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

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
qPCR quantative polymerase chain reaction
T thymin
HAoSMC human arortic smooth muscle cell
IF immunoflourescence

1

Introduction

1.1 Coronary artery disease

- CAD is serious. - Describe the pheontype - Describe prevalence - Check intro of Anja, papers for more info - Go into treatment and Risk - herditary parts in disease

1.2 GWAS

GWAS

What are GWAS. Describe how it works, why we do it.

Post GWAS

And their limitations. Possible follow up studies. Focus on computational and cell based assays.

1.3 Muscle Cells in CAD

- We now that smooth muscle cells play a key role - It is widely accepted that there is not only one type of smooth muscle cell - Go into the contractile phenotype and synthetic phenotype - TGFb or PDGF used to induce them - contractile phenotype is thought to be protective

1.4 PDGF Signaling and Oxidative Stress

PDGF Signaling

ROS

- ROS in general and the role of ROS in disease

ROS in PDGF Signaling

- ROS as a second messenger in PDGF signaling

1.5 Aim of the thesis

- Build tool for visual exploration of the CAD GWAS data. - Establish a system to test the role of ROS in CAD.

2

Material

2.1 Manufactors

Just get me the list. lol -> first I need to write the methods part.

2.2 Antibodies

fibronektion 8oxoG primary ones

2.3 Celllines

HAoSMCs

2.4 Primer

GAPDH CNN1 MMP9

2.5 Chemicals

NAC, stuff for IF, ...

2.6 Media, Supplements & Cytokines

2.7 Kits

Jena RNA Extraction

2.8 Consumables

tubes and stuff

2.9 Devises

microscope, zentrifuge cellculture, centrifuge RNA iso, centrifuge plate, taqman, seahorse,

2.10 Programs & Modules

- python 3.9 conda - software on the keyence - lualatex and template of Till Tantau
    - python packages: get a list

3

Methods

3.1 Cultivation and differentiation of HAoSMCs

For the following experiments human arortic smooth muscle cells were used. [Describe some properties of the cells]. Cells were kept at 37°C and 5% CO2 when ever possibile.

Cells were differentiated using Cytokines -> reference to the introdcution and cited literature.

Thawing & Cultivation

Cells were cultivated to a maximum passage of 10, after that new passage cells were thawed. For long time storage cells were kept in liquid nitrogen in [cryo medium]. When required new cells (6th passage) were need cells were thawed at 37°C in the water bath and transfered to a falcon. After centrifugation for 2 min at 300xg the supernatant was remoced and the cell pellet taken up in 14 mL of M231 + SMSG and cultivated in a T75 flask. 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 PBS. Then the cells were incubation with 3 mL trypsin for 4 min at 37°C. After 7 mL M231 were added to the deattched cells and the cells were transfered to a falcon and pelleted for 4 min at 300xg. The supernatant was removed and the pellet resuspenden in M231 + SMGS, seeding 500×10^3 cells per T75 flask.

Preparation of Collagen I matrix

Have a look at the protocol

Differentitation of cells

Differentiation was carried out in 24 wells plates with 1 mL M231 supplemented with 1 % FBS and different cytokines:

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

3.2 mRNA Quantification

Some sentence ragarding the cells and primers and the method.

RNA Isolation

RNA was isolated using the kit and extraction was performed according to the corresponding protocol, using an extra washing step with Ethanol and eluting with 20 µL of RNase-free water. Determination of nucleic acid concentration was carried out with the NanoDrop.

Reverse Transcription

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

The reaction 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

qPCR was performed for CNN1 and MMP9, using GAPDH mRNA levels as a reference (see table for primer pairs). The samples were prepared in a 384-well plate using SYBR® Green Master Mix:

Wells were thoroughly mixed the assay performed using follwing programme on the TaqMan:

Threshhold cycle values were 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 ct of the reference gene GAPDH was substracted from the mean ct of the gene of interest:

$$\Delta ct = ct(CNN1 or MMP9) - ct(GAPDH)$$

The Δct can then be transformed into an relative expression level:

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

Where 10×10^6 is just a constant to yield values that are easier to work with.

Processing of data

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

Short description of the method and again the cells that were used.

Seahorse Assay

- See the protocol written by Tobi and adapted by me. Also website?

