
Original Paper

Bioinformatics-based Analysis of the Role of Sialylation Inhibition in Intervertebral Disc Degeneration of Humans

Aung Myat Phy 1,* , Kieran Joyce 2 , Barry Digby 1 and Abhay Pandit 2

¹ School of Mathematical and Statistical Sciences, University of Galway, H91 TK33, Ireland

² CÚRAM, SFI Research Centre for Medical Devices, Biomedical Science Building, University of Galway, H91 TK33, Ireland.

*To whom correspondence should be addressed.

Abstract

Motivation: Intervertebral disc degeneration (IVDD) is one of the causative agents for lower back pain, which is already a global health issue and creates a socioeconomic burden. Nowadays, the popular treatments for IVDD are painkillers, rehabilitation and surgery, which mainly emphasise decreasing the symptoms of disc degeneration while they do not offer to regenerate the normal anatomy and physiology of the disc. In order to fulfil this objective, there is still needed for many research which will assist to apprehend the pathophysiology of IVDD and the regenerative function of the disc. In this research, 3Fax-Peracetyl Neu5Ac (Neu5Ac-inhib) is applied to degenerative nucleus pulposus (NP) cells to inhibit the sialyltransferase enzymes regarding glycosylation inhibition. The advance in bioinformatic tools helps to understand more about the pathophysiological process of IVDD. This bioinformatics-based analysis of the role of sialylation inhibition in IVDD of humans will validate that pro-inflammatory cytokine (neutrophils) involvement and metabolic impairment in the process of IVDD, reduction of oxidative stress in IVDD and maintenance of extracellular matrix by application of 3Fax-Peracetyl Neu5Ac.

Results: In healthy nucleus pulposus (NP) cells with the cytokine cocktail, neutrophil migration and cytokine activity are enriched. In the mixture of healthy NP cells and cytokine cocktail treated with Neu5Ac-inhib, cytosolic ribosome and eukaryotic translation pathways are significant. In the degenerative NP cells treated with Neu5Ac-inhib, ECM constituent proteoglycans pathways are upregulated, while oxidative phosphorylation and respiratory electron transport pathways are empowered in the degenerative control cells. Moreover, SRP-dependent cotranslational protein targeting to membrane and selenoamino acid metabolism pathways are augmented with the sialylation inhibition in healthy NP cells and cytokine cocktail. These may be interesting areas to investigate in the future. Neu5Ac-inhib benefits for nucleus pulposus cells by balancing ECM metabolism and preventing ribosomal dysfunction, oxidative stress and apoptosis.

Contact: a.phy01@nuigalway.ie

Supplementary information: Supplementary data and code applied to perform the analysis are accessible at https://github.com/aungmyatphyogalway/MSc_Project

1 Introduction

Intervertebral disc degeneration (IVDD), recognised as the majority prevailing initiator of low back pain and disability, is a ubiquitous worldwide health concern and produces a socioeconomic encumbrance (Katz et al., 2006). The occurrence of IVDD rises significantly, showing that over 90% of individuals over the age of 50 suffer (Andersson, 1999;

Cheung et al., 2009; Teraguchi et al., 2014). The remaining confounding factors are physical injury, genetically inherited susceptibility, and lifestyle, such as obesity and smoking (Clouet et al., 2009; Hadjipavlou et al., 2008; Urban and Roberts, 2003). The existing primary treatments consist of conservative strategies such as painkillers, bed rest, and rehabilitation and non-conservative approaches such as interventional and surgical therapy (Cloyd et al., 2007). However, the limitation of these approaches is that they focus on reducing the symptoms and do not offer to restore the normal anatomy and homeostasis of the disc. Therefore, a

