INSTITUT FÜR KARDIOGENETIK

Identification of genetic risk variants for atherosclerosis using oxidative stress assays in vascular smooth muscle cells and bioinformatic approaches

Identification genetischer Risikovarianten für Artheriosklerose via oxidativem Stress Assay in glatten Muskulaturzellen und bioinformatische Ansätze

Masterarbeit

verfasst am
Institut für Kardiogenetik

im Rahmen des Studiengangs
Molecular Life Science
der Universität zu Lübeck

vorgelegt von
Torben Falk

ausgegeben und betreut von
Prof. Dr. Jeanette Erdmann

mit Unterstützung von
Dr. Tobias Reinberger

Lübeck, den 21. Juli 2022

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

_____

Torben Falk

Zusammenfassung

Ich muss das Ding wohl irgendwann auch noch in Deutsch schreiben...

Abstract

Placeholder, this is more or less what I am doing:

I am currently writing my Master thesis at the university of Lübeck at the Institute for Cardiogenetics on the topic of "Identification of genetic risk variants for atherosclerosis using oxidative stress assays in vascular smooth muscle cells and bioinformatic approaches":

Coronary artery disease (CAD) describes the arterial build-up of fatty deposits to a point where the blood supply to the heart gets interrupted. It is one of the major causes of death worldwide. Risk factors for CAD are typical lifestyle factors like smoking or physical inactivity, but also include hereditary factors (**cdcCoronaryArteryDisease2021**; **CoronaryHeartDisease2018**). These can provide access to the molecular pathology of the disease. One amazing resource for studying these interactions are genome wide association studies (GWAS). Unfortunately, GWAS are just the first step in a longer journey of establishing causal loci to gene links, uncovering the molecular basis of disease, and implementing tools for clinical risk prediction. A plethora of follow-up analyses (postGWAS) can and need to be performed (**lichouFunctionalStudiesGWAS2020a**).

We hypothesize that oxidative stress in smooth muscle cells plays a role in stability of atherosclerotic plaques. For this reason, I am cultivating and differentiating primary human smooth muscle cells and characterizing them using oxidative stress assay, qPCR, seahorse assay & immunofluorescence (IF).

Additionally, I am working with GWAS data on CAD (**aragamDiscoverySystematicCharacterization2021a**). Curating further publicly available data that can be used for bioinformatic follow-up analyses like the enrichment for involved tissues. Further, I am using the data to build a web application that allows co-visualization and visual exploration.

Acknowledgements

Daaanke an alle!

Contents

1

Introduction

1.1 Coronary artery disease

Coronay artery disease (CAD) is among the leading causes of death in the (western) worldwide, being prevalent in about 6.7 % of American adults and killing more than 350'000 people in the USA in 2019 alone (Disease Control and Prevention, 2022; Fryar, 2012). CAD is characterized by the build up of fatty plaques in the arteries leading to the heart. This process, also called atherosclerosis, can interrupt the blood supply to the heart (National Health Service, 24 Oct 2017, 4:45 p.m.). Its most common complication is myocardial infarction (MI), which usually manifests as chest pain (angina) and can cause serious damage to the heart muscle. Next to common and well known life-style factors like tobacco use or physical inactivity, CAD risk additionally has a hereditary component (Task Force Members et al., 2013).

1.2 Muscle Cells in CAD

The lumen of a typical blood vessel is surrounded by three distinct layers. The outer adventitia is rich in connective tissue, shapes the vessel and wraps the media. The media contains vascular smooth muscle cells (VSMCs). VSMCs of the media are required to mediate vasodilation and vasoconstriction according to signaled requirements. The inner layer consists of endothelial cells that define the lumen of the vessel. (Tucker, Arora, and Mahajan, 2022; Yap et al., 2021)

For the longest time, the role of VSMCs in the development and progression of atherosclerosis has been underestimated and over simplified. They have simply been considered to be either promoting of arteriosclerosis or beneficial for plaque stability. Only with the emergence of novel and exciting technologies like single cell transcriptomics and lineage tracking, this view is changing to a more differentiated one. (Grootaert and Bennett, 2021; Yap et al., 2021) The study of VSMCs in arteriosclerosis is rapidly evolving, and the underlying models being adjusted accordingly. The black and white idea of VSMCs in arteriosclerosis existing either as a differentiated (contractile) phenotype or as a dedifferentiated (synthetic) phenotype, is making place for a model that considers a diverse set of dedifferentiated phenotypes (Grootaert and Bennett, 2021; Yap et al., 2021). The phenotypic switch describes the down regulation of contractile markers and can give rise to a diverse bouquet of different

phenotypes which can be found in the fibrous cap and plaque core (Grootaert and Bennett, 2021). Thier number as well as their impact on disease progression are still the subject of intensive research.

Two external stimuli that seem to play central roles as cytokines determining the fate of VSMCs in artherogenesis are TGFβ & platelet-derived growth factor-BB (PDGF-BB).

1.3 TGFβ Signaling

TGFβ Signaling in General

TGFβ is a summarizing term for a super family of cytokines, the most prominent of which is TGFβ1. After secretion and activation, the active TGFβ dimer binds to a heteromeric receptor complex. The intracellular signaling is mainly implemented via Smad transcription factors. The effects of TGFβ is highly dependent on the cell type and can even be pleiotropic for cells of the same type. The most prominent function of TGFβ is its role in anti-inflammatory regulation of immune cells. (Goumans and Dijke, 2018; Batlle and Massagué, 2019)

TGFβ Signaling in VSMCs & atherosclerosis

In the context of VSMCs, TGFβ promotes proliferation and hypertrophy. Further, it promotes VSMC differentiation, via elevation of contractile gene expression as well as the down regulation of Kruppel-like factor 4 (KLF4) (Davis-Dusenbery et al., 2011), a transcription factor (TF) prominent for its application in inducing pluripotency (Takahashi et al., 2007) that is also required for phenotype switching. This way hindering (Davis-Dusenbery et al., 2011) or potentially reversing phenotype switching (Pan et al., 2020).

1.4 PDGF Signaling

PDGF Signaling in General

Five different PDGF isoforms have been identified that form as dimeric combination of four distinct polypeptide chains (PDGF-AA, PDGF-AB, PDGF-BB, PDGF-CC & PDGF-DD). All five isoforms bind to tyrosine kinase receptors (Platelet-derived growth factor receptor (PDGFR)α& PDGFRβ). Upon activation, the receptor dimerizes, allowing autophosphorylation which activates the kinase domain and creates binding sites for signaling molecules. The active receptor is involved in a plethora of prominent messaging pathways like the mitogen activated protein (MAP)-kinase pathway, phosphatidylinositol 3'-kinase (PI3K)-signaling or signal transducers and activators of transcription (STAT)-signaling. All these pathways are ultimately involved in the promotion of cellular proliferation, survival and migration (Chen, Chen, and He, 2013; Heldin, 2013; Hu and Huang, 2015).

The predominantly expressed by endothelial cells seems to be PDGF-BB (Andrae, Gallini, and Betsholtz, 2008; Heldin, 2013) which acts as a paracrine activator for VSMCs and other mesenchymal cells (Heldin, 2013). Signaling via PDGF-BB and the PDGFRβplays an important role in development of multiple tissues, e.g. in the development of the cardio

vascular system (Levéen et al., 1994). After completed development, PDGF-BB picks up an important role in wound healing processes (Robson et al., 1992). The role of PDGFRβsignaling in pathologic processes like cancer or cardio vascular disease has been a subject of extensive study for decades (Heldin, 2013; Raines, 2004).

PDGF Signaling in VSMCs & atherosclerosis

In the context of VSMCs, PDGF-BB was shown to increase KLF4 levels, which results in up-regulation of mesenchymal markers as well as the loss of contractile markers. Ultimately, serving as an external stimulus for proliferation and phenotype switching (Yap et al., 2021).

Similarly to the overall role of VSMCs in arteriosclerosis, the role of PDGF-BB is still the subject of extensive study. All PDGF isoforms are abundantly found in arteriosclerotic cell walls, and further PDGFR expression is elevated in affected vessels (Hu and Huang, 2015). For a long time PDGF signaling and inflammation has been assumed to be disease promoting (Andrae, Gallini, and Betsholtz, 2008; Chen, Chen, and He, 2013; He et al., 2015; Hu and Huang, 2015) and in 2015 He et al. (2015) showed that PDGFRβsignaling in mouse model leads to inflammation and increased plaque formation. In contrast to this consensus, Newman et al. (2021) were recently able to demonstrate, that sustained signaling via PDGFR$\beta$is required for VSMC involvement. They Further observed, again in mouse model, that lack of VSMC involvement during plaque formation, can be temporarily compensated by non-VSMC-derived cells, but long-term leads to instability of arteriosclerotic lesions.

ROS in PDGF Signaling

reactive oxygen species (ROS) is a broad term for a class of highly reactive molecules derived from elemental oxygen (O_2). They are traditionally infamous for the damage they can do to proteins and nucleic acids when not kept in check, potentially causing irreparable damage leading to cell death. Recently, this perception has been shifting, and specially hydrogen peroxide (H_2O_2) and superoxide anion radical ($O_2^{\cdot -}$), are being recognized for their role in cellular signaling. (Sies and Jones, 2020)

Human cells contain dozens of enzymes, which are capable of generating ROS and enzymatically maintain a steady redox state (Sies and Jones, 2020). H_2O_2 and $O_2^{\cdot -}$ serve as important second messengers in the central nervous system (**nayerniaNewInsightsNOX2014**) or in the reapir of vascular lesions (Andrae, Gallini, and Betsholtz, 2008). Interestingly, the generation of ROS as a second messenger gets triggered by stimulation withPDGF-BB (Sundaresan et al., 1995; Bouzigues et al., 2014).

1.5 GWAS

The hereditary components of disease onset and progression can provide access to its pathology on a molecular level.

GWAS

An amazing resource for getting a first glance into these interactions are genome wide association study (GWAS), a method that allows for the identification of genetic variants

associated with a phenotype.

While GWAS were initially an extraordinary endeavor, requiring the evaluation of hundreds or thousands of participants, they have gotten a lot more accessible with the availability of genetic data from public biobanks. After profiling the cohort on a genomic level (usually via microarrays but whole genome sequencing (WGS) is probably the future) and phenotypically, the collected data needs to pass through several steps of quality control, e.g. for the removal of rare variants, miss matched phenotypes, etc. Afterwards, variants which were not directly analyzed are inferred from a reference. The final step of the initial analysis is the statistical mode, where a regression model is used to test for association of all variants with the phenotype in question. It is crucial to be completely aware of potential biases, that might have introduced in this process, some of which (like age, sex or ancestry) can and need to be included as covariant in the used model. (Uffelmann et al., 2021; Flint, 2013) The model will output a list of p-values, effect sizes (and their direction) for all tested variants. A GWAS is the first important step in determining causal variants for disease and therefore a first glimpse into the molecular biology of the observed phenotype (Uffelmann et al., 2021).

postGWAS

Unfortunately, GWAS are just the first step in a longer journey of establishing causal loci to gene links, uncovering the molecular basis of disease, and implementing tools for clinical risk prediction. A plethora of follow-up analyses (postGWAS) can and need to be performed to determine a set of credible variants and to assess their molecular mechanism.

The first important follow-up, that is usually done immediately, is fine-mapping. Due to the complex linage disequilibrium (LD) of variants in the human genome (see sec. 1.6), identified loci in GWAS unusually do not contain a single significant variant, but are made up of a potentially large set of linked variants. Fine-mapping describes the process of identifying the actually causal variant in this mess. Multiple very sophisticated statistical methods have been developed, the most popular of which is Bayesian modelling, which yields variant specific posterior inclusion probabilitys (PIPs) that form a credible set of potentially causal variants. It is important to remember, that fine-mapping is not a solved problem, available methods are continuously improving and will most likely keep getting more complex with increasing complexity of the studied phenotypes. Further fine-mapping is a statistical approach which will never be able to determine causality! (Schaid, Chen, and Larson, 2018; Uffelmann et al., 2021)

After the identification of likely causal variants, the next steps aim to gain information on their effect in determining the analyzed phenotype. Variants still require mapping to impacted genes, associated pathways and relevant tissues to get a glance of the complete image. For these steps no standard protocols exists and the procedure highly depends on the genomic context of the variant. Coding variants are and offer themselves to be immediately studied on a protein level, while non-coding variants are greatly benefit from the consultation of more high throughput data in the form of e.g. expression quantitative trait locis (eQTLs) (Uffelmann et al., 2021).

Finally, all the previous results can and need to be taken back to the wet lab, to verify and extend the ideas derived from statistical models. Utilizing all the recent great advances in molecular and cellular biology, such as the development of increasingly comprehensive *in vitro* models as well and their manipulation via methods like Clustered Regularly Interspaced

Short Palindromic Repeats (CRISPR)-Cas gene-editing (Lichou and Trynka, 2020).

1.6 Complementary High Through Put Methods

The development of high trough put methods as well as the great increase in computing power over the last few years have spawned a plethora of incredible datasets that already have been and can be further utilized for postGWAS analyses. A short overview of some definitions and methods mentioned in this thesis can be found in the following paragraphs:

Linkage Disequilibrium

LD is a parameter from populations genetics that describes the non-random association of two or more alleles. The LD is often quantified using the correlation coefficient r^2 (Slatkin, 2008).

$$D_{AB} = p_{AB} - p_A p_B$$

$$r^2 = \frac{D_{AB}^2}{p_A(1-p_A) \times p_B(1-p_B)}$$

Where p_A and p_B is the frequency of the alleles A and B respectively. p_{AB} is the frequency of the AB haplotype.

The LD becomes important in the context of GWAS because identified SNPs often do not occur in isolation, but a network of linked and significant variants can span large haplotype blocks in the genome (Slatkin, 2008).

Locus To Gene Scores

Problems of interpretation of GWAS data are already described in section 1.5. link to gene (L2G) scores are an attempt at overcoming the challenges of establishing causal relationships between variants and genes. The authors employed a machine learning-model to integrate fine-mapping with functional genomics data and *in silico* predictions to link GWAS loci to their target genes. The output L2G scores are calibrated to represent the probability (0, 1). (Mountjoy et al., 2021)

Regulatory Build

The ensembl regulatory build compiles a summary of putative regulatory regions found in the human genome. It is constructed from publically available data on epigenetic marks and TF binding sides. It contains promotors, proximal enhancers, distal enhancers and CCCTC binding factor (CTCF) binding sites. (Zerbino et al., 2015)

ENCODE candidate cis-regulatroy elements (cCREs)

Very similarly, the ENCyclopedia Of DNA Elements project (ENCODE) project summarizes DNA accessibility and chromatin modification data into cCREs. Regions showing high DNase signal are further annotated to be proximal enhancer-like elements (pELS) or distal enhancer-like elements (dELS), promoter-like elements (PLS), other regions with high histone 3 lysine 4 trimethylathion (H3K4me3) signal (which might represent poised or non-canonical promotors), or CTCF-only elements based on the existence of H3K4me3, histone 3 lysine 27 acetylation (H3K27ac) and CTCF signals. (Moore et al., 2020)

ATAC-seq

assay for transposase-accessible chromatin using sequencing (ATAC-seq) is a method to access chromatin accessibility in the genome. ATAC-seq utilizes the hyperactive Tn5 transponase to insert sequencing adapters into accessible regions of chromosome. DNA is purified and amplified via polymerase chain reaction (PCR) and then sequenced. Mapping sites with insertions to the genome allows for the identification of highly accessible genomic regions. (Buenrostro et al., 2013; Buenrostro et al., 2015a)

PCR amplification of the DNA makes this method extremely sensitive. Pushing the requirement of biomaterial to the minimum, ATAC-seq is applicable on a single cell level. For scATAC-seq, individual cells are isolated and their DNA tagged with barcoded primers during the PCR. These barcodes allow mapping of ATAC-seq data to the isolated cells. (Buenrostro et al., 2015b)

ABC Model

The activity by contect (ABC) model grants insights into potential cell specific enhancer-gene interactions based on chromatin state, outperforming previously used methods (Fulco et al., 2019; Nasser et al., 2021).

$$ABC\ score_{E,G} = \frac{A_E \times C_{E,G}}{\sum_{all\ elements\ e\ within\ 5\ Mb\ of\ G} A_e \times C_{e,G}}$$

Generally speaking, the model incorporates the activity of an enhancer A_E as well as contacts with the gene of interest $C_{E,G}$, normalized by the total effect of all elements in the area (Fulco et al., 2019; Nasser et al., 2021).