Processing of data

- Describe and link the script

3.4 Oxidative Stress Assay

How does it work in general? Used with different conditions. First just the boost, than titration, finally the quench.

CellROX Assay

- Wash - Add stuff - Wait - Image

Boost with PDGF-BB

Describe the different conentrations and evaliation of data.

Time Titration and PDGF-BB Titration

Describe the different conditions and evaluation of data.

Quench with NAC

Short descrription of NAC and description of conditions and evaluation of data

3.5 Immunoflourescence

Fibronektion as marker of matrix. Used cells. Maybe also the anti-8-oxoG AB?

Protocol

Fixate, do a lot of steps.

Processing of data

Me counting pixels.

3.6 Curation of Data for postGWAS analyses

Describe how we get the data. REST APIs, FTP server. Describe a relational database (sqlite3)

Describe all the data that we have, where we get it and how we process it before entering it into the database.

3.7 Visualization of GWAS data

Some general words regarding bokeh webserver. Everything we have is split into the main file that takes care of all the visualization and the data provider that serves as a kind of backend and provides then main file with the required data.

3.8 Enrichment analysis

- Theory, what is happening - The Analysis - Visualization

4

Results

4.1 Differentitaion

Characterization of the phenotype we are inducing.

Expression of CNN1 & MMP9

Markers CNN1 and MMP9.

Further getting a quick glimps into the energy profile of the cells.

Energy profile

Looking at the extracellulat matrix.

4.2 Evaluation of oxidative Stress

PDGF boost of out cells indcues oxidative stress

Characterization of the CellROX Assay

Rescue of ROS production using NAC

4.3 Visualization of GWAS data

After tinkering around for quite a bit we decided to build a webapplication for a multitude of reasons. After tinkering around for even more time we decided to build a backend in the for of sqlite database because it is the most sustainable. Relational databases are used so much for a reason.

Curation of Data for postGWAS analyses

Describe all the data that is in the database and how we processed it.

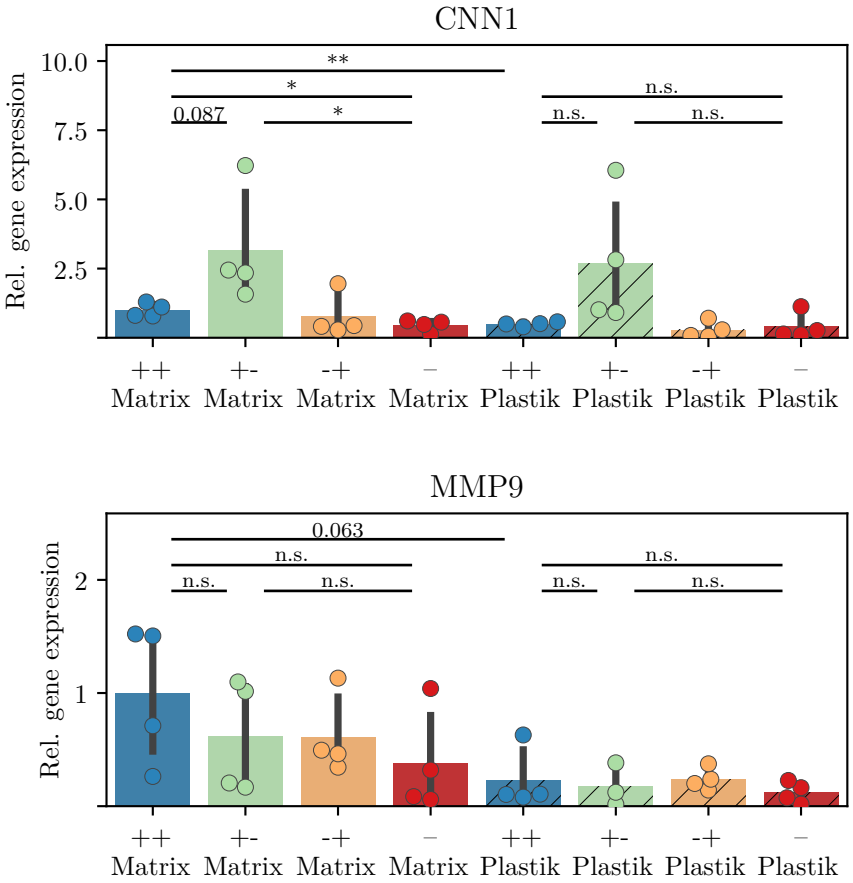

Figure 4.1: Expression of CNN1 & MMP9 in HAoSMCs
WUHU!. Meine erste Abbildung!