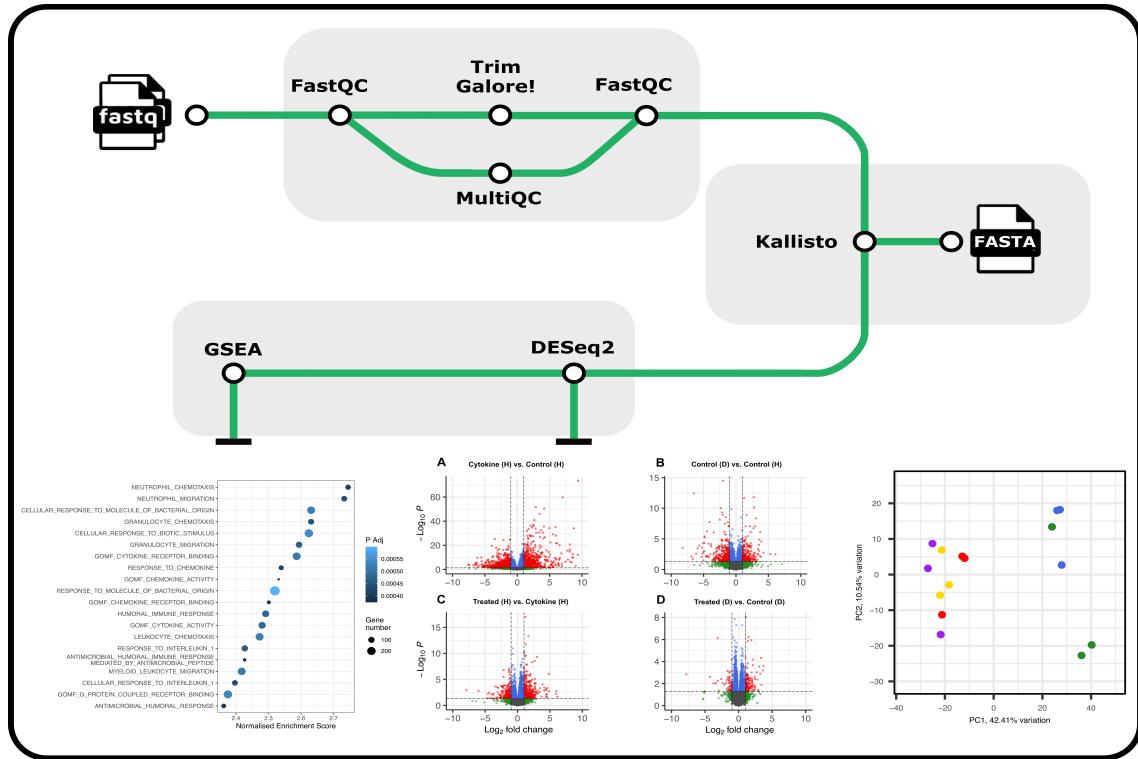


Fig. 1: Overview of project workflow: Firstly, fastqc, multiqc and trim-galore tools are utilised for quality control step of raw fastq files. Kallisto tools is also used for alignment, assembly and quantification steps. For differential gene expression analysis, DESeq2 tools is applied and gene set enrichment analysis is carried out for the downstream analysis

lot of research and understanding of the pathophysiology of IVDD will be needed to explore the regeneration of the degenerative disc.

The structure of the intervertebral disc (IVD) belongs to the internal nucleus pulposus (NP), encircled by the external annulus fibrosus (AF), and hyaline cartilaginous endplates which surround the NP and AF at the vertebral junction, termed as the zygapophyseal joint. Because of an avascular organisation, the disc depends on the neighbouring endplates to acquire nutrients and gas exchange (De Geer, 2018). The core constituents of the NP are proteoglycans, aggrecans and type II collagen, which conserve water to withstand compressive loads on the IVD. The AF offers more load bearing because it is more rigid than the NP. It comprises type I collagen, organised in concentric rings (lamellae). The extracellular matrix (ECM) is vital and responsible for the mechanical functionality of the disc. In terms of the avascular character of the disc, these constituents may innately concentrate and corrupt with time. The bulk noteworthy conversion is the deprivation of proteoglycans, which reduces the disc matrix's osmotic pressure (Urban and Roberts, 2003).

IVDD is represented by the dysregulation of the disc's metabolism. The imbalance of NP cells produces a reduction in the synthesis of ECM components and an escalation in the production of deteriorative molecules of ECM (Roberts et al., 2000; Vo et al., 2013). There will gradually evolve disorganised discs and irregular lamellae of the AF (Buckwalter, 1995). The progression of the degeneration results in the tiny, calcified pores of the endplate, which produces the irregular process of taking oxygen and nutrients, facilitating the degeneration cascade (Matta and Erwin, 2020).

Inflammation, an essential biological role and a feature of the manifestation of the degenerated disc, exacerbate IVDD (Johnson et al., 2015; Wuertz and Haglund, 2013). Despite knowing the function of transformed glycosylation patterns in animal studies with inflamed IVD, the aspect of the transformed glycosylation representation in humans is still unknown. Critical intracellular processes (protein assembly (Komekado et al., 2007), receptor expression (da Silva Correia and Ulevitch, 2002;

Liu et al., 2011) and enzymatic function (Kim and Varki, 1997)) are intervened by glycosylation structures. The glycan synthesis begins on the rough endoplasmic reticulum synchronised by *glycosyltransferases* and *glycohydrolases*. Sialylated glycans through *sialyltransferases* (STs) are crucial in cellular communication (Varki, 2007; Chen and Varki, 2010). *Sialyltransferases* are obligated to the Golgi apparatus by a transmembrane anchor aligned with the operative location in the lumen (Li and Chen, 2012). The addition of the sialic acids by STs to a terminal galactose at *N*- and *O*-glycans, known as glyco-modification, is significant in IVDD (Collin et al., 2016; Kazenzian et al., 2017; Mohd Isa et al., 2018). The glycoprofile of the disc reports the unique distortions in global sialylation and fucosylation (Joyce et al., 2018).

