EC subtypes. We observed that ECs involved in angiogenesis were predominantly found in males, whereas ECs undergoing EndMT were significantly more prevalent in females. We will focus our future research on EndMT, which is increasingly recognized as a key contributor to plaque erosion. This form of plaque instability is a significant cause of heart attacks, especially in younger women. By elucidating the mechanisms by which EndMT drives plaque erosion, we aim to uncover novel, sex-specific therapeutic targets to prevent acute cardiovascular events in women.

8. CONCLUSIONS

Study 1: Understanding the mechanisms driving sex as a biological variable in atherosclerotic disease is critical for mitigating life-threatening symptoms. To this end, we performed single-cell RNA sequencing on both male and female patients undergoing carotid endarterectomy. While major cell types showed no sex-biased differences, their cellular subtypes were sex-biased. Specifically, in females, we observed an osteogenic phenotype in smooth muscle cells, immunomodulating macrophages, and endothelial cells undergoing EndoMT/EndICLT. In males, we found SMCs with a chondrocyte phenotype, MPs involved in tissue remodeling, and ECs with angiogenic activity. In Study 1 we found that the ECs originated from two vascular beds: the vasa vasorum within the carotid plaques and the carotid artery lumen. To determine key drivers, clinical contexts, heritability contributions, and pathophysiological mechanisms associated with the fifteen identified sex-biased subcellular clusters in human carotid plaques we integrated with 135 tissue-specific gene regulatory networks from the human cardiometabolic-focused, multi-omics STARNET dataset. We then robustly confirmed the cellular origin and clinical implications of GRNs 33, 122, and 195 across multiple datasets. Functional validation of GRN 195 suggested potential male-specific therapeutic targets.

Study 2: In this study we investigated cellular plasticity in atherosclerosis progression using $Ldlr^{-/-}Apob^{100/100}$ mice and human scRNA sequencing data. We reidentified and expanded the gene content of key SMC and macrophage MP subtypes associated with advanced and symptomatic plaques. By integrating with STARNET GRNs, we highlighted three critical arterial wall GRNs: MP-specific GRN33 and GRN122, and SMC-specific GRN39. GRN39, validated across multiple datasets and experiments, was shown to drive the transformation of contractile SMCs into an osteogenic phenotype, promoting advanced and symptomatic plaques. All scRNA-seq data are fully available at https://sicof.medh.ki.se/athero.

Study 3: In this study we analyzed genetic, clinical, and RNA-seq data from 600 CAD and 250 CAD- free patients across seven tissues, identifying 224 GRNs that explain over 54% of CAD heritability. Notably, 89 cross-tissue GRNs linked to CAD severity revealed a key adipose-to-liver endocrine axis, involving 374 endocrine factors, validated in mice where recombinant adipose factors significantly altered lipid and glucose levels. STARNET study, the largest multi-tissue transcriptional analysis of CMDs and CAD to date, provides a comprehensive view of gene-regulatory interplay, crucial for understanding these complex disorders. This novel pathophysiological data will aid in interpreting GWAS findings, which often lack multi-organ context, and ultimately accelerate the development of early, personalized diagnostics and therapeutics in precision medicine. The STARNET database and GRN browser http://starnet.mssm.edu offer a multi-

organ framework for understanding the molecular interplay between cardio-metabolic disorders and CAD.

Declaration of interests
All the authors declare that there is no conflict of interest.

9. SUMMARY IN ESTONIAN

Inimese geeniregulatsioonivõrgustikega integreeritud üheraku RNA järjestuse analüüs annab mehhanistlikke teadmisi ateroskleroosi kohta nii meestel kui naistel

Sisukirjeldus

Südame-veresoonkonna haigused (CVD) on ülemaailmne tervisekriis, mis nõuab igal aastal miljoneid inimelusid, millest kolmveerand leiab aset madala ja keskmise sissetulekuga riikides. Maailma Terviseorganisatsiooni (WHO) 2019. aasta andmetel suri südame-veresoonkonna haigustesse 17,9 miljonit inimest, mis moodustab 32% kõigist ülemaailmsetest surmajuhtumitest. Neist surmajuhtumitest 85% olid tingitud südamelihaseinfarktist ja insuldist. 2019. aasta globaalse haiguskoormuse (GBD) uuring näitas, et levinud südame-veresoonkonna haiguste arv peaaegu kahekordistus 271 miljonilt 1990. aastal 523 miljonile 2019. aastal. Südame- ja veresoonkonnahaigused hõlmavad mitmesuguseid häireid, mis mõjutavad südant või veresooni. Veresooni haaravat CVD tüüpi nimetatakse vaskulaarhaiguseks, ateroskleroos (arterite kõvenemine) on koronaararterite haiguse (CAD), ajuveresoonkonna haiguste (CVD) ja alajäseme arterite haiguse (PAD) aluseks. Vaskulaarsed haigused on tavalised keerulised degeneratiivsed haigused, mis arenevad aja jooksul, muutes vaskulaarset homeostaasi ja põhjustades aterosklerootilise naastu moodustumist arteri seinas. Ateroskleroosi arengu kiirus sõltub nii keskkonna survest kui ka inimese geneetilisest eelsoodumusest. Aastate jooksul muudavad need tegurid geeniekspressiooni koronaararterite ateroskleroosi jaoks olulistes organites, mis peegelduvad vereringes nt. plasma lipiidide tasemes. Kandidaatgeeni otsimine, mis keskendub eelmääratletud geenide geneetiliste variatsioonide seostele, on efektiivne pärilike ühe geeni häirete korral, samas kui nende roll keerukate mitme geeni haiguste jaoks on ebapiisav.

Komplekshaigusi on nende keeruka olemuse tõttu raske uurida, kuna neil ei ole lihtsaid Mendeli pärilikkuse mustreid. Nende arengule aitavad kaasa mitmed tegurid, sageli nii geneetilised kui ka keskkonnast tulenevad, mistõttu on raske ühe teguri mõju eristada. Geneetika ja tehnoloogia edusammud viisid genoomi hõlmavate assotsiatsiooniuuringuteni (GWAS), mis põhinevad piiratud arvul geneetilistel variantidel, mille esinemissagedus (tavaliselt üle 5%) üldpopulatsioonis aitab kaasa haiguste vastuvõtlikkusele. Ühe nukleotiidi polümorfismi (SNP) uurimine on laialdaselt kasutatav meetod SNP-de genotüpiseerimiseks, et tuvastada haigusega seotud geneetilisi variante. Samas on sellel meetodil piirangud, nagu kindlakstegemise eelarvamus, väiksem eraldusvõime ja ainult genotüübiga eelnevalt määratletud SNP-d. Selle ületamiseks oli kogu genoomi järjestamisest (WGS) saanud GWAS-i juhtiv tehnoloogia, kus saab sekveneerida kogu organismi DNA-sisaldust. Alates 2007. aastast on GWAS-i uuringud CAD ja müokardiinfarkti (MI) kohta andnud olulisi teadmisi, nagu kromosoomi 9p21 riskilookuse tuvastamine mitmete uurimisrühmade poolt. Praeguseks on tuvastatud üle 241 sõltumatu geneetilise variandi 198 CAD-haigusega seotud lookuses.

Eeldatakse, et GWAS-i tuvastatud CAD-i riskivariandid asuvad haiguse arengu aeglases või varases faasis, samas kui viimast etappi mõjutavad variandid jäetakse vahele, kuna need sõltuvad keskkonnamõjudest ja põletikulistest seisunditest, mida GWAS ei kontrolli.

CAD-l on märkimisväärne geneetiline komponent, perekonnauuringud hindavad pärilikkust 40–60%. Kuigi GWAS on tuvastanud arvukalt CAD-ga seotud SNP-sid, selgitavad need SNP-d vaid väikest osa üldisest pärilikkusest. See "puuduv pärilikkus" viitab sellele, et muud geneetilised tegurid, nagu haruldased variandid, struktuursed variandid või geeni-geeni interaktsioonid, võivad mängida rolli CAD riskis. Geeniekspressiooni geneetilised uuringud, mida sageli nimetatakse ekspressioonikvantitatiivsete tunnuste lookuste (eQTL) uuringuteks, laiendavad GWAS-i, lisades vahekihina RNA (geeni) ekspressiooni. Transkriptoom on rakus või organismis toodetud RNA molekulide täielik komplekt. RNA-sekveneerimine (RNA-Seq) on võimas tehnoloogia, mis on muutnud transkriptoomianalüüsi, ületades mikrokiibi tehnoloogia, mis võimaldab ühe baasi eraldusvõimet. Kombineerides RNA-Seq genotüpiseerimisandmetega, saame tuvastada SNP-sid, mis moduleerivad transkriptsioonitasemeid, mida nimetatakse eQTL-ideks või eSNP-deks ja mida saab seostada geeniekspressiooni tasemetega kas lokaalselt (cis) või eemal (trans, st rohkem kui 5 Mb kaugusel seotud variandist). See andmepõhine lähenemine ei tuvasta mitte ainult haiguse geenipõhist teket, vaid lisaks ka CAD-oluliste signaalirajadu geeniregulatsioonivõrgustike näol. Geeniregulatsioonivõrgud (GRN-id) geenide interaktsioonide suundumust nendes GRN-ides järeldatakse Bayesi võrgualgoritmi abil, kus teatud tugevalt ühendatud geenid, mida nimetatakse võtmejuhtgeenideks (KDR-id), pakuvad väärtuslikku mehhanismi keeruka haigusbioloogia dešifreerimiseks.

Meie CAD-uuringute käigus töötati välja biopank STARNET (Stockholm Tartu Atherosclerosis Reverse Network Engineering Task), mis on suurim vere, ainevahetusorganite ja arterite seinte transkriptoomiline uuring. Juhtudeks olid patsiendid, kellele tehti koronaararterite šunteerimine, ja kontrollideks olid patsiendid, kellel ei olnud ateroskleroosi ega CAD-d (kinnitatud operatsioonieelse angiograafiaga), kellele tehti muid südameoperatsioone (nt isoleeritud aordiklapi parandamine). Tuvastati 224 geeniregulatoorset koekspressioonivõrku (GRN), integreerides koronaarhaigusega (n = 600) ja ilma (n = 250) patsientide geneetilised ja kliinilised andmed seitsme haigusega seotud koe RNA-järjestuste andmetega, nimelt aterosklerootiliste kahjustusteta sisemise rinnakasti arteri (MAM), aterosklerooosiga arteri seinaga (AOR), nahaalune rasvkude (SF), vistseraalne kõhurasv (VAF), skeletilihas (SKLM) ja maks (LIV). 224-st GRN-ist, mis selgitavad >54% CAD-i pärilikkusest, 89 on kudedevahelised GRN-id, mis on seotud CAD-i raskusastmega. Mass RNA-Seq annab keskmise geeniekspressiooni profiili, jättes tähelepanuta rakkudevahelise varieeruvuse. Üherakuline RNA-Seq (scRNA-Seq) ületab selle piirangu, analüüsides geeniekspressiooni üksiku raku tasemel. See võimaldab tuvastada haruldasi rakupopulatsioone, iseloomustada rakutüüpe ja terviklikumalt mõista rakkude mitmekesisust koes. On olemas erinevad scRNA-Seq meetodid, millest igaühel on ainulaadsed kriteeriumid rakkude püüdmiseks ja sekveneerimise sügavuseks. Tavalised etapid hõlmavad rakkude

eraldamist, ühe raku püüdmist, RNA amplifikatsiooni, „raamatukogu“ ettevalmistamist, sekveneerimist ja andmete analüüsi. Peamised scRNA-Seq tehnoloogiad hõlmavad „plate-based“, „microfluidic-based“ ja „droplet-based“ meetodeid. Switch Mechanism at the 5' End of RNA Template Sequencing (SMART-seq2) on plate-based meetod, mis tagab parema tundlikkuse, täpsuse ja transkriptide täispika katvuse. Praegused piirangud on ahela spetsiifilisuse puudumine ja võimetus tuvastada polüadenüülimata (polüA-) RNA-d. Üherakuliste geenieksпressioonimustrite mõistmine vaskulaarsetes rakkudes, nimelt endoteeli-, silelihas- ja makrofaagides ateroskleroosi progresseerumise ajal, sealhulgas soopõhised erinevused, annab olulise ülevaate haiguse mehhanismidest ja võimalikest terapeutilistest sihtmärkidest.

Eesmärk ja ülesanded

Käesoleva töö üldeesmärk on paremini mõista ateroskleroosi protsessi erinevates vaskulaarsetes piirkondades ning täpsemalt peamiste rakutüüpide dediferentseerumist ja transdiferentseerumist ning soolisi erinevusi naastude subtsellulaarses koostises.

1) Uurida peamiste rakutüüpide subtsellulaarseid koostisi, millel on soolised erinevused, ning pakkuda kliinilist ja patofüsioloogilist konteksti, integreerides STARNETi GRN-idega (artikkel 1).
2) Mõista ateroskleroosi progresseerumise ajal vaskulaarsete rakutüüpide transformatsiooni aluseks olevaid üheraku geenieksпressiooni mustreid (artikkel 2).
3) Selgitada geneetiliste ja mittegeneetiliste riskitegurite delikaatset koosmõju, integreerides seitsme metaboolse koe geneetilised ja kliinilised andmed ning RNA-sekveneerimise andmed GRN-ide kujul (artikkel 3).

Materjal ja meetodid

Kaasatud olid Tartu Ülikooli Kliinikumi patsiendid, kellel oli näidustatud unearteri endarterektoomia operatsioon veresoontekirurgia osakonnas ning koronaararterite šunteerimisoperatsiooni või muu südameoperatsioon kardiokirurgia osakonnas.

Uuringud 1 ja 2 (unearteri endarterektoomia): kaasati sisemiste unearterite proksimaalosa olulise stenoosiga (>70%) patsiendid, kellel oli näidustatud endarterektoomia. Patsiendid liigitati ajuisheemia kliiniliste tunnuste olemasolu põhjal asümptomaatiliseks või sümptomaatiliseks. Oluline on see, et tervete unearteritega patsientidel proove ei võetud, kuna operatsioon teostatakse ainult ahenenud unearteritel ja eemaldatakse kogu aterosklerootiline naast. 7 naispatsiendilt ja 8 meespatsiendilt (uuring 1) koguti kokku 15 unearteri naastu, mis hõlmasid 7 asümptomaatilist ja 8 sümptomaatilist isikut (uuring 2), kellele tehti unearteri endarterektoomia. Need naastud lõigati ensümaatiliselt, kasutades I/II tüüpi kollagenaasi ja elastaasi, et saada üherakuline suspensioon. Seejärel see suspensioon filtreeriti, tsentrifuugiti ja resuspendeeriti edasiseks analüüsiks FACS puhvris.

Uuring 3 (koronaararterite šunteerimise / klapi proteesimine): uuringusse on kaasatud isikud, kellel on koronaararterite ateroskleroos (CAD), mis on sobilik koronaararterite šunteerimise operatsiooniks, ja isikud, kes on sobilikud muudele südameoperatsioonidele (nt klapivahetus) ilma obstruktiivse CAD-ita (kinnitatud operatsioonieelse angiogrammiga). Täisveri (EDTA katseklaasides) ja seitse erinevat koebiopsiat koguti 600 CAD-ga isikult ja 250 kontrollisikult, kellele tehti südameoperatsioon. Biopsiad hõlmasid aordi seina (AOR), seesmist rinnakasti arterit (MAM – ainult CAD patsientidel), vistseraalset kõhurasva (VAF), nahaalust rasva (SF), skeletilihast (SKLM) ja maksa (LIV) proovi. Kõiki biopsiaid säilitati RNAlateris temperatuuril -80 °C kuni RNA ekstraheerimiseni.

Üldised kaasamise/välistamise kriteeriumid: kõik osalejad täitsid küsimustiku oma haigusloo, ravimite ja elustiili kohta. Raskete süsteemsete haigustega, nagu aktiivsed põletikulised seisundid või vähk, patsiendid jäeti välja. Kõigilt osalejatelt saadi teadlik kirjalik nõusolek.

Ldlr-/-Apob100/100 hiired, mudel, millel on inimesele sarnane kõrge LDL-kolesterooli tase ja spontaanne ateroskleroos, keda hoiti standardtingimustes. 46 hiirt jagati algtaseme (mitteateroskleroOtilised, n = 8), varajase (n = 18) ja kaugelearenenud (n = 20) ateroskleroosi rühmadesse. Aordikaared koguti igas etapis vanuse ja sooga sobitatud paaridest, perfuseeriti PBS-iga ja töödeldi endoteelirakkude isoleerimiseks kaheastmelise ensümaatilise seedimisega 10 mg/ml I, II ja IV tüüpi kollagenaasiga ning 2 mg/ml elastaasiga. Saadud rakususpensioonid filtreeriti, tsentrifuugiti ja resuspendeeriti FACS puhvris edasiseks analüüsiks. Inimese ja hiire üherakuliste suspensioonide spetsiifilisi rakupopulatsioone (SMC-d, EC-d, immuunrakud) rikastati voolutsütomeetria jaoks, kasutades fluorofooriga konjugeeritud antikehi (PDGFRβ SMC-de jaoks, CD31/CD144 või CD31 EC-de jaoks ja CD45 AM-i roheliste immuunrakkude jaoks) ja Calceini roheliste rakkude jaoks. Seejärel lahjendati märgistatud rakud ja jahutati enne FACS analüüsi.

Geeniregulatsioonivõrgustike (GRN-ide) – GRN 33, GRN 39, GRN122 ja GRN195 – valideerimine, kasutades mitmeid sõltumatuid inimandmekogumeid, nagu isheemilised ja mitteisheemilised inimese koronaararterite koed nõusoleku saanud südamesiirdamise doonoritelt ja normaalsed koronaararterite bioproovid mittesobivatelt doonor südametelt Stanfordi Ülikooli haiglas, GTEx koronaararterite, GTEx aordi sein, STARNET primaarsed monotsüüdid (MP), NGS-PREDICT primaarsed MP, inimese aordi endoteelirakud (HAEC) meesdoonorilt ja üheraku sekveneerimisandmed. Võrgustiku säilimise hindamine: NetRep R paketti kasutati kolme GRN-i säilimise hindamiseks mitme välise andmestiku vahel.

Athero-Expressi uuringu valideerimine: Athero-Expressi (AE) uuringus 20 naise ja 26 mehe unearteri naastude üherakuliste RNA-seq andmete DEG-sid kasutati rikastatud GRN-ide valideerimiseks konkreetsetes rakutüüpides. AE uuring hõlmab suurt hulka patsiente, kellele tehakse unearteri endarterektoomia ja kellel on ulatuslikud kliinilised ja multi-omika andmed.