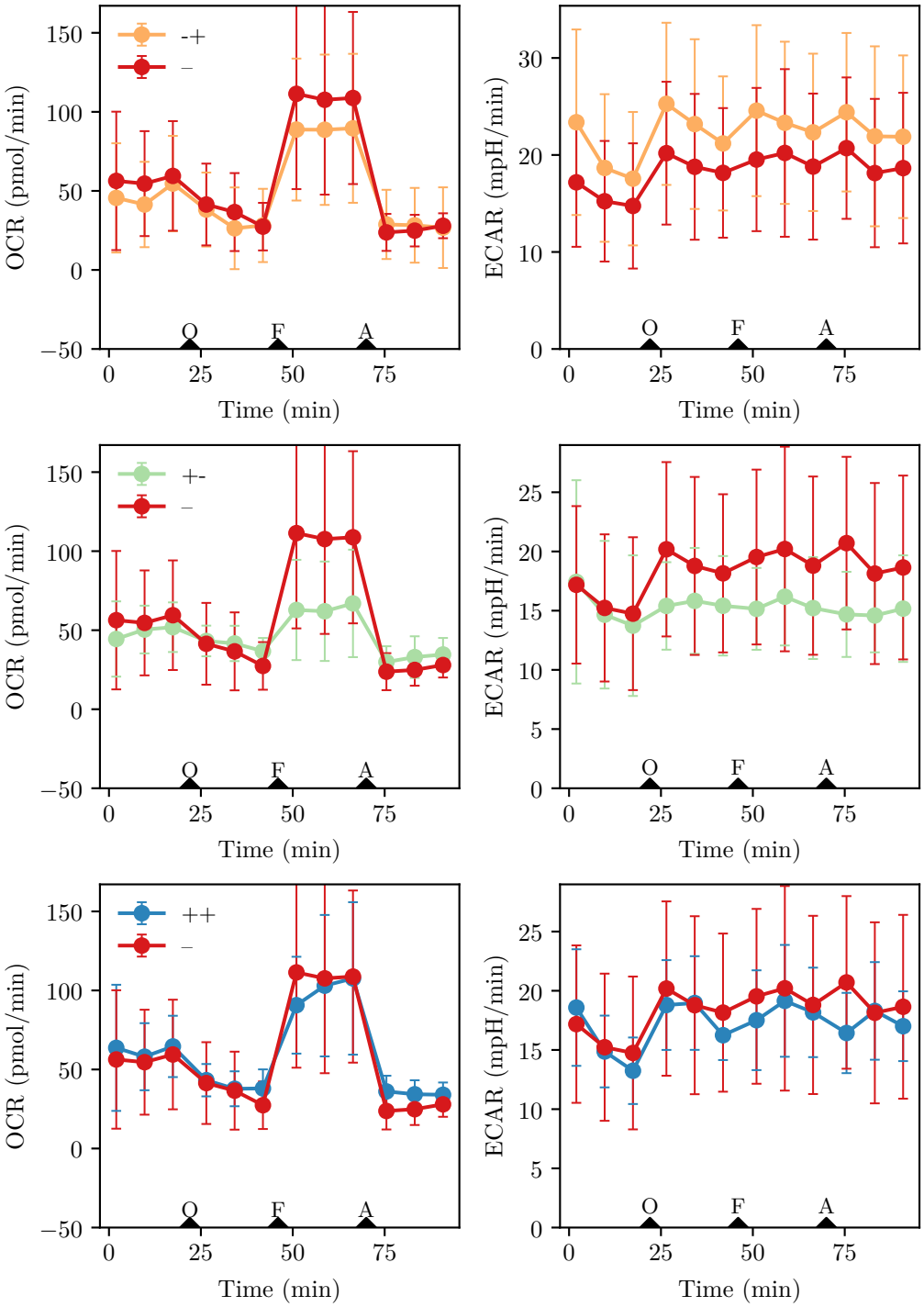

Figure 4.2: Seahorse tracks The seahorse tracks I did record.

