EVIDENCE FOR OXIDATIVE STRESS IN LUNG TISSUE OF MICE OVEREXPRESSING THE ENZYME TARTRATE-RESISTANT ACID PHOSPHATASE (TRAP) USING GENE EXPRESSION ANALYSIS

Name: Silke Oosting
Student number: S3975770
Supervisor: Prof. B. N. Melgert
Submission date: 22/11/2022

Rijksuniversiteit Groningen | Science and Engineering, department of Molecular Pharmacology

Abstract

Chronic Obstructive Pulmonary Disease (COPD) and asthma are very common chronic diseases characterized by inflamed lung tissue and affect many people worldwide. There is no cure for both diseases yet. Inflammation in the lung walls can cause oxidative stress. In lung tissue of mice with induced COPD or asthma there is also an increase in TRAP values measured. TRAP is an metalloenzyme that potentially causes oxidative stress since a redox reaction can take place at its active site. Oxidants can be detoxified by activation of the Nrf2-Keap1 pathway. The RNA in lung tissue of mice overexpressing Acp5 and wildtype mice was isolated, and the gene expression profile was identified. Differentially expressed genes(DEGs) were obtained using R and pathway analysis and Gene Ontology was done using GSEA and Metascape. First, RNA quality of the samples was confirmed. 11 out of 467 significantly upregulated genes were linked to responses to oxidative stress. Aldh3b1, Btk, Rgs14, Thg1l, PLCy2, Trim30a, Acp5, Syk, Ncf1 and Ncf4 and Hvcn1 were upregulated. Unbiased analysis of the DEGs indicated enrichment of pathways related to B-cell migration, activation and proliferation. In summary, this study found that Acp5 has a link to oxygen radical formation, but no substantial evidence for the induction of oxidative stress in lung tissue was found. What we did discover was that an overexpression of Acp5 gives an enrichment in the pathways involved in B-cell proliferation, activation and survival. The link between Acp5 and B-cells requires further research.

Introduction

Chronic Obstructive Pulmonary Disease (COPD) and asthma are very common chronic diseases characterized by inflamed lung tissue and affect many people all over the world. More than 200 million people globally were diagnosed with COPD in 2019 and have to deal with the disease that is gradually getting worse over the years [1]. In that same year, even more cases of asthma identified - over 700 million cases worldwide. A third of those cases were defined as: ‘poorly controlled asthma’; in this stage medication will not be very effective [2]. There is no cure for either COPD or Asthma, yet.. Despite the existence of reliever and controller medication for both diseases, an increasing number of cases are being observed over time. Studies have shown that the inflammation on the inner surfaces of the lung walls is what causes oxidative stress [3]. Extensive research must be done on pathophysiology and therapies that will decrease inflammation and minimize scar tissue in the lung [4]. Boorsma et al. found an increased quantity of the enzyme TRAP in patients with COPD and asthma [5]. There could be a relationship between the two diseases, the enzyme and oxidative stress.

Tartrate-Resistant Acid Phosphate (TRAP) is an enzyme encoded by the Acid Phosphatase 5(Acp5) gene. TRAP is a metalloenzyme with an iron center at its active site. A Fe2+/Fe3+ centre enables the hydrolysis of phosphatase esters via a redox reaction. [6]. In lung tissue, the active centre of TRAP could provide catalyzation of the redox-reaction where electrons can be exchanged [7]. When electrons are gained or lost reactive oxygen species (ROS) are generated such as peroxides and superoxides. Electrons ‘normally’ come paired but in the valency shell of free oxygen radicals they are

unpaired which makes them unstable. Unstable oxygen radicals tend to react with surrounding (macro)molecules and damage these molecules, such as cellular membrane lipids. Oxygen radical homeostasis is disturbed and detoxification is impaired. [8–10].

The Nrf2-Keap1 pathway plays an essential role in detoxification of free oxygen radicals. Transcription factor-erythroid 2-related factor (Nrf2) induces transcription for genes that for instance produce antioxidants or induces anti-inflammatory responses. Antioxidants donate an electron to the oxidant and homeostasis is maintained. Normally, when no extra antioxidants are needed, the pathway goes as follows: Nrf2 is bound in cytosol to the dimer of Kelch-like ECH-associated protein 1 (Keap1)(Fig. 1). Keap1 is an inhibitor protein that has another binding site for Cullin 3 (Cul3) which generates an 'ubiquitin E3 ligase complex' with the help of Ubiquitin. Ubiquitin "labels" the complex and makes it ready for proteasomal degradation [11]. No antioxidants are made. When there is an imbalance in antioxidant-oxidant levels, and oxidants dominate a protective pathway is enabled. The bond between Nrf2 and Keap1 weakens and eventually breaks if increasing ROS are present. Cytosolic Nrf2 is then transported to the nucleus where it produces more and stable Nrf2. Nucleic Nrf2 merges to musculoaponeurotic fibrosarcoma protein (MAF) and binds to the promotor region of DNA that encodes for antioxidants: antioxidant response element (ARE). Nrf2 is a transcription factor thus the transcription of specific genes against oxidative stress is initiated [12].

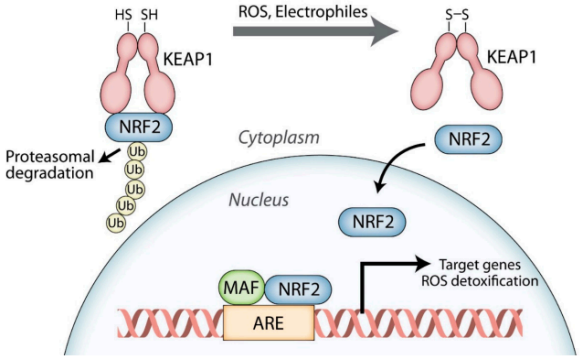


Figure 1. Nrf2-Keap1 pathway where Nrf2 binds to MAF and ARE in the nucleus. The complex is the start of transcription of important genes for detoxification, anti-oxidation and metabolism of oxidants [12].

Important genes involved in the response on oxidative stress by Nrf2 are among others Aldehyde Dehydrogenase 3 Family Member B1 (Aldh3B1) and NADPH quinone oxidoreductase 1 (NQO1). Marchitti et al. have indicated that Aldh3b1 is overexpressed when ROS are increased in the cell. With the inhibitor Keap1 detached from Nrf2, the Aldh3b1 protein responds to oxidants and eventually causes activation of ARE [13]. ARE detoxifies oxidants by producing antioxidants. NQO1 is another gene that is a response on Nrf2. The antioxidant mechanism of action leads to the Nrf2-Keap1 pathway being activated and when ARE starts to produce antioxidants, NQO1 speeds up the process of electron exchange resulting in less reactive oxygen species[14]. Bruton's Tyrosine Kinase (Btk) uses another pathway than Nrf2-Keap1, it induces ROS via an increase in oxidized Low-Density lipoprotein (ox-LDL) which is shown to be higher in COPD patients [15]. But when there is a knock-down of Btk there is an upregulation of cytosolic Nrf2 transporting to the nucleus [16].

Mice with induced asthma/COPD express more Acp5, thus Acp5 may be an inducer for oxidative stress in their lungs. We hypothesize that TRAP is clearly linked to production of oxygen radicals and is expected to contribute to oxidative stress in the lungs. To investigate this hypothesis, we calculated the amount of Differential Expressed Genes (DEGs) in lung tissue of transgenic mice that are characterized by an overexpression of TRAP relative to wildtype mice using statistical program R and subsequently looked if specific pathways were enriched. This could give answers to the question what the implications are of increased TRAP expression in lung tissue of people with COPD or asthma. It is from great importance that fewer individuals worldwide have to cope with chronic lung diseases. ACP5 inhibitors might help in this.

Material and methods:

Animals and dataset

Lung tissue of fourteen male and female mice were bought from the Karolinska Institute of Stockholm. Eight of these samples were genetically modified and thereafter bred at the institute so that they overexpress the gene Acp5. Subsequently more TRAP was obtained in these mice. After the sacrifice date the lung tissue of those mice were obtained and sent to Groningen for subsequent research. RNA was isolated by Marina Visser via the Maxwell method. Promega obtained the 16 LEV simplyRNA Tissue Kit.

R Statistics

The version of Rstudio that was used was 4.2.1, 2022.07.1 Build 554© 2009-2022 RStudio, PBC. With studio R insight in the dataset was obtained. First by doing a Quality Control (QC) using normalized counts from the dataset. A Differential Expression Analysis was done with the RNA sequence. A technique called DESeq2 employs a negative binomial model. The raw counts of the dataset together with the metadata (phenotype) were combined and the DESeq2 model was created. Differentially expressed genes were obtained and grouped by LFC (Log Fold Change). LFC < 0 were downregulated genes and LFC > 0 were upregulated genes. Not every gene has the same number of reads and is thus not proportional to the RNA in each sample, therefore the raw counts were normalized. Unsupervised clustering analysis was done by obtaining a correlation heatmap and a PCA plot assessing how much all mice differ genetically from one another. The relationship between mean and variance was displayed using DESeq2 which calculated a dispersion plot. How we visualized significant genes was by means of a volcano plot. Log2 FoldChanges (2^n where n = FoldChange) were plotted against the p-value. The last way we mapped down results of the DESeq2 was by summarizing the top 20 most up- or downregulated genes, a biased analysis was done on these genes.

GSEA (analysis)

The Gene Set Enrichment Analysis (GSEA) is an unbiased way to identify gene sets and whether they are enriched in a dataset when compared to the control. The version that was used is 3.4.1, 2003-2019. We used the normalized counts of the total gene set to run the GSEA that was obtained via DESeq2 with R. GSEA ranks the data set on log2 FoldChange and generates a 'gene set'(S) which is linked to a typical pathway. An Enrichment Score (ES) is calculated from S and a graph will follow. The higher the enrichment score the more genes of our dataset were present in the pathway and the more likely it is from being significantly different compared to the control group. [17] False discovery rate (FDR) q-value < 0.05 was considered significant.

Metascape (analysis)

The pathways involved with the upregulated genes were visualized and mapped using Metascape (https://metascape.org/gp/index.html#/main/step1)[18]. Metascape is a website where a gene symbol list can be uploaded, and different specific pathways were highlighted. Only genes with p-values < 0.05 were stated as significant. The log2 FoldChange under -2 and above 2 was approved since, this is a fold change that has a minimum of being quadrupled or quartered. This standard was used since Metascape was to detect the most significant pathways when Acp5 was upregulated.

BioJupies (analysis)

BioJupies is a website (https://maayanlab.cloud/biojupies/) that automatically analyzes RNA-sequence data in the form of a notebook. 'Gene IDs' were transferred to 'Gene Names' at the start of the analysis. Gene Names were obtained from GenomeScan B.V. Gene Ontology Enrichment analysis was used to calculate the library size of the read counts.