Hi-C & TADs

Hi-C is a method for mapping chromosomal conformation. For achieve this, genome associated proteins are cross-linked with formaldehyde, DNA is digested with restriction enzymes and generated overhangs are filled in with biotinylated nucleotides. The resulting fragments are ligated to covalently link DNA fragments, which were originally in close spatial proximity. The DNA is purified and fragmented, allowing the pulldown of fragments containing junctions sites via the filled in biotin tags. After sequencing of the enriched fragments, their

sequences are mapped to the genome, identifying interacting DNA regions. (Lieberman-Aiden et al., 2009; Wit and Laat, 2012)

Looking at Hi-C data, topologically associated domains (TADs) were identified to be a basic feature of genome organization with an average size of 880 kb (Dixon et al., 2012; Wang et al., 2018b). What makes TADs of such high interest is the fact that interactions of DNA sequences are usually confined within TADs. Tissue-specific genes and their enhancers are usually found in the middle of TADs, while the edges enrich for housekeeping genes and CTCF binding sides which might serves as insulators between different domains (Pombo and Dillon, 2015).

1.7 Aim of the thesis

The aims of this thesis are split into two quite distinct projects that both ultimately aim to contribute to a better understanding of artheriosclerosis and CAD:

- The split role of PDGF-BB during progression of artheriosclerosis (see section 1.4), indicates that PDGF-BB signaling is neither completely benefitial nor disadvantageous to diseases but there is an optimal dose. Combining this theory with the fact that **ROS!** (**ROS!**) are involved in PDGF-BB signaling and also highly associated with artheriosclerosis (Burtenshaw et al., 2019), we hypothesized, that PDGF-BB signaling may cause oxidative stress, this way contributing to disease progression. The first part of this thesis will deal with the *in vitro* characterization of PDGF-BB stimulated VSMCs and the establishment of a robust assay for oxidative stress in VSMCs.
- The summary statistics from **aragamDiscoverySystematicCharacterization2021** are a great resource. To make this data more accessible we are building an interactive web-based visualization tool, co-visualizing the data with different kinds of annotations like gene products, associated phenotypes from other GWAS or putative regulatory elements.
- Finally, we'll use the data curated for the visualization tool for postGWAS studies.

  Have fun with my thesis, this still is a mess...

2

Material

2.1 Manufactors

Manufacturer	Seat
Agilent Technologies, Inc.	Santa Clara, CA, USA
Assistent - look this up	?!, DE
BRAND GMBH & Co. KG	Wertheim, DE
ChemoMetec A/S	Allerod, DK
Eppendorf AG	Hamburg, DE
Heraeus Holding GmbH	Hanau, DE
Merck KGaA	Darmstadt, DE
Keyence Corporation	Osaka, JP
Kisker Biotech GmbH & Co. KG	?!, DE
Sarstedt AG & Co.	Nürnberg, DE
Sigma-Aldrich Co. LLC.	St. Louis, MO, USA
Thermo Fisher Scientific Inc.	Waltham, MA, USA
Pepro Tech -> look this up	?!
Pechiney Plastic Packaging, Inc.	Chicago, IL, USA
J.T. Baker -> look up	?!
GFL mbH	Burgwedel, DE
AB -> look this up	Warrington, UK
Gibco BRL (Brand of Thermo Fisher Scientific Inc.	Gaithersburg, MD, USA
Ibidi -> loo this up	

Continued on next page

(Continued)

Manufacturer	Seat
Invitrogen™ (Brand of Thermo Fisher Scientific Inc.)	
Lonza Group AG	Basel, CHE
Biosell -> look this up	
New England Biolabs GmbH	Ipswich, MA, USA
Nikon Corporation	Minato, JP
Sartorius ?!	Göttingen, DE
SensoQuest GmbH	G¨ottingen, DE
HANNA Instruments	??
Heidolph Instruments Labortechnik	Schwabach, DE
Ziegra Eis Machinen	?!, DE

2.2 Antibodies

Name	Species	Manufacturer
8oxoG	?	?
other oxStress	?	?
Fibronectin	?	?
secondary ones	?	?

2.3 Celllines

Name	Celltype	Manufacturer
Human Aortic Smooth Muscle Cell (HAoSMC)	prim. human cell	

2.4 Primer

Target	Name	Sequence
CNN1	Fw	5'-seq-3'
	Rv	5'-seq-3'
GAPDH	Fw	5'-seq-3'

Continued on next page

(Continued)

Target	Name	Sequence
MMP9	Rv	5'-seq-3'
	Fw	5'-seq-3'
	Rv	5'-seq-3'

2.5 Chemicals

Name	Manufacturer
5X First Strand Buffer	Invitrogen™
Antimycin A	Sigma-Aldrich Co. LLC.
BSA	?!?
CellROX	Thermo Fisher Scientific Inc.
Collagen Type I, rat tail	Ibidi -> loo this up
dNTP Mix	AB -> look this up
DTT	Invitrogen™
Ethanol (concentration?)	J.T. Baker -> look up
FCCP	Sigma-Aldrich Co. LLC.
D-Glucose	Sigma-Aldrich Co. LLC.
GlutaMAX™-I	Gibco BRL
Hoechst	?!?
IL-1	Pepro Tech -> look this up
M-MLV RT	Thermo Fisher Scientific Inc.
NAC	Sigma-Aldrich Co. LLC.
NaHCO3	Carl Roth GmbH + Co. KG
NaOH, 1 N	Carl Roth GmbH + Co. KG
Oligomycin	Sigma-Aldrich Co. LLC.
Roth Hexanukleotid Random-Primer	Carl Roth GmbH + Co. KG
PBS	Lonza Group AG
PDGF-BB	Pepro Tech -> look this up
Sodium Pyruvate	Gibco BRL
RiboLock R	Thermo Fisher Scientific Inc.
Seahorse XF calibrant	Agilent Technologies, Inc.
PowerUp™SYBR™GREEN Master Mix	Thermo Fisher Scientific Inc.

Continued on next page

(Continued)

Name	Manufacturer
TGFβ	Pepro Tech -> look this up

2.6 Media, Supplements

Name	Manufacturer
FBS Gold Plus	Biosell -> look this up
Medium 231 (M231)	Gibco BRL
Smooth Muscle Cell Growth Supplement	?!?
XF Base Medium	Agilent Technologies, Inc.

2.7 Solutions

Name	Manufacturer
Interleukin 1 beta (IL-1)	IL-1 0.1 % BSA in PBS
N-Acetylcystein (NAC)	0.25 M NAC in water, ~pH 7
Platelet-derived growth factor-BB (PDGF-BB)	PDGF-BB 0.1 % BSA in PBS
Transforming Growth Factor beta (TGF)	TGFβ 0.1 % BSA in PBS

2.8 Kits

Kit	Manufacturer
Total RNA Purification Kit	Jena Bioscience GmbH

2.9 Consumables

Name	Manufacturer
Quali-PCR-Tubes 0,2 mL	Kisker Biotech GmbH & Co. KG
Quali-PCR-Tubes 0,5 mL	Kisker Biotech GmbH & Co. KG
SafeSeal Gefäß 1,5 mL	Sarstedt AG & Co.
SafeSeal Gefäß 1,5 mL	Sarstedt AG & Co.
SafeSeal Gefäß 5 mL	Sarstedt AG & Co.
Nunc Cell-Culture Treated Multidish 24	Thermo Fisher Scientific Inc.
Agilent Seahorse XF24 Cell Culture Microplate	Agilent Technologies, Inc.
Agilent Seahorse XF24 Extracellular Flux Assy Kit	Agilent Technologies, Inc.
384 Well Multiply PCR plates	
Pasteurpipetten ISO 7712	Assistent - look this up
Pipette tip 20 µL	Sarstedt AG & Co.
Pipette tip 200 µL	Sarstedt AG & Co.
Pipette tip 1000 µL	Sarstedt AG & Co.
Filter tip 20 µL	Sarstedt AG & Co.
Filter tip 200 µL	Sarstedt AG & Co.
Filter tip 1000 µL	Sarstedt AG & Co.
BD Discardit™II	?!?
0,20 µm filter	?!?
Via1-Casette™	ChemoMetec A/S
Tube 15 ml	Sarstedt AG & Co.
Tube 50 ml	Sarstedt AG & Co.
Serological pipette 5 mL	Sarstedt AG & Co.
Serological pipette 10 mL	Sarstedt AG & Co.
Serological pipette 25 mL	Sarstedt AG & Co.
Serological pipette 50 mL	Sarstedt AG & Co.
Parafilm®M	Pechiney Plastic Packaging, Inc.
TC Flask T75, Cell+, Vented Cap	Sarstedt AG & Co.
CRYSTAL qPCR-Folie	New England Biolabs GmbH

2.10 Devices

Name	Manufacturer
Bench Heraus Bench 2	Heraeus Holding GmbH
Research pipettes (2.5µL, 10 µL, 100 µL, 1000µL)	Eppendorf AG
Hera Cell	Heraeus Holding GmbH
Hera Cell 150	Heraeus Holding GmbH
NucleoCounter NC-200	ChemoMetec A/S
Incubation/Inactivation bath 1083	GFL mbH
Centrifuge 5702 R	Eppendorf AG
Eclipse TS100	Nikon Corporation
pipet X rainin	?!?
Centrifuge 5415 R	Eppendorf AG
Rotana 460 R	Andreas Hettich GmbH & Co. KG
XF24 Extracellular Flux Analyzer	Agilent Technologies, Inc.
7900HT Fast Real-Time PCR System	Thermo Fisher Scientific Inc.
LA 120 S	Sartorius ?!
Reax Top	Heidolph Instruments Labortechnik
MR 3001	Heidolph Instruments Labortechnik
BZ-X810 All-in-One Flourescence Microscope	Keyence Corporation
BZ-X800 All-in-One Flourescence Microscope POWER	Keyence Corporation
SensoQuest labcycler	SensoQuest GmbH
pH 221 Mircoprocessor pH Meter	HANNA Instruments
NanoDrop 2000	Thermo Fisher Scientific Inc.
Ice maschine	Ziegra Eis Machinen

2.11 Programs & Modules

Programs

Program	Version	Manufacturer
Affinity Designer	1.10	Serif (Europe) Ltd.
Excel	Version 2205	Microsoft Corporation
GitHub		GitHub, Inc

Continued on next page

(Continued)

Program	Version	Manufacturer
keyence sftware?!		
MiKTeX	2.9	Christian Schenk
python	3.9	Python Software Foundation
PyCharm (Community edition)	2021.2.2	JetBrains s.r.o.
SchemaSpy	5.0.0	John Currier
SDS	2.2.2	Thermo Fisher Scientific GmbH Im
sqlite3_analyzer	3.38.5.	The SQLite Consortium
Wave Controller	2.6.3	Agilent Technologies

Python Modules

Module	Version	Info
beautifulsoup4	4.11.1	crummy.com/software/BeautifulSoup
bokeh	2.4.1	bokeh.org
numpy	1.21.4	numpy.org
pandas	1.3.4	pandas.pydata.org
Pillow	8.4.0	python-pillow.org
pyliftover	0.4	github.com/konstantint/pyliftover
python standard library	3.9	docs.python.org
matplotlib	3.4.3	matplotlib.org
requests	2.26.0	requests.readthedocs.io
scipy	1.7.3	scipy.org
seaborn	0.11.2	seaborn.pydata.org
urllib3	1.26.7	urllib3.readthedocs.io
wget	3.2	bitbucket.org/licface/pywget

Frameworks

- This thesis was generated with the uzl-thesis class have been written and kindly provided by Till Tantau.
- Styling of the GWAS Visualizer was done with the CSS Framework Bootstrap.

2.12 Public Data

3

Methods

3.1 Cultivation and differentiation of HAoSMCs

For the following experiments human arortic smooth muscle cells (HAoSMCs) were used. A cell type commonly used for the study of cardiovascular function and disease [Reference for this claim]. Cells were kept at 37°C and 5% CO2 when ever possibile.For differentiation cells were treated first with TGFβ and then with Interleukin 1 beta (IL-1) & PDGF-BB to induce a synthetic phenotype. For more imformation please check the section 1.3 & 1.4 as well as the referenced literature.

Thawing & Cultivation

Cells were cultivated to a maximum passage of 10, after that new cells were thawed. For long time storage cells were stored in liquid nitrogen. When required, new cells (6th passage) were thawed at 37°C in the water bath and transfered to a 15 mL tube. After centrifugation for 2 min at 300xg the supernatant was removed and the cell pellet taken up in 14 mL of M231 + Smooth Muscle Cell Growth Supplement (SMGS) for cultivation in a TC Flask T75. Every other day 2/3 of the medium were removed and replaced by fresh.

Passaging

When reaching a maximum of 80% confluency (approx. once a week) the medium was removed completely and cells were washed once with 5 mL of phosphate buffered solution (PBS). The washed cells were incubation with 3 mL trypsin for 4 min at 37°C before 7 mL M231 were added to the deattched cells. The cell suspension was transfered to a 15 mL tube and pelleted for 4 min at 300xg. Finally, supernatant was removed and the pellet resuspended in M231 + SMGS, seeding 500×10^3 cells per TC Flask T75.

Preparation of Collagen I matrix

For preparation of the collagen type I (col I) matrix (1.8 mg/mL) all the components were mixed, adding the collagen last. All components were stored at 4°C and all pipetting steps were carried out on ice:

Table 3.1: col I Matrix Composition

component	concentration	volume (µL)
H20	-	38.9
M231	-	53.3
SMGS	20x	5,3
NaOH	1 M	2,7
NaHCO3	7.5 %	2.1
Col I	5 mg/mL	57.6
total	-	160

160 µL of matrix mix were transfered in each used well of a Nunc Cell-Culture Treated Multidish 24, fully coating the bottom of the well. For polymerization the matrix was incubated at 37°C for at least 60 min.