GRN195 funktsionaalne valideerimine in vitro katsetega, kasutades inimese aordi endoteelirakke (HAEC): TeloHAEC ja primaarsed HAEC-d transdutseeriti

lentiviiruse vektoritega, et üleekspresseerida PLVAP-i või FAM110D, kahte geeni GRN195 sees või mittesihtivat kontrolli (CTL). Rakke kultiveeriti spetsiifilises endoteelirakkude kasvusöötmes koos antibiootikumidega ja selekteeriti blastitsidiini abil. Viidi läbi geeniekspressiooni analüüs, proliferatsioonianalüüs ja rakkude värvimise analüüs. GRN39 funktsionaalset valideerimist in vitro katsetega, kasutades primaarseid vaskulaarseid SMC-sid, mis olid eraldatud tervetelt südame siirdamise doonoritelt (UCLA) ja ostetud ettevõttelt Lonza / PromoCell, kultiveeriti ja iseloomustati migratsiooni, proliferatsiooni ja lupjumise suhtes. Need rakud allutati RNA sekveneerimisele vaiksetes ja proliferatiivsetes tingimustes.

Immunofluorestsents: Hiire aordid fikseeriti, külmutati sahharoosi/PBS-is, sisestati krüo-söötmesse ja lõigati 14 μm. Sektsioonid kuivatati, pesti ja blokeeriti enne üleöö inkubeerimist primaarsete antikehadega (anti-alfa-SMA-Cy3, anti-CD31, anti-CD68) ja järgnevat inkubeerimist sekundaarsete antikehadega (Alexa Fluor-488 ja Alexa Fluor-647). Paigaldatud sektsioonid pildistati Leica TCS SP8 konfokaalse mikroskoobiga ja pilte töödeldi ImageJ / FIJI-ga, tavaliselt maksimaalse intensiivsusega projektsioonidena.

GRN78 endokriinsete teguritega hiirte süstimine mõjutab GRN98 (uuring 3): Rasva ja maksa endokriinsete suhete uurimiseks korreleeriti STARNETi koekspressioonimoodulite omageenid siht maksa moodulitega. Seda analüüsi laiendati HMDP hiirtele, kasutades kaardistatud inimese geenihomolooge ja rasva / maksa mikrokiibi andmeid erinevate dieetide lõikes, apoE-Leideni rasvandmetega, mis saadi RNA-seq abil. Rekombinantsed endokriinsed valgud (FCN2, LBP, FSTL3, EPDR1) toodeti HEK293 rakkudes, puhastati ja süstiti isastele C57BL/6J hiirtele iga päev nelja päeva jooksul. Seejärel hiired surmati ning maksa RNA ekstraheeriti ja sekveneeriti. Diferentsiaalne ekspressioonianalüüs viidi läbi, kasutades DESeq2, et hinnata nende endokriinsete tegurite mõju maksa geeniekspressioonile. Valitud valguefektid valideeriti kaubanduslikult ostetud valkude abil.

Peamised tulemused ja järeldused

Uuring 1. Soo kui aterosklerootilise haiguse bioloogilise muutuja mõju mõistmine on eluohtlike sümptomite leevendamiseks ülioluline. Sel eesmärgil teostasime üherakulise RNA sekveneerimise nii mees- kui ka naissoost patsientidel, kellele tehti unearteri endarterektoomia. Kuigi peamised rakutüübid ei näidanud soopõhiseid erinevusi siis nende rakkude alatüübid näitasid. Täpsemalt, naistel täheldasime osteogeenset fenotüüpi silelihasrakkudes, immunomoduleerivates makrofaagides ja endoteelirakkudes, läbides endoteeli-mesenhümaalset üleminekut. Meestel leidsime kondrotsüütide fenotüübiga SMC-sid, kudede ümberkujundamises osalevaid MP-sid ja angiogeense aktiivsusega EC-sid. EC-d pärinesid kahest vaskulaarsest lokalisatsioonist: vasa vasorum unearteri naastudes ja unearteri luumenist. Integreerisime oma viisteist soopõhist alamklastrit 135 koespetsiifilise inimese kardiometaboolse fookusega geeniregulatsioonivõrguga, multi-oomika STARNETi andmekogust, et määrata kindlaks peamised tegurid,

kliinilised kontekstid, pärilikkuse panused ja patofüsioloogilised mehhanismid, mis on seotud tuvastatud soopõhiste subtsellulaarsete klastritega inimese unearteri naastudes. Kinnitasime kindlalt GRN-ide 33, 122 ja 195 rakulist päritolu ja kliinilisi tagajärgi mitmes andmekogumis. GRN 195 funktsionaalne valideerimine näitas potentsiaalseid meesspetsiifilisi terapeutilisi sihtmärke.

Uuring 2: selles uuringus uuriti rakulist plastilisust ateroskleroosi progresseerumisel, kasutades Ldlr−/−Apob100/100 hiiri ja inimese RNA sekveneerimise andmeid. Tuvastasime ja laiendasime geenide osalust suurte ja sümptomaatiliste karotiidarterite naastudega seotud peamiste SMC ja MP alatüüpides. STARNETi GRN-idega integreerides tõstsime esile kolm kriitilist arteriseina GRN-i: MP-spetsiifiline GRN33 ja GRN122 ning SMC-spetsiifiline GRN39. Näidati, et GRN39, mis on valideeritud mitmes andmekogumis ja katses, soodustas kontraktiilsete SMC-de muutumist osteogeenseks fenotüübiks, soodustades naastude kasvamist ja sümptomite teket. Kõik scRNAseqi andmed on täielikult saadaval aadressil https://sicof.medh.ki.se/athero.

Uuring 3. Selles uuringus analüüsiti 600 CAD-i ja 250 CAD-vaba patsiendi geneetilisi, kliinilisi ja RNA-seq andmeid seitsmes koes, tuvastades 224 GRN-i, mis selgitavad üle 54% CAD-i pärilikkusest. Nimelt näitasid 89 CAD raskusastmega seotud ristkudede GRN-i peamist rasva ja maksa vahelist endokriinset telge, mis hõlmas 374 endokriinset faktorit, mis kinnitati hiirtel, kus rekombinantsed rasvfaktorid muutsid oluliselt lipiidide ja glükoosi taset. STARNET andmebaas ja GRN-brauser http://starnet.mssm.edu pakuvad mitme organi raamistikku kardiometaboolsete häirete ja CAD-i molekulaarse koosmõju mõistmiseks. Kuna meie uuring on seni suurim CMD-de ja CAD-i mitmekoeline transkriptsioonianalüüs, annab meie uuring tervikliku ülevaate geeniregulatsiooni koosmõjust, mis on nende keeruliste häirete mõistmiseks ülioluline. Need uudsed patofüsioloogilised andmed aitavad tõlgendada GWAS-i leide, millel sageli puudub mitme elundi kontekst, ning kiirendavad lõppkokkuvõttes täppismeditsiinis varajase, isikupärastatud diagnostika ja ravi väljatöötamist.

10. REFERENCES

Albanese, I., Yu, B., Al-Kindi, H., Barratt, B., Ott, L., Al-Refai, M., de Varennes, B., Shum-Tim, D., Cerruti, M., Gourgas, O., Rheaume, E., Tardif, J. C., & Schwertani, A. (2017). Role of Noncanonical Wnt Signaling Pathway in Human Aortic Valve Calcification. *Arterioscler Thromb Vasc Biol*, *37*(3), 543–552. https://doi.org/10.1161/ATVBAHA.116.308394

Alencar, G. F., Owsiany, K. M., Karnewar, S., Sukhavasi, K., Mocci, G., Nguyen, A. T., Williams, C. M., Shamsuzzaman, S., Mokry, M., Henderson, C. A., Haskins, R., Baylis, R. A., Finn, A. V., McNamara, C. A., Zunder, E. R., Venkata, V., Pasterkamp, G., Bjorkegren, J., Bekiranov, S., & Owens, G. K. (2020). Stem Cell Pluripotency Genes Klf4 and Oct4 Regulate Complex SMC Phenotypic Changes Critical in Late-Stage Atherosclerotic Lesion Pathogenesis. *Circulation*, *142*(21), 2045–2059. https://doi.org/10.1161/CIRCULATIONAHA.120.046672

Algabri, Y. A., Li, L., & Liu, Z. P. (2022). scGENA: A Single-Cell Gene Coexpression Network Analysis Framework for Clustering Cell Types and Revealing Biological Mechanisms. *Bioengineering (Basel)*, *9*(8). https://doi.org/10.3390/bioengineering9080353

Aragam, K. G., Jiang, T., Goel, A., Kanoni, S., Wolford, B. N., Atri, D. S., Weeks, E. M., Wang, M., Hindy, G., Zhou, W., Grace, C., Roselli, C., Marston, N. A., Kamanu, F. K., Surakka, I., Venegas, L. M., Sherliker, P., Koyama, S., Ishigaki, K., . . . Consortium, C. A. D. (2022). Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nat Genet*, *54*(12), 1803–1815. https://doi.org/10.1038/s41588-022-01233-6

Astudillo, P., & Larraín, J. (2014). Wnt signaling and cell-matrix adhesion. *Curr Mol Med*, *14*(2), 209–220. https://doi.org/10.2174/1566524014666140128105352

Baaten, C., Nagy, M., Bergmeier, W., Spronk, H. M. H., & van der Meijden, P. E. J. (2024). Platelet biology and function: plaque erosion vs. rupture. *Eur Heart J*, *45*(1), 18–31. https://doi.org/10.1093/eurheartj/ehad720

Bagnoli, J. W., Ziegenhain, C., Janjic, A., Wange, L. E., Vieth, B., Parekh, S., Geuder, J., Hellmann, I., & Enard, W. (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nat Commun*, *9*(1), 2937. https://doi.org/10.1038/s41467-018-05347-6

Barabási, A.-L., Gulbahce, N., & Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. *Nature Reviews Genetics*, *12*(1), 56–68. https://doi.org/10.1038/nrg2918

Becht, E., McInnes, L., Healy, J., Dutertre, C. A., Kwok, I. W. H., Ng, L. G., Ginhoux, F., & Newell, E. W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. https://doi.org/10.1038/nbt.4314

Bjorkegren, J. L., Hagg, S., Talukdar, H. A., Foroughi Asl, H., Jain, R. K., Cedergren, C., Shang, M. M., Rossignoli, A., Takolander, R., Melander, O., Hamsten, A., Michoel, T., & Skogsberg, J. (2014). Plasma cholesterol-induced lesion networks activated before regression of early, mature, and advanced atherosclerosis. *PLoS Genet*, *10*(2), e1004201. https://doi.org/10.1371/journal.pgen.1004201

Björkegren, J. L. M., Kovacic, J. C., Dudley, J. T., & Schadt, E. E. (2015). Genome-wide significant loci: how important are they? Systems genetics to understand heritability of coronary artery disease and other common complex disorders. *J Am Coll Cardiol*, *65*(8), 830–845. https://doi.org/10.1016/j.jacc.2014.12.033

Bjorkegren, J. L. M., & Lusis, A. J. (2022). Atherosclerosis: Recent developments. *Cell*. https://doi.org/10.1016/j.cell.2022.04.004

Blanco-Gomez, A., Castillo-Lluva, S., Del Mar Saez-Freire, M., Hontecillas-Prieto, L., Mao, J. H., Castellanos-Martin, A., & Perez-Losada, J. (2016). Missing heritability of complex diseases: Enlightenment by genetic variants from intermediate phenotypes. *Bioessays*, *38*(7), 664–673. https://doi.org/10.1002/bies.201600084

Bodmer, W., & Bonilla, C. (2008). Common and rare variants in multifactorial susceptibility to common diseases. *Nature Genetics*, *40*(6), 695–701. https://doi.org/10.1038/ng.f.136

Botstein, D., & Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet*, *33 Suppl*, 228–237. https://doi.org/10.1038/ng1090

Boucher, P., Matz, R. L., & Terrand, J. (2020). atherosclerosis: gone with the Wnt? *Atherosclerosis*, *301*, 15–22. https://doi.org/10.1016/j.atherosclerosis.2020.03.024

Brembeck, F. H., Rosario, M., & Birchmeier, W. (2006). Balancing cell adhesion and Wnt signaling, the key role of beta-catenin. *Curr Opin Genet Dev*, *16*(1), 51–59. https://doi.org/10.1016/j.gde.2005.12.007

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*, *36*(5), 411–420. https://doi.org/10.1038/nbt.4096

Chen, W., Wang, S., Tithi, S. S., Ellison, D. W., Schaid, D. J., & Wu, G. (2022). A rare variant analysis framework using public genotype summary counts to prioritize disease-predisposition genes. *Nat Commun*, *13*(1), 2592. https://doi.org/10.1038/s41467-022-30248-0

Cochain, C., Vafadarnejad, E., Arampatzi, P., Pelisek, J., Winkels, H., Ley, K., Wolf, D., Saliba, A. E., & Zernecke, A. (2018). Single-Cell RNA-Seq Reveals the Transcriptional Landscape and Heterogeneity of Aortic Macrophages in Murine Atherosclerosis. *Circ Res*, *122*(12), 1661–1674. https://doi.org/10.1161/CIRCRESAHA.117.312509

Consortium, T. G., Ardlie, K. G., Deluca, D. S., Segrè, A. V., Sullivan, T. J., Young, T. R., Gelfand, E. T., Trowbridge, C. A., Maller, J. B., Tukiainen, T., Lek, M., Ward, L. D., Kheradpour, P., Iriarte, B., Meng, Y., Palmer, C. D., Esko, T., Winckler, W., Hirschhorn, J. N., . . . Dermitzakis, E. T. (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*, *348*(6235), 648–660. https://doi.org/doi:10.1126/science.1262110

Dalager, S., Paaske, W. P., Kristensen, I. B., Laurberg, J. M., & Falk, E. (2007). Artery-related differences in atherosclerosis expression: implications for atherogenesis and dynamics in intima-media thickness. *Stroke*, *38*(10), 2698–2705. https://doi.org/10.1161/strokeaha.107.486480

Demer, L. L., & Tintut, Y. (2008). Vascular calcification: pathobiology of a multifaceted disease. *Circulation*, *117*(22), 2938–2948. https://doi.org/10.1161/CIRCULATIONAHA.107.743161

Depuydt, M. A. C., Prange, K. H. M., Slenders, L., Ord, T., Elbersen, D., Boltjes, A., de Jager, S. C. A., Asselbergs, F. W., de Borst, G. J., Aavik, E., Lonnberg, T., Lutgens, E., Glass, C. K., den Ruijter, H. M., Kaikkonen, M. U., Bot, I., Slutter, B., van der Laan, S. W., Yla-Herttuala, S., . . . Pasterkamp, G. (2020). Microanatomy of the Human Atherosclerotic Plaque by Single-Cell Transcriptomics. *Circ Res*, *127*(11), 1437–1455. https://doi.org/10.1161/CIRCRESAHA.120.316770

Ding, J., Adiconis, X., Simmons, S. K., Kowalczyk, M. S., Hession, C. C., Marjanovic, N. D., Hughes, T. K., Wadsworth, M. H., Burks, T., Nguyen, L. T., Kwon, J. Y. H., Barak, B., Ge, W., Kedaigle, A. J., Carroll, S., Li, S., Hacohen, N., Rozenblatt-Rosen, O., Shalek, A. K., . . . Levin, J. Z. (2020). Author Correction: Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nat Biotechnol*, *38*(6), 756. https://doi.org/10.1038/s41587-020-0534-z

Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*, *5*(6), 435-445. https://doi.org/10.1038/nrg1348

Fernandez, D. M., Rahman, A. H., Fernandez, N. F., Chudnovskiy, A., Amir, E. D., Amadori, L., Khan, N. S., Wong, C. K., Shamailova, R., Hill, C. A., Wang, Z., Remark, R., Li, J. R., Pina, C., Faries, C., Awad, A. J., Moss, N., Bjorkegren, J. L. M., Kim-Schulze, S., . . . Giannarelli, C. (2019). Single-cell immune landscape of human atherosclerotic plaques. *Nat Med*, *25*(10), 1576-1588. https://doi.org/10.1038/s41591-019-0590-4

Förstermann, U., & Sessa, W. C. (2011). Nitric oxide synthases: regulation and function. *European Heart Journal*, *33*(7), 829–837. https://doi.org/10.1093/eurheartj/ehr304

Fox, C. S., Polak, J. F., Chazaro, I., Cupples, A., Wolf, P. A., D'Agostino, R. A., O'Donnell, C. J., & Framingham Heart, S. (2003). Genetic and environmental contributions to atherosclerosis phenotypes in men and women: heritability of carotid intima-media thickness in the Framingham Heart Study. *Stroke*, *34*(2), 397–401. https://doi.org/10.1161/01.str.0000048214.56981.6f

Franzen, O., Ermel, R., Cohain, A., Akers, N. K., Di Narzo, A., Talukdar, H. A., Foroughi-Asl, H., Giambartolomei, C., Fullard, J. F., Sukhavasi, K., Koks, S., Gan, L. M., Giannarelli, C., Kovacic, J. C., Betsholtz, C., Losic, B., Michoel, T., Hao, K., Roussos, P., . . . Bjorkegren, J. L. (2016). Cardiometabolic risk loci share downstream cis- and trans-gene regulation across tissues and diseases. *Science*, *353*(6301), 827–830. https://doi.org/10.1126/science.aad6970

Freund, M. K., Burch, K. S., Shi, H., Mancuso, N., Kichaev, G., Garske, K. M., Pan, D. Z., Miao, Z., Mohlke, K. L., Laakso, M., Pajukanta, P., Pasaniuc, B., & Arboleda, V. A. (2018). Phenotype-Specific Enrichment of Mendelian Disorder Genes near GWAS Regions across 62 Complex Traits. *Am J Hum Genet*, *103*(4), 535–552. https://doi.org/10.1016/j.ajhg.2018.08.017

Furie, B., & Furie, B. C. (2008). Mechanisms of Thrombus Formation. *New England Journal of Medicine*, *359*(9), 938–949. https://doi.org/doi:10.1056/NEJMra0801082

George, S. J. (2008). Wnt pathway: a new role in regulation of inflammation. *Arterioscler Thromb Vasc Biol*, *28*(3), 400–402. https://doi.org/10.1161/ATVBAHA.107.160952

Ghazalpour, A., Rau, C. D., Farber, C. R., Bennett, B. J., Orozco, L. D., van Nas, A., Pan, C., Allayee, H., Beaven, S. W., Civelek, M., Davis, R. C., Drake, T. A., Friedman, R. A., Furlotte, N., Hui, S. T., Jentsch, J. D., Kostem, E., Kang, H. M., Kang, E. Y., . . . LeBoeuf, R. C. (2012). Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits. *Mamm Genome*, *23*(9–10), 680–692. https://doi.org/10.1007/s00335-012-9411-5