Statistical analysis

Statistical analyses were conducted using R (version 4.2.1). A Mann-Whitney U test was used to compare two independent groups and define the statistical differences. The data presented as mean ± standard deviation (SD). A gene was determined as a significantly differential expressed when a double-sided P-value < 0.05 was true.

Excluded data

No data was excluded from this study.

Results

Quality control

Normalization and unsupervised clustering analysis

A correlation heatmap was created to show how biologically different the transgenic mice were from wildtype mice. The transgenic mice did not significantly cluster apart from the normal mice (figure 2) and there was no clear correlation between replicates separated on condition. This may indicate that both groups did not differ that much from each other.

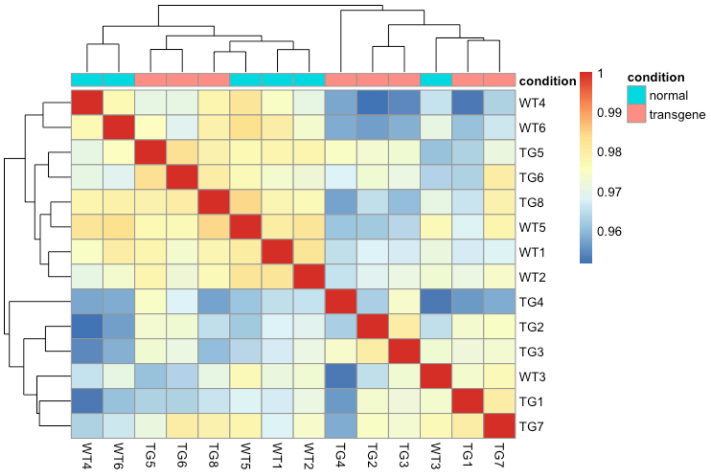


Figure 2: Hierarchical heatmap using correlation values between ± 0.95 and 1. Where 1 (red) means perfect correlation.

Principal Component Analysis (PCA)(Fig. 3) is another method by which we expected to see clustering in conditions. The two groups of mice were expected to be separated on the first principal component (PC1). PC1 shows the biggest variations between the groups. 31% of the mice overexpressed with TRAP corresponds to PC1. The wildtype mice cluster apart from the transgene mice except TG8 and WT2, these were clustered in the other condition. This may indicate an unintentional data exchange between TG8 and WT2. In PC2 no clear separation was seen thus PC2 was not explained by this experiment.

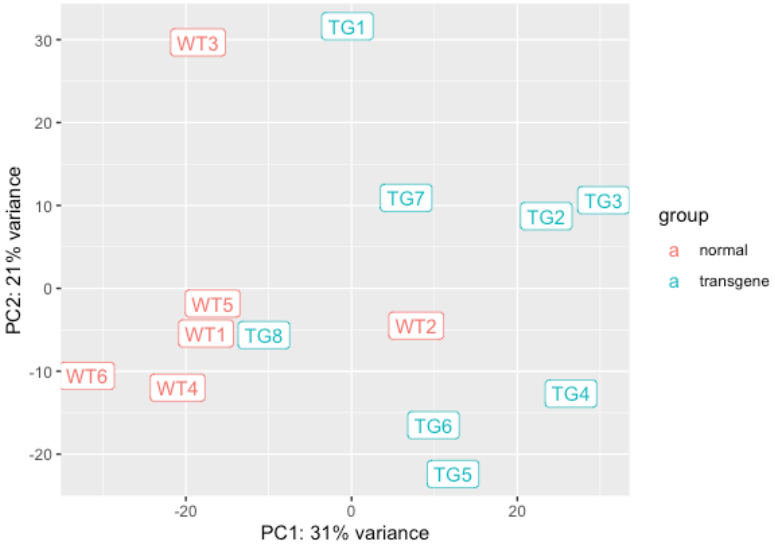


Figure 3: Principal Component Analysis based on groups. Transgene vs. Normal is displayed in the legenda.

To verify that indeed WT2 and TG8 were mixed-up, the gene Acp5 is plotted against the normalized counts which represented an expression graph (Fig. 4). The blue dot that clustered between the wildtype condition is TG4, this indicated that a mix-up of TG8 and WT2 did not seem likely. Figure 4 was also a control for if those transgene mice were successful genetically modified, and thus overexpress ACP5. The different conditions clustered together except for TG4.

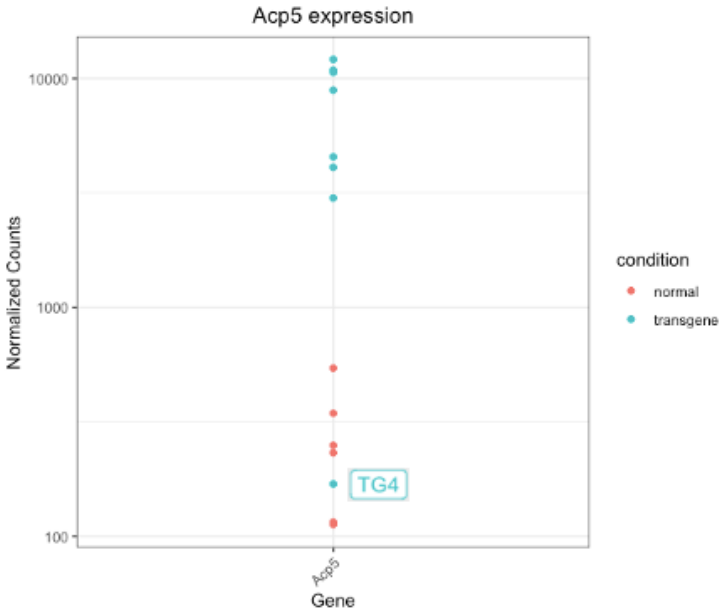


Figure 4: Expression graph of the Acp5 gene in normal and transgenic mice. TG4 is labeled since the counts are lower than expected.

The library size in million reads per sample obtained from Metascape (Fig. 5). WT3 and TG2 are a threefold higher than the other reads. The number of reads was smaller for TG4 and WT4. This could have something to do with the low reads counts of TG4 that clusters between the counts of the wildtype mice.

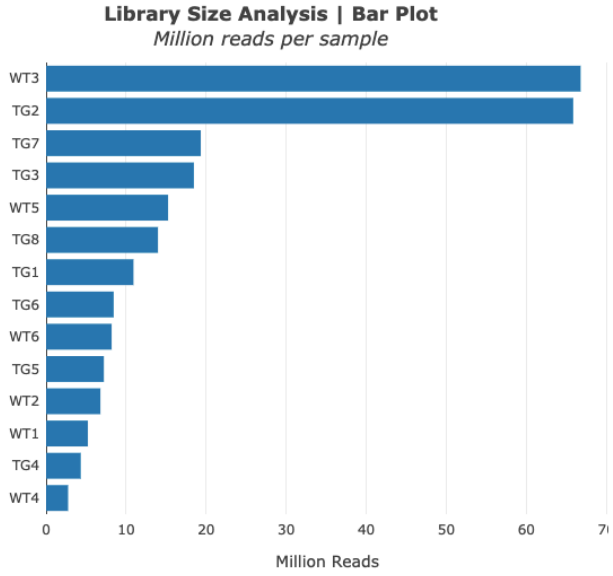


Figure 5. The total number of reads mapped for each RNA-seq sample in the dataset is shown in a bar chart. The samples are displayed on the y-axis while the reads (in millions) are on the x-axis.

To see whether the genes in our data set show enough variation from the mean a dispersion plot is made. Each black dot in the dispersion plot (Fig. 6) encodes for a gene. Variances were supposed to increase with an increasing mean value. The dispersion is displayed by mean or variance and gives information about how far the data is separated from the mean value. The increase in mean counts per gene in figure 6 gives a decrease in dispersion.

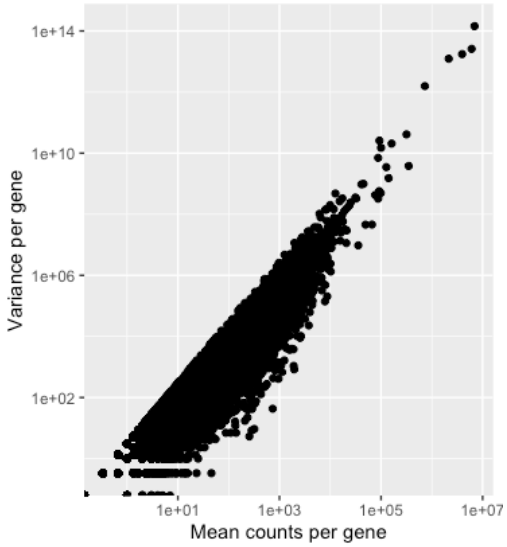
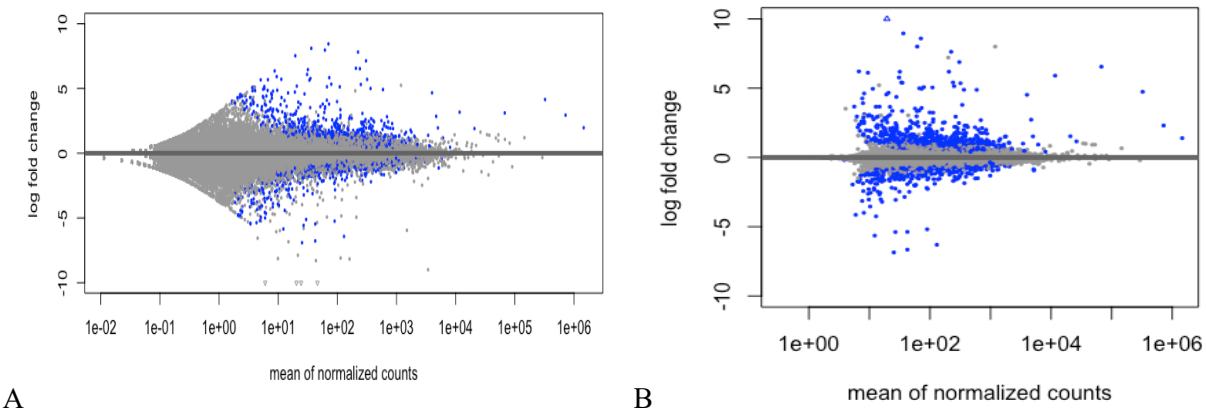


Figure 6: mean counts per gene plotted against the variance per gene in dataset.

Up- and downregulated genes

The MA plot represents the mean of the normalized counts versus the log2FoldChanges for all genes tested using DESeq2 model. The significantly DEGs are shown in blue. The large log2FoldChanges with low mean of normalized counts (Figure 7a) still needs to be corrected for because these FoldChanges are hardly accurate due to low number of counts. After shrinkage (Figure 7b) the genes with a low gene count and high log2 fold changes were deleted.