Differentitation of HAoSMCs

Differentiation was carried out over a total of seven days in 24 wells plates. 1 mL M231 was used as the medium, supplemented with 1 % fetal bovine serum (FBS) and different cytokines:

- **Day 0:** Matrix and cells were prepareed as described in the sections Preparation of col I matrix and Passaging. Seeding of 40×10^3 in M231 + SMGS on plastic or on 160 µL col I matrix.
- **Day 1:** After 24 h the medium was replaced with 1 mL M231 + 1% FBS + 5 ng/mL TGFβ (or 1 mL M231 + 1% FBS).
- **Day 5:** The medium was replaced with 1 mL M231 + 1% FBS + 10 ng/mL IL-1 + 10 ng/mL PDGF-BB (or just 1 mL M231 + 1% FBS).
- **Day 7:** Potentially further stimulation described in section of the corresponding assay.

3.2 mRNA Quantification

SYBR® Green is an intercalating DNA dye that allows for the monitoring of DNA amplification. Flourescence is measured after every amplification cycle of the PCR yielding a crossing point when signal reaches a certain threshhold. A lower quantification cycle (Cq) corresponds to an higher inital DNA concentration. (Huggett and Bustin, 2011) quantative PCR (qPCR) was utilized to assess the mribonucleic acid (RNA) concentration of the two reporter genes Calponin 1 (CNN1) and Matrix metallopeptidase 9 (MMP9) in HAoSMCs differentiated as described in section 3.1. Using the house keeping gene Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) for reference.

RNA Isolation

RNA was isolated using Total RNA Purification Kit and extraction was performed according to the corresponding protocol, using the extra washing step with 700 µL 80 % ethanol and eluting with 30 µL of RNase-free water. Determination of nucleic acid concentration was carried out with the NanoDrop.

Reverse Transcription

For reverse transcription (RT), RNA samples were diluted to yield 10 µL of 10 ng/µL RNA. The samples were heated for 5 min at 68°C before adding 10 µL of the RT reaction mix:

Table 3.2: Master Mix for RT

component	concentration	volume (µL)
First Strand Buffer	5x	4
DTT		2
dNTP		1
Oligos		1
RiboLock		1
M-MLVRT		1

The reverse transcription was carried out for 60 min at 37°C, before inactivating the enzyme for 5 min at 95°C. cDNA was used for qPCR or stored at -20°C.

qPCR

Table 3.3: Sample Composition for qPCR

component	conentration	volume (µL)
SYBR GREEN Master Mix	1:2	3.75
Primer (forward + reverse)	5 pM (each)	1.125
H20	-	1.125
cDNA	-	1.5

Samples were prepared in a 384-well Multiply PCR plate, the wells were sealed, thoroughly mixed by invertation of the plate and the assay performed with 7900HT Fast Real-Time PCR System:

Table 3.4: qPCR Cycle

step	time (s)	temperature (°C)	loop to	passes
1	120	50		1
2	600	95		1
3	15	60		40
4	60	60	3	40
5	600	95		1
6	-	16		1

Processing of Data

The Cq was automatically calculated by the software SDS2.2.2 and exported for further analysis. The arithmetic mean of three 3 technical was calculated for each sample, disregarding values that are obvious outliers. For normalization the mean Cq of the reference gene GAPDH was substracted from the mean Cq of the gene of interest:

$$\Delta ct = ct(\text{geneofinterest}) - ct(\text{GAPDH})$$

Taking into account the exponential amplification of DNA in PCR, the Δct can then be transformed into an relative expression level. Where 10×10^6 is just a constant to yield values that are easier to work with:

$$\text{rel.expr.} = 2^{-\Delta ct 10 \times 10^6}$$

In total four biological replicates were done. Data visualization and statistical analysis was done in python using the modules: pandas, numpy, scipy as well as pyplot and seaborn. Assuming a normal distribution, student's t-test was used, a p-value of 0.05 is considered as significant. For detailed information please check the script.

3.3 Energy Profiling

The Seahorse XF Analyzer allows real time measurement of dissolved oxygen and protons in a confined small volume by using solid state sensor probes. These are used to calculate the oxygen consumption rate (OCR) and **eacr!** (**eacr!**) of a cell monolayer. TheOCR and extracellular acidification rate (ECAR) are indicators for mitochondrial respiration and glycolisis respectively and can be used to assess the metabolic function of cells. (**HowAgilentSeahorse**) Seahorse Assay was utilized to assess the energy profile of HAoSMCs differentiated as described in section 3.1. For this assay cells were not differentiated in a multidish but a XF24 Cell Culture Microplate, using 5 technical repeats and on control well for the 4 tested conditions. Since the confined volume required for the assay would not fit the matrix, cells were cultivated in plastic!

Seahorse Assay

On the day before the assay the Seahorse XF Analyzer was turned on to calibrate and the XF24 Extracellular Flux Assay Kit catridge was left to equilibrate in Seahorse XF calibrant over night at 37 °C (in non-CO2 environment).
On the day of the assay, cells were washed with 500 µL PBS each and afterwards 500 µL XF BASE medium, supplemented with 1 mM Pyrucate, 10 mM Glucose, 2 mM Glutamine & 90 µM NaOH, was added. The cells were left to incubate for 1 h at 37°C in non-CO2 environment. During this time toxins for disruption of the respiratory chain were prepared and loaded into the XF24 Extracellular Flux Assay catridge:

The catridge was loaded into the XF Analyser for calibration, after successful calibration the hydration catridge was replaced with the cell plate. Measurement was programmed out as following:

- Calibration of the probes.
- Equilibration
- 3 Repeats of:
  - Mixing (1 min)
  - Pause (2 min)
  - Detection of OCR and ECAR (4 min)

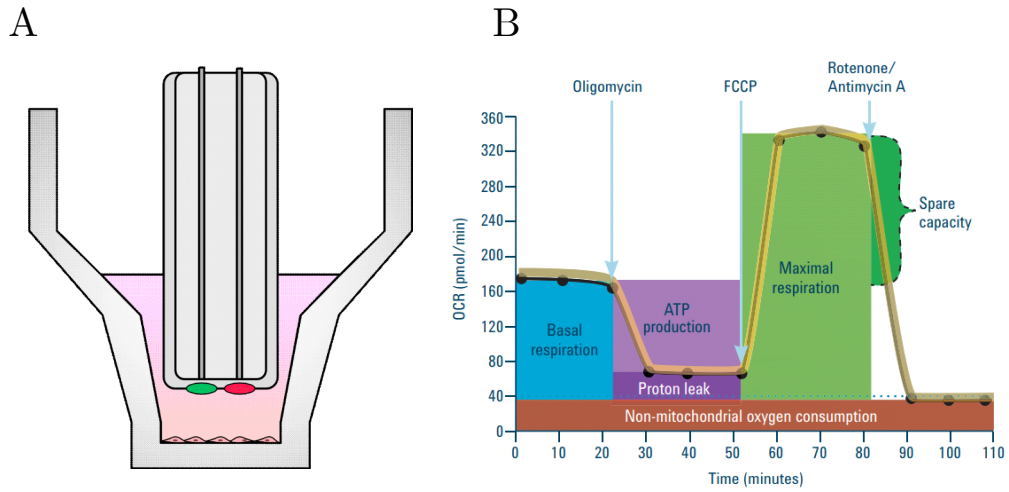

Figure 3.5: Basics of Seahorse Assay (a placeholder)
(A) Schematic of a well used for Seahorse Assay. For the measurement the piston in the middle lowers to the bottom, this way defining a restricted space at the bottom. OCR and **eacr!** in this volume are measured via two porbes (red and green). **(B)** Examplary curve for OCR recorded over time and extractable properties of the respiratory chain.

Table 3.6: Toxin Concentrations for XF24 Extracellular Flux Assay

component	concentration in cartridge(µM)	volume in cartridge(µL)	concentration in well (µM)
Oligomycin	14	55	1.4
FCCP	10	60	2.0
Antimycin	50	65	5.0

- Pause (2 min)
- Injection of 55 µL Oligomycin
- 3 Repeats of:
  - Mixing (1 min)
  - Pause (2 min)
  - Detection of OCR and ECAR (4 min)
- Pause (2 min)
- Injection of 60 µL FCCP
  - Mixing (1 min)
  - Pause (2 min)
  - Detection of OCR and ECAR (4 min)
- Pause (2 min)
- Injection of 55 µL Antimycin
- 3 Repeats of:
  - Mixing (1 min)
  - Pause (2 min)

– Detection of OCR and ECAR (4 min)

Finally the medium was removed and cells were stained for 15 min with 1 µg/mL Hoechst in PBS and photographed to determine cell count for normalization.

Processing of Data

Cells were quantified using a python script provided by my supervisor Dr. Tobias Reinberger. OCR and ECAR calculated by the XF Analyzer were normalize in using this cell count and the signal of the control wells. In total three biological were recorded of which on was excluded because no changes in OCR and ECAR could be detected and cells deacttched from the bottom of the wells during Hoechst staining. For the remaining two replicates the least fitting of the 5 technical repeats for each condition was manually excluded. Further initial OCR and ECAR as well as the characteristics of the respiratory chain displayed in figure 3.5 B were calculated, again using a modified python script provided by Dr. Tobias Reinberger. Assuming a normal distribution, student's t-test was used, a p-value of 0.05 is considered as significant. For detailed information please check the script.

3.4 Oxidative Stress Assay

CellROX Green is a flourescent dye that gets oxidized in a situation of oxidative stress and then binds to DNA, showing bright green flourescence (**CellROXGreenReagent**). CellROX Green assay was used to assess generation of ROS in HAoSMCs differentiated as described in section 3.1. After further differentiation, further stimulation (from here on refered to as 'boost') with PDGF-BB was carried put. Additionally, a recovery experiment was performed using N-acetylcystein (NAC), a potent anioxidant, to quench generation of ROS.

CellROX Assay

For the assay cells were washed with PBS, then the boost was performed using variable concentrations of PDGF-BB in 300 µL Hanks balanced salt solution (HBSS). For ROS quenching with NAC, 0.25 M NAC solution was added to the wells 2 h prior to the experient and also added to HBSS during the experiment.

Table 3.7: Composition for Seahorse Assay Boost

component	concentration	final concentration	volume (µL)
HBSS	-	-	300
PDGF	?	variable (0 - 400 ng/mL)	variable
Hoechst	1 mg/mL	1 µg/mL	0.3
CellROX (1:500)	2.5 mM	5 µM	0.6
NAC	0.25 M	variable (0 - 8 mM)	variable
total	-	-	~300

Cells were kept at 37°C in 5 % CO2 environment during the boost, the incubation time is indicated with the results of the respective experiment. Imaging was done with the BZ-X810 All-in-One Flourescence Microscope, using standard sensitivity. Images for

the NAC quench were recorded as a z-stack and merged into one image using [KEYENCE SOFTWARE].

Processing of Data

For time resolved PDGF-BB boost titration, 7 biological repeats were performed of which one was excluded because of extremly high signal in the negative control. For NAC quench, 4 biological repeats were performed of which one was excluded because no signal in the positive control. For quantification of signal itensity, pixels with a green value higher than 90 were counted. Differences in cell count were adjusted by divison through the number of pixels with a blue value bigger than 80. To adjust for large variance in total signal intensity between biological repeats, values were adjusted by divison through the total signal of all recorded conditions. Mann Whitney U Test was used, a p-value of 0.05 is considered as significant. For detailed information please check the scripts.

3.5 Curation of Data for postGWAS Analyses

Dta for postGWAS analyses and co-visualization with the GWAS data, was downloaded from public resources. Processing of the data and further annotation is briefly described in the follwing listing. The finallay generated tables are summarized in figure 4.8 and table **??**. For a complete view please consult the the download scripts.

- **GWAS Summary Statistics:** The CAD GWAS summary statistics as well as a list of indetified proxy single nucleotid polymorphisms (SNPs) from the study were annotated via the ensembl representational state transfer (REST) application programming interface (API) by Dr. Tobias Reinberger.
- **HGNC Gene List** The newest quarterly update to the complete human gene nomenclature consortium (HGNC) dataset was downloaded via the HGNC file transfer protocol (FTP) server. The dataset was used to generate a list of all 43135 approved symbols, mapping to their HGNC id as well as a list of all 98723 symbols (approved, alias and previous), mapping to their HGNC id.
- **Linked SNPs** LD r^2 values for variants in a 500 kb window around all variants in the list of CAD GWAS proxy variants, were computed and downloaded via the ensembl REST API. For humans ensembl calculates the LD with data from the 1000 Genomes project (see table 3.8). In the same process linked SNPs were annotated with their most severe consequence by the ensembl variant effect predictor (VEP). In total information for 449770 relationships was downloaded.

Table 3.8: 1000 Genomes Populations

Name	Size (individuals)	Description
1000GENOMES:phase3:ALL	2504	All phase 3 individuals
1000GENOMES:phase3:AMR	347	Americans
1000GENOMES:phase3:EAS	504	East Asians
1000GENOMES:phase3:EUR	503	European
1000GENOMES:phase3:SAS	489	South Asian

- **Ensembl Genome Annotatation** The newest ensembl build (ensembl release 106) was downloaded via the ensembl FTP server. Features annotated as genes of the type protein coding (19994), lncRNA (17734) or miRNA (1877) were extracted, further gene symbols were mapped to their HGNC id if possible.
- **Ensembl Regulatory Build** The newest ensembl regulatory build (ensembl release 106) was downloaded via the ensembl FTP server, containing 110623 open chromatin regions, 30873 TF binding sites, 175885 CTCF bindsing sites, 127935 enhancers, 36597 promotors & 140548 promotor flanking regions.
- **Open Target Genetics l2g Scores** The lattest list of Open Target Genetics L2G Scores was downloaded via the open target geneitcs FTP server. Entries were annotated with their HGNC ID when ever possible, 655 entries that do not map to a gene that is approved by the HGNC were dropped, yielding a total of 3580206 database entries.
- **TSS** 35160 transcription start sites (TSS) for protein coding genes were extracted from a university of california santa cruz (USCS) Genome Brower dump -> I know this needs improvment, but this is literally the last point on my agenda.
- **Associated traits from GWAS catalog** The SNP trait associations from the lattest release of the GWAS catalog as well as the accompanying list of studies was downloaded via the GWAS catalog FTP server. 14892 SNP-trait correlations missing a the position or a p-value for the association were dropped from the data set. Further the column for Odds Ration or beta was seperated in to two columns. In total 370002 assocciations from 5831 distinct studies were collected.
- **TADs** TADs predicted by software adapted from Dixon et al. (2012) were downloaded via the 3D genome browser. In total TADs in 40 distinct biosamples were downloaded.
- **scATAC-seq from newmanMultipleCellTypes2021a** Processed scATAC-seq data for 8 celltypes [SOME MORE INFO] were scraped from the Miller Lab GitHub repository.
- **scATAC-seq from CATlas** Processed scATAC-seq data was scraped from the Ren Labs website for 222 biosamples.
- **ABC model** The ABC model data for 131 biosamples was downloaded from the Engreitz Lab FTP server. The data was further translated from Hg19 to Hg38 using pyliftover.
- **ENCODE cCREs** cCREs in distinct biosamples were downloaded by Dr. Tobias Reinberger, filtering out elements that were annotated *unclassified*.

3.6 Visualization of GWAS data

For visualization of the data, an bokeh application was build, that fetches the data from the database and renders it to a webbrowser.
Bokeh is a python module that allow easy and interactive visualization of data. It combines the powerful data processing tools of python with the interactivity of JavaScript running in the browser. The python side of bokeh creates python objects which are serialized into JSON data and handed over to bokehJS which deserializes them into JavaScript objects that are rendered to the browser. The intergrated bokeh server additionally offers the possiblity to synchronize data between the underlying python environment and brower side JavaScript library, allowing real time updates to the displayed data.