Gierahn, T. M., Wadsworth, M. H., Hughes, T. K., Bryson, B. D., Butler, A., Satija, R., Fortune, S., Love, J. C., & Shalek, A. K. (2017). Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nature Methods*, *14*(4), 395–398. https://doi.org/10.1038/nmeth.4179

Gole, S., Tkachenko, S., Masannat, T., Baylis, R. A., & Cherepanova, O. A. (2022). Endothelial-to-Mesenchymal Transition in Atherosclerosis: Friend or Foe? *Cells*, *11*(19). https://doi.org/10.3390/cells11192946

Gomez, D., & Owens, G. K. (2012). Smooth muscle cell phenotypic switching in atherosclerosis. *Cardiovasc Res*, *95*(2), 156–164. https://doi.org/10.1093/cvr/cvs115

Greenawalt, D. M., Dobrin, R., Chudin, E., Hatoum, I. J., Suver, C., Beaulaurier, J., Zhang, B., Castro, V., Zhu, J., Sieberts, S. K., Wang, S., Molony, C., Heymsfield, S. B., Kemp, D. M., Reitman, M. L., Lum, P. Y., Schadt, E. E., & Kaplan, L. M. (2011). A survey of the genetics of stomach, liver, and adipose gene expression from a morbidly obese cohort. *Genome Res*, *21*(7), 1008–1016. https://doi.org/10.1101/gr.112821.110

Grundberg, E., Small, K. S., Hedman, A. K., Nica, A. C., Buil, A., Keildson, S., Bell, J. T., Yang, T. P., Meduri, E., Barrett, A., Nisbett, J., Sekowska, M., Wilk, A., Shin, S. Y., Glass, D., Travers, M., Min, J. L., Ring, S., Ho, K., . . . Multiple Tissue Human Expression Resource, C. (2012). Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet*, *44*(10), 1084–1089. https://doi.org/10.1038/ng.2394

Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., Ottina, K., Wallace, M. R., Sakaguchi, A. Y., & et al. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Nature*, *306*(5940), 234–238. https://doi.org/10.1038/306234a0

Gustafsson, J., Held, F., Robinson, J. L., Björnson, E., Jörnsten, R., & Nielsen, J. (2020). Sources of variation in cell-type RNA-Seq profiles. *PLoS One*, *15*(9), e0239495. https://doi.org/10.1371/journal.pone.0239495

Hagg, S., Skogsberg, J., Lundstrom, J., Noori, P., Nilsson, R., Zhong, H., Maleki, S., Shang, M. M., Brinne, B., Bradshaw, M., Bajic, V. B., Samnegard, A., Silveira, A., Kaplan, L. M., Gigante, B., Leander, K., de Faire, U., Rosfors, S., Lockowandt, U., . . . Bjorkegren, J. (2009). Multi-organ expression profiling uncovers a gene module in coronary artery disease involving transendothelial migration of leukocytes and LIM domain binding 2: the Stockholm Atherosclerosis Gene Expression (STAGE) study. *PLoS Genet*, *5*(12), e1000754. https://doi.org/10.1371/journal.pgen.1000754

Haghverdi, L., Buettner, F., & Theis, F. J. (2015). Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics*, *31*(18), 2989–2998. https://doi.org/10.1093/bioinformatics/btv325

Haines, J. L., & Pericak-Vance, M. A. (2007). *Genetic analysis of complex disease*. John Wiley & Sons.

Hartman, R. J. G., Owsiany, K., Ma, L., Koplev, S., Hao, K., Slenders, L., Civelek, M., Mokry, M., Kovacic, J. C., Pasterkamp, G., Owens, G., Bjorkegren, J. L. M., & den Ruijter, H. M. (2021). Sex-Stratified Gene Regulatory Networks Reveal Female Key Driver Genes of Atherosclerosis Involved in Smooth Muscle Cell Phenotype Switching. *Circulation*, *143*(7), 713–726. https://doi.org/10.1161/CIRCULATIONAHA.120.051231

Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., de Leeuw, Y., Anavy, L., Gennert, D., Li, S., Livak, K. J., Rozenblatt-Rosen, O., Dor, Y., Regev, A., & Yanai, I. (2016). CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol*, *17*, 77. https://doi.org/10.1186/s13059-016-0938-8

Hashimshony, T., Wagner, F., Sher, N., & Yanai, I. (2012). CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep*, *2*(3), 666–673. https://doi.org/10.1016/j.celrep.2012.08.003

Hedlund, E., & Deng, Q. (2018). Single-cell RNA sequencing: Technical advancements and biological applications. *Mol Aspects Med*, *59*, 36–46. https://doi.org/10.1016/j.mam.2017.07.003

Helgadottir, A., Thorleifsson, G., Manolescu, A., Gretarsdottir, S., Blondal, T., Jonasdottir, A., Jonasdottir, A., Sigurdsson, A., Baker, A., Palsson, A., Masson, G., Gudbjartsson, D. F., Magnusson, K. P., Andersen, K., Levey, A. I., Backman, V. M., Matthiasdottir, S., Jonsdottir, T., Palsson, S., . . . Stefansson, K. (2007). A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*, *316*(5830), 1491–1493. https://doi.org/10.1126/science.1142842

Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet*, *6*(2), 95–108. https://doi.org/10.1038/nrg1521

Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics*, *6*(2), 95–108. https://doi.org/10.1038/nrg1521

Hochgerner, H., Lonnerberg, P., Hodge, R., Mikes, J., Heskol, A., Hubschle, H., Lin, P., Picelli, S., La Manno, G., Ratz, M., Dunne, J., Husain, S., Lein, E., Srinivasan, M., Zeisel, A., & Linnarsson, S. (2017). STRT-seq-2i: dual-index 5' single cell and nucleus RNA-seq on an addressable microwell array. *Sci Rep*, *7*(1), 16327. https://doi.org/10.1038/s41598-017-16546-4

Huynh-Thu, V. A., Irrthum, A., Wehenkel, L., & Geurts, P. (2010). Inferring Regulatory Networks from Expression Data Using Tree-Based Methods. *PLoS One*, *5*(9), e12776. https://doi.org/10.1371/journal.pone.0012776

Hwang, B., Lee, J. H., & Bang, D. (2021). Author Correction: Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp Mol Med*, *53*(5), 1005. https://doi.org/10.1038/s12276-021-00615-w

Islam, S., Kjallquist, U., Moliner, A., Zajac, P., Fan, J. B., Lonnerberg, P., & Linnarsson, S. (2012). Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nat Protoc*, *7*(5), 813–828. https://doi.org/10.1038/nprot.2012.022

Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., Lonnerberg, P., & Linnarsson, S. (2014). Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods*, *11*(2), 163–166. https://doi.org/10.1038/nmeth.2772

Jaitin, D. A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., & Amit, I. (2014). Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science*, *343*(6172), 776–779. https://doi.org/10.1126/science.1247651

Joiret, M., Mahachie John, J. M., Gusareva, E. S., & Van Steen, K. (2019). Confounding of linkage disequilibrium patterns in large scale DNA based gene-gene interaction studies. *BioData Min*, *12*, 11. https://doi.org/10.1186/s13040-019-0199-7

Joshi, A., Rienks, M., Theofilatos, K., & Mayr, M. (2021). Systems biology in cardiovascular disease: a multiomics approach. *Nat Rev Cardiol*, *18*(5), 313–330. https://doi.org/10.1038/s41569-020-00477-1

Kannel, W. B. (1994). Risk factors for atherosclerotic cardiovascular outcomes in different arterial territories. *J Cardiovasc Risk*, *1*(4), 333–339.

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., . . . MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, *581*(7809), 434–443. https://doi.org/10.1038/s41586-020-2308-7

Kawai, K., Finn, A. V., & Virmani, R. (2024). Subclinical Atherosclerosis: Part 1: What Is it? Can it Be Defined at the Histological Level? *Arterioscler Thromb Vasc Biol*, *44*(1), 12–23. https://doi.org/10.1161/atvbaha.123.319932

Khan, S. U., Saeed, S., Alsuhaibani, A. M., Fatima, S., Ur Rehman, K., Zaman, U., Ullah, M., Refat, M. S., & Lu, K. (2023). Advances and Challenges for GWAS Analysis in Cardiac Diseases: A Focus on Coronary Artery Disease (CAD). *Curr Probl Cardiol*, *48*(9), 101821. https://doi.org/10.1016/j.cpcardiol.2023.101821

Kiselev, V. Y., Andrews, T. S., & Hemberg, M. (2019). Publisher Correction: Challenges in unsupervised clustering of single-cell RNA-seq data. *Nat Rev Genet*, *20*(5), 310. https://doi.org/10.1038/s41576-019-0095-5

Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A., & Kirschner, M. W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*, *161*(5), 1187–1201. https://doi.org/10.1016/j.cell.2015.04.044

Klug, W. S. (2012). Cummings, Charlotte A. Spencer, Michael A. Palladino: Essentials of Genetics, 8. izdaja. In: Pearson.

Kobak, D., & Berens, P. (2019). The art of using t-SNE for single-cell transcriptomics. *Nat Commun*, *10*(1), 5416. https://doi.org/10.1038/s41467-019-13056-x

Koplev, S., Seldin, M., Sukhavasi, K., Ermel, R., Pang, S., Zeng, L., Bankier, S., Di Narzo, A., Cheng, H., & Meda, V. (2022). A mechanistic framework for cardiometabolic and coronary artery diseases. *Nature Cardiovascular Research*, *1*(1), 85–100.

Koplev, S., Seldin, M., Sukhavasi, K., Ermel, R., Pang, S., Zeng, L., Bankier, S., Di Narzo, A., Cheng, H., Meda, V., Ma, A., Talukdar, H., Cohain, A., Amadori, L., Argmann, C., Houten, S. M., Franzen, O., Mocci, G., Meelu, O. A., . . . Bjorkegren, J. L. M. (2022). A mechanistic framework for cardiometabolic and coronary artery diseases. *Nat Cardiovasc Res*, *1*(1), 85–100. https://doi.org/10.1038/s44161-021-00009-1

Kruglyak, L., & Nickerson, D. A. (2001). Variation is the spice of life. *Nat Genet*, *27*(3), 234–236. https://doi.org/10.1038/85776

Ku, C. S., Pawitan, Y., Sim, X., Ong, R. T., Seielstad, M., Lee, E. J., Teo, Y. Y., Chia, K. S., & Salim, A. (2010). Genomic copy number variations in three Southeast Asian populations. *Hum Mutat*, *31*(7), 851–857. https://doi.org/10.1002/humu.21287

Lafzi, A., Moutinho, C., Picelli, S., & Heyn, H. (2018). Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies. *Nat Protoc*, *13*(12), 2742–2757. https://doi.org/10.1038/s41596-018-0073-y

Lai, H. Y., Hsu, L. W., Tsai, H. H., Lo, Y. C., Yang, S. H., Liu, P. Y., & Wang, J. M. (2017). CCAAT/enhancer-binding protein delta promotes intracellular lipid accumulation in M1 macrophages of vascular lesions. *Cardiovasc Res*, *113*(11), 1376–1388. https://doi.org/10.1093/cvr/cvx134

Lander, E. S., & Botstein, D. (1986). Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms. *Proc Natl Acad Sci U S A*, *83*(19), 7353–7357. https://doi.org/10.1073/pnas.83.19.7353

Li, Z., Li, X., Zhou, H., Gaynor, S. M., Selvaraj, M. S., Arapoglou, T., Quick, C., Liu, Y., Chen, H., Sun, R., Dey, R., Arnett, D. K., Auer, P. L., Bielak, L. F., Bis, J. C., Blackwell, T. W., Blangero, J., Boerwinkle, E., Bowden, D. W., . . . Lin, X. (2022). A framework for detecting noncoding rare-variant associations of large-scale whole-genome sequencing studies. *Nat Methods*, *19*(12), 1599–1611. https://doi.org/10.1038/s41592-022-01640-x

Libby, P. (2021). The changing landscape of atherosclerosis. *Nature*, *592*(7855), 524–533. https://doi.org/10.1038/s41586-021-03392-8

Libby, P., & Theroux, P. (2005). Pathophysiology of Coronary Artery Disease. *Circulation*, *111*(25), 3481–3488. https://doi.org/doi:10.1161/CIRCULATIONAHA.105.537878

Libby, P., & Theroux, P. (2005). Pathophysiology of coronary artery disease. *Circulation*, *111*(25), 3481–3488. https://doi.org/10.1161/circulationaha.105.537878

Lorenz, M. W., Markus, H. S., Bots, M. L., Rosvall, M., & Sitzer, M. (2007). Prediction of clinical cardiovascular events with carotid intima-media thickness: a systematic review and meta-analysis. *Circulation*, *115*(4), 459–467. https://doi.org/10.1161/circulationaha.106.628875

Lusis, A. J., Fogelman, A. M., & Fonarow, G. C. (2004). Genetic basis of atherosclerosis: part I: new genes and pathways. *Circulation*, *110*(13), 1868–1873. https://doi.org/10.1161/01.CIR.0000143041.58692.CC

Lusis, A. J., & Weiss, J. N. (2010). Cardiovascular networks: systems-based approaches to cardiovascular disease. *Circulation*, *121*(1), 157–170. https://doi.org/10.1161/CIRCULATIONAHA.108.847699

Maas, A. H., & Appelman, Y. E. (2010). Gender differences in coronary heart disease. *Neth Heart J*, *18*(12), 598–602. https://doi.org/10.1007/s12471-010-0841-y

Mach, F., Baigent, C., Catapano, A. L., Koskinas, K. C., Casula, M., Badimon, L., Chapman, M. J., De Backer, G. G., Delgado, V., Ference, B. A., Graham, I. M., Halliday, A., Landmesser, U., Mihaylova, B., Pedersen, T. R., Riccardi, G., Richter, D. J., Sabatine, M. S., Taskinen, M.-R., . . . Group, E. S. D. (2019). 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk: The Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS). *European Heart Journal*, *41*(1), 111–188. https://doi.org/10.1093/eurheartj/ehz455

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A., & McCarroll, S. A. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, *161*(5), 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002

Man, J. J., Beckman, J. A., & Jaffe, I. Z. (2020). Sex as a Biological Variable in Atherosclerosis. *Circ Res*, *126*(9), 1297–1319. https://doi.org/10.1161/circresaha.120.315930

Manolio, T. A. (2010). Genomewide association studies and assessment of the risk of disease. *N Engl J Med*, *363*(2), 166–176. https://doi.org/10.1056/NEJMra0905980

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., McCarthy, M. I., Ramos, E. M., Cardon, L. R., Chakravarti, A., Cho, J. H., Guttmacher, A. E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C. N., Slatkin, M., Valle, D., Whittemore, A. S., . . . Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature*, *461*(7265), 747–753. https://doi.org/10.1038/nature08494

McPherson, R., Pertsemlidis, A., Kavaslar, N., Stewart, A., Roberts, R., Cox, D. R., Hinds, D. A., Pennacchio, L. A., Tybjaerg-Hansen, A., Folsom, A. R., Boerwinkle, E., Hobbs, H. H., & Cohen, J. C. (2007). A common allele on chromosome 9 associated with coronary heart disease. *Science*, *316*(5830), 1488–1491. https://doi.org/10.1126/science.1142447

Mereu, E., Lafzi, A., Moutinho, C., Ziegenhain, C., McCarthy, D. J., Alvarez-Varela, A., Batlle, E., Sagar, Grun, D., Lau, J. K., Boutet, S. C., Sanada, C., Ooi, A., Jones, R. C.,

Kaihara, K., Brampton, C., Talaga, Y., Sasagawa, Y., Tanaka, K., . . . Heyn, H. (2020). Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. *Nat Biotechnol*, *38*(6), 747–755. https://doi.org/10.1038/s41587-020-0469-4

Minelli, S., Minelli, P., & Montinari, M. R. (2020). Reflections on Atherosclerosis: Lesson from the Past and Future Research Directions. *J Multidiscip Healthc*, *13*, 621–633. https://doi.org/10.2147/jmdh.S254016

Moghtaderi, A., Sanei-Sistani, S., Abdollahi, G., & Dahmardeh, H. (2014). Comparison of intima-media thickness of common and internal carotid arteries of patients with ischemic stroke and intracerebral hemorrhage. *Iran J Neurol*, *13*(4), 226–230.