A B

Figure 7a/b: MAplot where the mean of the normalized counts is plotted against the FoldChange. Figure 7a shows all the genes where the blue dots are significant up- or downregulated genes. In figure 7b a shrinkage is applied.

Differential Expressed Genes

Data of Acp5 overexpressing and wildtype mice

A total of 55,487 genes were identified and selected from the Acp5 and wildtype lung tissue specimen. The analyses were based on the P-value, the adjusted p-value and the FoldChange. The log2FoldChange was set <-1 for downregulated genes and >1 for upregulated genes. Where -1 means that the gene expression was halved and 1 means it doubled. DEGs were stated to be significant when the (adjusted) p-value was <0.05. Subsequent research (DESeq2) showed that there was a total of 715 significantly DEGs from which 476 genes were upregulated and 239 genes downregulated. Of the 476 upregulated genes 11 genes play a role in the context of this research. Various processes involved in the response on oxidative stress were examined.

Pathway Analysis

DEGs related to oxidative stress

We investigated the effect of overexpression of Acp5 in mice and if it had something to do with specific pathways that are a response of oxidative stress or inflammation. There was mainly looked at upregulated genes in transgenic and wildtype mice (see table 1). In '(Cellular) responses to oxidative stress'; Aldehyde Dehydrogenase 3 Family Member B1 (Aldh3b1), Bruton tyrosine kinase (Btk), Regulator of G Protein Signaling 14 (Rgs14) and TRNA-Histidine Guanylyltransferase 1 Like (Thg1l) were upregulated (Fig. 8a-d).

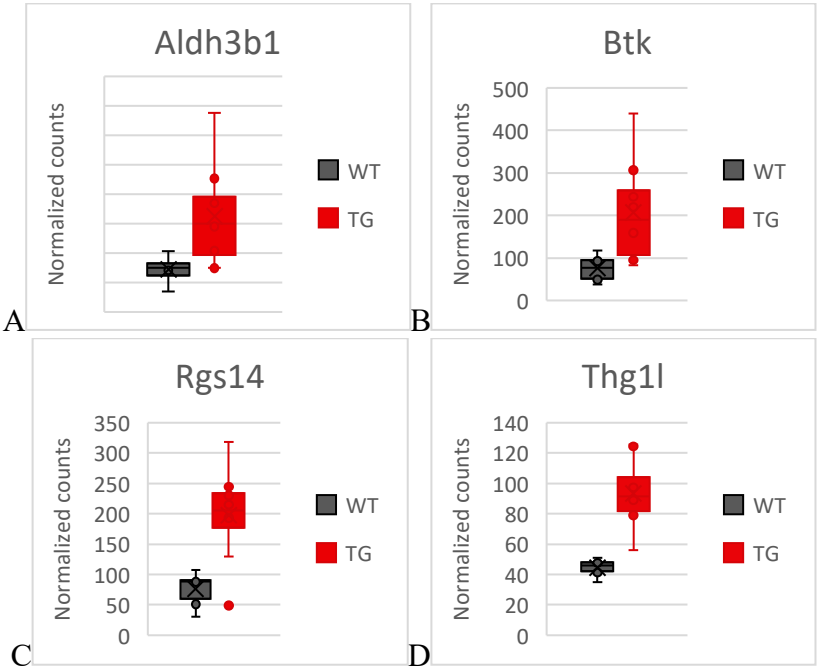
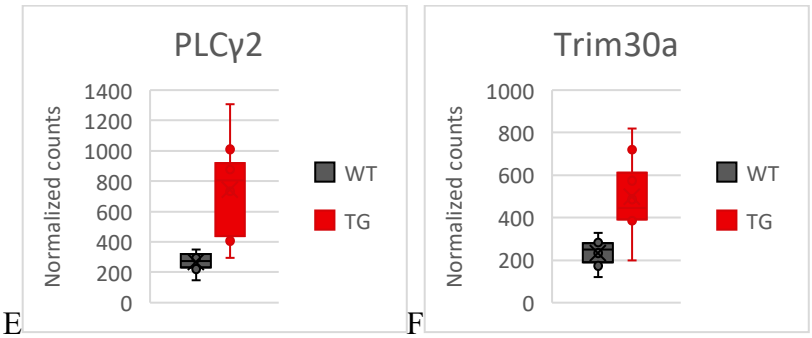


Figure 8a-d: Whisker boxplots obtained from the different genes in the pathway of 'responses to oxidative stress'. The normalized counts were taken for comparison between wildtype mice (WT) indicated in grey and transgenic mice (TG) in red. The inner points show of the different samples show.

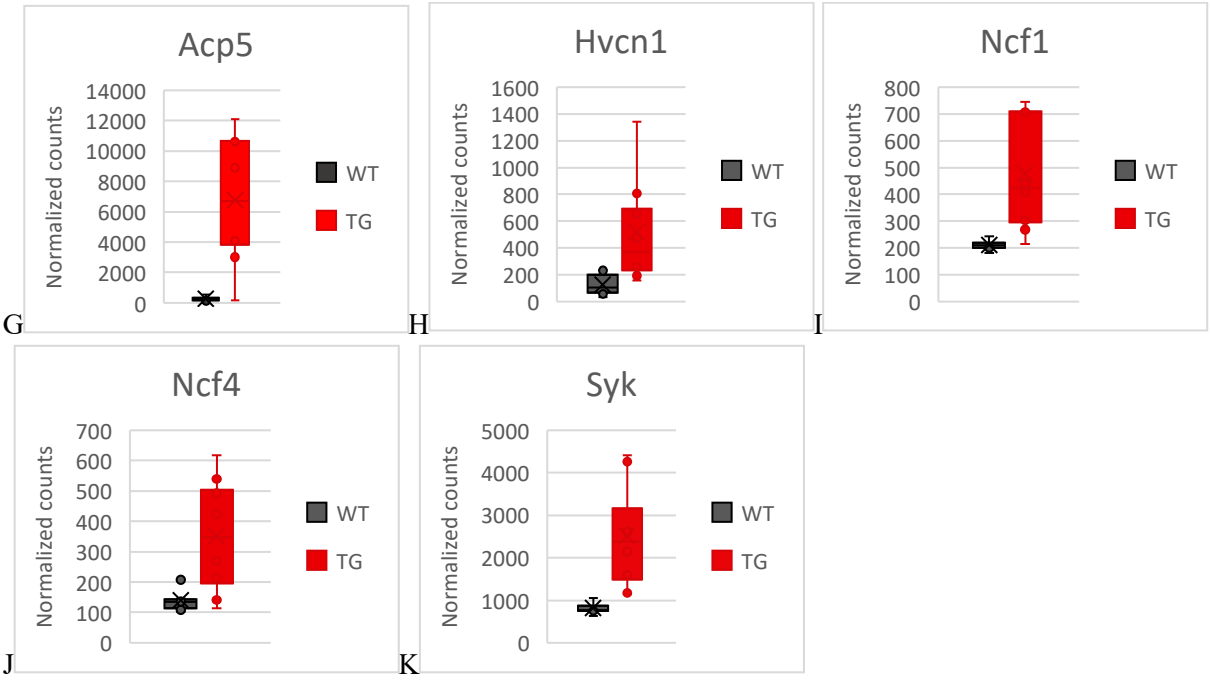
The NLRP3 inflammasome is activated by ROS and other inflammation factors[19]. It is therefore a reaction on the oxidative stress that Acp5 was expected to induce. In the pathway 'regulation of NLRP3 inflammasome complex assembly' two genes were upregulated (Fig. 8e/f) namely phospholipase C gamma 2 (PLCγ2) which stimulates the cytokine production of IL-1b via calcium-activated NLRP3[20,21]. and tripartite motif-containing 30 alpha (Trim30a). Trim30a negatively regulates the NLRP3 inflammasome via inhibition of ROS[22].



E F

Figure 8e/f. Whisker boxplots obtained from genes enriched in the pathway: 'regulation of NLRP3 inflammasome complex assembly'. The normalized counts were taken for comparison between wildtype mice (WT) indicated in grey and transgenic mice (TG) in red. The inner points show of the different samples show.

Superoxides are products that contain a free oxygen radical. If Acp5 induces ROS, there is expected to also have metabolism of species containing free oxygen radicals and therefore there was also looked at the pathway: 'superoxide metabolic processes. Upregulated genes are TRAP encoded by the Acp5 gene, this was not unexpected. The Hvcn1 gene which encodes for Hydrogen voltage-gated protein channel (1). Neutrophil Cytosolic Factor 1 and 4 (Ncf1, Ncf4) genes encodes for $p47^{phox}$ and $p40^{phox}$. Spleen Tyrosine Kinase (Syk) were upregulated (Fig. 8g-j).



G H I

J K

Figure 8g-k: Whisker boxplots obtained from the different genes in the pathway of 'superoxide metabolic processes. The normalized counts were taken for comparison between wildtype mice (WT) indicated in grey and transgenic mice (TG) in red. The inner points show of the different samples show.

The genes involved in the responses to inflammation and oxidative stress were significantly expressed but none of the genes of those pathways came back in the top 20 most upregulated (Tab. 2) or downregulated genes (Supplemental data Tab. 3). Whether inhibition of Acp5 would have been sufficient to markedly reduce oxidative stress in mice or humans diagnosed with COPD is up for debate.

Table 1. Significant upregulated genes with their GeneID, Gene symbol and the pathway they are enriched in. log2FoldChanges and (adjusted) p-values are also given.

GeneID	Gene symbol	Pathway*	log2FoldChange	p-value	padj
ENSMUSG00000001348	Acp5	3	4.67	1.17E-13	1.26E-10
ENSMUSG00000024885	Aldh3b1	1	1.17	7.18E-04	2.04E-02
ENSMUSG00000031264	Btk	1	1.47	2.16E-04	8.87E-03
ENSMUSG00000052087	Rgs14	1	1.39	6.87E-05	4.20E-03
ENSMUSG00000011254	Thg1l	1	1.04	1.51E-06	2.05E-04
ENSMUSG00000034330	PLCγ2	2	1.49	3.36E-06	3.99E-04
ENSMUSG00000030921	Trim30a	2	1.09	2.62E-04	1.02E-02
ENSMUSG00000015950	Ncf1	3	1.18	4.07E-05	2.91E-03
ENSMUSG00000071715	Ncf4	3	1.35	2.00E-04	8.49E-03
ENSMUSG00000021457	Syk	3	1.63	1.70E-07	3.21E-05
ENSMUSG00000064267	Hvnc1	3	2.03	1.08E-04	5.80E-03

*

1	(Cellular) responses to oxidative stress
2	regulation of NLRP3 inflammasome complex assembly
3	superoxide metabolic processes

The volcano plot provided a global overview of the differentially expressed genes (Figure 8). The chosen thresholds (also indicated with the dashed line) determined whether the genes were significantly expressed. A log2 FoldChange under -1 and above 1 is specified, and a p-value less than 0.05 was accepted. The most significant up- and downregulated are labeled and will be further discussed later on. A close-up of Fig. 9a is given in Fig. 9b. Significant genes related to inflammation and oxidative stress were displayed in red in this figure. The x-axis ranges from 0-2.5 and significant log2FoldChange ranges from 1 to 2.5 thus all genes, except for Acp5, ranged in this area.