According to good design principles the concerns of the application are split into two

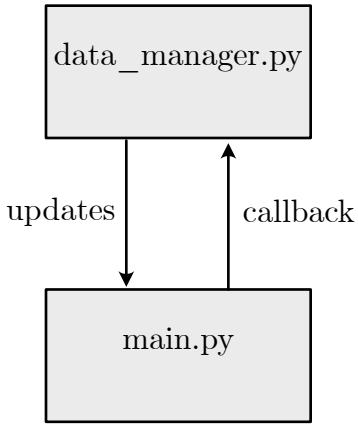

Figure 3.9: Architecture of the vis tool

sections shown in fig. 3.9. Reading of data from the database and further processing steps are managed by a data provide and enclosed in one class. In contrast to the model-controler-view architecture, a popular architectural pattern for the design of user interfaces, there is no partition between a view and a controler. Since data visualization as well as the control widgets are created by bokeh, it is convienient to use the build in event listeners of the library to handle the required callbacks. Therefore the main file is responsilbe for the creation of all plots and widgets as well as listening for inputs.

3.7 Enrichment analysis

Based on the data in the database an initial postGWAS studies was run. Annotation enrichment analyses are a popular tool for the identification of terms that are over-represent in a list of interest. The most prominent application probably being their application as **gsea!** (**gsea!**). gene set enrichment analysiss (GESAs) are used to check for the over-representtion of a candidate gene list in a predefined set of genes (Tipney and Hunter, 2010). In this case the method is used to determine if cCREs overlap with CAD associted SNPs is enriched in a biosample, using fisher's exact test. For the analysis cCREs annotated as unclassified were excluded. As a list of CAD associted SNPs the list of 241 proxy variants from the database was used, as well all linked varinats (r2 >= 0.6) in the 1000 Genomes European Population. The the following parameters were calculated for all biosamples:

- The number of distinct cCREs among all biosamples (m)
- The number of distinct cCREs that are annotated in the biosample of interest (mt)
- The Number of distinct cCREs that overlap with a SNP in the SNP list in any biosample (n)
- The Number of distinct cCREs that overlap with a SNP in the SNP list in the biosample of interest (nt)

The p-value for the number of overlaps to be greater than or equal to the observation can be calculated as the cumulative distribution function of the hypergeometric distribution.

$$P(\sigma_t \geq n_t) = \sum_{k=n_t}^{min(m_t,n)} \frac{\binom{n}{k}\binom{m-n}{m_t-k}}{\binom{m}{m_t}}$$

To account for the multiple comparisons problem, p-values were adjusted with Bonferroni correction where n is the number of tests (≡ number of biosamples):

$$p_{ajd.} = p * n$$

The analysis and visualization were done in python. A p-value of 0.05 is considered as significant. For detailed information please check the analysis script and the visualization script.

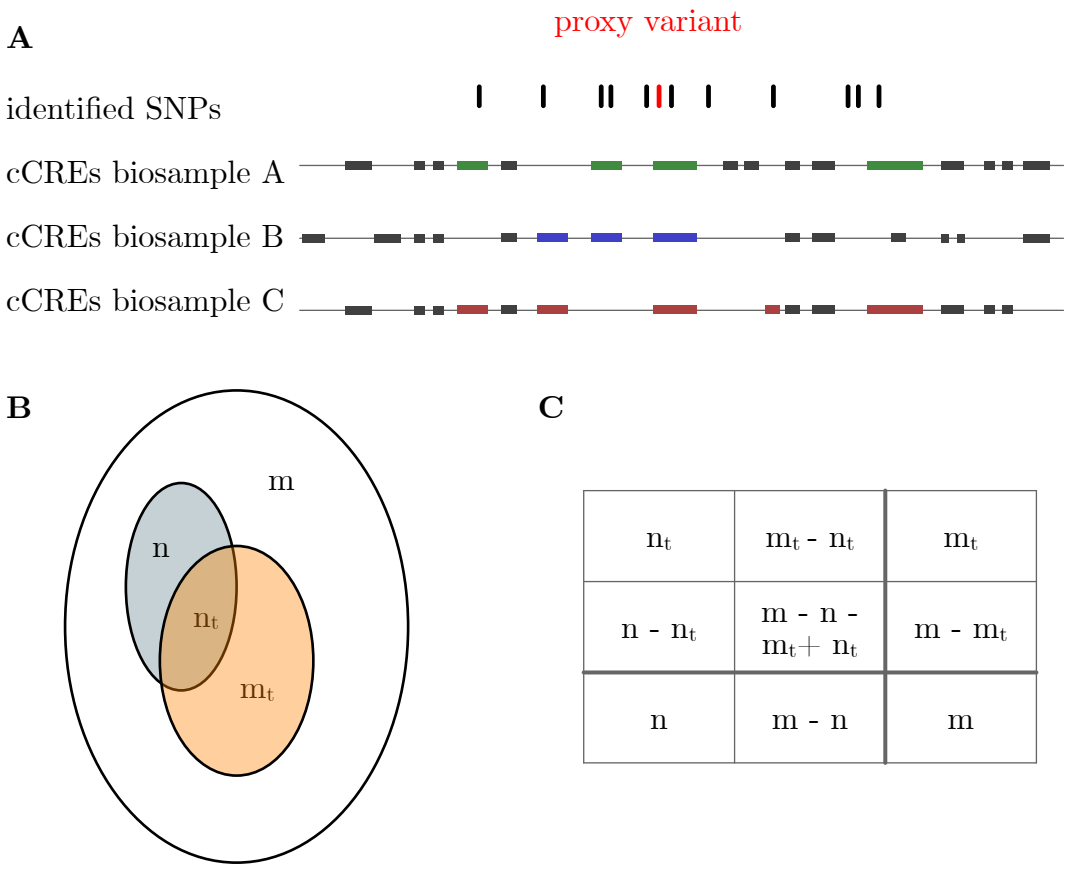

Figure 3.10: Enrichment analysis for cCREs!s overlapping with CAD risk SNPs

(A) Visual representation of the overlap calculation for enrichment calculation. The proxy variant is indicated as a red line, variants in LD are indicated as black lines. cCRE are shown as boxes, those that are overlapping with a SNP were colored according to the biosample they were annotated in. **(B)** Venn diagram of these values for a biosample. **(C)** Schematic contingency table for a biosample. (m) is the number of distinct cCREs found among all biosamples (23 in this example); (mt) the number of distinct cCREs annotated in the biosample of interest (16 for biosample A, 14 for biosample b, 14 for biosample C); (n) the number of distinct cCREs overlapping with a SNP (6 in this example); the number of distinct cCREs overlapping with a SNP in the biosample of interest (4 for biosample A (green), 3 for biosample B (blue), 5 for biosample C (red))

4

Results

4.1 Differentitaion

To chacarcterize the influence of PDGF-BB on **haomc!**s (**haomc!**s), the cells were first treated with TGFβ for two days to push them from a probably dedifferentiated phenotype during cultivation and lack of stimulation towards one that resambles the contractile phenotpye as a standardized starting point. Further cells were stimulated for four days with IL-1 and PDGF-BB. The induced phenptypes were then characterized via qPCR and Seahorse Assay.

Expression of CNN1 & MMP9

To track the differentatiation and confirm that the HAoSMCs first adopt a chontractile phenotype after TGFβ stimulation and then further dedifferentiate after stimulation with PDGF-BB, the mRNA levels of the marker genes CNN1 as well as MMP9 were determined using qPCR. CNN1 as a contractivle marker and MMP9 as a marker for a synthetic phenotype. For better compariability mRNA levels are considered in realtion to the house keeping gene GAPDH.

As seen in figure 4.1 (top panel), stimulation of HAoSMCs cultivated on a col I-Matrix with TGFβ causes a significant increase in CNN1 expression (+– vs. – –). After further stimulation with PDGF-BB & IL-1 , while not significant, **cnn!** (**cnn!**) expression declines again (+– vs. ++) but is still significantly higher than in HAoSMCs which were not stimulated at all (– – vs. ++). A similar trend seems to take place for HAoSMCs cultivated on plastic, while not significant with four biological repeats repeats. Further stimulation of HAoSMCs on plastic with TGFβ followed by stimulation with PDGF-BB & IL-1 , yields a significantly higher expression of CNN1 (++ Matrix vs. ++ Plastic).

As seen in the bottom panel of figure 4.1, within 4 biological repeats, no statistically significant trends can be observed for the expression of MMP9. Still the avarage expression of MMP9 seems to approx. doubled for all conditions, the most prominent difference being between HAoSMCs treated first with TGFβ, then with PDGF-BB & IL-1 (++ Matrix vs. ++ Plastic, p = 0.063).

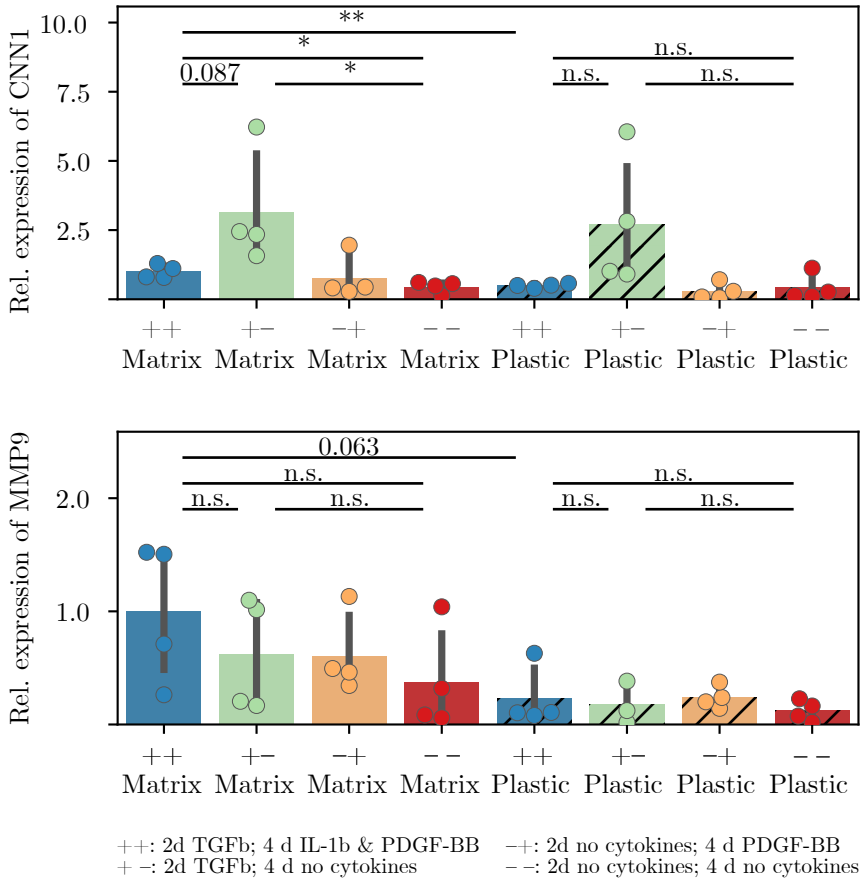

Figure 4.1: Relative Expression of CNN1 & MMP9 in HAoSMCs
qPCR analysis of expression for contractile marker CNN1 (top) and synthetic marker MMP9 (bottom) for HAoSMCs differentiated with different combinations of cytokines: **++:** 2 d with TGFβ followed by 4 d with IL-1 & PDGF-BB; **+−:** 2 d with TGFβ followed by 4 d without stimulation; **−+:** 2 d without stimulation followed by 4 d with IL-1 & PDGF-BB; **—:** 6 d without stimulation. All four conditions were tested on two differen surfaces (plastic vs. col I matrix). Expression levels are in relation to expression of housekeeping gene GAPDH. Statistical analysis for (n = 4) biological repeats was performed using student's T-test: $* : p < 0.05; ** : p < 0.01$

Energy profile

In addition to the expression of CNN1 & MMP9, the energy profiles of HAoSMCs were assesed via Seahorse Assay. It is important to note, that the assay was carried out on plastic because the col I matrix does not fit into the confined compartment created by the pistion for detection of OCR & ECAR. Further only two biological repeats were carried out, because it became increasingly clear, that all other experiments would be cariied out on a col I. Therefore all following results should be considered under these circumstances.

The readout paramters of the Seahorse assay are OCR as a representation of mitochon-

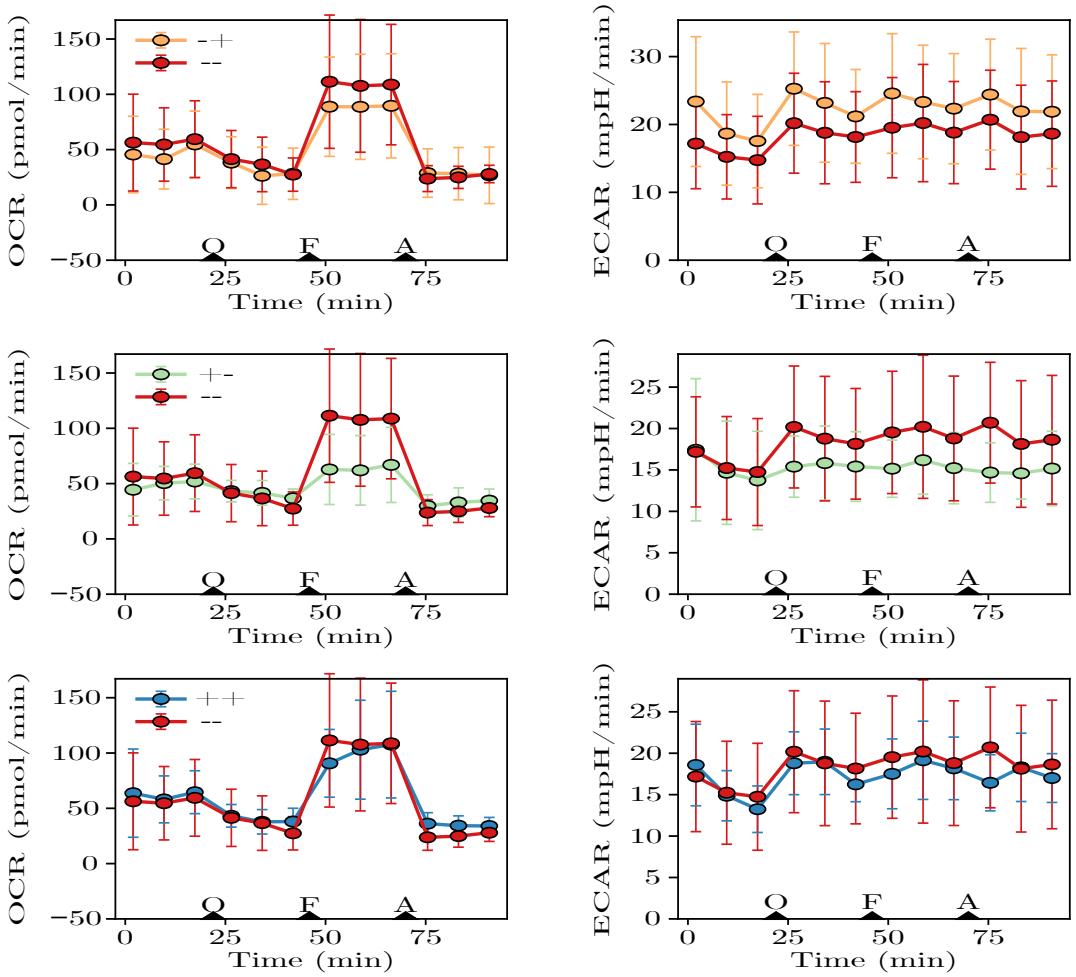

Figure 4.2: OCR & ECAR of HAoSMCs
Seahorse assay for HAoSMCs differentiated with different combinations of cytokines. **++:** 2 d with TGFβ followed by 4 d with IL-1 & PDGF-BB; **+−:** 2 d with TGFβ followed by 4 d without stimulation; **−+:** 2 d without stimulation followed by 4 d with IL-1 & PDGF-BB; **—:** 6 d without stimulation. OCR & ECAR are shown for −+ (top), +− (middle) and ++ (bottom) in comparison to —. Injectiontimes for toxins (O: Oligomycin; F: FCCP; A: Antimycin A) are marked as triangles. All tracks were recorded for cells cultivated on plastic. Shown datapoints are the average of (n = 2) biological repeats.

drial activity and the ECAR, representing glycolytic activity of the cells. OCR and ECAR for HAoSMCs displayed in figure 4.2. All cells show characteristic changes in OCR after addition of toxins impacting the respiratory chain (compare to figure 3.5 B). After inhibition of the ATP synthase with with Oligomycin, the basal OCR drops, this way making the proportion of the OCR accessible that was required for adenosine triphosphate (ATP) production. Further addition of Carbonyl cyanide-p-trifluoromethoxyphenylhydrazone (FCCP) decouples the respiratory chain, destroying the proton gradiant over the mitochondiral membran and letting the cells reach their maximal respiratory capacity. Finally, the inhibition of

coenzyme Q-cytochrome c reductase (complex III) with Antimycin A, stops all mitochondrial respiratory activity. The ECAR shows a mild increase after addition of Oligionmycin, most likely because the cells are compensating the loss of mitochondiral ATP production via increased glycolysis.