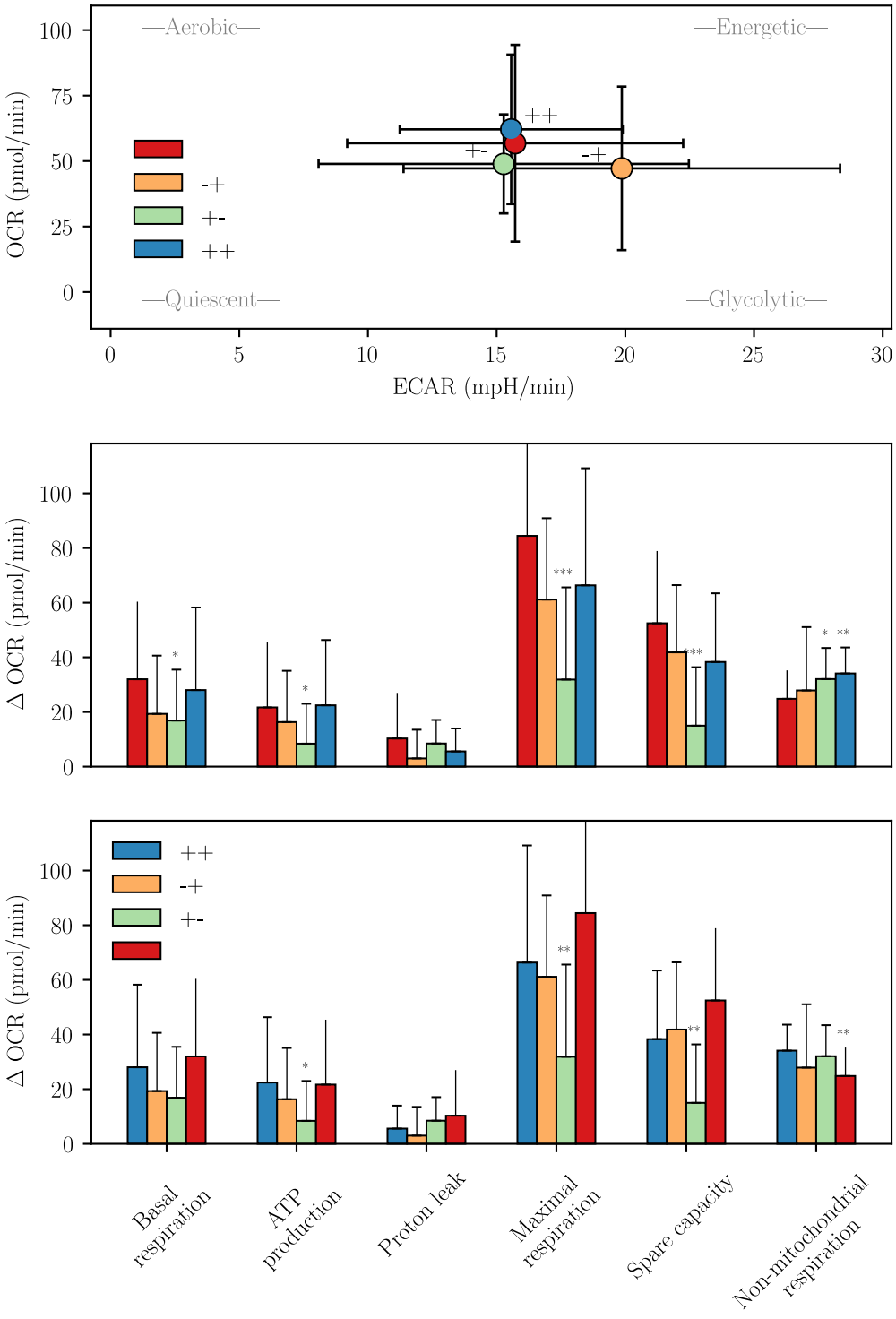

Figure 4.3: Energy profile Evaluation of those seahorse tracks.

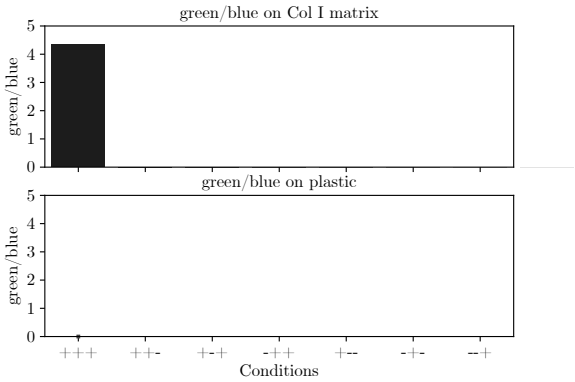

Figure 4.4: Stimulation with PDGF induces oxidative stress.
Repeat of the result already shown by Tobi.