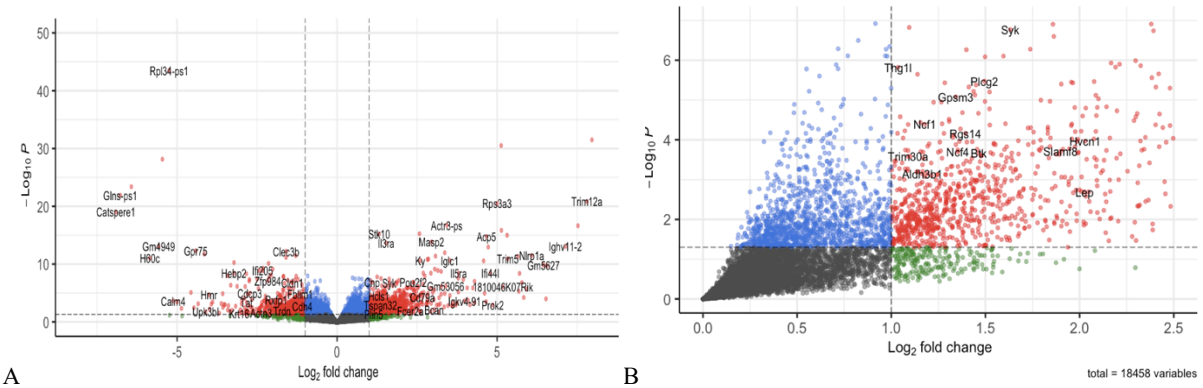


Figure 9a. Volcano plot where the FoldChange is plotted on the y-axis against the p-value on the x-axis. Significant up- and downregulated genes are shown in red where with -1 > LFC > 1 and a p-value < 0.05. Non-significant genes are shown in blue, where the p-value is <0.05 but FC is too low; between -1 and 1. Green dots have a good fold change under -1 and above 1 but have a p-value > 0.05. Grey dots are not significant at all because they have a FC between -1 and 1 and a p-value > 0.05.
In 8b. a zoomed-in volcano plot labeled with significant upregulated genes for a variety of pathways that have to do with oxidative stress.

In addition, a Gene Ontology Enrichment Analysis was done (Fig. 10) and a network of the most significant pathways was obtained. Clusters were based and colored on p-values. An overarching theme was seen namely B-cells and everything that has to do with this. Activation of immune cells (red), B cell activation (blue), regulation of B cell proliferation (orange), CD22 mediated BCR regulation (purple) and immunoglobulin production involved in immunoglobulin- mediated immune response (yellow) all had to do something with B-cells. Next to these 'B-cell' clusters there is also a smaller but not negligible cluster that has to do with inflammatory responses (brown).

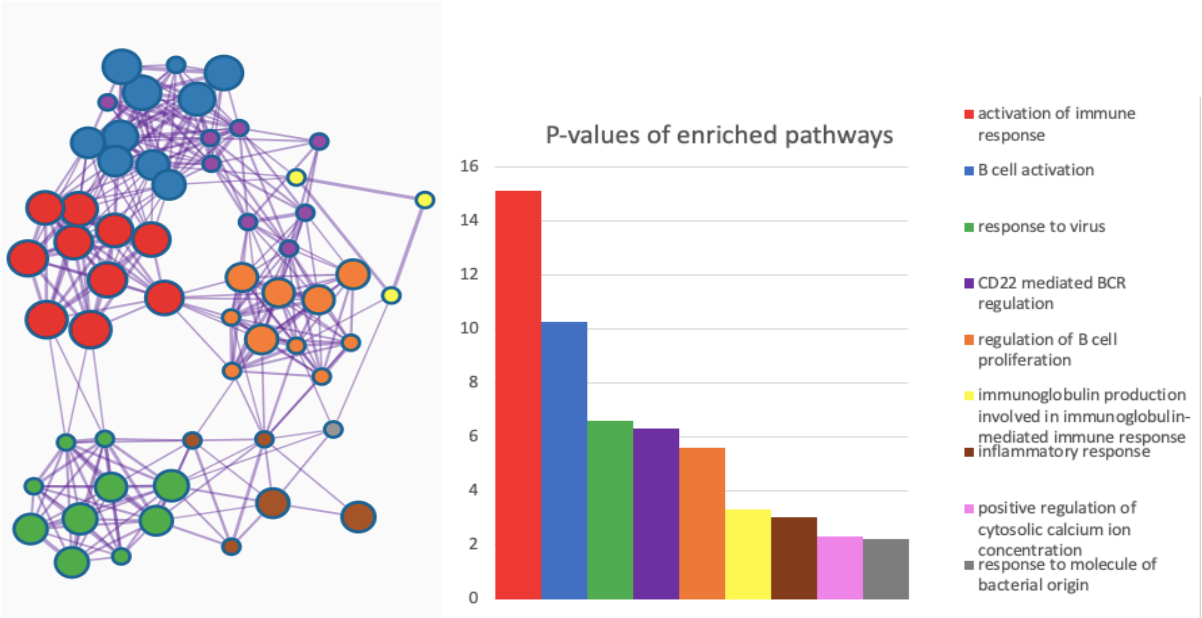


Figure 10: Gene Ontology Enrichment Analysis demonstrated in Metascape. The most upregulated genes with a log2 FoldChange under -2 and above 2 were imported in Metascape. The system in colored on p-value. P values can be found in Supplemental data 2 (S2).)

According to the GO Enrichment analysis oxidative stress is not the most abundant and significant pathway that has been changed by the overexpression of the enzyme TRAP. Other pathways play a more significant role, therefore table 2 will give an insight in the most significant genes based on

the log2FoldChange. Different kinds of immunoglobulins fill the top 20 and other genes than genes related to oxidative stress are present such as Zfp990 and Dynlt1. Zinc finger protein 990 for example was involved in multiple processes, including cell proliferation [23].

Table 2. top 20 upregulated genes based on log2FoldChange. All values have a $p < 0.05$.

Gene symbol	GeneID	log2FoldChange	pvalue	padj
Zfp990	ENSMUSG00000078503	8.439170432	2.97E-21	7.75E-18
Dynlt1	ENSMUSG00000096255	8.102625051	1.16E-13	1.26E-10
Gm10020	ENSMUSG00000057262	7.963116596	8.03E-30	4.89E-26
Trim12a	ENSMUSG00000066258	7.810137938	5.93E-20	1.20E-16
Rpl15-ps2	ENSMUSG00000098915	7.52517511	4.76E-16	7.25E-13
Ighv11-2	ENSMUSG00000096108	7.1340572	1.01E-12	1.03E-09
Tmem181b-ps	ENSMUSG00000096780	6.562123231	1.43E-09	7.45E-07
Gm22513	ENSMUSG00000096349	6.52715865	0.00021809	0.02699858
Gm5627	ENSMUSG00000093812	6.452131552	1.07E-09	5.94E-07
Nlrp1a	ENSMUSG00000069830	6.082168856	4.07E-11	3.10E-08
Ighv2-9	ENSMUSG00000096638	5.916672894	7.54E-06	0.0017225
Vaultrc5	ENSMUSG00000065145	5.83274664	0.00014574	0.01929755
Ighv9-3	ENSMUSG00000096459	5.774960627	7.28E-07	0.00021112
Igkv14-126	ENSMUSG00000094345	5.699241442	2.83E-08	1.23E-05
Ighv2-5	ENSMUSG00000096498	5.680707457	1.65E-05	0.00338895
Tmem181c-ps	ENSMUSG00000093880	5.646236963	6.82E-11	4.98E-08
Trim5	ENSMUSG00000060441	5.332748819	1.86E-10	1.17E-07
Gm15772	ENSMUSG00000062353	5.311439586	4.62E-14	5.63E-11
1810046K07Rik	ENSMUSG00000036027	5.182274087	3.70E-06	0.00096545
Trim30d	ENSMUSG00000057596	5.134251872	9.48E-15	1.33E-11

GSEA

To identify which specific pathways are mostly enriched when uploading the entire gene set, we performed a GSEA. We found that the significantly downregulated genes from supplemental data S3 were not significantly enriched in pathways. Therefore, downregulated genes by overexpression of Acp5 relative to wildtype mice were not further examined. We found that 32 upregulated pathways were significantly enriched in the Reactome. Enrichment plots were made from the top 5 most significant pathways (Fig. 11a-e). For each gene that is enriched in a pathway the green line will move upwards and the enrichment score (ES) will increase. The eventual ES is the height of the peak in the figures. All five pathways have an FDR q-value of 0.0. 'Antigen activates B cell receptor (BCR) leading to generation of second messengers' has an ES of 0.74 and was ranked first. 89/136 genes in the pathway were enriched. All genes that were enriched in the first five pathways can be found in Supplemental data S4. From the top 20 most upregulated genes in table 2 the gene Ighv11-2, which encodes for immunoglobulin heavy variable 11-2, is second most enriched in this pathway.

Second, 'Role of LAT2, NTAL and LAB on calcium mobilization' has an ES of 0.71 with 78/128 enriched genes (Sup. Data S4). Linker of activator of T-cells 2 (LAT2), non-T-cell activation linker (NTAL) and Linker for B-cell activation (LAB) are messengers for the B-cell activation

cascade[24]. Since FCERI mediated Ca^{2+} mobilization is the next figure (Fig. 11c) which is highly enriched there could be assumed that the FCERI is also stimulated.

Both 'FCERI mediated Ca^{2+} mobilization' and 'Role of phospholipids in phagocytosis' had an ES of 0.69 but subsequently different amounts of enriched genes (84/142 and 81/137). Lastly, 'Signaling by the B cell receptor BCR' had an ES of 0.64 and 123/216 enriched genes.