Nelson, C. P., Goel, A., Butterworth, A. S., Kanoni, S., Webb, T. R., Marouli, E., Zeng, L., Ntalla, I., Lai, F. Y., Hopewell, J. C., Giannakopoulou, O., Jiang, T., Hamby, S. E., Di Angelantonio, E., Assimes, T. L., Bottinger, E. P., Chambers, J. C., Clarke, R., Palmer, C. N. A., . . . Deloukas, P. (2017). Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet*, *49*(9), 1385–1391. https://doi.org/10.1038/ng.3913

Owens, G. K., Kumar, M. S., & Wamhoff, B. R. (2004). Molecular Regulation of Vascular Smooth Muscle Cell Differentiation in Development and Disease. *Physiological Reviews*, *84*(3), 767–801. https://doi.org/10.1152/physrev.00041.2003

Ozsolak, F., & Milos, P. M. (2011). RNA sequencing: advances, challenges and opportunities. *Nature Reviews Genetics*, *12*(2), 87–98. https://doi.org/10.1038/nrg2934

Pan, H., Xue, C., Auerbach, B. J., Fan, J., Bashore, A. C., Cui, J., Yang, D. Y., Trignano, S. B., Liu, W., Shi, J., Ihuegbu, C. O., Bush, E. C., Worley, J., Vlahos, L., Laise, P., Solomon, R. A., Connolly, E. S., Califano, A., Sims, P. A., . . . Reilly, M. P. (2020). Single-Cell Genomics Reveals a Novel Cell State During Smooth Muscle Cell Phenotypic Switching and Potential Therapeutic Targets for Atherosclerosis in Mouse and Human. *Circulation*, *142*(21), 2060–2075. https://doi.org/10.1161/CIRCULATIONAHA.120.048378

Pang, S., Yengo, L., Nelson, C. P., Bourier, F., Zeng, L., Li, L., Kessler, T., Erdmann, J., Magi, R., Lall, K., Metspalu, A., Mueller-Myhsok, B., Samani, N. J., Visscher, P. M., & Schunkert, H. (2023). Genetic and modifiable risk factors combine multiplicatively in common disease. *Clin Res Cardiol*, *112*(2), 247–257. https://doi.org/10.1007/s00392-022-02081-4

Pathan, N., Deng, W. Q., Di Scipio, M., Khan, M., Mao, S., Morton, R. W., Lali, R., Pigeyre, M., Chong, M. R., & Pare, G. (2024). A method to estimate the contribution of rare coding variants to complex trait heritability. *Nat Commun*, *15*(1), 1245. https://doi.org/10.1038/s41467-024-45407-8

Picelli, S., Bjorklund, A. K., Faridani, O. R., Sagasser, S., Winberg, G., & Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods*, *10*(11), 1096–1098. https://doi.org/10.1038/nmeth.2639

Picelli, S., Faridani, O. R., Bjorklund, A. K., Winberg, G., Sagasser, S., & Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc*, *9*(1), 171–181. https://doi.org/10.1038/nprot.2014.006

Potter, S. S. (2018). Single-cell RNA sequencing for the study of development, physiology and disease. *Nat Rev Nephrol*, *14*(8), 479–492. https://doi.org/10.1038/s41581-018-0021-7

Pritchard, J. K., & Cox, N. J. (2002). The allelic architecture of human disease genes: common disease-common variant...or not? *Hum Mol Genet*, *11*(20), 2417–2423. https://doi.org/10.1093/hmg/11.20.2417

Rajendran, P., Rengarajan, T., Thangavel, J., Nishigaki, Y., Sakthisekaran, D., Sethi, G., & Nishigaki, I. (2013). The vascular endothelium and human diseases. *Int J Biol Sci*, *9*(10), 1057–1069. https://doi.org/10.7150/ijbs.7502

Ramskold, D., Luo, S., Wang, Y. C., Li, R., Deng, Q., Faridani, O. R., Daniels, G. A., Khrebtukova, I., Loring, J. F., Laurent, L. C., Schroth, G. P., & Sandberg, R. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol*, *30*(8), 777–782. https://doi.org/10.1038/nbt.2282

Ramskold, D., Luo, S., Wang, Y. C., Li, R., Deng, Q., Faridani, O. R., Daniels, G. A., Khrebtukova, I., Loring, J. F., Laurent, L. C., Schroth, G. P., & Sandberg, R. (2020). Author Correction: Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol*, *38*(3), 374. https://doi.org/10.1038/s41587-020-0427-1

Regan, C., & Preall, J. (2022). Practical Considerations for Single-Cell Genomics. *Curr Protoc*, *2*(8), e498. https://doi.org/10.1002/cpz1.498

Reynolds, J. L., Joannides, A. J., Skepper, J. N., McNair, R., Schurgers, L. J., Proudfoot, D., Jahnen-Dechent, W., Weissberg, P. L., & Shanahan, C. M. (2004). Human vascular smooth muscle cells undergo vesicle-mediated calcification in response to changes in extracellular calcium and phosphate concentrations: a potential mechanism for accelerated vascular calcification in ESRD. *J Am Soc Nephrol*, *15*(11), 2857–2867. https://doi.org/10.1097/01.ASN.0000141960.01035.28

Rocheleau, G., Clarke, S. L., Auguste, G., Hasbani, N. R., Morrison, A. C., Heath, A. S., Bielak, L. F., Iyer, K. R., Young, E. P., Stitziel, N. O., Jun, G., Laurie, C., Broome, J. G., Khan, A. T., Arnett, D. K., Becker, L. C., Bis, J. C., Boerwinkle, E., Bowden, D. W., . . . Do, R. (2024). Rare variant contribution to the heritability of coronary artery disease. *Nat Commun*, *15*(1), 8741. https://doi.org/10.1038/s41467-024-52939-6

Rostom, R., Svensson, V., Teichmann, S. A., & Kar, G. (2017). Computational approaches for interpreting scRNA-seq data. *FEBS Lett*, *591*(15), 2213–2225. https://doi.org/10.1002/1873-3468.12684

Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Baddour, L. M., Barengo, N. C., Beaton, A. Z., Benjamin, E. J., Benziger, C. P., Bonny, A., Brauer, M., Brodmann, M., Cahill, T. J., Carapetis, J., Catapano, A. L., Chugh, S. S., Cooper, L. T., Coresh, J., . . . Group, G.-N.-J. G. B. o. C. D. W. (2020). Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019: Update From the GBD 2019 Study. *J Am Coll Cardiol*, *76*(25), 2982–3021. https://doi.org/10.1016/j.jacc.2020.11.010

Saliba, A. E., Westermann, A. J., Gorski, S. A., & Vogel, J. (2014). Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res*, *42*(14), 8845–8860. https://doi.org/10.1093/nar/gku555

Samani, N. J., Erdmann, J., Hall, A. S., Hengstenberg, C., Mangino, M., Mayer, B., Dixon, R. J., Meitinger, T., Braund, P., Wichmann, H. E., Barrett, J. H., Konig, I. R., Stevens, S. E., Szymczak, S., Tregouet, D. A., Iles, M. M., Pahlke, F., Pollard, H., Lieb, W., . . . the Cardiogenics, C. (2007). Genomewide association analysis of coronary artery disease. *N Engl J Med*, *357*(5), 443–453. https://doi.org/10.1056/NEJMoa072366

Sasagawa, Y., Danno, H., Takada, H., Ebisawa, M., Tanaka, K., Hayashi, T., Kurisaki, A., & Nikaido, I. (2018). Quartz-Seq2: a high-throughput single-cell RNA-sequencing method that effectively uses limited sequence reads. *Genome Biology*, *19*(1), 29. https://doi.org/10.1186/s13059-018-1407-3

Sasagawa, Y., Nikaido, I., Hayashi, T., Danno, H., Uno, K. D., Imai, T., & Ueda, H. R. (2013). Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol*, *14*(4), R31. https://doi.org/10.1186/gb-2013-14-4-r31

Schadt, E. E. (2009). Molecular networks as sensors and drivers of common human diseases. *Nature*, *461*(7261), 218-223. https://doi.org/10.1038/nature08454

Shapiro, E., Biezuner, T., & Linnarsson, S. (2013). Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet*, *14*(9), 618–630. https://doi.org/10.1038/nrg3542

Silva, S., Nitsch, D., & Fatumo, S. (2023). Genome-wide association studies on coronary artery disease: A systematic review and implications for populations of different ancestries. *PLoS One*, *18*(11), e0294341. https://doi.org/10.1371/journal.pone.0294341

Sivapalaratnam, S., Motazacker, M. M., Maiwald, S., Hovingh, G. K., Kastelein, J. J., Levi, M., Trip, M. D., & Dallinga-Thie, G. M. (2011). Genome-wide association studies in atherosclerosis. *Curr Atheroscler Rep*, *13*(3), 225–232. https://doi.org/10.1007/s11883-011-0173-4

Skogsberg, J., Dicker, A., Rydén, M., Aström, G., Nilsson, R., Bhuiyan, H., Vitols, S., Mairal, A., Langin, D., Alberts, P., Walum, E., Tegnér, J., Hamsten, A., Arner, P., & Björkegren, J. (2008). ApoB100-LDL acts as a metabolic signal from liver to peripheral fat causing inhibition of lipolysis in adipocytes. *PLoS One*, *3*(11), e3771. https://doi.org/10.1371/journal.pone.0003771

Skogsberg, J., Lundstrom, J., Kovacs, A., Nilsson, R., Noori, P., Maleki, S., Kohler, M., Hamsten, A., Tegner, J., & Bjorkegren, J. (2008). Transcriptional profiling uncovers a network of cholesterol-responsive atherosclerosis target genes. *PLoS Genet*, *4*(3), e1000036. https://doi.org/10.1371/journal.pgen.1000036

Slenders, L., Landsmeer, L. P. L., Cui, K., Depuydt, M. A. C., Verwer, M., Mekke, J., Timmerman, N., van den Dungen, N. A. M., Kuiper, J., de Winther, M. P. J., Prange, K. H. M., Ma, W. F., Miller, C. L., Aherrahrou, R., Civelek, M., de Borst, G. J., de Kleijn, D. P. V., Asselbergs, F. W., den Ruijter, H. M., . . . Mokry, M. (2022). Intersecting single-cell transcriptomics and genome-wide association studies identifies crucial cell populations and candidate genes for atherosclerosis. *Eur Heart J Open*, *2*(1), oeab043. https://doi.org/10.1093/ehjopen/oeab043

Strong, J. P., Malcom, G. T., McMahan, C. A., Tracy, R. E., Newman III, W. P., Herderick, E. E., Cornhill, J. F., & Group, f. t. P. D. o. A. i. Y. R. (1999). Prevalence and Extent of Atherosclerosis in Adolescents and Young AdultsImplications for Prevention From the Pathobiological Determinants of Atherosclerosis in Youth Study. *JAMA*, *281*(8), 727–735. https://doi.org/10.1001/jama.281.8.727

Svensson, V., Vento-Tormo, R., & Teichmann, S. A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. *Nat Protoc*, *13*(4), 599–604. https://doi.org/10.1038/nprot.2017.149

Talukdar, H. A., Foroughi Asl, H., Jain, R. K., Ermel, R., Ruusalepp, A., Franzén, O., Kidd, B. A., Readhead, B., Giannarelli, C., Kovacic, J. C., Ivert, T., Dudley, J. T., Civelek, M., Lusis, A. J., Schadt, E. E., Skogsberg, J., Michoel, T., & Björkegren, J. L. (2016). Cross-Tissue Regulatory Gene Networks in Coronary Artery Disease. *Cell Syst*, *2*(3), 196–208. https://doi.org/10.1016/j.cels.2016.02.002

Tegner, J., & Bjorkegren, J. (2007). Perturbations to uncover gene networks. *Trends Genet*, *23*(1), 34–41. https://doi.org/10.1016/j.tig.2006.11.003

Tegner, J., Skogsberg, J., & Bjorkegren, J. (2007). Thematic review series: systems biology approaches to metabolic and cardiovascular disorders. Multi-organ whole-genome measurements and reverse engineering to uncover gene networks underlying complex traits. *J Lipid Res*, *48*(2), 267–277. https://doi.org/10.1194/jlr.R600030-JLR200

Trapnell, C. (2015). Defining cell types and states with single-cell genomics. *Genome Res*, *25*(10), 1491–1498. https://doi.org/10.1101/gr.190595.115

Tsui, L.-C., Buchwald, M., Barker, D., Braman, J. C., Knowlton, R., Schumm, J. W., Eiberg, H., Mohr, J., Kennedy, D., Plavsic, N., Zsiga, M., Markiewicz, D., Akots, G., Brown, V., Helms, C., Gravius, T., Parker, C., Rediker, K., & Donis-Keller, H. (1985). Cystic Fibrosis Locus Defined by a Genetically Linked Polymorphic DNA Marker. *Science*, *230*(4729), 1054–1057. https://doi.org/doi:10.1126/science.2997931

van den Beucken, T. (2019). Systems biology approaches to interpreting genomic data. *Current Opinion in Toxicology*, *18*, 1–7. https://doi.org/https://doi.org/10.1016/j.cotox.2019.02.004

van den Brink, S. C., Sage, F., Vertesy, A., Spanjaard, B., Peterson-Maduro, J., Baron, C. S., Robin, C., & van Oudenaarden, A. (2017). Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat Methods*, *14*(10), 935–936. https://doi.org/10.1038/nmeth.4437

Virmani, R., Burke, A. P., Farb, A., & Kolodgie, F. D. (2006). Pathology of the Vulnerable Plaque. *JACC*, *47*(8_Supplement), C13-C18. https://doi.org/doi:10.1016/j.jacc.2005.10.065

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *Am J Hum Genet*, *90*(1), 7–24. https://doi.org/10.1016/j.ajhg.2011.11.029

Visscher, P. M., & Goddard, M. E. (2019). From R.A. Fisher's 1918 Paper to GWAS a Century Later. *Genetics*, *211*(4), 1125–1130. https://doi.org/10.1534/genetics.118.301594

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era – concepts and misconceptions. *Nat Rev Genet*, *9*(4), 255–266. https://doi.org/10.1038/nrg2322

Vrijenhoek, J. E. P., Haitjema, S., de Borst, G. J., de Vries, J.-P. P. M., Vaartjes, I., Moll, F. L., Pasterkamp, G., & den Ruijter, H. M. (2014). The impact of female sex on long-term survival of patients with severe atherosclerosis undergoing endarterectomy. *Atherosclerosis*, *237*(2), 521–527. https://doi.org/10.1016/j.atherosclerosis.2014.10.010

Wang, B., Yang, X., Sun, X., Liu, J., Fu, Y., Liu, B., Qiu, J., Lian, J., & Zhou, J. (2022). ATF3 in atherosclerosis: a controversial transcription factor. *J Mol Med (Berl)*, *100*(11), 1557–1568. https://doi.org/10.1007/s00109-022-02263-7

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), 57–63. https://doi.org/10.1038/nrg2484

Watson, J. D., & Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, *171*(4356), 737–738. https://doi.org/10.1038/171737a0

Webb, R. C. (2003). Smooth muscle contraction and relaxation. *Adv Physiol Educ*, *27*(1–4), 201–206. https://doi.org/10.1152/advan.00025.2003

Williams, J. W., Winkels, H., Durant, C. P., Zaitsev, K., Ghosheh, Y., & Ley, K. (2020). Single Cell RNA Sequencing in Atherosclerosis Research. *Circ Res*, *126*(9), 1112–1126. https://doi.org/10.1161/circresaha.119.315940

Winkels, H., Ehinger, E., Vassallo, M., Buscher, K., Dinh, H. Q., Kobiyama, K., Hamers, A. A. J., Cochain, C., Vafadarnejad, E., Saliba, A. E., Zernecke, A., Pramod, A. B.,

Ghosh, A. K., Anto Michel, N., Hoppe, N., Hilgendorf, I., Zirlik, A., Hedrick, C. C., Ley, K., & Wolf, D. (2018). Atlas of the Immune Cell Repertoire in Mouse Atherosclerosis Defined by Single-Cell RNA-Sequencing and Mass Cytometry. *Circ Res*, *122*(12), 1675–1688. https://doi.org/10.1161/CIRCRESAHA.117.312513

Wirka, R. C., Wagh, D., Paik, D. T., Pjanic, M., Nguyen, T., Miller, C. L., Kundu, R., Nagao, M., Coller, J., Koyano, T. K., Fong, R., Woo, Y. J., Liu, B., Montgomery, S. B., Wu, J. C., Zhu, K., Chang, R., Alamprese, M., Tallquist, M. D., . . . Quertermous, T. (2019). Atheroprotective roles of smooth muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell analysis. *Nat Med*, *25*(8), 1280–1289. https://doi.org/10.1038/s41591-019-0512-5

Witteman, J. C., Grobbee, D. E., Kok, F. J., Hofman, A., & Valkenburg, H. A. (1989). Increased risk of atherosclerosis in women after the menopause. *BMJ*, *298*(6674), 642–644. https://doi.org/10.1136/bmj.298.6674.642

Wu, M. C., Lee, S., Cai, T., Li, Y., Boehnke, M., & Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet*, *89*(1), 82–93. https://doi.org/10.1016/j.ajhg.2011.05.029

Xu, J., Lu, X., & Shi, G.-P. (2015). Vasa Vasorum in Atherosclerosis and Clinical Significance. *International Journal of Molecular Sciences*, *16*(5), 11574–11608. https://www.mdpi.com/1422-0067/16/5/11574

Yao, D. W., O'Connor, L. J., Price, A. L., & Gusev, A. (2020). Quantifying genetic effects on disease mediated by assayed gene expression levels. *Nat Genet*, *52*(6), 626–633. https://doi.org/10.1038/s41588-020-0625-2

Yavorska, O. O., & Burgess, S. (2017). MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int J Epidemiol*, *46*(6), 1734–1739. https://doi.org/10.1093/ije/dyx034

Zeng, L., Talukdar, H. A., Koplev, S., Giannarelli, C., Ivert, T., Gan, L. M., Ruusalepp, A., Schadt, E. E., Kovacic, J. C., Lusis, A. J., Michoel, T., Schunkert, H., & Bjorkegren, J. L. M. (2019). Contribution of Gene Regulatory Networks to Heritability of Coronary Artery Disease. *Journal of the American College of Cardiology*, *73*(23), 2946–2957. https://doi.org/10.1016/j.jacc.2019.03.520

Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., & Enard, W. (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mol Cell*, *65*(4), 631–643 e634. https://doi.org/10.1016/j.molcel.2017.01.023

ACKNOWLEDGEMENTS

The current doctoral research was conducted at the Department of Cardiology, Institute of Clinical Medicine, University of Tartu. First and foremost, I extend my deepest gratitude to my dedicated supervisors, Professor Johan Björkegren and Associate Professor Arno Ruusalepp, for their invaluable guidance and unwavering motivation throughout this endeavor. My intellectual journey into this field commenced when I joined the STARNET Biobank as a laboratory assistant. The extensive and intriguing questions this biobank held the potential to answer not only ignited my interest but also inspired me to embark on this doctoral degree. Professor Björkegren's profound expertise in the field, coupled with Associate Professor Ruusalepp's consistent and readily available support, fostered a supportive and intellectually stimulating environment that proved essential to the successful completion of this research.

I am particularly grateful for the pivotal support extended by Dr. Heli Järve, especially during the critical single-cell protocol standardization stage, where her guidance in facilitating the crucial tissue collection proved essential for the smooth execution of the experiment. I also wish to thank Dr. Raili Tagen for her invaluable assistance during the initial stages of understanding the STARNET biobank and for her constant friendship and support whenever needed. Furthermore, I am deeply indebted to Professor Marika Väli for generously providing me with both lab and comfortable office space, and for her consistent care and concern for my well-being throughout this journey.

This research work was a collaborative endeavor that necessitated travel to Karolinska Institutet in Sweden to learn SMART-seq2 technology. This constant travel not only facilitated my acquisition of new skills but also led to the formation of numerous cherished friendships. I particularly wish to emphasize my gratitude to Dr. Giuseppe Mocci for his exceptional patience, especially during challenging times, and for the much-needed comfort and solace provided in our weekly Zoom and occasional WhatsApp calls. His heartfelt offer to consider him an "elder brother" is something I deeply appreciate. I also extend my sincere thanks to Dr. Byambajav Buyandelger and Sonja Gustaffson for their efforts in making me feel at home during these trips.

I am immensely thankful to both the internal reviewers of this thesis, Professor Dr. Jaak Kals, Department of Surgery, Institute of Clinical Medicine, University of Tartu and Associate Professor Dr. Sander Pajusalu, Department of Genetics and Personalized Medicine, Institute of Clinical Medicine, University of Tartu for their time, helpful comments, suggestions, and constructive criticism to improve the quality of the thesis draft.

Recognizing the importance of balance – because all work and no play makes Jack a dull boy, I am immensely blessed and grateful to everyone who brought joy and laughter to my life throughout this challenging journey. In no particular order, I wish to acknowledge my cherished coffee buddies (Dipti Kotwal, Saoni Banerjee, and Qurat), supportive family friends (Gayatri, Hina, Nelli, Lena,

Pinky, Ranjana, Nisha, Shubham, Chitra and Late. Mr. and Mrs. Mody), wonderful school and university buddies (Manju, Arpitha, Deepthi, Jyothi, Suman Banik, Ravi Kanth, Anirudh and Shravanthi), fantastic lab/floor friends (Mahvish, Ruby, Maili and Nina), and the ever-vibrant JAGRATA GROUP (Aditya, Arpan, Deepika, Dharmendra, Harleen, Jhalak, Kajal, Rohish, Sanu, Shrikant, Sudhichan, Surabhi and Upasana).