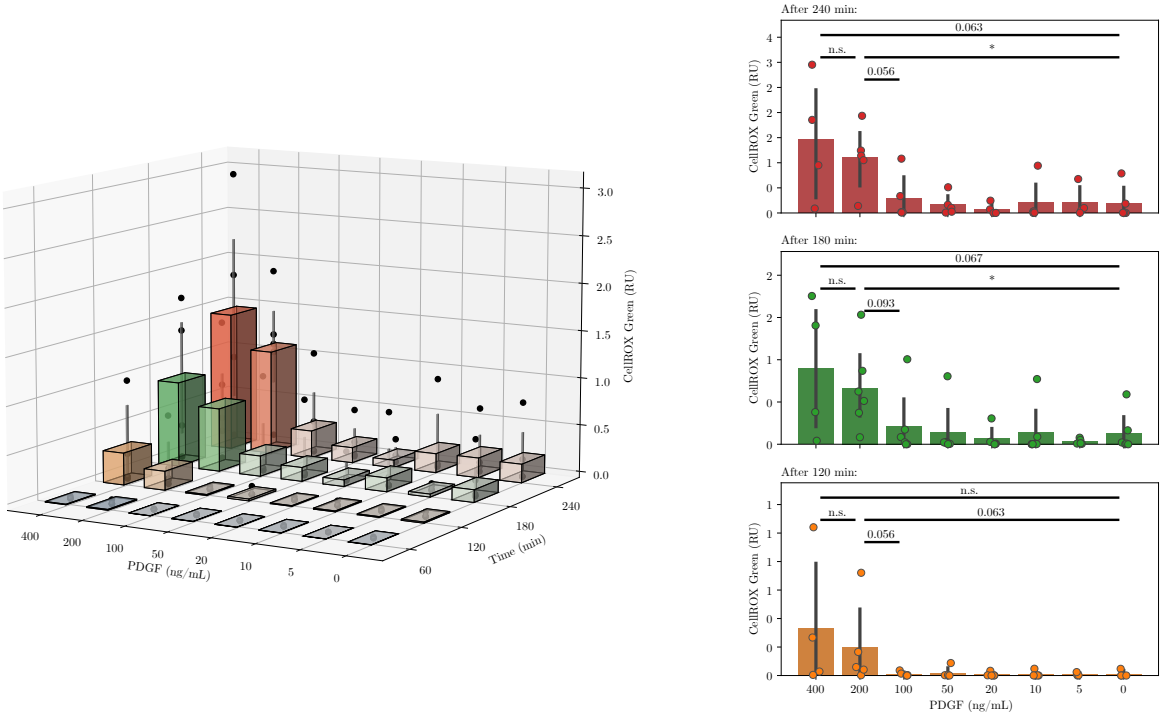

Figure 4.5: CellROX titration Titration

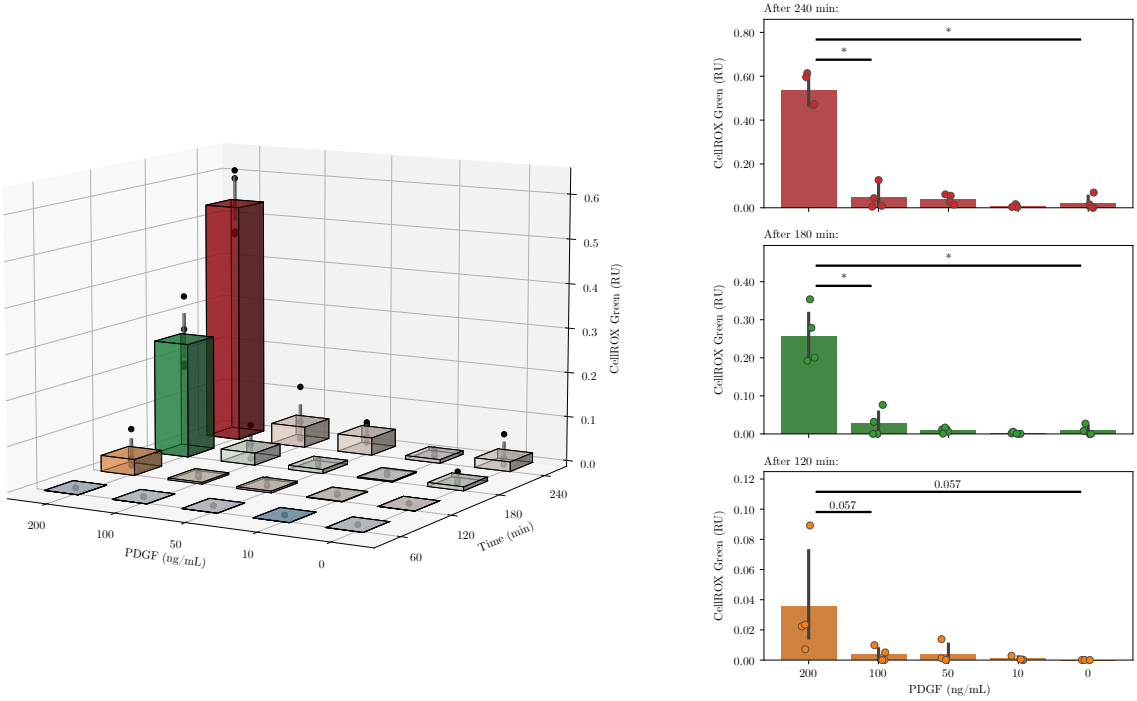

Figure 4.6: Stimulation with PDGF