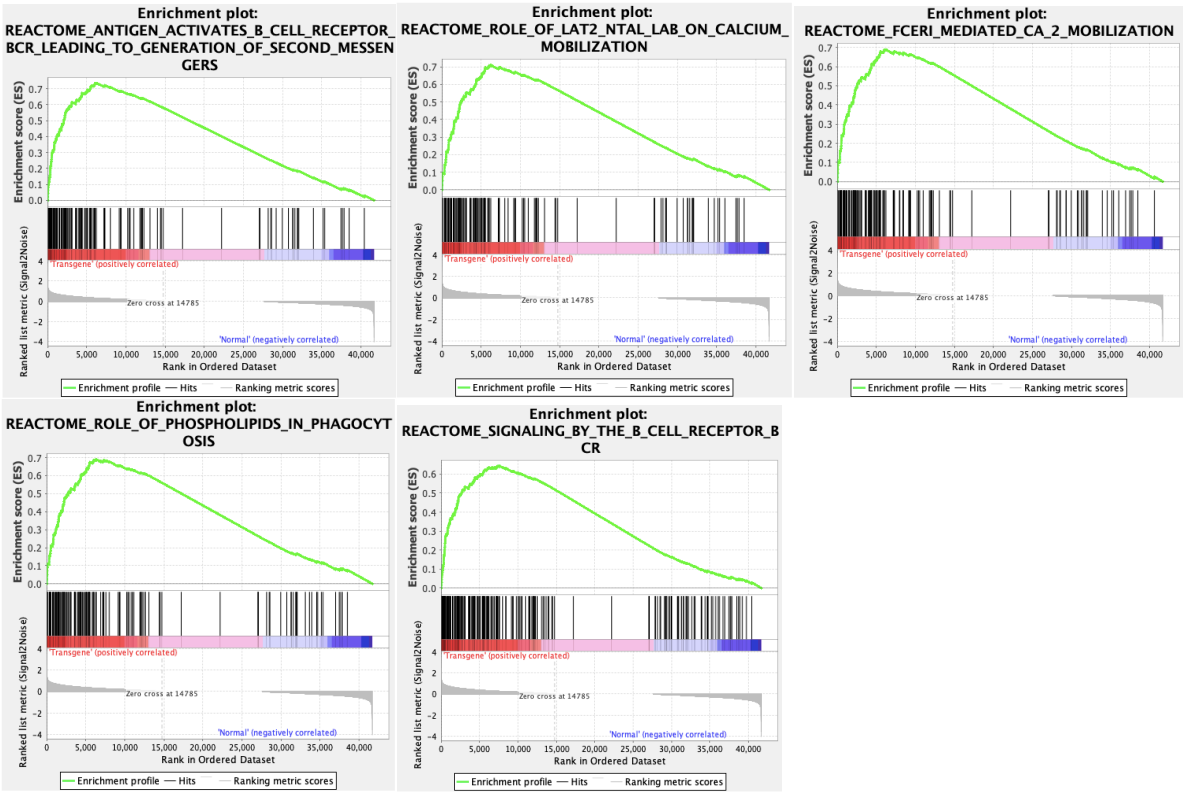


Figure 11a-e. illustrates the enrichment plots of the most significant upregulated pathways. The green line shows the progression of the enrichment score with ES as a peak. The black stripes indicate different genes enriched in the pathway.

All ES are above zero which means that the genes at the left side of the graph are highly expressed in transgenic mice compared to control mice. This illustrates that there are no significant genes upregulated in normal mice. This means that the genes enriched in the blue area, at the right, are the downregulated genes in transgenic mice but are not significantly downregulated in this current study.

Discussion

Oxidative stress is an ongoing process in every part of the body. There is no concern when this process is in homeostasis, this changes when homeostasis is disturbed and more reactive oxygen species are detected. Once free oxygen radicals are generated, a response follows. We look at this response because no evidence for oxygen radicals is encoded in our DNA [25,26]. In the tissue of the lungs of mice with induced COPD or Asthma, this ROS quantity is elevated. This response is a process that leads to detoxification and anti-oxidation. To find out whether TRAP stimulates the response to oxidative stress - and therefore the antioxidant production - RNA was isolated from lung tissue of mice overexpressing Acp5 gene and wildtype mice. Subsequent research was done to study the effect of Acp5 on cellular pathways.

Before conducting the present study, a QC was done on the dataset. No clear correlation between all mice could be obtained from the correlation heatmap due to fact that the transgene mice overall do not differ that significantly from wildtype mice. As a result, no clustering of TG vs WT was obtained. For our study this means that only based on the quality control, an Acp5 inhibitor is likely to have little effect on COPD. Because no clear correlation was found a PCA plot and a gene expression graph were made. It turned out that there was a clear segregation of the transgenic mice and the wildtype except for TG4. The low counts of TG4 may be proven due to the small library size seen in figure 5 but since there was corrected for it, this was not very likely. The Acp5 expression of TG4 may be low but in other genes TG4 does not cluster apart from its condition. Therefore, we assume that the Acp5 expression was, by coincidence, low in the section of the lung tissue that was taken for extraction. Quality of the dataset was guaranteed, and subsequent research was done.

In the current study we did a biased analysis and searched for upregulated genes in (cellular) responses to oxidative stress (GO:0034599, GO:0006979). We found higher expression of Aldh3b1, Btk, Rgs14 and Thg1l in mice overexpressing ACP5 compared to wildtype mice. We postulated that more ACP5 would lead to more oxidative stress and a higher expression of these genes fit this hypothesis. Aldh3b1 is a member of the Aldehyde dehydrogenases (ALDHs), the family plays an important role in the defense against toxification via generation of reactive aldehydes and oxygen radicals [13]. Generation of these radicals leads to interferes in the Nrf2-Keap1 pathway and upregulation of the metabolism of ROS is obtained. Therefore, more oxidative stress would lead to more expression of this enzyme. Btk is a tyrosine kinase (TK) gene that, unlike the ALDH-family, induces oxidative stress. This gene is markedly elevated and essential for the toxicity caused by ox-LDL[16]. Rgs14 is, what the name indicates, a regulator of G protein signaling and is involved in many processes. A study on Rgs14 indicated that Rgs14 induces oxidative stress via a cascade of different pathways [27]. Many different responses to oxidative stress are possible, Thg1l was reported to mediate the survival of cells by protecting them from ROS-induced apoptosis [28].

We found some highly expressed genes leading to the (de)activation of the inflammasome NLRP3 complex assembly (GO:0044546) in mice overexpressing ACP5 compared to wildtype mice.

To make the NLRP3 active, a cascade of reactions takes place where the cytokine Il-1b is released in the cytosol (Fig. 12). Upregulation of IL-1b, and thus upregulation of NLRP3 inflammasome, corresponds to the hypothesis since it is expected to be linked to Acp5-induced ROS. Higher amounts of IL-1b induces the infiltration of immune cells to places in the body where help against inflammation is needed.

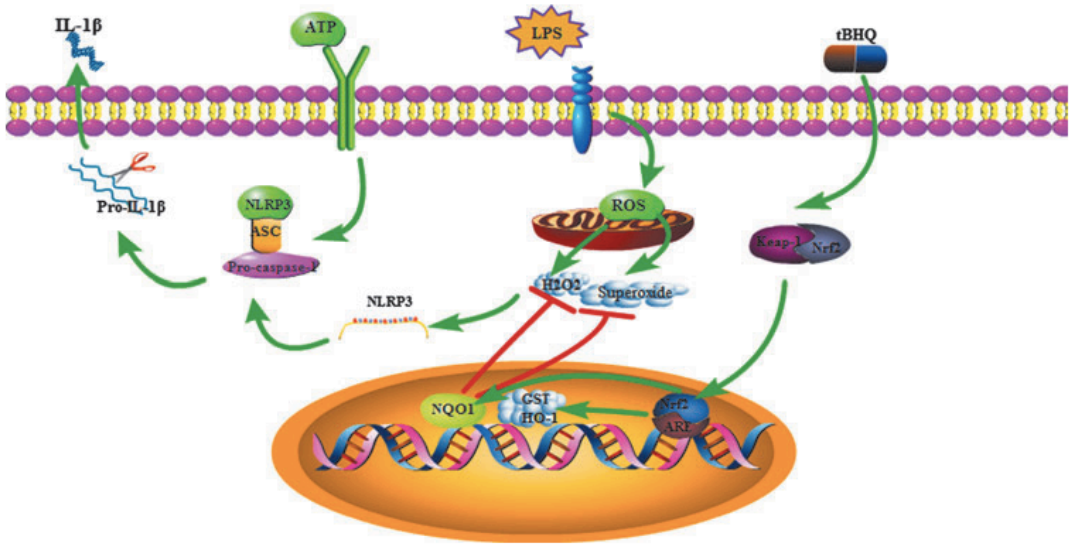


Figure 12. Mechanism of action on how Nrf2-Keap1 pathway is involved in activation of NLRP3 inflammasome assembly, which subsequently cleaves Pro-IL1b into the pro-inflammatory cytokine: IL-1b[19] .

Studies investigated that ROS triggers the pro-inflammatory complex NLRP3 [19,21,29]. Kelley et al. showed that the Nrf2-complex counteracts the activation of NLRP3 in the presence of NADPH Quinone Dehydrogenase 1 (NQO1)[19]. Genes such as PLCγ2 and Trim30a are involved in this mechanism and indirectly inhibit or induce ROS. When ROS is present, PLCγ2 activates calcium release in the cell[30,31]. Lee et al. proved that Ca^{2+} is a trigger for NLRP3 activation and in the end IL-1b is produced[20]. According to Hu et al. Trim30a had a more direct effect on ROS than PLCγ2. It does not increase calcium levels nor interfere with cascades connected to NLRP3. Trim30a was shown that it directly diminished the amount of ROS and thus negatively regulate NLRP3 inflammasome[22].

Superoxides ($O_2^{\bullet-}$) are minus charged free radicals that mildly react in surrounding cells. Just because they are not very reactive does not mean they cannot do harm to parts in the body. Superoxide generation happens over the cell membrane by exchanging electrons by NADPH oxidases (NOXs), and is therefore characterized by superoxide metabolic process(GO:0006801). NOXs convert NADPH and oxygen into NADP+ and a superoxide. When ROS is upregulated, more NOX complexes are generated. Syk is such a second messenger that is involved in this stimulating process [32]. Our hypothesis is supported by the increased Syk levels in the Acp5 overexpressing mice relative to wildtype mice. Other than the complexes being generated, they also must be activated. When ROS is induced, more activated NOXs are present. Enzymes that are involved in this process are $P47^{phox}$ and $P40^{phox}$ (Ncf1 and Ncf4 gene), they work closely together. The effect of Ncf4 is to stimulate Ncf1 and lead NADPH oxidases to

phagosomal vesicles transporting them to the membrane to further regulate ROS. When the different Ncf forms are bound to the membrane, NOX is activated [33]. When mice are Ncf4 knockdown, an infiltration of innate immune cells occurs which on their turn activate a cascade of responses [34]. To facilitate this conversion, a current over the membrane must be maintained. H+ channels pump protons over the membrane which are required in the cell to balance the current so less depolarization will happen because of the electron exchange. As a result, ROS generation is optimized [35,36]. Expected is that Acp5 induces oxidative stress and therefore it is essential that these genes are indeed upregulated in the pathway.