I am also appreciative of the numerous individuals I encountered on this journey. Each one, in their own way, imparted invaluable life lessons and enriched my understanding in ways I never anticipated.

My deepest gratitude and respect goes to my parents (Shri M. Sambasiva Rao, Smt. M. Aruna Kumari, Shri S. Siva Prasad Rao, Smt. S. Usha Rani), and my in-laws (Late. Shri Ram Murti Pathak and Smt. Kalawati Pathak) for their immense belief in my abilities. My appreciation extends to my dear sister D. Sailaja and my brothers, especially Dr. M. Vamsi Madhav, for their constant encouragement. I want to offer a heartfelt thank you to my life partner, Dr. Ajai K Pathak, and our darling son, Akshith. Their profound understanding and steadfast support during demanding times have been my rock and my stability.

This thesis represents not just years of research but also a personal journey of resilience and self-discovery. I am grateful for the inner strength that allowed me to persist through the most challenging moments.

I have always intended to conclude my acknowledgement section with this self-affirmation: I want to thank myself for always looking for the sunshine at the end of the torturous tunnel and for having a firm belief in karma.

PUBLICATIONS

CURRICULUM VITAE

Name: Katyayani Sukhavasi
Date of Birth: June 5, 1984, India
Address: University of Tartu, Faculty of medicine,
Institute of Clinical Medicine,
Department of Cardiology
8, Ludvig Puusepa, Tartu, 50406, Estonia
Telephone: +372 5689 1030
E-mail: katyayani.sukhavasi@ut.ee; katyayani.sukhavasi@gmail.com

Education:

2016–…	Ph.D. student, Faculty of Medicine, University of Tartu.
2004–2006	Master of Science (M.Sc.), Osmania University, Hyderabad, India.
2001–2004	Bachelor of Science (B.Sc.), Osmania University, Hyderabad, India.

Professional employment:

2018–present	Specialist, University of Tartu, Department of Cardiac Surgery, Tartu University Hospital.
2018–2021	Junior Research Fellow of Cardiovascular Genetics, Estonian Genome Center, University of Tartu.
2016–2017	Specialist, Department of Pathophysiology, University of Tartu.
2013–2015	Laboratory assistant, Department of Pathological Anatomy and Forensic Sciences, University of Tartu.

Publications, and Book chapters:

Publications:

1. AE Hartley, **Sukhavasi. K**, S Hu, M Traylor… & Y.Jamshidi (2025). Multi-Tissue Profiling Reveals tissue-specific protein regulation and relationships Between Protein Quantitative Trait Loci (pQTLs) and Cardiometabolic Disease. BioRxiv.
2. Ma, L., Tamis-Holland, J. E., Mocci, G., Wolhuter, K., Bryce, N. S., Sajja, S., Amadori, L., Pradhan, P., Chong, P. S., **Sukhavasi, K**., … & Björkegren, J. L. M. (2025). A macrophage gene-regulatory network linked to clinical severity of coronary artery disease: The STARNET and NGS-PREDICT primary blood macrophage studies. *Basic research in cardiology*, *120*(4), 799–814.
3. **Sukhavasi, K.,** Mocci, G., Ma, L., Hodonsky, C. J., Diez Benevante, E., Muhl, L., Liu, J., Gustafsson, S., Buyandelger, B., & Koplev, S. (2025). Single-cell RNA sequencing reveals sex differences in the subcellular

composition and associated gene-regulatory network activity of human carotid plaques. *Nature Cardiovascular Research*, 1–21.
4. Bankier, S., Talukdar, H., Khan, M., Mocci, G., **Sukhavasi, K.,** Hao, K., Ma, A., Ruusalepp, A., Schadt, E. E., & Kovacic, J. C. (2025). Plasma proteins are integral to gene-regulatory networks acting within and across blood cells, the arterial wall and major metabolic organs. *medRxiv*.
5. Mocci, G., **Sukhavasi, K.,** Örd, T., Bankier, S., Singha, P., Arasu, U. T., Agbabiaje, O. O., Mäkinen, P., Ma, L., & Hodonsky, C. J. (2024). Single-cell gene-regulatory networks of advanced symptomatic atherosclerosis. *Circulation Research*, *134*(11), 1405–1423.
6. Bijla, M., Saini, S. K., Pathak, A. K., Bharadwaj, K. P., **Sukhavasi, K.,** Patil, A., Saini, D., Yadav, R., Singh, S., & Leeuwenburgh, C. (2024). Microbiome interactions with different risk factors in development of myocardial infarction. *Experimental Gerontology*, *189*, 112409.
7. Savić, R., Yang, J., Koplev, S., An, M. C., Patel, P. L., O'Brien, R. N., Dubose, B. N., Dodatko, T., Rogatsky, E., & **Sukhavasi, K**. (2023). Integration of transcriptomes of senescent cell models with multi-tissue patient samples reveals reduced COL6A3 as an inducer of senescence. *Cell reports*, *42*(11).
8. Ma, L., Bryce, N. S., Turner, A. W., Di Narzo, A. F., Rahman, K., Xu, Y., Ermel, R., **Sukhavasi, K**., d'Escamard, V., & Chandel, N. (2022). The HDAC9-associated risk locus promotes coronary artery disease by governing TWIST1. *PLoS Genetics*, *18*(6), e1010261.
9. Li, L., Chen, Z., von Scheidt, M., Li, S., Steiner, A., Güldener, U., Koplev, S., Ma, A., Hao, K., Pan, C., **Sukhavasi, K**., … & Schunkert H. (2022). Transcriptome-wide association study of coronary artery disease identifies novel susceptibility genes. *Basic research in cardiology*, *117*(1), 6.
10. Koplev, S., Seldin, M., **Sukhavasi, K**., Ermel, R., Pang, S., Zeng, L., Bankier, S., Di Narzo, A., Cheng, H., & Meda, V. (2022). A mechanistic framework for cardiometabolic and coronary artery diseases. *Nature Cardiovascular Research*, *1*(1), 85–100.
11. Hao, K., Ermel, R., **Sukhavasi, K**., Cheng, H., Ma, L., Li, L., Amadori, L., Koplev, S., Franzén, O., & d'Escamard, V. (2022). Integrative prioritization of causal genes for coronary artery disease. *Circulation: Genomic and Precision Medicine*, *15*(1), e003365.
12. Crawford, A. A., Bankier, S., Altmaier, E., Barnes, C. L., Clark, D. W., Ermel, R., Friedrich, N., Van Der Harst, P., Joshi, P. K., Karhunen, V., **Sukhavasi, K**., … & Walker, B. R. (2021). Variation in the SERPINA6/SERPINA1 locus alters morning plasma cortisol, hepatic corticosteroid binding globulin expression, gene expression in peripheral tissues, and risk of cardiovascular disease. *Journal of human genetics*, *66*(6), 625–636.
13. Cohain, A. T., Barrington, W. T., Jordan, D. M., Beckmann, N. D., Argmann, C. A., Houten, S. M., Charney, A. W., Ermel, R., **Sukhavasi, K**., & Franzen, O. (2021). An integrative multiomic network model links lipid metabolism

to glucose regulation in coronary artery disease. *Nature communications*, *12*(1), 547.
14. Ma, L., Chandel, N., Ermel, R., **Sukhavasi, K**., Hao, K., Ruusalepp, A., Björkegren, J. L., & Kovacic, J. C. (2020). Multiple independent mechanisms link gene polymorphisms in the region of ZEB2 with risk of coronary artery disease. *Atherosclerosis*, *311*, 20–29.
15. Alencar, G. F., Owsiany, K. M., Karnewar, S., **Sukhavasi, K**., Mocci, G., Nguyen, A. T., Williams, C. M., Shamsuzzaman, S., Mokry, M., & Henderson, C. A. (2020). Stem cell pluripotency genes Klf4 and Oct4 regulate complex SMC phenotypic changes critical in late-stage atherosclerotic lesion pathogenesis. *Circulation*, *142*(21), 2045–2059.
16. Glicksberg, B. S., Amadori, L., Akers, N. K., **Sukhavasi, K**., Franzén, O., Li, L., Belbin, G. M., Akers, K. L., Shameer, K., & Badgeley, M. A. (2019). Integrative analysis of loss-of-function variants in clinical and genomic data reveals novel genes associated with cardiovascular traits. *BMC Medical Genomics*, *12*, 1–16.
17. Franzén, O., Ermel, R., **Sukhavasi, K**., Jain, R., Jain, A., Betsholtz, C., Giannarelli, C., Kovacic, J. C., Ruusalepp, A., & Skogsberg, J. (2018). Global analysis of A-to-I RNA editing reveals association with common disease variants. *PeerJ*, *6*, e4466.
18. Franzén, O., Ermel, R., Cohain, A., Akers, N. K., Di Narzo, A., Talukdar, H. A., Foroughi-Asl, H., Giambartolomei, C., Fullard, J. F., & **Sukhavasi, K**. (2016). Cardiometabolic risk loci share downstream cis-and trans-gene regulation across tissues and diseases. *Science*, *353*(6301), 827–830.

Book chapters:

1. Pathak, A. K., **Sukhavasi, K**., Marnetto, D., Chaubey, G., & Pandey, A. K. (2022). Human population genomics approach in food metabolism. In *Future Foods* (pp. 433–449). Elsevier.

ELULOOKIRJELDUS

Nimi: Katyayani Sukhavasi
Sünniaeg: 5 juuni 1984
Kontakt: Tartu Ülikool, Meditsiiniteaduskond,
Kliinilise Meditsiini Instituut, Kardioloogia Osakond
Ludvig Puusepa 8, Tartu, 50406, Eesti
Telefon: +372 5689 1030
E-meil: katyayani.sukhavasi@ut.ee; katyayani.sukhavasi@gmail.com

Hariduskäik:

2016–	PhD (Meditsiin), Meditsiiniteaduste valdkond, Tartu Ülikool, Eesti.
2004–2006	Teadusmagister (M.Sc), Osmania Ülikool, India.
2001–2004	Bakalaureus loodusteadustes (B.Sc.), Osmania Ülikool, India.

Teenistuskäik:

2018–present	Spetsialist, Tartu Ülikooli Kliinikum.
2018–2021	Nooremteadur, kardiovaskulaargeneetika, Tartu Ülikooli genoomika instituut.
2016–2017	Spetsialist, Patofüsioloogia osakond, Tartu Ülikool.
2013–2015	Laborant-spetsialist, Patoloogilise anatoomia ja Kohtumeditsiini osakond, Tartu Ülikool.

Publikatsioonid, ja Raamatute peatükid:
Publikatsioonid:

1. AE Hartley, **Sukhavasi. K**, S Hu, M Traylor… & Y.Jamshidi (2025). Multi-Tissue Profiling Reveals tissue-specific protein regulation and relationships Between Protein Quantitative Trait Loci (pQTLs) and Cardiometabolic Disease. BioRxiv.
2. Ma, L., Tamis-Holland, J. E., Mocci, G., Wolhuter, K., Bryce, N. S., Sajja, S., Amadori, L., Pradhan, P., Chong, P. S., **Sukhavasi, K**., … & Björkegren, J. L. M. (2025). A macrophage gene-regulatory network linked to clinical severity of coronary artery disease: The STARNET and NGS-PREDICT primary blood macrophage studies. *Basic research in cardiology*, *120*(4), 799–814.
3. **Sukhavasi, K.,** Mocci, G., Ma, L., Hodonsky, C. J., Diez Benevante, E., Muhl, L., Liu, J., Gustafsson, S., Buyandelger, B., & Koplev, S. (2025). Single-cell RNA sequencing reveals sex differences in the subcellular composition and associated gene-regulatory network activity of human carotid plaques. *Nature Cardiovascular Research*, 1–21.
4. Bankier, S., Talukdar, H., Khan, M., Mocci, G., **Sukhavasi, K.,** Hao, K., Ma, A., Ruusalepp, A., Schadt, E. E., & Kovacic, J. C. (2025). Plasma proteins

are integral to gene-regulatory networks acting within and across blood cells, the arterial wall and major metabolic organs. *medRxiv*.

5. Mocci, G., **Sukhavasi, K.,** Örd, T., Bankier, S., Singha, P., Arasu, U. T., Agbabiaje, O. O., Mäkinen, P., Ma, L., & Hodonsky, C. J. (2024). Single-cell gene-regulatory networks of advanced symptomatic atherosclerosis. *Circulation Research*, *134*(11), 1405–1423.
6. Bijla, M., Saini, S. K., Pathak, A. K., Bharadwaj, K. P., **Sukhavasi, K.,** Patil, A., Saini, D., Yadav, R., Singh, S., & Leeuwenburgh, C. (2024). Microbiome interactions with different risk factors in development of myocardial infarction. *Experimental Gerontology*, *189*, 112409.
7. Savić, R., Yang, J., Koplev, S., An, M. C., Patel, P. L., O'Brien, R. N., Dubose, B. N., Dodatko, T., Rogatsky, E., & **Sukhavasi, K**. (2023). Integration of transcriptomes of senescent cell models with multi-tissue patient samples reveals reduced COL6A3 as an inducer of senescence. *Cell reports*, *42*(11).
8. Ma, L., Bryce, N. S., Turner, A. W., Di Narzo, A. F., Rahman, K., Xu, Y., Ermel, R., **Sukhavasi, K**., d'Escamard, V., & Chandel, N. (2022). The HDAC9-associated risk locus promotes coronary artery disease by governing TWIST1. *PLoS Genetics*, *18*(6), e1010261.
9. Li, L., Chen, Z., von Scheidt, M., Li, S., Steiner, A., Güldener, U., Koplev, S., Ma, A., Hao, K., Pan, C., **Sukhavasi, K**., …& Schunkert H. (2022). Transcriptome-wide association study of coronary artery disease identifies novel susceptibility genes. *Basic research in cardiology*, *117*(1), 6.
10. Koplev, S., Seldin, M., **Sukhavasi, K**., Ermel, R., Pang, S., Zeng, L., Bankier, S., Di Narzo, A., Cheng, H., & Meda, V. (2022). A mechanistic framework for cardiometabolic and coronary artery diseases. *Nature Cardiovascular Research*, *1*(1), 85–100.
11. Hao, K., Ermel, R., **Sukhavasi, K**., Cheng, H., Ma, L., Li, L., Amadori, L., Koplev, S., Franzén, O., & d'Escamard, V. (2022). Integrative prioritization of causal genes for coronary artery disease. *Circulation: Genomic and Precision Medicine*, *15*(1), e003365.
12. Crawford, A. A., Bankier, S., Altmaier, E., Barnes, C. L., Clark, D. W., Ermel, R., Friedrich, N., Van Der Harst, P., Joshi, P. K., Karhunen, V., **Sukhavasi, K**., … & Walker, B. R. (2021). Variation in the SERPINA6/ SERPINA1 locus alters morning plasma cortisol, hepatic corticosteroid binding globulin expression, gene expression in peripheral tissues, and risk of cardiovascular disease. *Journal of human genetics*, *66*(6), 625–636.
13. Cohain, A. T., Barrington, W. T., Jordan, D. M., Beckmann, N. D., Argmann, C. A., Houten, S. M., Charney, A. W., Ermel, R., **Sukhavasi, K**., & Franzen, O. (2021). An integrative multiomic network model links lipid metabolism to glucose regulation in coronary artery disease. *Nature communications*, *12*(1), 547.
14. Ma, L., Chandel, N., Ermel, R., **Sukhavasi, K**., Hao, K., Ruusalepp, A., Björkegren, J. L., & Kovacic, J. C. (2020). Multiple independent

mechanisms link gene polymorphisms in the region of ZEB2 with risk of coronary artery disease. *Atherosclerosis*, *311*, 20–29.
15. Alencar, G. F., Owsiany, K. M., Karnewar, S., **Sukhavasi, K**., Mocci, G., Nguyen, A. T., Williams, C. M., Shamsuzzaman, S., Mokry, M., & Henderson, C. A. (2020). Stem cell pluripotency genes Klf4 and Oct4 regulate complex SMC phenotypic changes critical in late-stage atherosclerotic lesion pathogenesis. *Circulation*, *142*(21), 2045–2059.
16. Glicksberg, B. S., Amadori, L., Akers, N. K., **Sukhavasi, K**., Franzén, O., Li, L., Belbin, G. M., Akers, K. L., Shameer, K., & Badgeley, M. A. (2019). Integrative analysis of loss-of-function variants in clinical and genomic data reveals novel genes associated with cardiovascular traits. *BMC Medical Genomics*, *12*, 1–16.
17. Franzén, O., Ermel, R., **Sukhavasi, K**., Jain, R., Jain, A., Betsholtz, C., Giannarelli, C., Kovacic, J. C., Ruusalepp, A., & Skogsberg, J. (2018). Global analysis of A-to-I RNA editing reveals association with common disease variants. *PeerJ*, *6*, e4466.
18. Franzén, O., Ermel, R., Cohain, A., Akers, N. K., Di Narzo, A., Talukdar, H. A., Foroughi-Asl, H., Giambartolomei, C., Fullard, J. F., & **Sukhavasi, K**. (2016). Cardiometabolic risk loci share downstream cis-and trans-gene regulation across tissues and diseases. *Science*, *353*(6301), 827–830.

Raamatute peatükid:

1. Pathak, A. K., **Sukhavasi, K**., Marnetto, D., Chaubey, G., & Pandey, A. K. (2022). Human population genomics approach in food metabolism. In *Future Foods* (pp. 433–449). Elsevier.