induces oxidative stress - normalized.
My attempt at normalization.

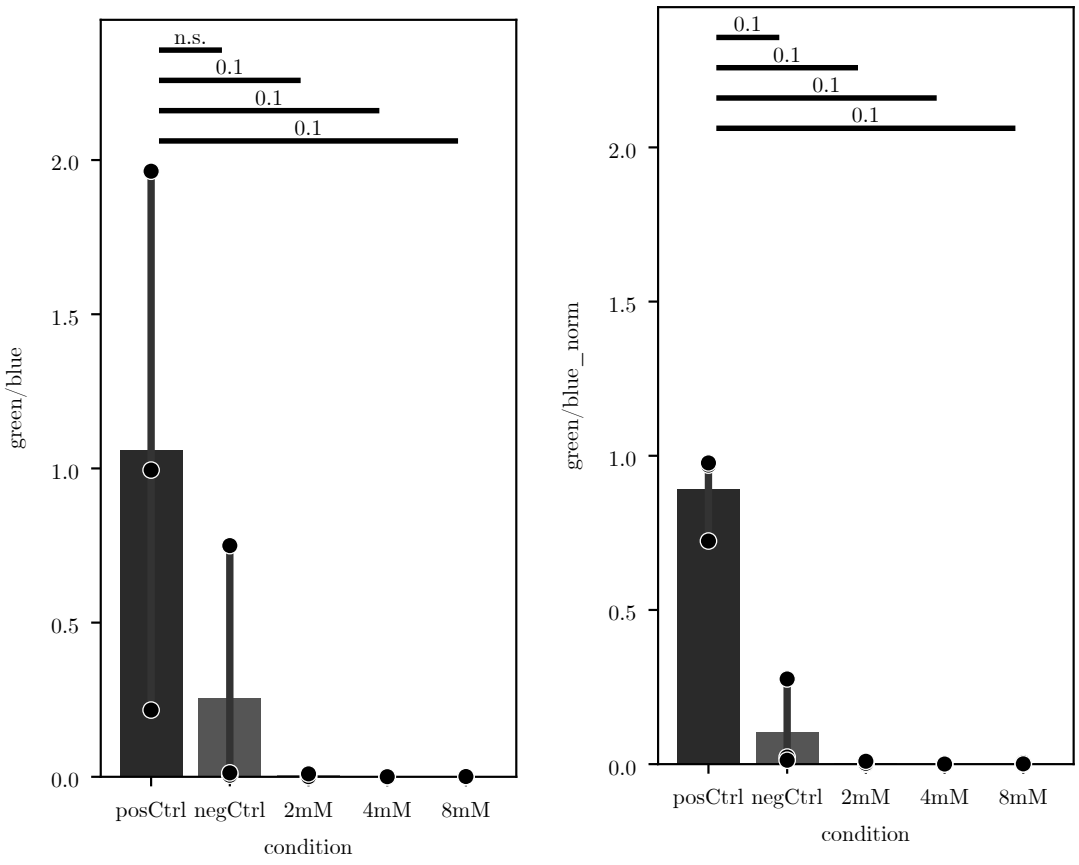

Figure 4.7: NAC quench The NAC quench. :).