In addition, we did a Gene Ontology Enrichment Analysis in Metascape which showed that the responses to oxidative stress were not the main difference of Acp5 overexpression relative to wildtype mice. Responses that had to do with B-Cell signaling took the lead. Therefore, we did an unbiased pathway analysis with all the genes in the gene set to see the biggest changes between the two conditions. The evidence for Acp5 inducing oxidative stress in lung tissue was not very strong. Which gene sets were enriched were made clear with a GSEA. Majority of the genes in the pathways are genes encoding for immunoglobulins (Igs) or specific Fc receptors (see supplementary data S4) involved in B cell signaling, activation and proliferation. Based on the pathways that have to do with responses on ROS there is a possibility that the B-cell infiltration is due to the body not recognizing the different products that are formed by the free oxygen radicals. The foreign proteins activate ROS production by interfering in the pathway PTP-ITAM-SYK pathway. Normally, PTP inhibits ITAM on the B-cell receptor (BCR). ROS deactivates PTP leading to the activated ITAM which on its term activate Syk [37]. Earlier on, Syk was also seen to be upregulated (Table 1) and does upregulate the expression of NOXs on the cell membrane. Potentially, there is B-cell infiltration to clear the cells because of the damage obtained from produced oxygen radicals.

Since those B-cell pathways are exceptionally enriched and a minor set of genes were enriched in the ROS pathways, it is more likely to presume that the tremendously upregulation of Igs and other receptors has to do with something else. We found that TRAP is a marker for a specific form of leukemia where B-cells phenotypically differ from the standard. This form of leukemia is called Hairy Cell Leukemia (HCL) and is a form of cancer where chronic proliferation of 'Hairy' B-cells occur and infiltrate in the blood and goes to other blood-rich tissue [38,39]. Typical for this disease is that HCs frequently express many Igs. The Igs that are involved have many immunoglobulin heavy variable (Ighv) isotypes and just for example one immunoglobulin kappa light variable (Igkv) in their gene expression profile [40]. BCR signaling pathway has a major effect on the survival and proliferation of the B-cells. Instead of more cell proliferation, extended cell survival is the primary driver of HCL clonal expansion. Important mediators in B-cell signaling are BTK and SYK [41]. Btk and Syk stimulate the PI3K/AKT pathway which also has a cascade of reactions [42]. Pro-apoptotic cells are inhibited and cell survival is stimulated [43]. Other enriched pathways obtained by GSEA are overall reactions on one another. When B-cells get activated and send second messengers, it is expected that calcium

mobilization takes place. LAT2/NTAL/LAB will provide a calcium influx but mainly this is provided by a FCERI aggregation [24]. FCERI aggregation indirectly induces calcium influx. The fact that no more conclusions can be derived from the lung tissue that was stained with Acp5 is a limitation of this finding. The lung tissue is already gone and there cannot be proved if it were indeed hairy cells.

In summary, an extensive database research was done on the effect of an overexpression of TRAP on lung tissue of mice. The RNA of genetically modified Acp5 overexpressing mice were compared to the RNA of wildtype mice by means of DESeq2 in R. Although there were some significant DEGs upregulated that had to do with the response to oxidative stress, no substantial evidence was found that Acp5 indeed induces oxidative stress in the lungs of mice. However, what we did find was at least as important. Unbiased analysis of DEGs indicated enrichment of pathways related to B-cell migration, activation and proliferation and induces an immune response. Next to enriched genes associated with B-cells, a tremendous upregulation of immunoglobulins were detected. We also proposed a new hypothesis for future research: that Acp5 (a marker of Hairy Cell Lymphoma) may potentially be an inducer of hairy cells. This database study did preliminary investigation for a new, interesting study on the effect of Acp5 on (hairy) B-cells in lung tissue of mice.

References

1. Safiri, S. *et al.* Burden of chronic obstructive pulmonary disease and its attributable risk factors in 204 countries and territories, 1990-2019: results from the Global Burden of Disease Study 2019. *BMJ* e069679 (2022) doi:10.1136/bmj-2021-069679.
2. Song, P. *et al.* Global, regional, and national prevalence of asthma in 2019: a systematic analysis and modelling study. *J. Glob. Health* **12**, 04052 (2022).
3. Saunders, R. M., Biddle, M., Amrani, Y. & Brightling, C. E. Stressed out - The role of oxidative stress in airway smooth muscle dysfunction in asthma and COPD. *Free Radic. Biol. Med.* **185**, 97–119 (2022).
4. Negewo, N. A., Gibson, P. G. & McDonald, V. M. COPD and its comorbidities: Impact, measurement and mechanisms: COPD and its comorbidities. *Respirology* **20**, 1160–1171 (2015).
5. Boorsma, C. E. *et al.* A Potent Tartrate Resistant Acid Phosphatase Inhibitor to Study the Function of TRAP in Alveolar Macrophages. *Sci. Rep.* **7**, 12570 (2017).
6. Oddie, G. W. *et al.* Structure, function, and regulation of tartrate-resistant acid phosphatase. *Bone* **27**, 575–584 (2000).
7. Sibille, J. C., Doi, K. & Aisen, P. Hydroxyl radical formation and iron-binding proteins. Stimulation by the purple acid phosphatases. *J. Biol. Chem.* **262**, 59–62 (1987).
8. Hecker, L. Mechanisms and consequences of oxidative stress in lung disease: therapeutic implications for an aging populace. *Am. J. Physiol.-Lung Cell. Mol. Physiol.* **314**, L642–L653 (2018).
9. Thimmulappa, R. K., Chattopadhyay, I. & Rajasekaran, S. Oxidative Stress Mechanisms in the Pathogenesis of Environmental Lung Diseases. in *Oxidative Stress in Lung Diseases* (eds. Chakraborti, S., Parinandi, N. L., Ghosh, R., Ganguly, N. K. & Chakraborti, T.) 103–137 (Springer Singapore, 2020). doi:10.1007/978-981-32-9366-3_5.
10. Walters, D. M., Cho, H.-Y. & Kleeberger, S. R. Oxidative Stress and Antioxidants in the Pathogenesis of Pulmonary Fibrosis: A Potential Role for Nrf2. *Antioxid. Redox Signal.* **10**, 321–332 (2008).

11. Suzuki, T. & Yamamoto, M. Stress-sensing mechanisms and the physiological roles of the Keap1–Nrf2 system during cellular stress. *J. Biol. Chem.* **292**, 16817–16824 (2017).
12. Hiebert, P. & Werner, S. Regulation of Wound Healing by the NRF2 Transcription Factor—More Than Cytoprotection. *Int. J. Mol. Sci.* **20**, 3856 (2019).
13. Marchitti, S. A., Brocker, C., Orlicky, D. J. & Vasiliou, V. Molecular characterization, expression analysis, and role of ALDH3B1 in the cellular protection against oxidative stress. *Free Radic. Biol. Med.* **49**, 1432–1443 (2010).
14. Weng, B., Zhang, X., Chu, X., Gong, X. & Cai, C. Nrf2-Keap1-ARE-NQO1 signaling attenuates hyperoxia-induced lung cell injury by inhibiting apoptosis. *Mol. Med. Rep.* **23**, 221 (2021).
15. Shen, Y. *et al.* Increased Serum ox-LDL Levels Correlated with Lung Function, Inflammation, and Oxidative Stress in COPD. *Mediators Inflamm.* **2013**, 1–5 (2013).
16. Qiu, J. *et al.* BTK Promotes Atherosclerosis by Regulating Oxidative Stress, Mitochondrial Injury, and ER Stress of Macrophages. *Oxid. Med. Cell. Longev.* **2021**, 1–15 (2021).
17. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).
18. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
19. Kelley, N., Jeltema, D., Duan, Y. & He, Y. The NLRP3 Inflammasome: An Overview of Mechanisms of Activation and Regulation. *Int. J. Mol. Sci.* **20**, 3328 (2019).
20. Lee, G.-S. *et al.* The calcium-sensing receptor regulates the NLRP3 inflammasome through Ca2+ and cAMP. *Nature* **492**, 123–127 (2012).
21. Zhao, S. *et al.* Reactive Oxygen Species Interact With NLRP3 Inflammasomes and Are Involved in the Inflammation of Sepsis: From Mechanism to Treatment of Progression. *Front. Physiol.* **11**, 571810 (2020).
22. Hu, Y. *et al.* Tripartite-Motif Protein 30 Negatively Regulates NLRP3 Inflammasome Activation by Modulating Reactive Oxygen Species Production. *J. Immunol.* **185**, 7699–7705 (2010).
23. Lei, C. *et al.* Isolation and characterization of a novel zinc finger gene, ZNFD, activating AP1(PMA) transcriptional activities. *Mol. Cell. Biochem.* **340**, 63–71 (2010).

24. Iwaki, S., Jensen, B. M. & Gilfillan, A. M. NTAL/LAB/LAT2. *Int. J. Biochem. Cell Biol.* **39**, 868–873 (2007).

25. Wang, Y., Branicky, R., Noë, A. & Hekimi, S. Superoxide dismutases: Dual roles in controlling ROS damage and regulating ROS signaling. *J. Cell Biol.* **217**, 1915–1928 (2018).

26. Hayyan, M., Hashim, M. A. & AlNashef, I. M. Superoxide Ion: Generation and Chemical Implications. *Chem. Rev.* **116**, 3029–3085 (2016).

27. Lin, Y.-R. *et al.* Regulation of longevity by regulator of G-protein signaling protein, Loco: Longevity regulation by RGS protein, Loco. *Aging Cell* **10**, 438–447 (2011).

28. Hickey, F. B. *et al.* IHG-1 Increases Mitochondrial Fusion and Bioenergetic Function. *Diabetes* **63**, 4314–4325 (2014).

29. Liu, X. *et al.* Nuclear Factor E2-Related Factor-2 Negatively Regulates NLRP3 Inflammasome Activity by Inhibiting Reactive Oxygen Species-Induced NLRP3 Priming. *Antioxid. Redox Signal.* **26**, 28–43 (2017).

30. Chae, J. J. *et al.* Brief Report: Connecting Two Pathways Through Ca^{2+} Signaling: NLRP3 Inflammasome Activation Induced by a Hypermorphic *PLCG2* Mutation. *Arthritis Rheumatol.* **67**, 563–567 (2015).

31. Luo, T., Gao, J., Lin, N. & Wang, J. Effects of Two Kinds of Iron Nanoparticles as Reactive Oxygen Species Inducer and Scavenger on the Transcriptomic Profiles of Two Human Leukemia Cells with Different Stemness. *Nanomaterials* **10**, 1951 (2020).