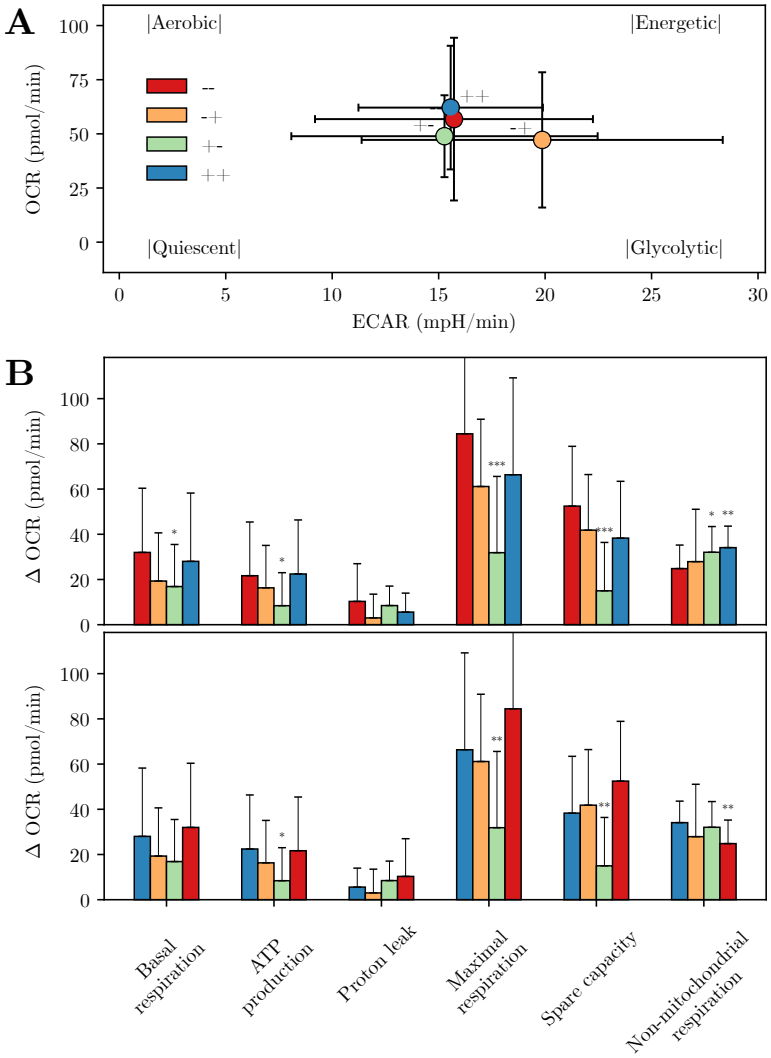

Figure 4.3: Energy profile of HAoSMCs
Seahorse assay for HAoSMCs differentiated with different combinations of cytokines as described in figure 4.2. (**A**) Initial OCR & ECAR of the four tested conditions. (**B**) Characteristics of the the respiratory chain calculated from the tracks shown in figure 4.2 as described in section **??**. Statistical analysis for (n = 2) biological repeats was performed using student's T-test: $* : p < 0.05; ** : p < 0.01, ; *** : p < 0.001$

Looking at the energy profile of the HAoSMCs it is easy to see that OCR & ECAR are quite similar for the conditions ++, +− and —. The only outlier showing a higher ECAR, are HAoSMCs only stimulated with IL-1 & PDGF-BB (fig. 4.3, A). More interesting difference can be observed when examining characterics of the respiratory chain. Stimulation with

only TGFβ causes a sigbnificant decrease in basal respiration, ATP production, maximal respiration as well as spare capacity (figure 4.3, B top). Further stimulation with IL-1 & PDGF-BB then causes again significant increase of these paramters to similar levels as in undifferentiated HAoSMCs (figure 4.3, B bottom).

4.2 Evaluation of oxidative Stress

Finally it was evaluated if further stimulation with PDGF-BB would yield generation of ROS to an extend that can not be compensated by the ROS defense and lead to oxidative stress.

PDGF boost of out cells indcues oxidative stress

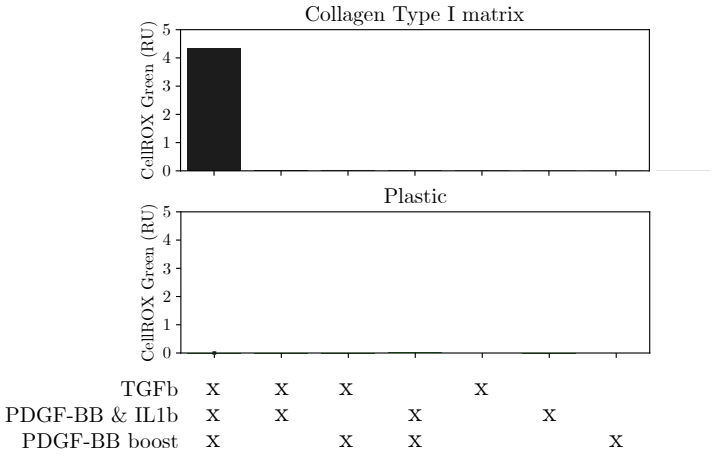

Figure 4.4: Boost with PDGF-BB induces generation of ROS.
CellROX assay for HAoSMCs differentiated with different combinations of cytokines: 2 d with TGFβ; followed by 4 d with IL-1 & PDGF-BB; follwoed by 2 h boost with 200 ng/mL PDGF-BB. Differentiation and assay carried out on col I matrix (top) or plastic (bottom). Shown signal was calculated according to section **??** as the CellROX Green signal, normalized by DAPI signal. No statistica analysis for (n = 1) biological repeats was performed.

At first an experiment already done in the group was repeated. Stimulating the four tested combinations of + 5 ng/mL TGFβ as well as 10 ng/mL IL-1 & 10 ng/mL PDGF-BB stimulated cells, for two more hours with 200 ng/mL As displayed in figure PDGF-BB in HBSS. As displayed in figure 4.4 only stimulation for two days with TGFβ, followed by 2 days with IL-1 & PDGF-BB, followed by 2 h boost with PDGF-BB was able to trigger noticable generation of ROS.

Characterization of the CellROX Assay

To get a better understanding of the assay and its limits a titration was carried out. For this HAoSMCs stimulated for 2 d with 5 ng/mL TGFβ as well as 4 d with 10 ng/mL IL-1

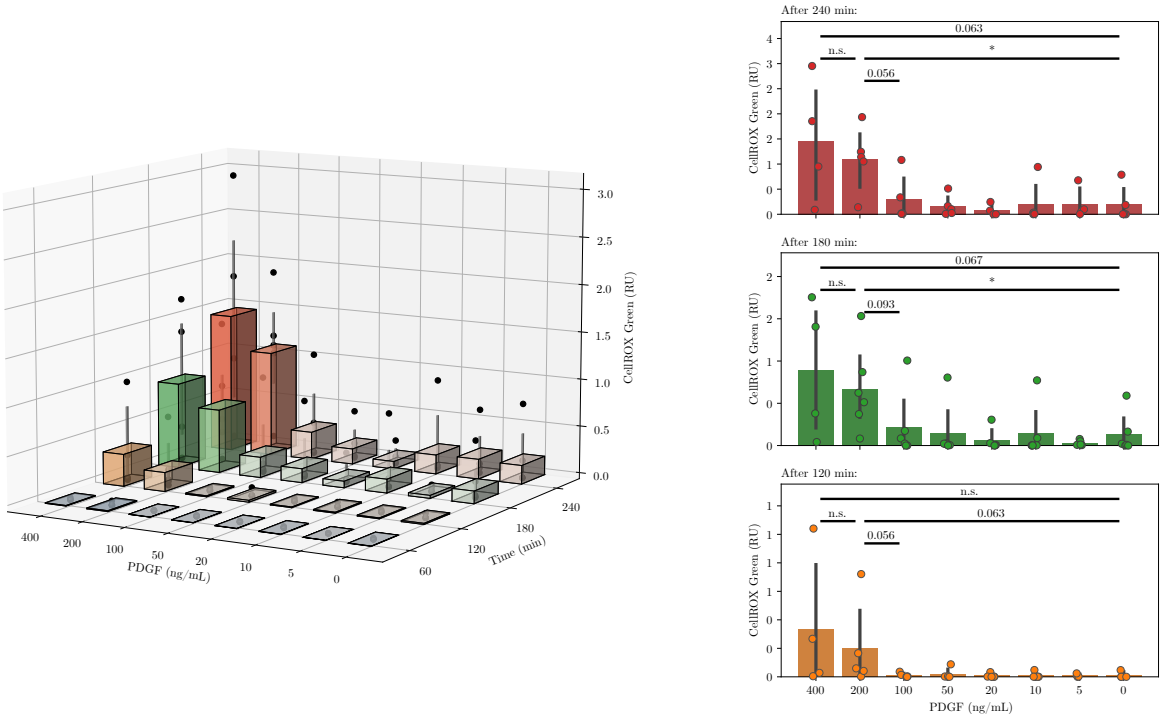

Figure 4.5: PDGF-BB boost titration
CellROX assay for HAoSMCs differentiated with different combinations of cytokines: 2 d with TGFβ; followed by 4 d with IL-1 & PDGF-BB; follwoed by 4 h boost with 0 - 400 ng/mL PDGF-BB. Differentiation and assay carried out on col I matrix. (**A**) 3D visualization: CellROX green signal as a function of PDGF-BB concnentration during the boost as well as incubation time. (**B**) 2D visualization: CellROX green signal as a function of PDGF-BB concnentration after 120 min, 180 min & 240 min. Shown signal was calculated according to section **??** as the CellROX Green signal, normalized by DAPI signal. Statistical analysis for (n = 6) biological repeats was performed using Mann-Whitney U test: $* : p < 0.05; ** : p < 0.01$. It is important to note that not for every biological repeat *all* PDGF-BB concentration were tested.

& 10 ng/mL PDGF-BB, were boosted with different concentrations of PDGF-BB (0 - 400 ng/mL). Signal was detected after 60, 120, 180 & 240 minutes in HBSS. As seen in figure 4.5, CellROX Green signal is neglable after 60 min and then increases with elongated boost times. Further CellROX Green signal stays negalble for boost concentrations 100 ng/mL PDGF-BB. Signal is significantly higher for boost with 200 ng/mL PDGF-BB than it is for no boost after 180 min (figure 4.5 B middle) and 240 min (figure 4.5 B top). While signal in wells boosted with 400 ng/mL PDGF-BB was further increased in regard to boost with 200 ng/mL PDGF-BB in two cases, the signal also collapsed in two repeats. For this reason the avarage signal for boost with 400 ng/mL PDGF-BB is the highest but the increase is not statistically significant.

Overall the trend of greatly increased CellROX signal for boost with 100 as well as 200 ng/mL PDGF-BB was consistent within biological repeats, however variance between

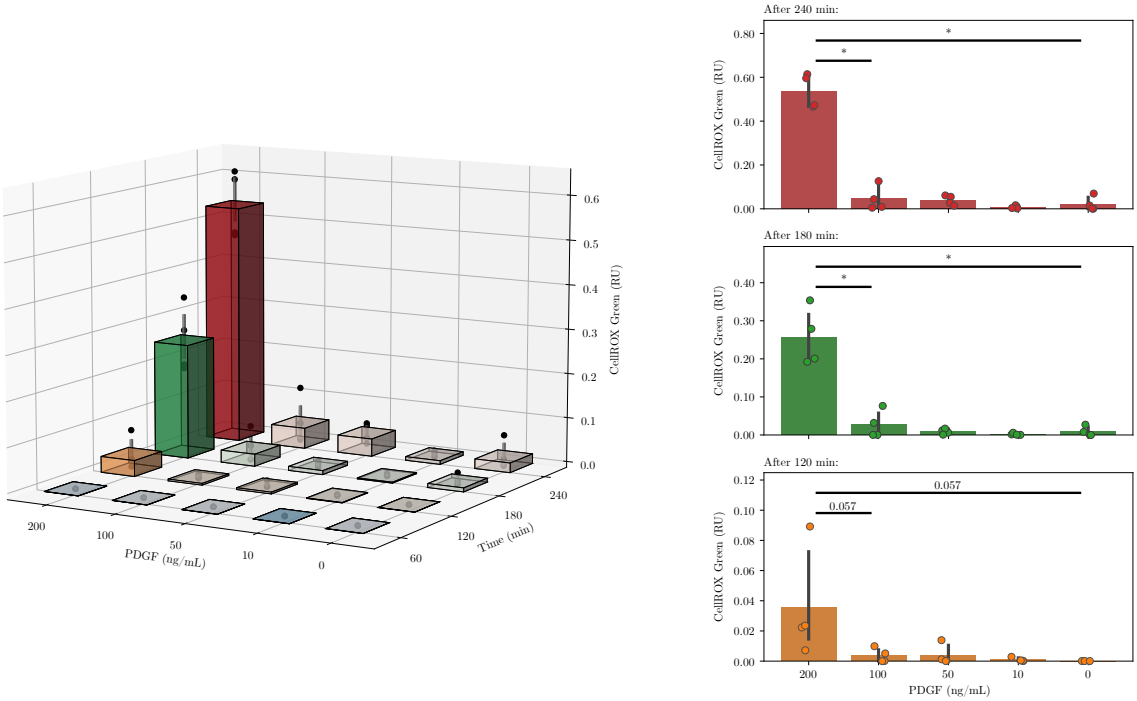

Figure 4.6: PDGF-BB boost titration - normalized
CellROX assay for HAoSMCs differentiated with different combinations of cytokines: 2 d with TGFβ; followed by 4 d with IL-1 & PDGF-BB; follwoed by 4 h boost with 0 - 200 ng/mL PDGF-BB. Differentiation and assay carried out on col I matrix. (**A**) 3D visualization: CellROX green signal as a function of PDGF-BB concnentration during the boost as well as incubation time. (**B**) 2D visualization: CellROX green signal as a function of PDGF-BB concnentration after 120 min, 180 min & 240 min. Shown signal was calculated according to section **??** as the CellROX Green signal, normalized by DAPI signal, further the signal was normalized via the total signal of the biological repeat. Statistical analysis for (n = 4) biological repeats was performed using Mann-Whitney U test: $* : p < 0.05; ** : p < 0.01$.

repeats was almostas high as differences between the conditions. Potential causes for this phenomenon are discussed in section **??**. To account for this large variation between biological repeats, the assay was reevaluated by selection of shared conditions among four biological repeats and normalized by the cummulative intensity of all conditions of the biological repeat (see figure 4.6). This way compensating for differences between biological repeats and resulting in the same observation as without normalization. CellROX signal after 180 min or 240 is signalifcantly higher for cells boostes with PDGF-BB than cell that were not boosted (0 ng/mL PDGF-BB).