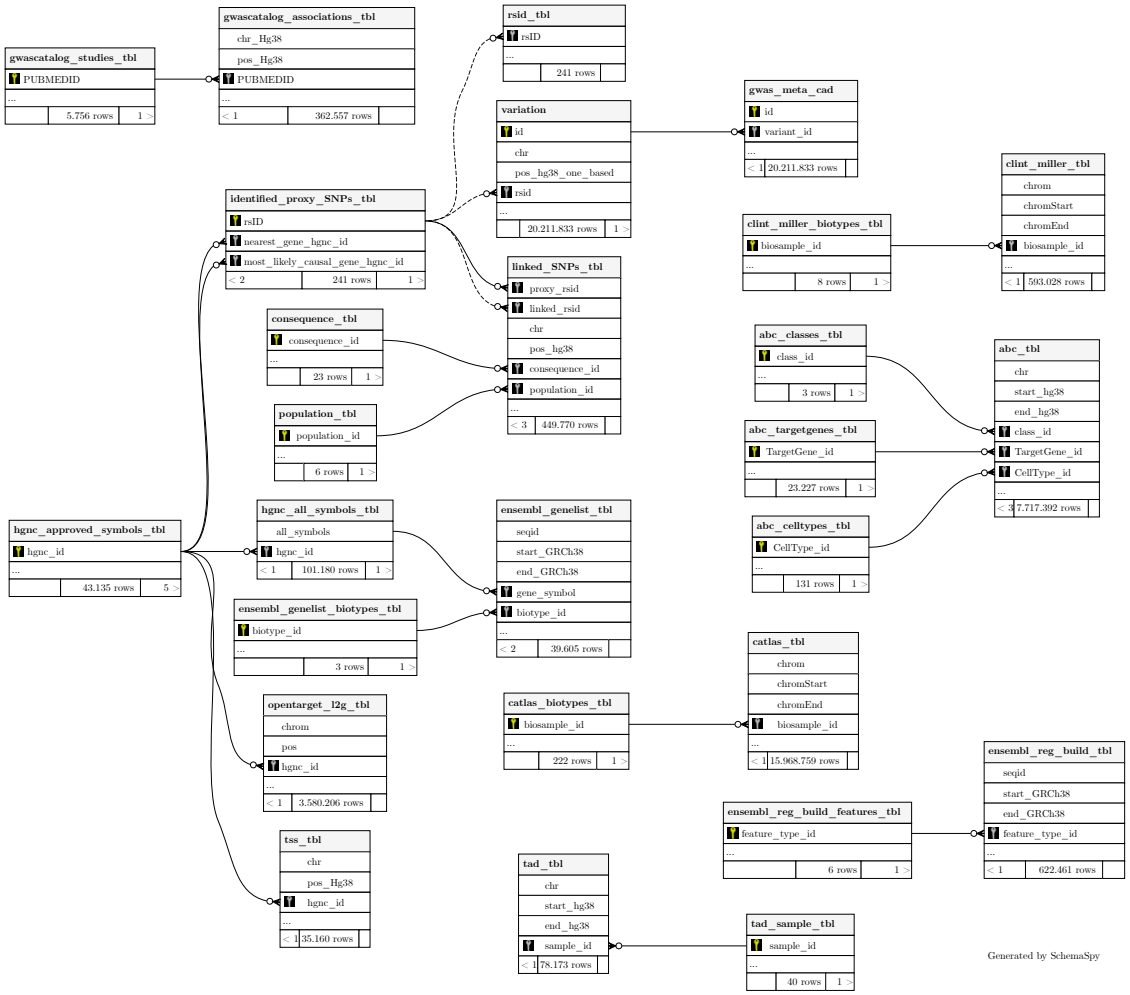

Figure 4.8: My super duper databaseDATABASE!!