32. Nguyen, G. T., Green, E. R. & Mecsas, J. Neutrophils to the ROScue: Mechanisms of NADPH Oxidase Activation and Bacterial Resistance. *Front. Cell. Infect. Microbiol.* **7**, 373 (2017).

33. Kanai, F. *et al.* The PX domains of p47phox and p40phox bind to lipid products of PI(3)K. *Nat. Cell Biol.* **3**, 675–678 (2001).

34. Winter, S., Hultqvist Hopkins, M., Laulund, F. & Holmdahl, R. A Reduction in Intracellular Reactive Oxygen Species Due to a Mutation in NCF4 Promotes Autoimmune Arthritis in Mice. *Antioxid. Redox Signal.* **25**, 983–996 (2016).

35. Murphy, R. & DeCoursey, T. E. Charge compensation during the phagocyte respiratory burst. *Biochim. Biophys. Acta BBA - Bioenerg.* **1757**, 996–1011 (2006).

36. Zhu, X., Mose, E. & Zimmermann, N. Proton channel HVCN1 is required for effector functions of mouse eosinophils. *BMC Immunol.* **14**, 24 (2013).
37. Yang, Y., Bazhin, A. V., Werner, J. & Karakhanova, S. Reactive Oxygen Species in the Immune System. *Int. Rev. Immunol.* **32**, 249–270 (2013).
38. Janckila, A., Cardwell, E., Yam, L. & Li, C. Hairy cell identification by immunohistochemistry of tartrate-resistant acid phosphatase. *Blood* **85**, 2839–2844 (1995).
39. Swords, R. & Giles, F. Hairy cell leukemia. *Med. Oncol.* **24**, 7–15 (2007).
40. Forconi, F. *et al.* Tumor cells of hairy cell leukemia express multiple clonally related immunoglobulin isotypes via RNA splicing. *Blood* **98**, 1174–1181 (2001).
41. Sivina, M., Kreitman, R. J., Arons, E., Ravandi, F. & Burger, J. A. The bruton tyrosine kinase inhibitor ibrutinib (PCI-32765) blocks hairy cell leukaemia survival, proliferation and B cell receptor signalling: a new therapeutic approach. *Br. J. Haematol.* **166**, 177–188 (2014).
42. Woyach, J. A., Johnson, A. J. & Byrd, J. C. The B-cell receptor signaling pathway as a therapeutic target in CLL. *Blood* **120**, 1175–1184 (2012).
43. Tiacci, E., Liso, A., Piris, M. & Falini, B. Evolving concepts in the pathogenesis of hairy-cell leukaemia. *Nat. Rev. Cancer* **6**, 437–448 (2006).

Supplemental Data

S1. R code

```
# First make sure all required packages are loaded
if (!require("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install("DESeq2")
BiocManager::install("RColorBrewer")
BiocManager::install("pheatmap")
BiocManager::install("tidyverse")
BiocManager::install("ggrepel")
BiocManager::install("dplyr")
BiocManager::install("apeglm")

# Load al libraries
library("DESeq2")
library("RColorBrewer")
library("pheatmap")
library("tidyverse")
library("ggrepel")
library("dplyr")
library("apeglm")

# Load and name transcript counts
counts<-read.csv("genecounts2.csv", row.names = 1)

# Load and name metadata
genotype <- c("tg", "tg", "tg", "tg", "tg", "tg", "tg", "tg", "wt", "wt", "wt", "wt", "wt", "wt")

condition <- c("transgene", "transgene", "transgene", "transgene", "transgene", "transgene",
"transgene", "transgene", "normal", "normal", "normal", "normal", "normal", "normal")

metadata<- data.frame(genotype, condition)

rownames(metadata) <- c("TG1", "TG2", "TG3", "TG4", "TG5", "TG6", "TG7", "TG8",
"WT1", "WT2", "WT3", "WT4", "WT5", "WT6")

# Check if order metadata = counts
all(rownames(metadata) == colnames(counts))

#Make a dispersion plot
# Calculate mean for each gene
mean_counts <- apply(counts[, 1:3], 1, mean)

# Calculate variance for each gene
variance_counts <- apply(counts[,1:3], 1, var)

# Creating data frame with mean and variance for every gene
ggplot(df) +
  geom_point(aes(x=mean_counts, y=variance_counts)) +
```

```
  scale_y_log10() +
  scale_x_log10() +
  xlab("Mean counts per gene") +
  ylab("Variance per gene")

# create DESeq2 object
dds <- DESeqDataSetFromMatrix(countData = reordered_counts,
                              colData = metadata,
                              design = ~ condition)

# Normalization of (raw) counts
dds <- estimateSizeFactors(dds)
sizeFactors(dds)

# Normalized counts extraction
normalized_counts <- counts(dds, normalized = TRUE)

# Log transformation of normalized counts, blind to sample info
vsd <- vst(dds, blind=TRUE)

# Extract vst matrix from object
vsd_mat <- assay(vsd)

# Compute pairwise correlation values
vsd_cor <- cor(vsd_mat)

# Correlation heatmap
pheatmap(vsd_cor, annotation = select(metadata, condition))

# plot PCA with labels
z<-plotPCA(vsd, intgroup="condition")
z + geom_label(aes(label = name))

# load DESeq2
dds <- DESeq(dds)

# results of DESeq2
results(dds, alpha = 0.05)

# Extract results of differential expression analysis
res <- results(dds,
               contrast = c("condition", "transgene", "normal"),
               alpha = 0.05)

# Shrink the log2 fold change estimates to be more accurate
res <- lfcShrink(dds = dds, coef = 2, res = res)

# Create dispersion plot
plotDispEsts(dds)
```

```
# Create MAplot
plotMA(res, ylim=c(-10,10))

# Descriptions for the columns in the results table (adjust p values)
mcols(res)
head(res, n=10)

# Prefiltering of all genes with a row total of at least 10 or more
keep <- rowMeans2(counts(dds)) >= 10
dds <- dds[keep,]

# The interaction, answering: is the wt effect different for treatment with TRAP?
res <- results(dds, contrast=c("condition","transgene","normal"), alpha = 0.05)

# summary of DESeq2 without threshold
summary(res)

# write gene counts to csv
write.csv(res, file = "results_genecounts2.csv")

#Save results as a data frame
res_all <- data.frame(res) %>%
  rownames_to_column(var="ensgene")

# Subset the results to only return the significant genes with p-adjusted values less than 0.05
res_sig <- subset(res_all, padj < 0.05)
res_sig <- res_sig %>% arrange(padj)

sig_norm_counts <- normalized_counts[res_sig$ensgene, ]

upregulated <- subset(res_sig, log2FoldChange > 1, row.names = TRUE)

downregulated <- subset(res_sig, log2FoldChange < -1, row.names = TRUE)

# Save up/downregulated genes
write.csv(upregulated, file = "upregulated.csv")
write.csv(downregulated, file = "downregulated.csv")

# Run pheatmap
heat_colors <- brewer.pal(n = 6, name = "YlOrRd")

pheatmap(sig_norm_counts,
         color = heat_colors,
         cluster_rows = TRUE,
         show_rownames = FALSE,
         annotation = select(metadata, condition),
         scale = "row")

# Create vulcano plot as in Datacamp
res_all <- data.frame(res) %>% mutate(threshold = padj < 0.05)
```

```
res_all <- res_all %>% rownames_to_column(var = "ensgene") %>%
  mutate(threshold = padj < 0.05)

ggplot(res_all) +
  geom_point(aes(x = log2FoldChange, y = -log10(padj), color = threshold)) +
  xlab("log2 fold change") +
  ylab("-log10 adjusted p-value") +
  theme(legend.position = "none",
        plot.title = element_text(size = rel(1.5), hjust = 0.5),
        axis.title = element_text(size = rel(1.25)))

# Install Mouse Musculus (Mm) gene symbols for labels in enhanced volcano plot
if (!require("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install("org.Mm.eg.db")
install.packages("textshaping")
BiocManager::install("EnhancedVolcano")

# Open library
library("org.Mm.eg.db")
library("textshaping")
library("EnhancedVolcano")

# make dataframe of results of DESeq2
sig.df <- as.data.frame(res)

# Make a new symbol column
sig.df$symbol <- mapIds(org.Mm.eg.db, keys = rownames(sig.df), keytype = "ENSEMBL",
column = "SYMBOL")
view(sig.df)

#Create enhanced volcano plot with labels that you need for e.g. significant genes

selected = c("Aldh3b1", "Acp5", "Btk", "Rgs14", "Thg1l", "Plcg2", "Trim30a", "Ncf1",
"Ncf4", "Syk", "Hvcn1")

EnhancedVolcano(sig.df, x = "log2FoldChange", y = "pvalue", lab = sig.df$symbol,
            pCutoff = 0.05, FCcutoff = 1, ylim = c(0,50), xlim = c(-8,8))

# Create right side of volcano with significant labels otherwise labels overlap and no labels at
all appear

selected = c("Aldh3b1", "Acp5", "Btk", "Rgs14", "Thg1l", "Plcg2", "Trim30a", "Ncf1",
"Ncf4", "Syk", "Hvcn1")

EnhancedVolcano(sig.df, x = "log2FoldChange", y = "pvalue", lab = sig.df$symbol,
            pCutoff = 0.05, FCcutoff = 1, ylim = c(0,7), xlim = c(0,2.5), select = selected)

# Top 20
```

```
top_20 <- data.frame(sig_norm_counts)[1:20,]%>% rownames_to_column(var
="ensgene")
top_20 <- gather(top_20, key ="samplename", value ="normalized_counts",2:15)

top_20 <- inner_join(top_20, rownames_to_column(metadata, var ="samplename"),
                     by ="samplename")

ggplot(top_20)+ geom_point(aes(x = ensgene,
                               y = normalized_counts, color = condition))+
  scale_y_log10()+ xlab("Genes")+ ylab("Normalized Counts")+
  ggtitle("Top 20 Significant DE Genes")+ theme_bw()+
  theme(axis.text.x = element_text(angle =45, hjust =1))+
  theme(plot.title = element_text(hjust =0.5))

# Acp5 expression as significant gene
Acp5 = "Acp5"

Acp5sig <- data.frame(sig_norm_counts)[18:18,]%>% rownames_to_column(var
="ensgene")
Acp5sig <- gather(Acp5sig, key ="samplename", value ="normalized_counts",2:15)

Acp5sig <- inner_join(Acp5sig, rownames_to_column(metadata, var
="samplename"),
                     by ="samplename")

ggplot(Acp5sig)+ geom_point(aes(x = Acp5,
                                y = normalized_counts, color = condition))+
  scale_y_log10()+ xlab("Gene")+ ylab("Normalized Counts")+
  ggtitle("Acp5 expression")+ theme_bw()+
  theme(axis.text.x = element_text(angle =45, hjust =1))+
  theme(plot.title = element_text(hjust =0.5))
```

S2. P-values of enriched pathways Metascape

Pathway:	-LOG$_{10}$(P)
activation of immune response	15.11
B cell activation	10.26
response to virus	6.6
CD22 mediated BCR regulation	6.29
regulation of B cell proliferation	5.58
immunoglobulin production involved in immun…	3.29
inflammatory response	3.04
positive regulation of cytosolic calcium ion concentration	2.29
response to molecule of bacterial origin	2.2

S3. Top 20 downregulated genes

Table 3. top 20 downregulated genes based on log2FoldChange. All values have a $p < 0.05$.