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS

1. **Heidi-Ingrid Maaroos**. The natural course of gastric ulcer in connection with chronic gastritis and *Helicobacter pylori*. Tartu, 1991.
2. **Mihkel Zilmer**. Na-pump in normal and tumorous brain tissues: Structural, functional and tumorigenesis aspects. Tartu, 1991.
3. **Eero Vasar**. Role of cholecystokinin receptors in the regulation of behaviour and in the action of haloperidol and diazepam. Tartu, 1992.
4. **Tiina Talvik**. Hypoxic-ischaemic brain damage in neonates (clinical, biochemical and brain computed tomographical investigation). Tartu, 1992.
5. **Ants Peetsalu**. Vagotomy in duodenal ulcer disease: A study of gastric acidity, serum pepsinogen I, gastric mucosal histology and *Helicobacter pylori*. Tartu, 1992.
6. **Marika Mikelsaar**. Evaluation of the gastrointestinal microbial ecosystem in health and disease. Tartu, 1992.
7. **Hele Everaus**. Immuno-hormonal interactions in chronic lymphocytic leukaemia and multiple myeloma. Tartu, 1993.
8. **Ruth Mikelsaar**. Etiological factors of diseases in genetically consulted children and newborn screening: dissertation for the commencement of the degree of doctor of medical sciences. Tartu, 1993.
9. **Agu Tamm**. On metabolic action of intestinal microflora: clinical aspects. Tartu, 1993.
10. **Katrin Gross**. Multiple sclerosis in South-Estonia (epidemiological and computed tomographical investigations). Tartu, 1993.
11. **Oivi Uibo**. Childhood coeliac disease in Estonia: occurrence, screening, diagnosis and clinical characterization. Tartu, 1994.
12. **Viiu Tuulik**. The functional disorders of central nervous system of chemistry workers. Tartu, 1994.
13. **Margus Viigimaa**. Primary haemostasis, antiaggregative and anticoagulant treatment of acute myocardial infarction. Tartu, 1994.
14. **Rein Kolk**. Atrial versus ventricular pacing in patients with sick sinus syndrome. Tartu, 1994.
15. **Toomas Podar**. Incidence of childhood onset type 1 diabetes mellitus in Estonia. Tartu, 1994.
16. **Kiira Subi**. The laboratory surveillance of the acute respiratory viral infections in Estonia. Tartu, 1995.
17. **Irja Lutsar**. Infections of the central nervous system in children (epidemiologic, diagnostic and therapeutic aspects, long term outcome). Tartu, 1995.
18. **Aavo Lang**. The role of dopamine, 5-hydroxytryptamine, sigma and NMDA receptors in the action of antipsychotic drugs. Tartu, 1995.
19. **Andrus Arak**. Factors influencing the survival of patients after radical surgery for gastric cancer. Tartu, 1996.

20. **Tõnis Karki**. Quantitative composition of the human lactoflora and method for its examination. Tartu, 1996.
21. **Reet Mändar**. Vaginal microflora during pregnancy and its transmission to newborn. Tartu, 1996.
22. **Triin Remmel**. Primary biliary cirrhosis in Estonia: epidemiology, clinical characterization and prognostication of the course of the disease. Tartu, 1996.
23. **Toomas Kivastik**. Mechanisms of drug addiction: focus on positive reinforcing properties of morphine. Tartu, 1996.
24. **Paavo Pokk**. Stress due to sleep deprivation: focus on $GABA_A$ receptor-chloride ionophore complex. Tartu, 1996.
25. **Kristina Allikmets**. Renin system activity in essential hypertension. Associations with atherothrombogenic cardiovascular risk factors and with the efficacy of calcium antagonist treatment. Tartu, 1996.
26. **Triin Parik**. Oxidative stress in essential hypertension: Associations with metabolic disturbances and the effects of calcium antagonist treatment. Tartu, 1996.
27. **Svetlana Päi**. Factors promoting heterogeneity of the course of rheumatoid arthritis. Tartu, 1997.
28. **Maarike Sallo**. Studies on habitual physical activity and aerobic fitness in 4 to 10 years old children. Tartu, 1997.
29. **Paul Naaber**. *Clostridium difficile* infection and intestinal microbial ecology. Tartu, 1997.
30. **Rein Pähkla**. Studies in pinoline pharmacology. Tartu, 1997.
31. **Andrus Juhan Voitk**. Outpatient laparoscopic cholecystectomy. Tartu, 1997.
32. **Joel Starkopf**. Oxidative stress and ischaemia-reperfusion of the heart. Tartu, 1997.
33. **Janika Kõrv**. Incidence, case-fatality and outcome of stroke. Tartu, 1998.
34. **Ülla Linnamägi**. Changes in local cerebral blood flow and lipid peroxidation following lead exposure in experiment. Tartu, 1998.
35. **Ave Minajeva**. Sarcoplasmic reticulum function: comparison of atrial and ventricular myocardium. Tartu, 1998.
36. **Oleg Milenin**. Reconstruction of cervical part of esophagus by revascularised ileal autografts in dogs. A new complex multistage method. Tartu, 1998.
37. **Sergei Pakriev**. Prevalence of depression, harmful use of alcohol and alcohol dependence among rural population in Udmurtia. Tartu, 1998.
38. **Allen Kaasik**. Thyroid hormone control over β-adrenergic signalling system in rat atria. Tartu, 1998.
39. **Vallo Matto**. Pharmacological studies on anxiogenic and antiaggressive properties of antidepressants. Tartu, 1998.
40. **Maire Vasar**. Allergic diseases and bronchial hyperreactivity in Estonian children in relation to environmental influences. Tartu, 1998.
41. **Kaja Julge**. Humoral immune responses to allergens in early childhood. Tartu, 1998.

42. **Heli Grünberg**. The cardiovascular risk of Estonian schoolchildren. A cross-sectional study of 9-, 12- and 15-year-old children. Tartu, 1998.
43. **Epp Sepp**. Formation of intestinal microbial ecosystem in children. Tartu, 1998.
44. **Mai Ots**. Characteristics of the progression of human and experimental glomerulopathies. Tartu, 1998.
45. **Tiina Ristimäe**. Heart rate variability in patients with coronary artery disease. Tartu, 1998.
46. **Leho Kõiv**. Reaction of the sympatho-adrenal and hypothalamo-pituitary-adrenocortical system in the acute stage of head injury. Tartu, 1998.
47. **Bela Adojaan**. Immune and genetic factors of childhood onset IDDM in Estonia. An epidemiological study. Tartu, 1999.
48. **Jakov Shlik**. Psychophysiological effects of cholecystokinin in humans. Tartu, 1999.
49. **Kai Kisand**. Autoantibodies against dehydrogenases of α-ketoacids. Tartu, 1999.
50. **Toomas Marandi**. Drug treatment of depression in Estonia. Tartu, 1999.
51. **Ants Kask**. Behavioural studies on neuropeptide Y. Tartu, 1999.
52. **Ello-Rahel Karelson**. Modulation of adenylate cyclase activity in the rat hippocampus by neuropeptide galanin and its chimeric analogs. Tartu, 1999.
53. **Tanel Laisaar**. Treatment of pleural empyema — special reference to intrapleural therapy with streptokinase and surgical treatment modalities. Tartu, 1999.
54. **Eve Pihl**. Cardiovascular risk factors in middle-aged former athletes. Tartu, 1999.
55. **Katrin Õunap**. Phenylketonuria in Estonia: incidence, newborn screening, diagnosis, clinical characterization and genotype/phenotype correlation. Tartu, 1999.
56. **Siiri Kõljalg**. *Acinetobacter* – an important nosocomial pathogen. Tartu, 1999.
57. **Helle Karro**. Reproductive health and pregnancy outcome in Estonia: association with different factors. Tartu, 1999.
58. **Heili Varendi**. Behavioral effects observed in human newborns during exposure to naturally occurring odors. Tartu, 1999.
59. **Anneli Beilmann**. Epidemiology of epilepsy in children and adolescents in Estonia. Prevalence, incidence, and clinical characteristics. Tartu, 1999.
60. **Vallo Volke**. Pharmacological and biochemical studies on nitric oxide in the regulation of behaviour. Tartu, 1999.
61. **Pilvi Ilves**. Hypoxic-ischaemic encephalopathy in asphyxiated term infants. A prospective clinical, biochemical, ultrasonographical study. Tartu, 1999.
62. **Anti Kalda**. Oxygen-glucose deprivation-induced neuronal death and its pharmacological prevention in cerebellar granule cells. Tartu, 1999.
63. **Eve-Irene Lepist**. Oral peptide prodrugs – studies on stability and absorption. Tartu, 2000.

64. **Jana Kivastik**. Lung function in Estonian schoolchildren: relationship with anthropometric indices and respiratory symptomas, reference values for dynamic spirometry. Tartu, 2000.
65. **Karin Kull**. Inflammatory bowel disease: an immunogenetic study. Tartu, 2000.
66. **Kaire Innos**. Epidemiological resources in Estonia: data sources, their quality and feasibility of cohort studies. Tartu, 2000.
67. **Tamara Vorobjova**. Immune response to *Helicobacter pylori* and its association with dynamics of chronic gastritis and epithelial cell turnover in antrum and corpus. Tartu, 2001.
68. **Ruth Kalda**. Structure and outcome of family practice quality in the changing health care system of Estonia. Tartu, 2001.
69. **Annika Krüüner**. *Mycobacterium tuberculosis* – spread and drug resistance in Estonia. Tartu, 2001.
70. **Marlit Veldi**. Obstructive Sleep Apnoea: Computerized Endopharyngeal Myotonometry of the Soft Palate and Lingual Musculature. Tartu, 2001.
71. **Anneli Uusküla**. Epidemiology of sexually transmitted diseases in Estonia in 1990–2000. Tartu, 2001.
72. **Ade Kallas**. Characterization of antibodies to coagulation factor VIII. Tartu, 2002.
73. **Heidi Annuk**. Selection of medicinal plants and intestinal lactobacilli as antimicrobil components for functional foods. Tartu, 2002.
74. **Aet Lukmann**. Early rehabilitation of patients with ischaemic heart disease after surgical revascularization of the myocardium: assessment of health-related quality of life, cardiopulmonary reserve and oxidative stress. A clinical study. Tartu, 2002.
75. **Maigi Eisen**. Pathogenesis of Contact Dermatitis: participation of Oxidative Stress. A clinical – biochemical study. Tartu, 2002.
76. **Piret Hussar**. Histology of the post-traumatic bone repair in rats. Elaboration and use of a new standardized experimental model – bicortical perforation of tibia compared to internal fracture and resection osteotomy. Tartu, 2002.
77. **Tõnu Rätsep**. Aneurysmal subarachnoid haemorrhage: Noninvasive monitoring of cerebral haemodynamics. Tartu, 2002.
78. **Marju Herodes**. Quality of life of people with epilepsy in Estonia. Tartu, 2003.
79. **Katre Maasalu**. Changes in bone quality due to age and genetic disorders and their clinical expressions in Estonia. Tartu, 2003.
80. **Toomas Sillakivi**. Perforated peptic ulcer in Estonia: epidemiology, risk factors and relations with *Helicobacter pylori*. Tartu, 2003.
81. **Leena Puksa**. Late responses in motor nerve conduction studies. F and A waves in normal subjects and patients with neuropathies. Tartu, 2003.
82. **Krista Lõivukene**. *Helicobacter pylori* in gastric microbial ecology and its antimicrobial susceptibility pattern. Tartu, 2003.

83. **Helgi Kolk**. Dyspepsia and *Helicobacter pylori* infection: the diagnostic value of symptoms, treatment and follow-up of patients referred for upper gastrointestinal endoscopy by family physicians. Tartu, 2003.
84. **Helena Soomer**. Validation of identification and age estimation methods in forensic odontology. Tartu, 2003.
85. **Kersti Oselin**. Studies on the human MDR1, MRP1, and MRP2 ABC transporters: functional relevance of the genetic polymorphisms in the *MDR1* and *MRP1* gene. Tartu, 2003.
86. **Jaan Soplepmann**. Peptic ulcer haemorrhage in Estonia: epidemiology, prognostic factors, treatment and outcome. Tartu, 2003.
87. **Margot Peetsalu**. Long-term follow-up after vagotomy in duodenal ulcer disease: recurrent ulcer, changes in the function, morphology and *Helicobacter pylori* colonisation of the gastric mucosa. Tartu, 2003.
88. **Kersti Klaamas**. Humoral immune response to *Helicobacter pylori* a study of host-dependent and microbial factors. Tartu, 2003.
89. **Pille Taba**. Epidemiology of Parkinson's disease in Tartu, Estonia. Prevalence, incidence, clinical characteristics, and pharmacoepidemiology. Tartu, 2003.
90. **Alar Veraksitš**. Characterization of behavioural and biochemical phenotype of cholecystokinin-2 receptor deficient mice: changes in the function of the dopamine and endopioidergic system. Tartu, 2003.
91. **Ingrid Kalev**. CC-chemokine receptor 5 (CCR5) gene polymorphism in Estonians and in patients with Type I and Type II diabetes mellitus. Tartu, 2003.
92. **Lumme Kadaja**. Molecular approach to the regulation of mitochondrial function in oxidative muscle cells. Tartu, 2003.
93. **Aive Liigant**. Epidemiology of primary central nervous system tumours in Estonia from 1986 to 1996. Clinical characteristics, incidence, survival and prognostic factors. Tartu, 2004.
94. **Andres, Kulla**. Molecular characteristics of mesenchymal stroma in human astrocytic gliomas. Tartu, 2004.
95. **Mari Järvelaid**. Health damaging risk behaviours in adolescence. Tartu, 2004.
96. **Ülle Pechter**. Progression prevention strategies in chronic renal failure and hypertension. An experimental and clinical study. Tartu, 2004.
97. **Gunnar Tasa**. Polymorphic glutathione S-transferases – biology and role in modifying genetic susceptibility to senile cataract and primary open angle glaucoma. Tartu, 2004.
98. **Tuuli Käämbre**. Intracellular energetic unit: structural and functional aspects. Tartu, 2004.
99. **Vitali Vassiljev**. Influence of nitric oxide syntase inhibitors on the effects of ethanol after acute and chronic ethanol administration and withdrawal. Tartu, 2004.

100. **Aune Rehema**. Assessment of nonhaem ferrous iron and glutathione redox ratio as markers of pathogeneticity of oxidative stress in different clinical groups. Tartu, 2004.
101. **Evelin Seppet**. Interaction of mitochondria and ATPases in oxidative muscle cells in normal and pathological conditions. Tartu, 2004.
102. **Eduard Maron**. Serotonin function in panic disorder: from clinical experiments to brain imaging and genetics. Tartu, 2004.
103. **Marje Oona**. *Helicobacter pylori* infection in children: epidemiological and therapeutic aspects. Tartu, 2004.
104. **Kersti Kokk**. Regulation of active and passive molecular transport in the testis. Tartu, 2005.
105. **Vladimir Järv**. Cross-sectional imaging for pretreatment evaluation and follow-up of pelvic malignant tumours. Tartu, 2005.
106. **Andre Õun**. Epidemiology of adult epilepsy in Tartu, Estonia. Incidence, prevalence and medical treatment. Tartu, 2005.
107. **Piibe Muda**. Homocysteine and hypertension: associations between homocysteine and essential hypertension in treated and untreated hypertensive patients with and without coronary artery disease. Tartu, 2005.
108. **Külli Kingo**. The interleukin-10 family cytokines gene polymorphisms in plaque psoriasis. Tartu, 2005.
109. **Mati Merila**. Anatomy and clinical relevance of the glenohumeral joint capsule and ligaments. Tartu, 2005.
110. **Epp Songisepp**. Evaluation of technological and functional properties of the new probiotic *Lactobacillus fermentum* ME-3. Tartu, 2005.
111. **Tiia Ainla**. Acute myocardial infarction in Estonia: clinical characteristics, management and outcome. Tartu, 2005.
112. **Andres Sell**. Determining the minimum local anaesthetic requirements for hip replacement surgery under spinal anaesthesia – a study employing a spinal catheter. Tartu, 2005.
113. **Tiia Tamme**. Epidemiology of odontogenic tumours in Estonia. Pathogenesis and clinical behaviour of ameloblastoma. Tartu, 2005.
114. **Triine Annus**. Allergy in Estonian schoolchildren: time trends and characteristics. Tartu, 2005.
115. **Tiia Voor**. Microorganisms in infancy and development of allergy: comparison of Estonian and Swedish children. Tartu, 2005.
116. **Priit Kasenõmm**. Indicators for tonsillectomy in adults with recurrent tonsillitis – clinical, microbiological and pathomorphological investigations. Tartu, 2005.
117. **Eva Zusinaite**. Hepatitis C virus: genotype identification and interactions between viral proteases. Tartu, 2005.
118. **Piret Kõll**. Oral lactoflora in chronic periodontitis and periodontal health. Tartu, 2006.
119. **Tiina Stelmach**. Epidemiology of cerebral palsy and unfavourable neurodevelopmental outcome in child population of Tartu city and county, Estonia Prevalence, clinical features and risk factors. Tartu, 2006.

120. **Katrin Pudersell**. Tropane alkaloid production and riboflavine excretion in the field and tissue cultures of henbane (*Hyoscyamus niger* L.). Tartu, 2006.
121. **Külli Jaako**. Studies on the role of neurogenesis in brain plasticity. Tartu, 2006.
122. **Aare Märtson**. Lower limb lengthening: experimental studies of bone regeneration and long-term clinical results. Tartu, 2006.
123. **Heli Tähepõld**. Patient consultation in family medicine. Tartu, 2006.
124. **Stanislav Liskmann**. Peri-implant disease: pathogenesis, diagnosis and treatment in view of both inflammation and oxidative stress profiling. Tartu, 2006.
125. **Ruth Rudissaar**. Neuropharmacology of atypical antipsychotics and an animal model of psychosis. Tartu, 2006.
126. **Helena Andreson**. Diversity of *Helicobacter pylori* genotypes in Estonian patients with chronic inflammatory gastric diseases. Tartu, 2006.
127. **Katrin Pruus**. Mechanism of action of antidepressants: aspects of serotoninergic system and its interaction with glutamate. Tartu, 2006.
128. **Priit Põder**. Clinical and experimental investigation: relationship of ischaemia/reperfusion injury with oxidative stress in abdominal aortic aneurysm repair and in extracranial brain artery endarterectomy and possibilities of protection against ischaemia using a glutathione analogue in a rat model of global brain ischaemia. Tartu, 2006.
129. **Marika Tammaru**. Patient-reported outcome measurement in rheumatoid arthritis. Tartu, 2006.
130. **Tiia Reimand**. Down syndrome in Estonia. Tartu, 2006.
131. **Diva Eensoo**. Risk-taking in traffic and Markers of Risk-Taking Behaviour in Schoolchildren and Car Drivers. Tartu, 2007.
132. **Riina Vibo**. The third stroke registry in Tartu, Estonia from 2001 to 2003: incidence, case-fatality, risk factors and long-term outcome. Tartu, 2007.
133. **Chris Pruunsild**. Juvenile idiopathic arthritis in children in Estonia. Tartu, 2007.
134. **Eve Õiglane-Šlik**. Angelman and Prader-Willi syndromes in Estonia. Tartu, 2007.
135. **Kadri Haller**. Antibodies to follicle stimulating hormone. Significance in female infertility. Tartu, 2007.
136. **Pille Ööpik**. Management of depression in family medicine. Tartu, 2007.
137. **Jaak Kals**. Endothelial function and arterial stiffness in patients with atherosclerosis and in healthy subjects. Tartu, 2007.
138. **Priit Kampus**. Impact of inflammation, oxidative stress and age on arterial stiffness and carotid artery intima-media thickness. Tartu, 2007.
139. **Margus Punab**. Male fertility and its risk factors in Estonia. Tartu, 2007.
140. **Alar Toom**. Heterotopic ossification after total hip arthroplasty: clinical and pathogenetic investigation. Tartu, 2007.