Visualization of GWAS data

The first use case of this database and our original goal is the bokeh app for visualization. Wuhu! Show some screenshots but easiest would just be to try it out.

Enrichment analysis

Second the data that we are curating can als be put to use for other scenarios. Running a first test analysis as a proof of concept.

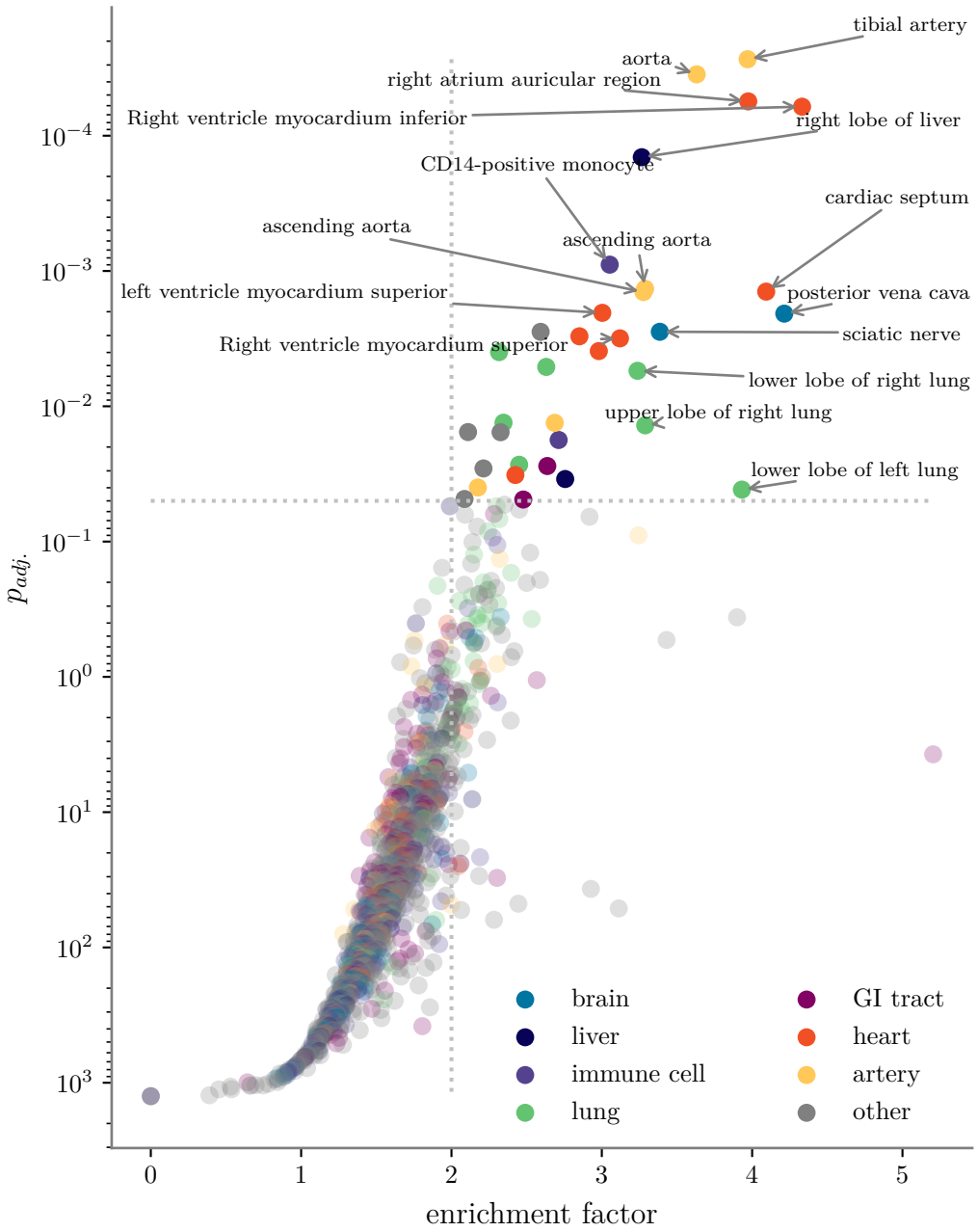

Figure 4.9: Enrichment Analysis This stuff actually seems to be working!!

5

Discussion

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

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.