GeneID	Gene symbol	log2FoldChange	pvalue	padj
Catspere	ENSMUSG00000102483	-6.909696346	4.09E-18	6.80E-15
Glns-ps1	ENSMUSG00000082100	-6.774180174	1.09E-20	2.50E-17
Gm4735	ENSMUSG00000064193	-6.428066952	6.29E-22	1.92E-18
H60c	ENSMUSG00000091618	-5.849625194	1.18E-10	7.69E-08
Gm4949	ENSMUSG00000118140	-5.571828504	2.03E-12	1.85E-09
AC160336.1	ENSMUSG00000115801	-5.460287409	8.85E-26	3.23E-22
Rpl34-ps1	ENSMUSG00000068396	-5.245159755	3.30E-39	3.01E-35
Lce3e	ENSMUSG00000074433	-4.567576352	3.29E-05	0.00583086
Gpr75	ENSMUSG00000043999	-4.415521017	1.99E-11	1.58E-08
Gm45844	ENSMUSG00000110105	-4.150270677	7.85E-11	5.52E-08
Hrnr	ENSMUSG00000041991	-4.004167813	8.50E-05	0.01262971
Gm16589	ENSMUSG00000074482	-3.574152355	7.15E-05	0.01107141
Gm46620	ENSMUSG00000118012	-3.419096479	2.11E-07	6.64E-05
Grm1	ENSMUSG00000019828	-3.324238715	4.98E-06	0.00121213
Hebp2	ENSMUSG00000019853	-3.224382186	1.11E-07	3.99E-05
Gm47547	ENSMUSG00000114196	-3.222055252	3.49E-09	1.64E-06
Rab3c	ENSMUSG00000021700	-3.124120348	8.44E-06	0.0018357
Spink5	ENSMUSG00000055561	-2.89271834	0.00046342	0.04666079
Ogdhl	ENSMUSG00000021913	-2.874344541	1.10E-05	0.00234631
Cldn34c1	ENSMUSG00000079450	-2.835954009	1.99E-07	6.46E-05

S4 GSEA upregulated genes in pathways

Pathway1*	**Pathway2***	**Pathway3***
Iglc1	Iglc1	Iglc1
Ighv11-2	Ighv11-2	Ighv11-2
Igkv14-126	Igkv14-126	Igkv14-126
Syk	Syk	Syk
Plcg2	Iglc2	Plcg2
Cd19	Ighv5-16	Iglc2
Cd79b	Ighv1-55	Ighv5-16
Pik3cd	Ighv1-76	Ighv1-55
Vav1	Ighv1-7	Ighv1-76
Iglc2	Ighv1-81	Ighv1-7
Ighv5-16	Ighv1-75	Ighv1-81
Ighv1-55	Lyn	Ighv1-75
Cd22	Igkv17-127	Igkv17-127
Cd79a	Ighv1-85	Ighv1-85
Ighv1-76	Ighv1-15	Ighv1-15
Ighv1-7	Ighv1-26	Ighv1-26
Btk	Ighv5-4	Ighv5-4
Ighv1-81	Ighv1-22	Ighv1-22
Ighv1-75	Ighv14-2	Ighv14-2
Blnk	Ighv5-12	Ighv5-12
Igkv17-127	Ighv5-9-1	Ighv5-9-1
Ighv1-85	Ighv1-52	Pld4
Ighv1-15	Ighv1-50	Ighv1-52
Ighv1-26	Ighv1-18	Ighv1-50
Ighv5-4	Fcer1g	Ighv1-18
Ptpn6	Ighv12-3	Pla2g6
Ighv1-22	Lat2	Ighv12-3
Ighv14-2	Igll1	Fcgr2b
Ighv5-12	Ighv1-72	Igll1
Ighv5-9-1	Ighv1-58	Ighv1-72
Ighv1-52	Ighv1-19	Ighv1-58
Ighv1-50	Ighv1-39	Ighv1-19
Ighv1-18	Ighv1-80	Ighv1-39
Ighv12-3	Ighv1-36	Ighv1-80
Pik3ap1	Ighv1-69	Ighv1-36
Igll1	Ighv5-6	Ighv1-69
Dapp1	Ighv1-53	Ighv5-6
Ighv1-72	Grb2	Ighv1-53
Ighv1-58	Ighv1-42	Ighv1-42
Ighv1-19	Ighv1-74	Ighv1-74
Ighv1-39	Ighv1-59	Ighv1-59

Ighv1-80	Ighv5-15	Ighv5-15
Itpr3	Ighv14-4	Ighv14-4
Ighv1-36	Ighv3-6	Ighv3-6
Ighv1-69	Ighv1-64	Ighv1-64
Ighv5-6	Igkv1-88	Igkv1-88
Ighv1-53	Igkv14-100	Fcgr1
Grb2	Ighv4-1	Igkv14-100
Ighv1-42	Ighv1-63	Ighv4-1
Ighv1-74	Ighv1-82	Ighv1-63
Ighv1-59	Igkv2-112	Ighv1-82
Ighv5-15	Ighv1-47	Igkv2-112
Ighv14-4	Ighv1-71	Ighv1-47
Ighv3-6	Ighv15-2	Ighv1-71
Ighv1-64	Ighv1-49	Ighv15-2
Sh3kbp1	Ighv3-3	Ighv1-49
Igkv1-88	Ighv1-4	Ighv3-3
Igkv14-100	Ighv1-5	Ighv1-4
Ighv4-1	Ighv1-61	Ighv1-5
Ighv1-63	Pik3r2	Ighv1-61
Ighv1-82	Igkv17-121	Pik3r2
Igkv2-112	Igkv1-132	Igkv17-121
Ighv1-47	Ighv5-17	Igkv1-132
Ighv1-71	Ighv1-66	Ighv5-17
Ighv15-2	Ighv1-20	Ighv1-66
Ighv1-49	Ighv1-78	Ighv1-20
Ighv3-3	Fcer1a	Ighv1-78
Ighv1-4	Ighv3-8	Fcgr4
Ighv1-5	Igkv15-103	Ighg3
Ighv1-61	Ighv1-62-2	Ighv3-8
Igkv17-121	Ighv1-34	Igkv15-103
Igkv1-132	Ighv14-1	Ighv1-62-2
Ighv5-17	Ighv1-37	Ighv1-34
Ighv1-66	Ighv1-31	Ighv14-1
Ighv1-20	Ighv8-2	Ighv1-37
Ighv1-78	Igkv13-84	Ighv1-31
Ighv3-8	Igkv18-36	Ighv8-2
Igkv15-103	Igkv1-133	Igkv13-84
Ighv1-62-2	Ighv5-2	Igkv18-36
Ighv1-34	Ighv1-77	Igkv1-133
Ighv14-1	Ighv11-1	Fcgr3
Ighv1-37	Igkv12-98	Pld3
Itpr1	Ms4a2	Ighv5-2
Ighv1-31	Ighv1-9	Ighv1-77

Ighv8-2	Ighv1-67	Ighv11-1
Igkv13-84	Ighv14-3	Cd247
Igkv18-36	Ighv1-54	Igkv12-98
Itpr2	Ighv8-4	Ighv1-9
Igkv1-133	Ighv3-4	Ighv1-67
Ighv5-2	Ighv8-6	Ighv14-3
Ighv1-77	Sos1	Ighv1-54
Ighv11-1	Ighv5-9	Ighv8-4
Igkv12-98	Igkv8-21	Ighv3-4
Ighv1-9	Igkv1-110	Cd3g
Ighv1-67	Ighv1-12	Ighv8-6
Ighv14-3	Ighv1-62-3	Plpp4
Ighv1-54	Igkv1-131	Ighv5-9
Ighv8-4	Ighv1-84	Igkv8-21
Ighv3-4	Ighv3-1	Igkv1-110
Ighv8-6	Igkv13-85	Ighv1-12
Sos1	Igkv20-101-2	Ighv1-62-3
Ighv5-9	Igkv1-35	Ighg2c
Igkv8-21	Ighv1-43	Ighg1
Igkv1-110	Ighv1-24	Prkce
Ighv1-12	Ighv1-16	Igkv1-131
Ighv1-62-3	Pdpk1	Ighv1-84
Igkv1-131	Igkv1-117	Ighv3-1
Stim1	Igkv2-137	Igkv13-85
Ighv1-84	Shc1	Igkv20-101-2
Ighv3-1	Pik3r1	Igkv1-35
Igkv13-85	Ighv1-11	Ighv1-43
Igkv20-101-2	Ighv8-9	Ighv1-24
Igkv1-35	Ighv8-8	Ighv1-16
Ighv1-43	Ighv1-56	Igkv1-117
Ighv1-24	Ighv8-12	Igkv2-137
Ighv1-16	Igkv14-111	Pik3r1
Igkv1-117	Ighv8-11	Ighv1-11
Igkv2-137	Pik3ca	Ighv8-9
Pik3r1	Igkv1-99	Ighv8-8
Nck1	Ighv8-13	Ighv1-56
Ighv1-11	Pik3cb	Ighv8-12
Ighv8-9	Igkv1-135	Igkv14-111
Ighv8-8	Ighv5-12-4	Ighv8-11
Ighv1-56	Ighe	Pik3ca
Ighv8-12	Igkv1-122	Igkv1-99
Igkv14-111	Ighv1-23	Prkcd

Ighv8-11	Igkv14-130	Pld2
Igkv1-99	Ighv8-5	Ighv8-13
Ighv8-13		Plcg1
Igkv1-135		Pik3cb
Ighv5-12-4		Igkv1-135
Igkv1-122		Ighv5-12-4
Ighv1-23		Plpp5
Igkv14-130		Igkv1-122
Ighv8-5		Ighv1-23
Trpc1		Igkv14-130
		Ighv8-5

*

1	(Cellular) responses to oxidative stress
2	regulation of NLRP3 inflammasome complex assembly
3	superoxide metabolic processes