Rescue of ROS production using NAC

Finally, a rescue experiment was performed, to verify that observed signal in the CellROX assay was indeed due to generation of ROS. For this ROS generation was quenched by the addition of 2, 4, or 8 mM of NAC. Indeed (while not statistically significant) a clear trend

can be observed, that HAoSMCs treated with NAC shows no signal. In the end, it should be noted, that the signal onyl build up over 15 - 20 minutes under the microscope after the cells were taken out of the incubator. This indicates that generateion of ROS might not exclusively triggered by PDGF-BB stimulation but also required additional contributors like the loss the optimized atmosphere of 37°C and 5 % CO2 in the incubator. This might not have been noted during the titration assay, because cells were taken out of the incubator after one hour anyway to image them for the first time.

4.3 Database and GWAS Visualizer

This was quite a long process, tinkering around with different designs to get a working solution. For this use case visualization in the browser using a database as the backend was the final decision. Compare to fetching everything online -> slow and just grabbing everything from files -> extremely stressful to maintain.

Curation of Data

Describe what actually happend to the data, probably in the methods section. This should be pretty self explanatory. Just have a look at the table and figure.

Visualization

Building on the initially intended use case for the data a visualization tool was build as briefly described in the methods section, for more details please check the source code or just ask.

Describe what the tool can do. Search function for SNPs, also for genes, providing info for associated SNPs. Visualization of the window SNPs, including linked SNPs annotated with important info such as most severe consequence. Further the data is aligned with important stuff like overlapping genes (integrating l2g scores). Also enesembl regulatory build and TADs. Further a lot of cell specific data in the form of scATAC seq tracks and promotor gene links from the ABC model. These are also shown and for easier navigation grouped into different classes using cellosaurus. To have a better look a individual variants these can be clicked, only highlighting tracks that overlap with the specific selected variant.

For more information regarding the aligned data please check the corresping paragraph in the intrdocution.

4.4 Enrichment analysis

The only data that is not displayed in the plot are cCRE elements which were used for an enrichment analysis. Checking different biosamples for significant enrichment of **ccre!**s (**ccre!**s) that are overlapping with proxy SNPs identified in the CAD GWAS or variants that are in LD with these. The procedure is described described in methods.

As seen in figure 4.10, statistical signifcant enrichment ($p_{adj.} < 0.05$) was observed for 34 biosamples. These biosamples were annoted using Cellosaurus data. The most prominent groups of associated tissues were heart (8), lung (7) and artery (6).

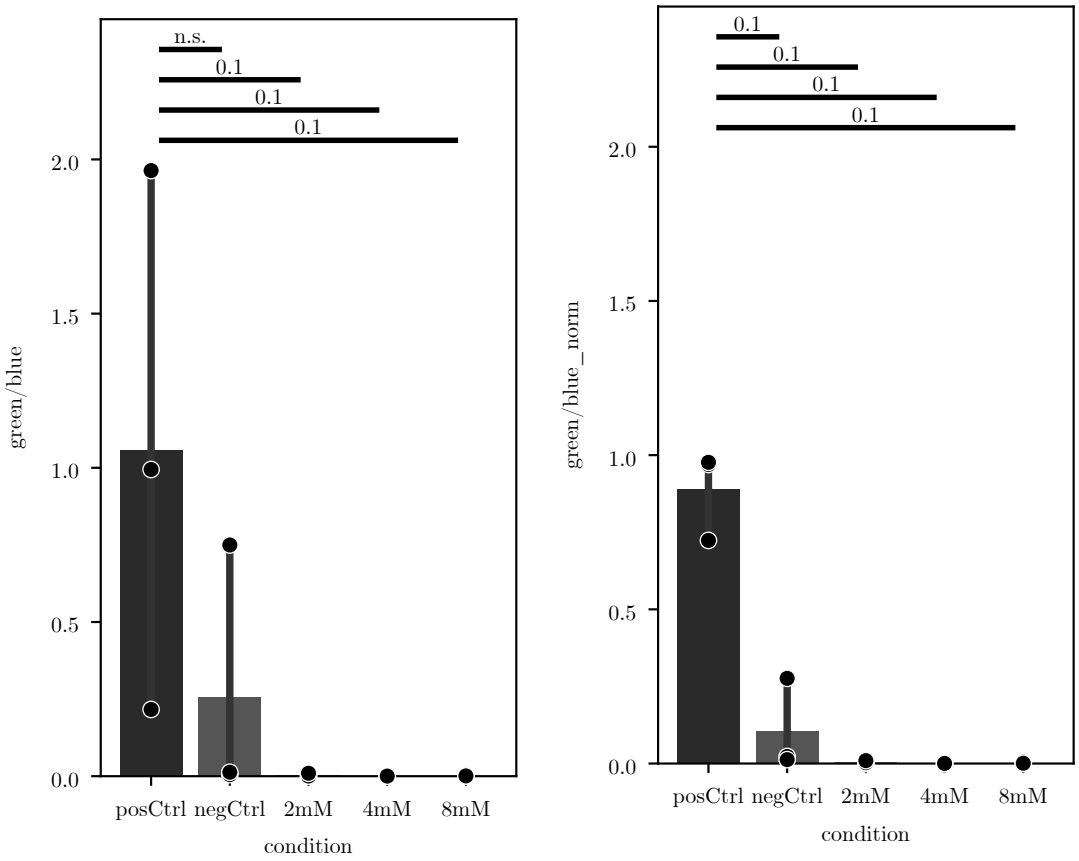

Figure 4.7: ROS generation due to PDGF-BB boost can be rescued with NAC

CellROX assay for HAoSMCs differentiated with different combinations of cytokines: 2 d with TGFβ; followed by 4 d with IL-1 & PDGF-BB; follwoed by 3 h boost with 200 ng/mL PDGF-BB. Differentiation and assay carried out on col I matrix. Cells were treated with 2, 4, or 8 mM of NAC 2 h before the assay. Shown signal was calculated according to section **??** as the CellROX Green signal, normalized by DAPI signal (**A**), further the signal was normalized via the total signal of the biological repeat (**B**). Statistical analysis for (n = 4) biological repeats was performed using Mann-Whitney U test: $* : p < 0.05; ** : p < 0.01$. pos Ctrl: not treated with NAC, negCtrl: no boost with PDGF-BB

Table 4.8: List of Database Tables
List of all the datasets and corresponding tables which were funneled into the database. The size of the tables (and accompanying indices) is indicated by the number of databank pages that are reserved for the data, each page fitting 4096 bytes.

Data	Tables	Page count (including indices), page size = 4096 bytes
GWAS Summary stats	variation	418318
	gwas_meta_cad	867025
	identified_proxy_SNPs_tbl	4
HGNC gene list	hgnc_all_symbols_tbl	826
	hgnc_approved_symbols_tbl	592
Linked SNPs	linked_SNPs_tbl	8819
	population_tbl	1
	consequence_tbl	1
Ensembl Genome Annotation	ensembl_genelist_tbl	613
	ensembl_genelist_biotypes_tbl	1
Ensembl Regulatory Build	ensembl_reg_build_tbl	8778
	ensembl_reg_build_features_tbl	1
TSS	tss_tbl	481
Open Target Genetics Scores	opentarget_l2g_tbl	40984
GWAS catalog	gwascatalog_associations_tbl	10569
	gwascatalog_studies_tbl	326
TADs	tad_tbl	902
	tad_sample_tbl	1
scATAC seq textcite{}	clint_miller_tbl	12370
	clint_miller_biotypes_tbl	1
scATAC seq CATlas	catlas_tbl	308574
	catlas_biotypes_tbl	3
ABC model	abc_tbl	153920
	abc_targetgenes_tbl	84
	abc_celltypes_tbl	3
	abc_classes_tbl	1
ENCODE cCREs	ENCODE_CCRE	4451476
	ENCODE_CCRE_META	107
total	-	6284781 $\equiv\approx$ 25.75 GB

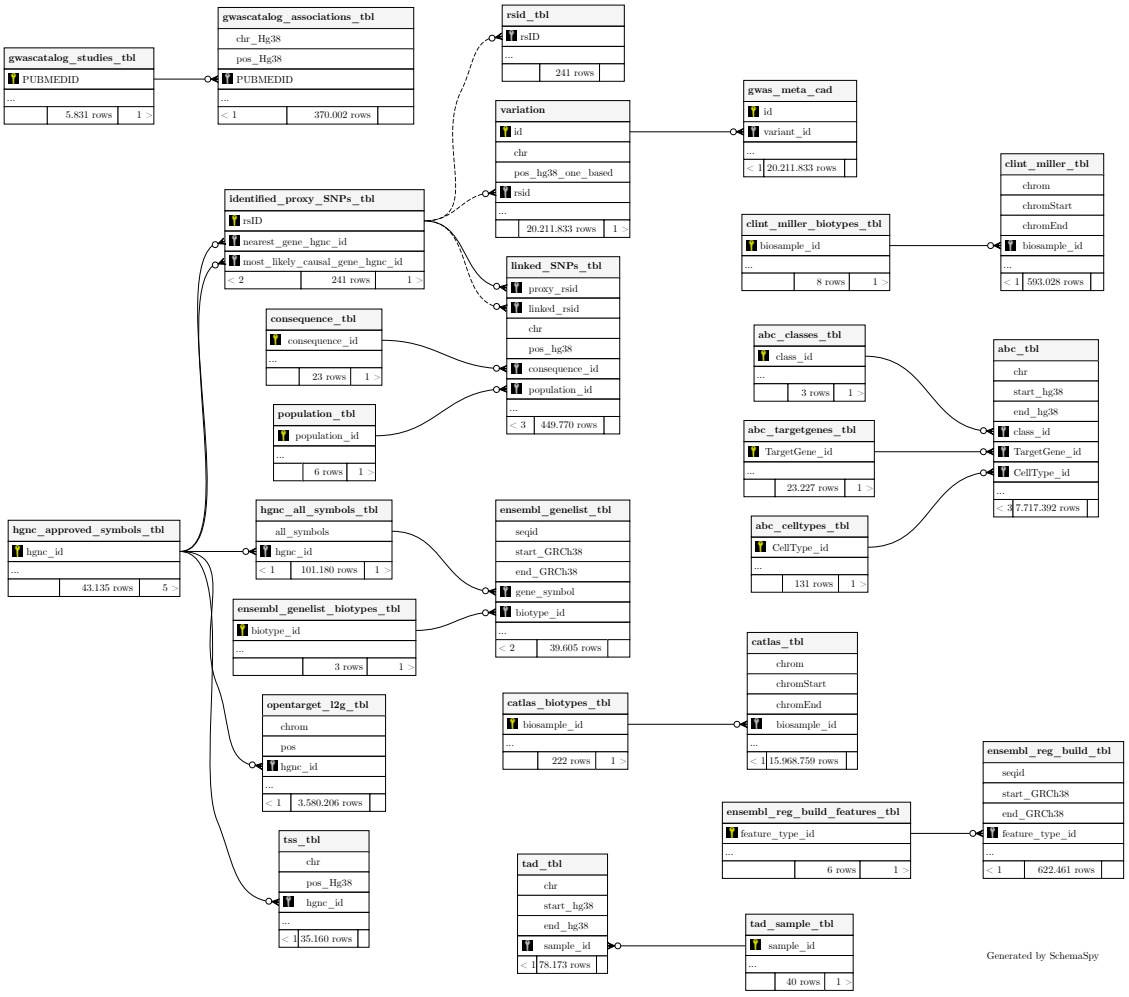

Figure 4.9: Entity-Relationship Diagram of the Database
Fields and relationships of the tables listed in table 4.9. The diagram was generated via SchemaSpy.

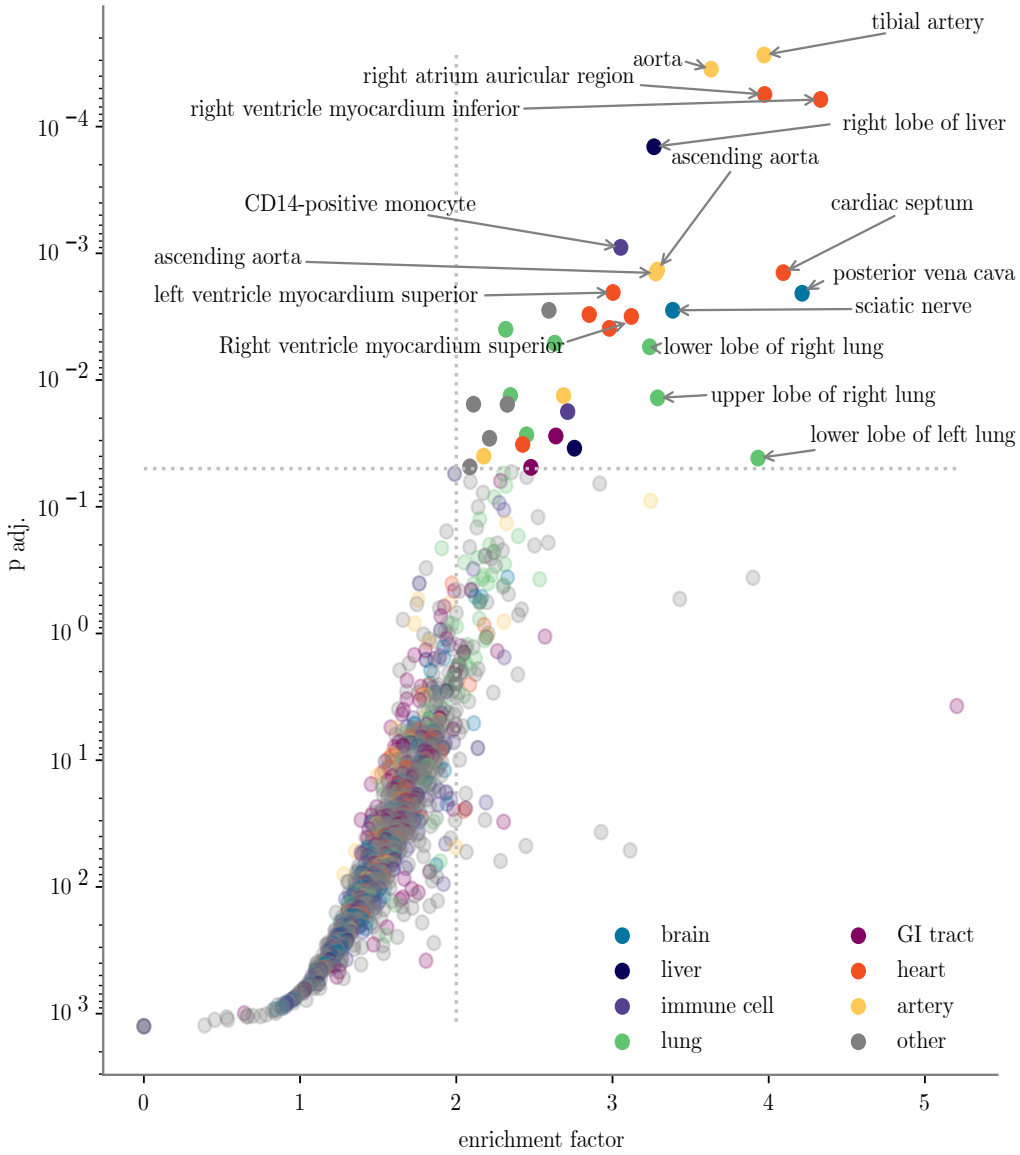

Figure 4.10: Enrichment Analysis This stuff actually seems to be working!!