141. **Lea Pehme**. Epidemiology of tuberculosis in Estonia 1991–2003 with special regard to extrapulmonary tuberculosis and delay in diagnosis of pulmonary tuberculosis. Tartu, 2007.
142. **Juri Karjagin**. The pharmacokinetics of metronidazole and meropenem in septic shock. Tartu, 2007.
143. **Inga Talvik**. Inflicted traumatic brain injury shaken baby syndrome in Estonia – epidemiology and outcome. Tartu, 2007.
144. **Tarvo Rajasalu**. Autoimmune diabetes: an immunological study of type 1 diabetes in humans and in a model of experimental diabetes (in RIP-B7.1 mice). Tartu, 2007.
145. **Inga Karu**. Ischaemia-reperfusion injury of the heart during coronary surgery: a clinical study investigating the effect of hyperoxia. Tartu, 2007.
146. **Peeter Padrik**. Renal cell carcinoma: Changes in natural history and treatment of metastatic disease. Tartu, 2007.
147. **Neve Vendt**. Iron deficiency and iron deficiency anaemia in infants aged 9 to 12 months in Estonia. Tartu, 2008.
148. **Lenne-Triin Heidmets**. The effects of neurotoxins on brain plasticity: focus on neural Cell Adhesion Molecule. Tartu, 2008.
149. **Paul Korrovits**. Asymptomatic inflammatory prostatitis: prevalence, etiological factors, diagnostic tools. Tartu, 2008.
150. **Annika Reintam**. Gastrointestinal failure in intensive care patients. Tartu, 2008.
151. **Kristiina Roots**. Cationic regulation of Na-pump in the normal, Alzheimer's and CCK_2 receptor-deficient brain. Tartu, 2008.
152. **Helen Puusepp**. The genetic causes of mental retardation in Estonia: fragile X syndrome and creatine transporter defect. Tartu, 2009.
153. **Kristiina Rull**. Human chorionic gonadotropin beta genes and recurrent miscarriage: expression and variation study. Tartu, 2009.
154. **Margus Eimre**. Organization of energy transfer and feedback regulation in oxidative muscle cells. Tartu, 2009.
155. **Maire Link**. Transcription factors FoxP3 and AIRE: autoantibody associations. Tartu, 2009.
156. **Kai Haldre**. Sexual health and behaviour of young women in Estonia. Tartu, 2009.
157. **Kaur Liivak**. Classical form of congenital adrenal hyperplasia due to 21-hydroxylase deficiency in Estonia: incidence, genotype and phenotype with special attention to short-term growth and 24-hour blood pressure. Tartu, 2009.
158. **Kersti Ehrlich**. Antioxidative glutathione analogues (UPF peptides) – molecular design, structure-activity relationships and testing the protective properties. Tartu, 2009.
159. **Anneli Rätsep**. Type 2 diabetes care in family medicine. Tartu, 2009.
160. **Silver Türk**. Etiopathogenetic aspects of chronic prostatitis: role of mycoplasmas, coryneform bacteria and oxidative stress. Tartu, 2009.

161. **Kaire Heilman**. Risk markers for cardiovascular disease and low bone mineral density in children with type 1 diabetes. Tartu, 2009.
162. **Kristi Rüütel**. HIV-epidemic in Estonia: injecting drug use and quality of life of people living with HIV. Tartu, 2009.
163. **Triin Eller**. Immune markers in major depression and in antidepressive treatment. Tartu, 2009.
164. **Siim Suutre**. The role of TGF-β isoforms and osteoprogenitor cells in the pathogenesis of heterotopic ossification. An experimental and clinical study of hip arthroplasty. Tartu, 2010.
165. **Kai Kliiman**. Highly drug-resistant tuberculosis in Estonia: Risk factors and predictors of poor treatment outcome. Tartu, 2010.
166. **Inga Villa**. Cardiovascular health-related nutrition, physical activity and fitness in Estonia. Tartu, 2010.
167. **Tõnis Org**. Molecular function of the first PHD finger domain of Autoimmune Regulator protein. Tartu, 2010.
168. **Tuuli Metsvaht**. Optimal antibacterial therapy of neonates at risk of early onset sepsis. Tartu, 2010.
169. **Jaanus Kahu**. Kidney transplantation: Studies on donor risk factors and mycophenolate mofetil. Tartu, 2010.
170. **Koit Reimand**. Autoimmunity in reproductive failure: A study on associated autoantibodies and autoantigens. Tartu, 2010.
171. **Mart Kull**. Impact of vitamin D and hypolactasia on bone mineral density: a population based study in Estonia. Tartu, 2010.
172. **Rael Laugesaar**. Stroke in children – epidemiology and risk factors. Tartu, 2010.
173. **Mark Braschinsky**. Epidemiology and quality of life issues of hereditary spastic paraplegia in Estonia and implemention of genetic analysis in everyday neurologic practice. Tartu, 2010.
174. **Kadri Suija**. Major depression in family medicine: associated factors, recurrence and possible intervention. Tartu, 2010.
175. **Jarno Habicht**. Health care utilisation in Estonia: socioeconomic determinants and financial burden of out-of-pocket payments. Tartu, 2010.
176. **Kristi Abram**. The prevalence and risk factors of rosacea. Subjective disease perception of rosacea patients. Tartu, 2010.
177. **Malle Kuum**. Mitochondrial and endoplasmic reticulum cation fluxes: Novel roles in cellular physiology. Tartu, 2010.
178. **Rita Teek**. The genetic causes of early onset hearing loss in Estonian children. Tartu, 2010.
179. **Daisy Volmer**. The development of community pharmacy services in Estonia – public and professional perceptions 1993–2006. Tartu, 2010.
180. **Jelena Lissitsina**. Cytogenetic causes in male infertility. Tartu, 2011.
181. **Delia Lepik**. Comparison of gunshot injuries caused from Tokarev, Makarov and Glock 19 pistols at different firing distances. Tartu, 2011.
182. **Ene-Renate Pähkla**. Factors related to the efficiency of treatment of advanced periodontitis. Tartu, 2011.

183. **Maarja Krass**. L-Arginine pathways and antidepressant action. Tartu, 2011.
184. **Taavi Lai**. Population health measures to support evidence-based health policy in Estonia. Tartu, 2011.
185. **Tiit Salum**. Similarity and difference of temperature-dependence of the brain sodium pump in normal, different neuropathological, and aberrant conditions and its possible reasons. Tartu, 2011.
186. **Tõnu Vooder**. Molecular differences and similarities between histological subtypes of non-small cell lung cancer. Tartu, 2011.
187. **Jelena Štšepetova**. The characterisation of intestinal lactic acid bacteria using bacteriological, biochemical and molecular approaches. Tartu, 2011.
188. **Radko Avi**. Natural polymorphisms and transmitted drug resistance in Estonian HIV-1 CRF06_cpx and its recombinant viruses. Tartu, 2011, 116 p.
189. **Edward Laane**. Multiparameter flow cytometry in haematological malignancies. Tartu, 2011, 152 p.
190. **Triin Jagomägi**. A study of the genetic etiology of nonsyndromic cleft lip and palate. Tartu, 2011, 158 p.
191. **Ivo Laidmäe**. Fibrin glue of fish (*Salmo salar*) origin: immunological study and development of new pharmaceutical preparation. Tartu, 2012, 150 p.
192. **Ülle Parm**. Early mucosal colonisation and its role in prediction of invasive infection in neonates at risk of early onset sepsis. Tartu, 2012, 168 p.
193. **Kaupo Teesalu**. Autoantibodies against desmin and transglutaminase 2 in celiac disease: diagnostic and functional significance. Tartu, 2012, 142 p.
194. **Maksim Zagura**. Biochemical, functional and structural profiling of arterial damage in atherosclerosis. Tartu, 2012, 162 p.
195. **Vivian Kont**. Autoimmune regulator: characterization of thymic gene regulation and promoter methylation. Tartu, 2012, 134 p.
196. **Pirje Hütt**. Functional properties, persistence, safety and efficacy of potential probiotic lactobacilli. Tartu, 2012, 246 p.
197. **Innar Tõru**. Serotonergic modulation of CCK-4- induced panic. Tartu, 2012, 132 p.
198. **Sigrid Vorobjov**. Drug use, related risk behaviour and harm reduction interventions utilization among injecting drug users in Estonia: implications for drug policy. Tartu, 2012, 120 p.
199. **Martin Serg**. Therapeutic aspects of central haemodynamics, arterial stiffness and oxidative stress in hypertension. Tartu, 2012, 156 p.
200. **Jaanika Kumm**. Molecular markers of articular tissues in early knee osteoarthritis: a population-based longitudinal study in middle-aged subjects. Tartu, 2012, 159 p.
201. **Kertu Rünkorg**. Functional changes of dopamine, endopioid and endocannabinoid systems in CCK2 receptor deficient mice. Tartu, 2012, 125 p.
202. **Mai Blöndal**. Changes in the baseline characteristics, management and outcomes of acute myocardial infarction in Estonia. Tartu, 2012, 127 p.

203. **Jana Lass**. Epidemiological and clinical aspects of medicines use in children in Estonia. Tartu, 2012, 170 p.
204. **Kai Truusalu**. Probiotic lactobacilli in experimental persistent *Salmonella* infection. Tartu, 2013, 139 p.
205. **Oksana Jagur**. Temporomandibular joint diagnostic imaging in relation to pain and bone characteristics. Long-term results of arthroscopic treatment. Tartu, 2013, 126 p.
206. **Katrin Sikk**. Manganese-ephedrone intoxication – pathogenesis of neurological damage and clinical symptomatology. Tartu, 2013, 125 p.
207. **Kai Blöndal**. Tuberculosis in Estonia with special emphasis on drug-resistant tuberculosis: Notification rate, disease recurrence and mortality. Tartu, 2013, 151 p.
208. **Marju Puurand**. Oxidative phosphorylation in different diseases of gastric mucosa. Tartu, 2013, 123 p.
209. **Aili Tagoma**. Immune activation in female infertility: Significance of autoantibodies and inflammatory mediators. Tartu, 2013, 135 p.
210. **Liis Sabre**. Epidemiology of traumatic spinal cord injury in Estonia. Brain activation in the acute phase of traumatic spinal cord injury. Tartu, 2013, 135 p.
211. **Merit Lamp**. Genetic susceptibility factors in endometriosis. Tartu, 2013, 125 p.
212. **Erik Salum**. Beneficial effects of vitamin D and angiotensin II receptor blocker on arterial damage. Tartu, 2013, 167 p.
213. **Maire Karelson**. Vitiligo: clinical aspects, quality of life and the role of melanocortin system in pathogenesis. Tartu, 2013, 153 p.
214. **Kuldar Kaljurand**. Prevalence of exfoliation syndrome in Estonia and its clinical significance. Tartu, 2013, 113 p.
215. **Raido Paasma**. Clinical study of methanol poisoning: handling large outbreaks, treatment with antidotes, and long-term outcomes. Tartu, 2013, 96 p.
216. **Anne Kleinberg**. Major depression in Estonia: prevalence, associated factors, and use of health services. Tartu, 2013, 129 p.
217. **Triin Eglit**. Obesity, impaired glucose regulation, metabolic syndrome and their associations with high-molecular-weight adiponectin levels. Tartu, 2014, 115 p.
218. **Kristo Ausmees**. Reproductive function in middle-aged males: Associations with prostate, lifestyle and couple infertility status. Tartu, 2014, 125 p.
219. **Kristi Huik**. The influence of host genetic factors on the susceptibility to HIV and HCV infections among intravenous drug users. Tartu, 2014, 144 p.
220. **Liina Tserel**. Epigenetic profiles of monocytes, monocyte-derived macrophages and dendritic cells. Tartu, 2014, 143 p.
221. **Irina Kerna**. The contribution of *ADAM12* and *CILP* genes to the development of knee osteoarthritis. Tartu, 2014, 152 p.

222. **Ingrid Liiv**. Autoimmune regulator protein interaction with DNA-dependent protein kinase and its role in apoptosis. Tartu, 2014, 143 p.
223. **Liivi Maddison**. Tissue perfusion and metabolism during intra-abdominal hypertension. Tartu, 2014, 103 p.
224. **Krista Ress**. Childhood coeliac disease in Estonia, prevalence in atopic dermatitis and immunological characterisation of coexistence. Tartu, 2014, 124 p.
225. **Kai Muru**. Prenatal screening strategies, long-term outcome of children with marked changes in maternal screening tests and the most common syndromic heart anomalies in Estonia. Tartu, 2014, 189 p.
226. **Kaja Rahu**. Morbidity and mortality among Baltic Chernobyl cleanup workers: a register-based cohort study. Tartu, 2014, 155 p.
227. **Klari Noormets**. The development of diabetes mellitus, fertility and energy metabolism disturbances in a Wfs1-deficient mouse model of Wolfram syndrome. Tartu, 2014, 132 p.
228. **Liis Toome**. Very low gestational age infants in Estonia. Tartu, 2014, 183 p.
229. **Ceith Nikkolo**. Impact of different mesh parameters on chronic pain and foreign body feeling after open inguinal hernia repair. Tartu, 2014, 132 p.
230. **Vadim Brjalin**. Chronic hepatitis C: predictors of treatment response in Estonian patients. Tartu, 2014, 122 p.
231. **Vahur Metsna**. Anterior knee pain in patients following total knee arthroplasty: the prevalence, correlation with patellar cartilage impairment and aspects of patellofemoral congruence. Tartu, 2014, 130 p.
232. **Marju Kase**. Glioblastoma multiforme: possibilities to improve treatment efficacy. Tartu, 2015, 137 p.
233. **Riina Runnel**. Oral health among elementary school children and the effects of polyol candies on the prevention of dental caries. Tartu, 2015, 112 p.
234. **Made Laanpere**. Factors influencing women's sexual health and reproductive choices in Estonia. Tartu, 2015, 176 p.
235. **Andres Lust**. Water mediated solid state transformations of a polymorphic drug – effect on pharmaceutical product performance. Tartu, 2015, 134 p.
236. **Anna Klugman**. Functionality related characterization of pretreated wood lignin, cellulose and polyvinylpyrrolidone for pharmaceutical applications. Tartu, 2015, 156 p.
237. **Triin Laisk-Podar**. Genetic variation as a modulator of susceptibility to female infertility and a source for potential biomarkers. Tartu, 2015, 155 p.
238. **Mailis Tõnisson**. Clinical picture and biochemical changes in blood in children with acute alcohol intoxication. Tartu, 2015, 100 p.
239. **Kadri Tamme**. High volume haemodiafiltration in treatment of severe sepsis – impact on pharmacokinetics of antibiotics and inflammatory response. Tartu, 2015, 133 p.

240. **Kai Part**. Sexual health of young people in Estonia in a social context: the role of school-based sexuality education and youth-friendly counseling services. Tartu, 2015, 203 p.
241. **Urve Paaver**. New perspectives for the amorphization and physical stabilization of poorly water-soluble drugs and understanding their dissolution behavior. Tartu, 2015, 139 p.
242. **Aleksandr Peet**. Intrauterine and postnatal growth in children with HLA-conferred susceptibility to type 1 diabetes. Tartu. 2015, 146 p.
243. **Piret Mitt**. Healthcare-associated infections in Estonia – epidemiology and surveillance of bloodstream and surgical site infections. Tartu, 2015, 145 p.
244. **Merli Saare**. Molecular Profiling of Endometriotic Lesions and Endometria of Endometriosis Patients. Tartu, 2016, 129 p.
245. **Kaja-Triin Laisaar**. People living with HIV in Estonia: Engagement in medical care and methods of increasing adherence to antiretroviral therapy and safe sexual behavior. Tartu, 2016, 132 p.
246. **Eero Merilind**. Primary health care performance: impact of payment and practice-based characteristics. Tartu, 2016, 120 p.
247. **Jaanika Kärner**. Cytokine-specific autoantibodies in AIRE deficiency. Tartu, 2016, 182 p.
248. **Kaido Paapstel**. Metabolomic profile of arterial stiffness and early biomarkers of renal damage in atherosclerosis. Tartu, 2016, 173 p.
249. **Liidia Kiisk**. Long-term nutritional study: anthropometrical and clinico-laboratory assessments in renal replacement therapy patients after intensive nutritional counselling. Tartu, 2016, 207 p.
250. **Georgi Nellis**. The use of excipients in medicines administered to neonates in Europe. Tartu, 2017, 159 p.
251. **Aleksei Rakitin**. Metabolic effects of acute and chronic treatment with valproic acid in people with epilepsy. Tartu, 2017, 125 p.
252. **Eveli Kallas**. The influence of immunological markers to susceptibility to HIV, HBV, and HCV infections among persons who inject drugs. Tartu, 2017, 138 p.
253. **Tiina Freimann**. Musculoskeletal pain among nurses: prevalence, risk factors, and intervention. Tartu, 2017, 125 p.
254. **Evelyn Aaviksoo**. Sickness absence in Estonia: determinants and influence of the sick-pay cut reform. Tartu, 2017, 121 p.
255. **Kalev Nõupuu**. Autosomal-recessive Stargardt disease: phenotypic heterogeneity and genotype-phenotype associations. Tartu, 2017, 131 p.
256. **Ho Duy Binh**. Osteogenesis imperfecta in Vietnam. Tartu, 2017, 125 p.
257. **Uku Haljasorg**. Transcriptional mechanisms in thymic central tolerance. Tartu, 2017, 147 p.
258. **Živile Riispere**. IgA Nephropathy study according to the Oxford Classification: IgA Nephropathy clinical-morphological correlations, disease progression and the effect of renoprotective therapy. Tartu, 2017, 129 p.