3Fax-Peracetyl Neu5Ac (Neu5Ac-inhib) is the ideal structure to arrest *sialyltransferase* enzymes for interpretation of the subsequent effects of glycosylation inhibition (Rillahan et al., 2012). Neu5Ac-inhib, the fluorinated analogue of sialic acid, enters the cells by diffusion and transforms into an analogue of CMP (cytidine monophosphate)- Neu5Ac (Rillahan et al., 2012). The aggregation of the analogue disturbs the de novo synthesis of CMP-Neu5Ac via the negative feedback (Gerardy-Schahn et al., 2015). The competitive inhibitor agent of the *sialyltransferase* enzyme is a novel method for glycosylation inhibition for the purpose of reducing the role of sialylation in IVDD. Targeted inhibition of sialylation can prohibit cytokine-mediated inflammation of NP cells by modulating the glycomics response. In this research, Neu5Ac-inhib is utilised as the sialylation inhibition agent in *vitro* human NP cells model of IVDD (cytokine-induced) to detect the consequences on glycosignature, ECM synthesis and cellular metabolism. This initial research proves that Neu5Ac-inhib retards competently the hypersialylation with no negative effect on cells survival (Joyce et al., 2021). The inhibitor regenerates physiological glycosylation of NP cells (Fig.2). Neu5Ac-inhib may endeavor new molecular agent for the treatment of IVDD.

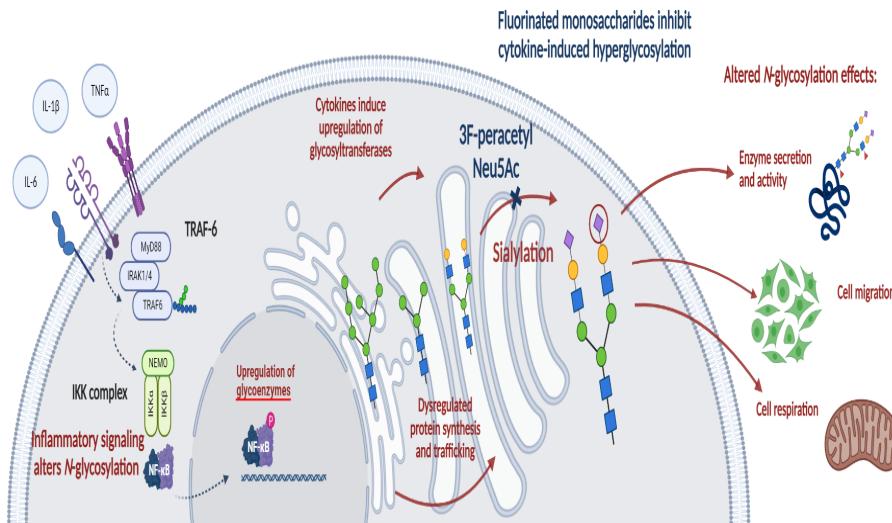


Fig. 2: Representative diagram of experimental results emphasising the mechanism of action for glycosylation inhibitors in *in vitro* human NP cells model of intervertebral disc degeneration

This research identified the differentially expressed genes (DEGs) between 5 groups of healthy and degenerative cases from the RNA-seq data set (Fig.1). Critical gene set enrichment and pathways are recognised through Gene Ontology (GOs), the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome pathways. This research will show new personalised therapeutic materials for IVD degeneration soon.

2 Methodology

2.1 Raw Data

The raw data are the compressed paired-end human mRNA-seq data in fastq files over 80 gigabytes from Illumina PE150. In these data files, there are five different experimental groups: Healthy Control (NP cells extracted from non-degenerative IVD), Degenerated Control (normal healthy NP cells extracted from degenerative IVD), Healthy Cytokine (normal healthy NP cells stimulated with cytokine cocktail: IL1 β , IL-6, TNF α), Healthy Treated (combination of normal healthy NP cells, cytokine cocktail and Neu5Ac-inhib), Degenerated Treated (degenerative IVD cells and Neu5Ac-inhib). The Degenerated Treated group does not contain the cytokine cocktail, as these cells have already been exposed to an inflammatory microenvironment *in vivo*.