Table 4.11: Enriched tissues

tissue	count in significant biosamples
heart	8
lung	7
artery	6
liver	2
GI tract	2
brain	2
immune cell	2
other	5
total	34

5

Discussion

Here I'll discuss my results should I ever finish the rest of my thesis.

SOME THING FOR ROS:

Recently, it was reported that STAT also has a cellular nongenomic function. STAT interacts with GRIM-19, a subunit of mitochondria respiratory complex I, which determines whether STAT3 is imported into mitochondria (12,13). Furthermore, there is direct evidence that STAT is present in the mitochondria of cultured cells and primary tissues. STAT mitochondrial importation selectively stabilizes and increases mitochondria respiratory complexes, allowing them to orchestrate responses to stimuli (14,15). As mitochondria respiratory complex I and III are thought to be the main source for ROS generation (16,17), it has been suggested that STAT1 facilitates ROS production and apoptosis (Lee et al., 2007). (Wang et al., 2018a)

Based on our results, STAT1, which is primarily induced by IFN- , is a potent effector responsible for ROS production and loss of $\Delta\Psi$m in LPS/D-GalNinduced fulminant liver injury, evidenced by the finding that ROS production and loss of $\Delta\Psi$m were almost completely abrogated in STAT1(−/−) and IFN- (−/−) mice (Figure 2). -> INF induced ROS production and apoptosis.

SMC PHENOTYPES:

Contractile vs. synthetic VSMCs: loss of contracile marker snad synthetic, migratory pehnotype. Historically, this phenotypic switch was viewed as the hallmark of vascular repair. VSMCs retain their noncontractile status during atherosclerosis due to the continuous exposure to phenotypic switching-inducing stimuli

Foam cells: he majority of foam cells in atherosclerotic plaques appear derived from VSMCs. lipid loaded-VSMCs in culture secrete a variety of pro-inflammatory mediators48 and undergo apoptosis from free cholesterol overload,49 likely compromising plaque stability.

Macrophage-like: up-regulate 'macrophage markers', such as LGALS3, CD68,51 and pro-inflammatory cytokines. difficult to predict their impact on atherosclerotic plaque growth and stability. their pro-inflammatory profile, macrophage-like VSMCs are likely detrimental for plaque stability

Adipocyte-like: These studies emphasize again the high level of plasticity of VSMCs, but the role and abundance of adipocyte-like VSMCs in atherogenesis is unclear, illustrated by a postulated neutral-to-negative position on the 'plaque-stability scale'.

Osteochondrogenic: Vascular calcification is a tightly regulated process, primarily driven by VSMCs developing an osteochondrogenic phenotype. In particular, micro-calcified de-

posits (<50 mm) are associated with increased inflammation and mostly observed in the fibrous cap of human lesions, and thus considered detrimental for plaque stability. In contrast, macrocalcifications (>200 mm) often accumulate in the deep intima or necrotic core in organized structures and may promote plaque stability.

Myofibroblast-like: loss of Tcf21 inhibits phenotypic modulation and attenuates fibromyocyte number in the fibrous cap, suggesting an athero-protective role (Figure 3).

EC-like: Direct evidence of endothelial-like VSMCs in human atherosclerosis is currently lacking.

MSC-like: The highly plastic nature of mesenchymal-like VSMCs offers the opportunity to therapeutically push these cells into a plaque-stabilising phenotype (Figure 3). The VSMCs undergo a 'reprogramming process' controlled by Kru¨ppel-like factor 4 (KLF4) and can transdifferentiate into different cell types, including macrophage-like, EC-like, chondrocyte-like, and adipocyte-like under defined in vitro conditions.

Synthetic VSMCs play a major role in forming and maintaining the fibrous cap by secreting ECM, yet their overall role in plaque stability could be detrimental. Synthetic VSMCs secrete a wide variety of proinflammatory molecules and matrix-degrading enzymes that can cause cell death of neighbouring cells.42,114 Pro-inflammatory cytokines, such as IL1b, IL6, and MCP1 promote atherogenesis by stimulating monocyte recruitment and cell death119 and synthetic VSMCs express a range of adhesion molecules and receptors (e.g. Toll-like-receptors) that promote monocyte recruitment and regulate intracellular inflammatory signalling, respectively. Moreover, synthetic VSMCs secrete extracellular vesicles (EVs) that can drive vascular calcification (see Section 5.2.4).59 Hence, synthetic VSMCs are beneficial for fibrous cap formation, but depending on the local environment and the disease stage, may promote inflammation, calcification, cell senescence, and plaque instability.

The contractile phenotype is positively controlled by the growth factor TGFb. PDGF-BB suppresses contractile gene expression via different mechanisms. (Grootaert and Bennett, 2021)

Vascular smooth muscle cells (VSMCs) play a key role in atherogenesis and have historically been considered beneficial for plaque stability.

Thinning of the fibrous cap in advanced plaques increases the risk of rupture, which triggers thrombus formation and subsequent clinical complications including heart attack and stroke (Libby et al., 2011; Tabas et al., 2015).

The model is rapdily evolving and adjusted: it isnt even clear if cells migrate and proliferate or if they proliferate and migrate then.

The synthetic VSMC phenotype is characterised by loss of contractile marker expression and up-regulation of selective gene sets, including pro-inflammatory cytokines and MMPs, leading to increased cell migration, proliferation, and secretion of pro-inflammatory cytokines.

Therefore, it has been proposed that VSMC-derived cells can both improve plaque stability and exacerbate plaque rupture

To develop efficient therapeutic strategies to limit cardiovascular risk, additional knowledge about how specific VSMC-derived cell types function in mature plaque is therefore needed. Additionally, mechanistic insight into the regulation of VSMC plasticity is required to enable specific interventions. (Harman and Jørgensen, 2019)

Up until recently, vSMCs were classified as either contractile or dedifferentiated (ie,

synthetic).

The central dedifferentiated vSMC type that we classified is the mesenchymal-like phenotype.

The 3 main cellular layers forming the vessel wall are the adventitia, media, and the intima, surrounding the lumen. In the media, the middle layer, vascular smooth muscle cells (vSMCs) are the major cellular component. These vSMCs contribute to the integrity of the vessels and are able to adequately respond to stimuli of vasoconstriction and vasodilation

For decades, vSMC activation and dedifferentiation has been regarded as the adoption of a single synthetic, proliferative phenotype. However, as revealed by recent scRNA-seq analyses, the diversity of vSMC phenotypes is far more sophisticated.

The combination of scRNA-seq and lineage tracing is extremely useful as it allows in-depth vSMC phenotypic characterization.

Contractile vSMCs are regarded as differentiated and quiescent cells under physiological conditions, expressing a panel of typical contractile proteins that is crucial to maintain vascular tension. Contractile vSMCs exhibit an elongated, spindleshaped morphology and express a well-characterized set of contractile markers including smooth muscle actin (ACTA2), smooth muscle myosin heavy chain (MYH11), smooth muscle protein 22-alpha (SM22 /TAGLN), smoothelin (SMTN), and calponin (CNN1). Expression of these proteins is controlled by the transcription factors MYOCD (myocardin) and SRF (serum response factor), both of which are involved in the regulation of differentiation to contractile vSMCs.

In addition, external stimuli, such as TGF- and heparin, play a pivotal role in promoting and maintaining the vSMCs contractile phenotype.

Once pathological processes in the vessel wall are initiated, vSMCs respond by changing their phenotype and function. A plethora of pathological cues induce these changes: factors from the circulation, compounds and proteins produced by activated endothelial cells, fibroblasts, perivascular adipocytes or inflammatory cells, (lack of) mechanical stress, damaged ECM protein fragments, or ECM-derived growth factors.

KLF4 is regulated by various signaling complexes at transcriptional and posttranslational levels.57,58 After exposure to PDGF-BB (platelet-derived growth factor BB), a stimulus of vSMC proliferation and phenotype switching, elevated levels of KLF4 were identified.

Induction of KLF4 in vSMCs results in a phenotypic switch from contractile to mesenchymal-like and initiates the expression of mesenchymal markers such as stem cell antigen-1 (SCA1)/LY6A, CD34, and CD44.40,41 During this transition, while gaining expression of mesenchymal markers, the contractile vSMCs lose expression of their contractile markers. (Yap et al., 2021)

To prevent or reverse the phenotypic transition to mesenchymal-like vSMCs, the expression of KLF4 can be suppressed by TGF- or miR-143/145 to maintain the contractile vSMC phenotype.

Alternatively, mesenchymal-like vSMC may undergo further changes into other vSMC phenotypes.31

In addition, we argued that the data point in the direction of the mesenchymal-like cells giving rise to the other vSMC phenotypes; however, further experimental validation is needed to fully support this hypothesis. (Yap et al., 2021)

This includes the activation of SMC marker genes by TGF 1 and contractile agonists such as angiotensin II as well as repression by PDGF BB

inflammatory cytokines such as IL-1 can induce rapid downregulation of expression of multiple SMC differentiation marker genes (Alexander and Owens, 2012)

Using single-cell and bulk RNA-sequencing analyses of the brachiocephalic artery region and in vitro models, we provide evidence that SMC-to-MF transitions are induced by PDGF and transforming growth factor- and dependent on aerobic glycolysis, while EndoMT is induced by interleukin-1 and transforming growth factor- . Together, we provide evidence that the ACTA2+ fibrous cap originates from a tapestry of cell types, which transition to an MF-like state through distinct signalling pathways that are either dependent on or associated with extensive metabolic reprogramming.

xacerbated lesion development, suggesting that dysregulated EndoMT was detrimental for atherogenesis

SMC PDGFR signalling plays a critical role in SMC investment within the lesion and the fibrous cap. no significant differences in any of the indices of stability examined, including collagen content of the lesion and fibrous cap.

Indeed, sustained PDGFR signalling is essential for retention of ACTA2+ SMCs and collagen content within the fibrous cap, suggesting a critical protective role of PDGFR signalling in fibrous cap development and maintenance. Therefore, contrary to dogma, we propose that augmenting46 rather than reducing PDGFR signalling in SMCs during late-stage atherosclerosis would be a beneficial therapeutic strategy for maintaining lesion stability.

We have further shown that a PDGF/TGF- -induced shift in bioenergetic pathway directly affects ECM synthesis in cultured SMCs that have phenotypically modulated to a MF-like state.

Indeed, sustained PDGFR signalling is essential for retention of ACTA2+ SMCs and collagen content within the fibrous cap, suggesting a critical protective role of PDGFR signalling in fibrous cap development and maintenance. Therefore, contrary to dogma, we propose that augmenting46 rather than reducing PDGFR signalling in SMCs during late-stage atherosclerosis would be a beneficial therapeutic strategy for maintaining lesion stability.

SMC PDGFR signalling plays a critical role in SMC investment within the lesion and the fibrous cap. no significant differences in any of the indices of stability examined, including collagen content of the lesion and fibrous cap.

Taken together, these data suggest that, in the absence of SMC investment, mesenchymal transitions of non-SMC-derived cells are capable of only temporarily maintaining indices of lesion stability. (Newman et al., 2021)

PDGF::

KLF4 is regulated by various signaling complexes at transcriptional and posttranslational levels.57,58 After exposure to PDGF-BB (platelet-derived growth factor BB), a stimulus of vSMC proliferation and phenotype switching, elevated levels of KLF4 were identified.

Induction of KLF4 in vSMCs results in a phenotypic switch from contractile to mesenchymal-like and initiates the expression of mesenchymal markers such as stem cell antigen-1 (SCA1)/LY6A, CD34, and CD44.40,41 During this transition, while gaining expression of mesenchymal markers, the contractile vSMCs lose expression of their contractile markers. (Yap et al., 2021)

- Through the PDGFRb receptor, PDGFBB stimulates ELK1 phosphorylation, which

competes with MYOCD for the same docking site on SRF, suppressing contractile gene expression. (Grootaert and Bennett, 2021)

In addition to its increasingly well described role in tumorgenesis, PDGFRβsignaling has also been implicated in the promotion of atherosclerosis (Andrae, Gallini, and Betsholtz, 2008; He et al., 2015).

cardiovascular disease: In general, two types of cells appear to respond in a pathological fashion to PDGFs—SMCs and fibroblasts—promoting vessel wall pathologies and fibrotic tissue scarring, respectively. Another general remark is that PDGFR- appears to be the dominant PDGFR involved in vascular pathology

PDGF signaling in vascular disorders: PDGF signaling, especially signaling of PDGF-BB via PDGFRβhas been implicated in the development of Artheriosklerose for a long time. But the molecular mechanism has escaped research until today. (Chen, Chen, and He, 2013) (Andrae, Gallini, and Betsholtz, 2008) <- Review 2

A number of adhesion molecules, cytokines and growth factors are implicated in the atherosclerotic lesion formation. These factors interact with each other and form an intricate network to affect tissue repair, cell proliferation and lipid metabolism. Among these factors that contribute to the development of atherosclerosis, an elevated PDGF expression has been detected in nearly all cell types of the atherosclerotic arterial wall and in the infiltrating inflammatory cells. All PDGFs, especially PDGF-A and PDGF-B, are detectable in atherosclerotic lesions. The PDGFR expression is also increased in the atherosclerotic vessel wall. (Hu and Huang, 2015)

Shown in mouse model that PDGF signanling promotes artherosclerosis: PDGFRb pathway activation has a profound effect on vascular disease and support the conclusion that inflammation in the outer arterial layers is a driving process for atherosclerosis.