259. **Hiie Soeorg**. Coagulase-negative staphylococci in gut of preterm neonates and in breast milk of their mothers. Tartu, 2017, 216 p.
260. **Anne-Mari Anton Willmore**. Silver nanoparticles for cancer research. Tartu, 2017, 132 p.
261. **Ott Laius**. Utilization of osteoporosis medicines, medication adherence and the trend in osteoporosis related hip fractures in Estonia. Tartu, 2017, 134 p.
262. **Alar Aab**. Insights into molecular mechanisms of asthma and atopic dermatitis. Tartu, 2017, 164 p.
263. **Sander Pajusalu**. Genome-wide diagnostics of Mendelian disorders: from chromosomal microarrays to next-generation sequencing. Tartu, 2017, 146 p.
264. **Mikk Jürisson**. Health and economic impact of hip fracture in Estonia. Tartu, 2017, 164 p.
265. **Kaspar Tootsi**. Cardiovascular and metabolomic profiling of osteo-arthritis. Tartu, 2017, 150 p.
266. **Mario Saare**. The influence of AIRE on gene expression – studies of transcriptional regulatory mechanisms in cell culture systems. Tartu, 2017, 172 p.
267. **Piia Jõgi**. Epidemiological and clinical characteristics of pertussis in Estonia. Tartu, 2018, 168 p.
268. **Elle Põldoja**. Structure and blood supply of the superior part of the shoulder joint capsule. Tartu, 2018, 116 p.
269. **Minh Son Nguyen**. Oral health status and prevalence of temporo-mandibular disorders in 65–74-year-olds in Vietnam. Tartu, 2018, 182 p.
270. **Kristian Semjonov**. Development of pharmaceutical quench-cooled molten and melt-electrospun solid dispersions for poorly water-soluble indomethacin. Tartu, 2018, 125 p.
271. **Janne Tiigimäe-Saar**. Botulinum neurotoxin type A treatment for sialorrhea in central nervous system diseases. Tartu, 2018, 109 p.
272. **Veiko Vengerfeldt**. Apical periodontitis: prevalence and etiopathogenetic aspects. Tartu, 2018, 150 p.
273. **Rudolf Bichele**. TNF superfamily and AIRE at the crossroads of thymic differentiation and host protection against *Candida albicans* infection. Tartu, 2018, 153 p.
274. **Olga Tšuiko**. Unravelling Chromosomal Instability in Mammalian Pre-implantation Embryos Using Single-Cell Genomics. Tartu, 2018, 169 p.
275. **Kärt Kriisa**. Profile of acylcarnitines, inflammation and oxidative stress in first-episode psychosis before and after antipsychotic treatment. Tartu, 2018, 145 p.
276. **Xuan Dung Ho**. Characterization of the genomic profile of osteosarcoma. Tartu, 2018, 144 p.
277. **Karit Reinson**. New Diagnostic Methods for Early Detection of Inborn Errors of Metabolism in Estonia. Tartu, 2018, 201 p.

278. **Mari-Anne Vals**. Congenital N-glycosylation Disorders in Estonia. Tartu, 2019, 148 p.
279. **Liis Kadastik-Eerme**. Parkinson's disease in Estonia: epidemiology, quality of life, clinical characteristics and pharmacotherapy. Tartu, 2019, 202 p.
280. **Hedi Hunt**. Precision targeting of intraperitoneal tumors with peptide-guided nanocarriers. Tartu, 2019, 179 p.
281. **Rando Porosk**. The role of oxidative stress in Wolfram syndrome 1 and hypothermia. Tartu, 2019, 123 p.
282. **Ene-Ly Jõgeda**. The influence of coinfections and host genetic factor on the susceptibility to HIV infection among people who inject drugs. Tartu, 2019, 126 p.
283. **Kristel Ehala-Aleksejev**. The associations between body composition, obesity and obesity-related health and lifestyle conditions with male reproductive function. Tartu, 2019, 138 p.
284. **Aigar Ottas**. The metabolomic profiling of psoriasis, atopic dermatitis and atherosclerosis. Tartu, 2019, 136 p.
285. **Elmira Gurbanova**. Specific characteristics of tuberculosis in low default, but high multidrug–resistance prison setting. Tartu, 2019, 129 p.
286. **Van Thai Nguyeni**. The first study of the treatment outcomes of patients with cleft lip and palate in Central Vietnam. Tartu, 2019, 144 p.
287. **Maria Yakoreva**. Imprinting Disorders in Estonia. Tartu, 2019, 187 p.
288. **Kadri Rekker**. The putative role of microRNAs in endometriosis pathogenesis and potential in diagnostics. Tartu, 2019, 140 p.
289. **Ülle Võhma**. Association between personality traits, clinical characteristics and pharmacological treatment response in panic disorder. Tartu, 2019, 121 p.
290. **Aet Saar**. Acute myocardial infarction in Estonia 2001–2014: towards risk-based prevention and management. Tartu, 2019, 124 p.
291. **Toomas Toomsoo**. Transcranial brain sonography in the Estonian cohort of Parkinson's disease. Tartu, 2019, 114 p.
292. **Lidiia Zhytnik**. Inter- and intrafamilial diversity based on genotype and phenotype correlations of Osteogenesis Imperfecta. Tartu, 2019, 224 p.
293. **Pilleriin Soodla**. Newly HIV-infected people in Estonia: estimation of incidence and transmitted drug resistance. Tartu, 2019, 194 p.
294. **Kristiina Ojamaa**. Epidemiology of gynecological cancer in Estonia. Tartu, 2020, 133 p.
295. **Marianne Saard**. Modern Cognitive and Social Intervention Techniques in Paediatric Neurorehabilitation for Children with Acquired Brain Injury. Tartu, 2020, 168 p.
296. **Julia Maslovskaja**. The importance of DNA binding and DNA breaks for AIRE-mediated transcriptional activation. Tartu, 2020, 162 p.
297. **Natalia Lobanovskaya**. The role of PSA-NCAM in the survival of retinal ganglion cells. Tartu, 2020, 105 p.

298. **Madis Rahu**. Structure and blood supply of the postero-superior part of the shoulder joint capsule with implementation of surgical treatment after anterior traumatic dislocation. Tartu, 2020, 104 p.
299. **Helen Zirnask**. Luteinizing hormone (LH) receptor expression in the penis and its possible role in pathogenesis of erectile disturbances. Tartu, 2020, 87 p.
300. **Kadri Toome**. Homing peptides for targeting of brain diseases. Tartu, 2020, 152 p.
301. **Maarja Hallik**. Pharmacokinetics and pharmacodynamics of inotropic drugs in neonates. Tartu, 2020, 172 p.
302. **Raili Müller**. Cardiometabolic risk profile and body composition in early rheumatoid arthritis. Tartu, 2020, 133 p.
303. **Sergo Kasvandik**. The role of proteomic changes in endometrial cells – from the perspective of fertility and endometriosis. Tartu, 2020, 191 p.
304. **Epp Kaleviste**. Genetic variants revealing the role of STAT1/STAT3 signaling cytokines in immune protection and pathology. Tartu, 2020, 189 p.
305. **Sten Saar**. Epidemiology of severe injuries in Estonia. Tartu, 2020, 104 p.
306. **Kati Braschinsky**. Epidemiology of primary headaches in Estonia and applicability of web-based solutions in headache epidemiology research. Tartu, 2020, 129 p.
307. **Helen Vaher**. MicroRNAs in the regulation of keratinocyte responses in *psoriasis vulgaris* and atopic dermatitis. Tartu, 2020, 242 p.
308. **Liisi Raam**. Molecular Alterations in the Pathogenesis of Two Chronic Dermatoses – Vitiligo and Psoriasis. Tartu, 2020, 164 p.
309. **Artur Vetkas**. Long-term quality of life, emotional health, and associated factors in patients after aneurysmal subarachnoid haemorrhage. Tartu, 2020, 127 p.
310. **Teele Kasepalu**. Effects of remote ischaemic preconditioning on organ damage and acylcarnitines' metabolism in vascular surgery. Tartu, 2020, 130 p.
311. **Prakash Lingasamy**. Development of multitargeted tumor penetrating peptides. Tartu, 2020, 246 p.
312. **Lille Kurvits.** Parkinson's disease as a multisystem disorder: whole transcriptome study in Parkinson's disease patients' skin and blood. Tartu, 2021, 142 p.
313. **Mariliis Põld.** Smoking, attitudes towards smoking behaviour, and nicotine dependence among physicians in Estonia: cross-sectional surveys 1982–2014. Tartu, 2021, 172 p.
314. **Triin Kikas**. Single nucleotide variants affecting placental gene expression and pregnancy outcome. Tartu, 2021, 160 p.
315. **Hedda Lippus-Metsaots**. Interpersonal violence in Estonia: prevalence, impact on health and health behaviour. Tartu, 2021, 172 p.

316. **Georgi Dzaparidze.** Quantification and evaluation of the diagnostic significance of adenocarcinoma-associated microenvironmental changes in the prostate using modern digital pathology solutions. Tartu, 2021, 132 p.
317. **Tuuli Sedman.** New avenues for GLP1 receptor agonists in the treatment of diabetes. Tartu, 2021, 118 p.
318. **Martin Padar.** Enteral nutrition, gastrointestinal dysfunction and intestinal biomarkers in critically ill patients. Tartu, 2021, 189 p.
319. **Siim Schneider.** Risk factors, etiology and long-term outcome in young ischemic stroke patients in Estonia. Tartu, 2021, 131 p.
320. **Konstantin Ridnõi.** Implementation and effectiveness of new prenatal diagnostic strategies in Estonia. Tartu, 2021, 191 p.
321. **Risto Vaikjärv.** Etiopathogenetic and clinical aspects of peritonsillar abscess. Tartu, 2021, 115 p.
322. **Liis Preem.** Design and characterization of antibacterial electrospun drug delivery systems for wound infections. Tartu, 2022, 220 p.
323. **Keerthie Dissanayake.** Preimplantation embryo-derived extracellular vesicles: potential as an embryo quality marker and their role during the embryo-maternal communication. Tartu, 2022, 203 p.
324. **Laura Viidik.** 3D printing in pharmaceutics: a new avenue for fabricating therapeutic drug delivery systems. Tartu, 2022, 139 p.
325. **Kasun Godakumara.** Extracellular vesicle mediated embryo-maternal communication – A tool for evaluating functional competency of preimplantation embryos. Tartu, 2022, 176 p.
326. **Hindrek Teder.** Developing computational methods and workflows for targeted and whole-genome sequencing based non-invasive prenatal testing. Tartu, 2022, 138 p.
327. **Jana Tuusov.** Deaths caused by alcohol, psychotropic and other substances in Estonia: evidence based on forensic autopsies. Tartu, 2022, 157 p.
328. **Heigo Reima.** Colorectal cancer care and outcomes – evaluation and possibilities for improvement in Estonia. Tartu, 2022, 146 p.
329. **Liisa Kuhi.** A contribution of biomarker collagen type II neoepitope C2C in urine to the diagnosis and prognosis of knee osteoarthritis. Tartu, 2022, 157 p.
330. **Reeli Tamme.** Associations between pubertal hormones and physical activity levels, and subsequent bone mineral characteristics: a longitudinal study of boys aged 12–18. Tartu, 2022, 118 p.
331. **Deniss Sõritsa.** The impact of endometriosis and physical activity on female reproduction. Tartu, 2022, 152 p.
332. **Mohammad Mehedi Hasan.** Characterization of follicular fluid-derived extracellular vesicles and their contribution to periconception environment. Tartu, 2022, 194 p.
333. **Priya Kulkarni.** Osteoarthritis pathogenesis: an immunological passage through synovium-synovial fluid axis. Tartu, 2022, 268 p.

334. **Nigul Ilves.** Brain plasticity and network reorganization in children with perinatal stroke: a functional magnetic resonance imaging study. Tartu, 2022, 169 p.
335. **Marko Murruste.** Short- and long-term outcomes of surgical management of chronic pancreatitis. Tartu, 2022, 180 p.
336. **Marilin Ivask.** Transcriptomic and metabolic changes in the WFS1-deficient mouse model. Tartu, 2022, 158 p.
337. **Jüri Lieberg.** Results of surgical treatment and role of biomarkers in pathogenesis and risk prediction in patients with abdominal aortic aneurysm and peripheral artery disease. Tartu, 2022, 160 p.
338. **Sanna Puusepp.** Comparison of molecular genetics and morphological findings of childhood-onset neuromuscular disorders. Tartu, 2022, 216 p.
339. **Khan Nguyen Viet.** Chemical composition and bioactivity of extracts and constituents isolated from the medicinal plants in Vietnam and their nanotechnology-based delivery systems. Tartu, 2023, 172 p.
340. **Getnet Balcha Midekessa.** Towards understanding the colloidal stability and detection of Extracellular Vesicles. Tartu, 2023, 172 p.
341. **Kristiina Sepp.** Competency-based and person-centred community pharmacy practice – development and implementation in Estonia. Tartu, 2023, 242 p.
342. **Linda Sõber.** Impact of thyroid disease and surgery on patient's quality of voice and swallowing. Tartu, 2023, 114 p.
343. **Anni Lepland.** Precision targeting of tumour-associated macrophages in triple negative breast cancer. Tartu, 2023, 160 p.
344. **Sirje Sammul.** Prevalence and risk factors of arterial hypertension and cardiovascular mortality: 13-year longitudinal study among 35- and 55-year-old adults in Estonia and Sweden. Tartu, 2023, 158 p.
345. **Maarjaliis Paavo**. Short-Wavelength and Near-Infrared Autofluorescence Imaging in Recessive Stargardt Disease, Choroideremia, *PROM1*-Macular Dystrophy and Ocular Albinism. Tartu, 2023, 202 p.
346. **Kaspar Ratnik.** Development of predictive multimarker test for preeclampsia in early and late pregnancy. Tartu, 2023, 134 p.
347. **Kärt Simre.** Development of coeliac disease in two populations with different environmental backgrounds. Tartu, 2023, 161 p.
348. **Qurat Ul Ain Reshi.** Characterization of the maternal reproductive tract and spermatozoa communication during periconception period via extracellular vesicles. Tartu, 2023, 182 p.
349. **Stanislav Tjagur.** *Mycoplasma genitalium* and other sexually transmitted infections causing urethritis – their prevalence, impact on male fertility parameters and prostate health. Tartu, 2023, 225 p.
350. **Lagle Lehes.** The first study of voice and resonance related treatment outcomes of Estonian cleft palate children. Tartu, 2023, 126 p.
351. **Liis Ilves.** Metabolomic profiling of chronic inflammatory skin diseases. Tartu, 2023, 146 p.

352. **Marina Šunina**. Flow cytometric analysis of T and B cell properties in healthy donors and subjects with vitiligo. Tartu, 2023, 164 p.
353. **Jaanus Suumann.** Gastric biomarkers and their dynamics as a less invasive method to evaluate stomach health in bariatric surgery patients. Tartu, 2023, 122 p.
354. **Ele Hanson.** Clinical and biochemical markers for the prediction and early diagnosis of pregnancy related complications. Tartu, 2023, 145 p.
355. **Priit Pauklin.** Hemodynamic and biochemical characteristics of patients with atrial fibrillation and anticoagulation of ≥65-year-old patients with atrial fibrillation in Estonia. Tartu, 2023, 144 p.
356. **Triinu Keskpaik.** Quality Indicators and Non-Ischemic Myocardial Injury in Emergency Medicine. Tartu, 2023, 121 p.
357. **Laura Roht.** Hereditary colorectal cancer syndromes in Estonia. Tartu, 2023, 178 p.
358. **Norman Ilves.** Risk factors and onset time of periventricular hemorrhagic infarction in preterm born children and periventricular venous infarction in term born children. Tartu, 2024, 177 p.
359. **Edgar Lipping.** Postoperative antibacterial therapy in complicated appendicitis and appendectomy in pregnancy. Tartu, 2024, 121 p.
360. **Celia Teresa Pozo Ramos.** Preparation and assessment of antimicrobial electrospun matrices for prospective applications in wound healing. Tartu, 2024, 203 p.
361. **Karl Kuusik.** Effects of remote ischaemic preconditioning on arterial stiffness, organ damage and metabolomic profile in patients with lower extremity artery disease. Tartu, 2024, 173 p.
362. **Kelli Somelar-Duracz.** The molecular and cellular mechanisms of brain plasticity impairing factors. Tartu, 2024, 245 p.
363. **Aleksei Baburin.** Breast cancer incidence, mortality and survival in Estonia in the context of health care system changes and screening. Tartu, 2024, 130 p.
364. **Marina Loid.** Molecular and cellular determinants of healthy receptive and aged endometrium. Tartu, 2024, 159 p.
365. **Ulvi Vaher.** Epilepsy after ischemic perinatal stroke in term born children: neuroimaging predictors, clinical course and cognitive outcome. Tartu, 2024, 160 p.
366. **Allan Tobi.** Development of Smart Nanoparticles for Experimental Treatment of Cancer. Tartu, 2024, 160 p.
367. **Leho Rips.** The influence of vitamin D on the physical performance of conscripts in the Estonian Defence Forces. Tartu, 2024, 147 p.
368. **Kati Kärberg.** Factors and markers predicting subclinical atherosclerosis in type 2 diabetes. Tartu, 2024, 161 p.
369. **Valeria Sidorenko.** Novel anthracycline-loaded nanoparticles for precision cancer therapy. Tartu, 2024, 183 p.

370. **Kadri Kõivumägi.** Acute gastroenteritis hospitalizations in Estonia after implementation of universal mass vaccination against rotavirus. Tartu, 2024, 150 p.
371. **Ingrid Oit-Wiscombe.** Genetic markers of enzymatics in the pathogenesis of chronic obstructive pulmonary disease as a systemic disease and the effects of antioxidant peptides. Tartu, 2025, 170 p.
372. **Gerli Mõts.** Ethical issues in nursing before and during the COVID-19 pandemic: a multi-method study. Tartu, 2025, 150 p.
373. **Annika Valner.** Changes in structure and function of extremities in early rheumatoid arthritis. Tartu, 2025, 129 p.
374. **Meruert Sarsenova.** Molecular and cellular landscape of endometriosis. Tartu, 2025, 157 p.
375. **Mailis Liiv.** Role of mitochondrial dynamics in Wolfram syndrome. Tartu, 2025, 180 p.
376. **Marta Velgan.** Addressing the family physician shortage: Career and migration intentions in Estonia and Europe. Tartu, 2025, 161 p.
377. **Ihor Filippov.** Single-cell data analysis in immunology: from technology to applications. Tartu, 2025, 204 p.
378. **Anna Tisler.** HPV-related cancers among people living with HIV and transition towards risk-based prevention. Tartu, 2025, 122 p.
379. **Priit Paluoja.** Computational methods for NIPT-based aneuploidy and microdeletion screening. Tartu, 2025, 154 p.
380. **Ere Uibu.** Utilisation and outcomes of patient safety incident reporting and learning in hospitals from a nursing perspective: a multi-method study. Tartu, 2025, 142 p.
381. **Jane Idavainu.** Health effects of environmental contamination in the oil shale industry region of Estonia. Tartu, 2025, 164 p.
382. **Taavi Torga.** Association of molecular markers CILP-2, DDR2 and C2C with the severity of tissue damage in knee osteoarthritis. Tartu, 2025, 156 p.
383. **Karl-Gunnar Isanda.** The impact of frailty on outcomes following emergency laparotomy: Enhancing risk prediction and clinical decision-making. Tartu, 2025, 142 p.
384. **Kristina Isanda.** Natural history of non-functioning pituitary microadenomas and venous thromboembolism in patients with pituitary adenomas and Cushing syndrome. Tartu, 2025, 156 p.
385. **Ankita Sunil Lawarde.** Integrative omics approaches for analyzing endometrial pathologies and cancer classification. Tartu, 2025, 142 p.