2.2 Quality Control

For checking the quality of high throughput raw mRNA-seq data, fastqc v.0.12.1 quality control tools (Andrews, 2010) developed by Babraham Bioinformatics are applied. Each raw data experimental group has three biological replicates (five experimental groups), and each biological replicate is divided into two files for paired-end. Finally, the raw data is into 30 subsets for quality control. Each fastqc report of the subsets is combined into one report via multiqc tools v.1.0 (Ewels et al., 2016). This multiqc report describes not only a general statistics table showing duplication reads, average GC content, and entire sequences but also sequence quality histograms, per sequence quality scores, per sequence, quality content and adapter content graphs. According to the adapter content graph result in this report, we decided to apply trim-galore v.0.6.10 tools (<https://github.com/FelixKrueger/TrimGalore>) to perform adapter trimming and low-quality read filtering. After completing these steps, each replicate has

around 99 % of the read-passing filter. Finally, another multiqc report of the HTML file is produced, and the data is ready for the next step.

2.3 Alignment, Assembly and Quantification

For these steps of the data analysis process, we utilised the Kallisto v.0.44.0 tools (<https://github.com/pachterlab/kallisto>). Firstly, the ENSEMBL Human cDNA file (the transcript FASTA file): (https://ftp.ensembl.org/pub/release-109/fasta/homo_sapiens/cdna/Homo_sapiens.GRCh38.cdna.all.fa.gz) is downloaded for the reference file of alignment. This reference file is used through the kallisto index command, and GRCh38.idx is named as the filename for constructing the kallisto index. For the quantification step, this GRCh38.idx file is applied for the index, and two CPU threads are applied through the kallisto quant command. In this command, –bias is used for an optional argument to avoid various sources of biases and support more precise estimates of gene expression levels of RNA-seq (Zheng et al., 2011). There are 15 output folders, each containing three files: abundance.h5, Jason files and abundance.tsv in which effective length, estimated counts and TPM data are included.

2.4 Summarizing Gene Counts

The remaining steps of RNA-seq data analysis are carried out via R studio. Firstly, the 15 kallisto quant output folder file is uploaded into R studio, and quantdir is given as a variable name. At the same time, we create metadata of CSV file type, in which the row names of this metadata file must match with 15 kallisto quantification directory names. In the metadata file, any columns of numeric value are converted to a factor to avoid incorrectly implying a linear increase in fold change per replicate. As part of staging Kallisto files, we develop a file handle object, including the row names of metadata and the Kallisto quantification.h5 files.

We applied the tximport v.1.24.0 package (Soneson et al., 2015) to import transcript-level abundances from Kallisto quantification tools and change them into gene counts for downstream analysis. We chose the biomart ensembl database and the hsapiens gene ensembl dataset for connecting and mapping transcript IDs to gene IDs through the application of biomaRt package (Durinck et al., 2005). Then, we retrieve information such as Ensembl transcript id version and hgnc symbol from the biomart database according to the reference file, Homo sapiens.GRCh38.cdna.all.fa.gz, which is used on

kallisto quanting. Finally, we produce a gene-level counts matrix from summed transcript counts using the tximport R package.

2.5 Differential Gene Expression Analysis

For that step, the DESeq2 v.1.36.0 (Love et al., 2014) is utilised, and design identification is the essential component of high performance data analysis. DESeq2 package applies negative binomial generalised linear models to individual genes, and the identification of differentially expressed genes is utilised by the Wald test for hypothesis testing. In this analysis, replicate+condition is used as a design based on the metadata columns mentioned earlier. We explored four different experiment data analyses between two groups: Healthy Cytokine Vs Healthy Control, Degenerated Control Vs Healthy Control, Healthy Treated Vs Healthy Cytokine and Degenerated Treated Vs Degenerated Control. By applying to relevel command, we change a reference level, and it will vary in 4 different data analyses according to our focused groups.

For quality control checks on the sample in the experiment, firstly, we extract gene-level counts from the differential expression variable and transform them using regularised logarithm transformation (RLD), minimising discrepancies between samples for genes with small counts and normalising to library size. Moreover, we test whether the samples are clustered as expected by producing a sample heatmap through pheatmap R package (Kolde, 2019). We also produce a PCA plot that identifies the samples as clusters, patterns and relationships in the data through PCAtools R package (Blighe and Lun, 2022).