By deletion of a STAT1-floxed allele in VSMCs, we show that PDGF-driven chemokine signalling and inflammation are STAT1 dependent. Finally, in the context of PDGF-driven atherosclerosis, deletion of STAT1 from VSMCs reduces plaque formation, demonstrating that inflammation of the arterial media and adventitia are important mechanisms by which PDGF signalling promotes atherosclerosis. (He et al., 2015)

Indeed, sustained PDGFR signalling is essential for retention of ACTA2+ SMCs and collagen content within the fibrous cap, suggesting a critical protective role of PDGFR signalling in fibrous cap development and maintenance. Therefore, contrary to dogma, we propose that augmenting46 rather than reducing PDGFR signalling in SMCs during late-stage atherosclerosis would be a beneficial therapeutic strategy for maintaining lesion stability.

SMC PDGFR signalling plays a critical role in SMC investment within the lesion and the fibrous cap. no significant differences in any of the indices of stability examined, including collagen content of the lesion and fibrous cap.

Taken together, these data suggest that, in the absence of SMC investment, mesenchymal transitions of non-SMC-derived cells are capable of only temporarily maintaining indices of lesion stability. (Newman et al., 2021)

6

Conclusion & Outlook

We are closer to doing postGWAS analyses, we really hope that the database makes everything smoother. And we have a system where we can functionally access these identified SNPs. We are close to a point were we can combine both part of the project.

integration of ensembl database

Abkürzungsverzeichnis

A adenin
AF488 Alexa Fluor® 488
bp basepairs
C cytosin
DNA deoxyribonucleic acid
dsDNA doublestranded DNA
E. coli *Eschericha coli*
EDTA Ethylenediaminetetraacetic acid
G guanine
HCl hydrogen chloride
PCR polymerase chain reaction
qPCR quantative PCR
T thymin
HAoSMC human arortic smooth muscle cell
IF immunoflourescence
TGFβ Transforming Growth Factor beta
SMGS Smooth Muscle Cell Growth Supplement
IL-1 Interleukin 1 beta
PDGF-BB platelet-derived growth factor-BB
TGFβ Transforming Growth Factor beta
DE Deutschland
MMP9 Matrix metallopeptidase 9
CNN1 Calponin 1
GAPDH Glyceraldehyde-3-phosphate dehydrogenase
col I collagen type I
OCR oxygen consumption rate
ECAR extracellular acidification rate
FCCP Carbonyl cyanide-p-trifluoromethoxyphenylhydrazone
CAD coronay artery disease
PBS phosphate buffered solution
FBS fetal bovine serum
Cq quantification cycle
RNA ribonucleic acid
RT reverse transcription
DTT reverse transcription
dNTP deoxyribose nucleoside triphosphate
ROS reactive oxygen species
NAC N-acetylcystein
LD linage disequilibrium

GWAS genome wide association study
L2G link to gene
H_2O_2 hydrogen peroxide
$O_2^{\cdot -}$ superoxide anion radical
O_2 elemental oxygen
MI myocardial infarction
VSMC vascular smooth muscle cell
KLF4 Kruppel-like factor 4
TF transcription factor
PDGFR Platelet-derived growth factor receptor
MAP mitogen activated protein
STAT signal transducers and activators of transcription
PI3K phosphatidylinositol 3'-kinase
ROS reactive oxygen species
eQTL expression quantitative trait loci
CRISPR Clustered Regularly Interspaced Short Palindromic Repeats
CTCF CCCTC binding factor
ENCODE ENCyclopedia Of DNA Elements project
cCRE candidate cis-regulatroy element
ABC activity by contect
ATAC-seq assay for transposase-accessible chromatin using sequencing
TAD topologically associated domain
WGS whole genome sequencing
PIP posterior inclusion probability
pELS proximal enhancer-like elements
dELS distal enhancer-like elements
PLS promoter-like elements
H3K4me3 histone 3 lysine 4 trimethylathion
H3K27ac histone 3 lysine 27 acetylation
HBSS Hanks balanced salt solution
GESA gene set enrichment analysis
REST representational state transfer
API application programming interface
FTP file transfer protocol
HGNC human gene nommenclature consortium
SNP single nucleotid polymorphism
VEP variant effect predictor
TSS transcription start sites
USCS university of california santa cruz
ATP adenosine triphosphate

Bibliography

Disease Control and Prevention, Centers for (2022). *Heart Disease Facts | Cdc.Gov*. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/heartdisease/facts.htm (visited on 06/07/2022).

Fryar, Cheryl D (2012). Prevalence of Uncontrolled Risk Factors for Cardiovascular Disease: United States, 1999–2010, 8.

National Health Service (24 Oct 2017, 4:45 p.m.). *Heart Attack*. URL: https://www.nhs.uk/conditions/heart-attack/ (visited on 06/07/2022).

Task Force Members et al. (2013). 2013 ESC Guidelines on the Management of Stable Coronary Artery Disease: The Task Force on the Management of Stable Coronary Artery Disease of the European Society of Cardiology. European Heart Journal *34*, 2949–3003. DOI: 10.1093/eurheartj/eht296.

Tucker, William D., Arora, Yingyot, and Mahajan, Kunal (2022). 1. 1, Treasure Island (FL): StatPearls Publishing.

Yap, Carmen et al. (2021). Six Shades of Vascular Smooth Muscle Cells Illuminated by KLF4 (Krüppel-Like Factor 4). Arteriosclerosis, Thrombosis, and Vascular Biology, ATVBAHA121316600. DOI: 10.1161/ATVBAHA.121.316600.

Grootaert, Mandy O J and Bennett, Martin R (2021). Vascular Smooth Muscle Cells in Atherosclerosis: Time for a Re-Assessment. Cardiovascular Research *117*, 2326–2339. DOI: 10.1093/cvr/cvab046.

Goumans, Marie-José and Dijke, Peter ten (2018). TGF- Signaling in Control of Cardiovascular Function. Cold Spring Harbor Perspectives in Biology *10*, a022210. DOI: 10.1101/cshperspect.a022210.

Batlle, Eduard and Massagué, Joan (2019). Transforming Growth Factor- Signaling in Immunity and Cancer. Immunity *50*, 924–940. DOI: 10.1016/j.immuni.2019.03.024.

Davis-Dusenbery, Brandi N. et al. (2011). Down-Regulation of Krüppel-like Factor-4 (KLF4) by MicroRNA-143/145 Is Critical for Modulation of Vascular Smooth Muscle Cell Phenotype by Transforming Growth Factor- and Bone Morphogenetic Protein 4. The Journal of Biological Chemistry *286*, 28097–28110. DOI: 10.1074/jbc.M111.236950.

Takahashi, Kazutoshi et al. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell *131*, 861–872. DOI: 10.1016/j.cell.2007.11.019.

Pan, Huize et al. (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*, 2060–2075. DOI: 10.1161/CIRCULATIONAHA.120.048378.

Chen, Po-Han, Chen, Xiaoyan, and He, Xiaolin (2013). Platelet-Derived Growth Factors and Their Receptors: Structural and Functional Perspectives. Biochimica et biophysica acta *1834*, 2176–2186. DOI: 10.1016/j.bbapap.2012.10.015.

Heldin, Carl-Henrik (2013). Targeting the PDGF Signaling Pathway in Tumor Treatment. Cell Communication and Signaling *11*, 97. DOI: 10.1186/1478-811X-11-97.

Hu, Weining and Huang, Yu (2015). Targeting the Platelet-Derived Growth Factor Signalling in Cardiovascular Disease. Clinical and Experimental Pharmacology and Physiology *42*, 1221–1224. DOI: 10.1111/1440-1681.12478.

Andrae, Johanna, Gallini, Radiosa, and Betsholtz, Christer (2008). Role of Platelet-Derived Growth Factors in Physiology and Medicine. Genes & Development *22*, 1276–1312. DOI: 10.1101/gad.1653708.

Levéen, P. et al. (1994). Mice Deficient for PDGF B Show Renal, Cardiovascular, and Hematological Abnormalities. Genes & Development *8*, 1875–1887. DOI: 10.1101/gad.8.16.1875.

Robson, M. C. et al. (1992). Platelet-Derived Growth Factor BB for the Treatment of Chronic Pressure Ulcers. Lancet (London, England) *339*, 23–25. DOI: 10.1016/0140-6736(92)90143-q.

Raines, Elaine W (2004). PDGF and Cardiovascular Disease. Cytokine & growth factor reviews *15*, 237–254. DOI: 10.1016/j.cytogfr.2004.03.004.

He, Chaoyong et al. (2015). PDGFR Signalling Regulates Local Inflammation and Synergizes with Hypercholesterolaemia to Promote Atherosclerosis. Nature Communications *6* (1), 7770. DOI: 10.1038/ncomms8770.

Newman, Alexandra A. C. et al. (2021). Multiple Cell Types Contribute to the Atherosclerotic Lesion Fibrous Cap by PDGFR and Bioenergetic Mechanisms. Nature Metabolism *3* (2), 166–181. DOI: 10.1038/s42255-020-00338-8.

Sies, Helmut and Jones, Dean P. (2020). Reactive Oxygen Species (ROS) as Pleiotropic Physiological Signalling Agents. Nature Reviews Molecular Cell Biology *21* (7), 363–383. DOI: 10.1038/s41580-020-0230-3.

Sundaresan, M. et al. (1995). Requirement for Generation of H2O2 for Platelet-Derived Growth Factor Signal Transduction. Science (New York, N.Y.) *270*, 296–299. DOI: 10.1126/science.270.5234.296.

Bouzigues, Cedric I. et al. (2014). Regulation of the ROS Response Dynamics and Organization to PDGF Motile Stimuli Revealed by Single Nanoparticle Imaging. Chemistry & Biology *21*, 647–656. DOI: 10.1016/j.chembiol.2014.02.020.

Uffelmann, Emil et al. (2021). Genome-Wide Association Studies. Nature Reviews Methods Primers *1* (1), 1–21. DOI: 10.1038/s43586-021-00056-9.

Flint, Jonathan (2013). GWAS. Current Biology *23*, R265–R266. DOI: 10.1016/j.cub.2013.01.040.

Schaid, Daniel J., Chen, Wenan, and Larson, Nicholas B. (2018). From Genome-Wide Associations to Candidate Causal Variants by Statistical Fine-Mapping. Nature reviews. Genetics *19*, 491–504. DOI: 10.1038/s41576-018-0016-z.

Lichou, Florence and Trynka, Gosia (2020). Functional Studies of GWAS Variants Are Gaining Momentum. Nature Communications *11*, 6283. DOI: 10.1038/s41467-020-20188-y.

Slatkin, Montgomery (2008). Linkage Disequilibrium — Understanding the Evolutionary Past and Mapping the Medical Future. Nature reviews. Genetics *9*, 477–485. DOI: 10.1038/nrg2361.

Mountjoy, Edward et al. (2021). An Open Approach to Systematically Prioritize Causal Variants and Genes at All Published Human GWAS Trait-Associated Loci. Nature Genetics *53* (11), 1527–1533. DOI: 10.1038/s41588-021-00945-5.

Zerbino, Daniel R. et al. (2015). The Ensembl Regulatory Build. Genome Biology *16*, 56. DOI: 10.1186/s13059-015-0621-5.

SCREEN: Search Candidate Regulatory Elements by ENCODE (n.d.). URL: https://screen.encodeproject.org/ (visited on 06/07/2022).

Moore, Jill E. et al. (2020). Expanded Encyclopaedias of DNA Elements in the Human and Mouse Genomes. Nature *583* (7818), 699–710. DOI: 10.1038/s41586-020-2493-4.

Buenrostro, Jason D. et al. (2013). Transposition of Native Chromatin for Fast and Sensitive Epigenomic Profiling of Open Chromatin, DNA-binding Proteins and Nucleosome Position. Nature Methods *10* (12), 1213–1218. DOI: 10.1038/nmeth.2688.

Buenrostro, Jason D. et al. (2015a). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Current Protocols in Molecular Biology *109*, 21.29.1–21.29.9. DOI: 10.1002/0471142727.mb2129s109.

Buenrostro, Jason D. et al. (2015b). Single-Cell Chromatin Accessibility Reveals Principles of Regulatory Variation. Nature *523* (7561), 486–490. DOI: 10.1038/nature14590.

Fulco, Charles P. et al. (2019). Activity-by-Contact Model of Enhancer–Promoter Regulation from Thousands of CRISPR Perturbations. Nature Genetics *51* (12), 1664–1669. DOI: 10.1038/s41588-019-0538-0.

Nasser, Joseph et al. (2021). Genome-Wide Enhancer Maps Link Risk Variants to Disease Genes. Nature *593* (7858), 238–243. DOI: 10.1038/s41586-021-03446-x.

Lieberman-Aiden, Erez et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science *326*, 289–293. DOI: 10.1126/science.1181369.

Wit, Elzo de and Laat, Wouter de (2012). A Decade of 3C Technologies: Insights into Nuclear Organization. Genes & Development *26*, 11–24. DOI: 10.1101/gad.179804.111.

Dixon, Jesse R. et al. (2012). Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions. Nature *485* (7398), 376–380. DOI: 10.1038/nature11082.

Wang, Yanli et al. (2018b). The 3D Genome Browser: A Web-Based Browser for Visualizing 3D Genome Organization and Long-Range Chromatin Interactions. Genome Biology *19*, 151. DOI: 10.1186/s13059-018-1519-9.

Pombo, Ana and Dillon, Niall (2015). Three-Dimensional Genome Architecture: Players and Mechanisms. Nature Reviews Molecular Cell Biology *16* (4), 245–257. DOI: 10.1038/nrm3965.

Burtenshaw, Denise et al. (2019). Reactive Oxygen Species (ROS), Intimal Thickening, and Subclinical Atherosclerotic Disease. Frontiers in Cardiovascular Medicine *6*, 89. DOI: 10.3389/fcvm.2019.00089.

Huggett, Jim and Bustin, Stephen A. (2011). Standardisation and Reporting for Nucleic Acid Quantification. Accreditation and Quality Assurance *16*, 399. DOI: 10.1007/s00769-011-0769-y.

Tipney, Hannah and Hunter, Lawrence (2010). An Introduction to Effective Use of Enrichment Analysis Software. Human Genomics *4*, 202. DOI: 10.1186/1479-7364-4-3-202.

Lee, Hyun Jung et al. (2007). The Role of STAT1/IRF-1 on Synergistic ROS Production and Loss of Mitochondrial Transmembrane Potential during Hepatic Cell Death Induced by LPS/d-GalN. Journal of Molecular Biology *369*, 967–984. DOI: 10.1016/j.jmb.2007.03.072.

Wang, Yan et al. (2018a). The STAT-ROS Cycle Extends IFN-induced Cancer Cell Apoptosis. International Journal of Oncology *52*, 305–313. DOI: 10.3892/ijo.2017.4196.

Harman, Jennifer L. and Jørgensen, Helle F. (2019). The Role of Smooth Muscle Cells in Plaque Stability: Therapeutic Targeting Potential. British Journal of Pharmacology *176*, 3741–3753. DOI: 10.1111/bph.14779.

Alexander, Matthew R. and Owens, Gary K. (2012). Epigenetic Control of Smooth Muscle Cell Differentiation and Phenotypic Switching in Vascular Development and Disease. Annual Review of Physiology *74*, 13–40. DOI: 10.1146/annurev-physiol-012110-142315.