In the differential gene analysis, we decided that the adjusted p-value cutoff and alpha value are 0.05, the Benjamini-Hochberg method is applied for adjusting p-values, and independent hypothesis weighting is chosen as the performance of independent filtering through IHW v.1.24.0 Bioconductor package (Ignatiadis and Huber, 2017). Moreover, we use a shrinkage method called " Adaptive Permutation based Empirical Bayes Gene-wise Linear Models" known as apegelm v.1.18.0 (Zhu et al., 2018) to alleviate the problem of unreliably large log fold change estimates. Biological and statistically significant differentially expressed upregulated and downregulated genes are determined by $| \log_{2}FC | > 1$ and adjusted p-value < 0.05 .

2.6 Gene Set Enrichment and Functional Enrichment Analysis

The final data analysis explored different gene set enrichment in these GO, KEGG and Reactome pathways. We downloaded three gmt files: Gene Ontology gene sets such as Biological Process, Cellular Component, Molecular Function as subsets, KEGG and Reactome subsets as Canonical pathways. In these gene set enrichment analyses, we utilise the fgsea v.1.24.0 R package (Korotkevich et al., 2019) with default specifications to decide functionally associated gene groups. We extracted the gene names and related columns from the differential gene expression dataset to produce the ranked gene catalogue. Moreover, various enrichment gene sets in each GO, KEGG and Reactome pathways are produced, in which the normalised enrichment score is more significant than 2.2 and less than -2.2. Enriched gene sets less than 0.05 p-value are considered significant. The results are described in the bubble plot, and the heatmap shows the differentially expressed genes in each specific enriched gene set. Finally, the over-representation methods of clusterprofiler package v.4.8.1 (Wu et al., 2021; Yu et al., 2012) are utilised to identify functionally enriched pathways with parameters of 0.05 adjusted p-value and 0.01 q-value cutoff and the Benjamini-Hochberg method as pAdjustmethod after changing gene-IDs into Entrez-IDs, producing dot plots and category plots.

3 Results

3.1 Quality Control

After running the raw data files in fastqc tools, the adapter content plot of this multiqc report describes that 2 % of the sequences of some raw data files possess adapter sequences, primers and poly-A tails (Fig.S1A). With trim-galore tools, the adaptor content declines around 0.5 % of the sequences (Fig.S1B) and around 99 % of reads of all raw data files passed the filter (Table.S1). Regarding total reads after filtering, the data range from 65.4 million reads in Healthy Cytokine replicate two to 86.6 million in Degenerated Treated replicate one. After filtering, the read-level GC content is approximately 50 % of each raw data subset, showing no technical biases and the quality of sequencing libraries. The mean quality scores of all subsets are over 35 Phred scores in the sequence quality histograms (Fig.S2) and the per sequence quality scores plot (Fig.S3), which depicts a better base call.

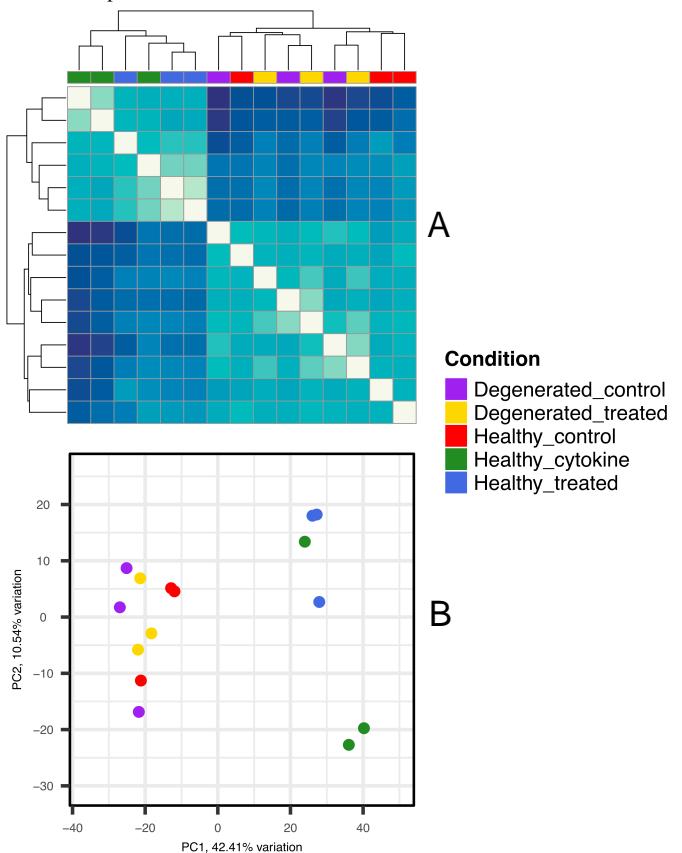


Fig. 3: (A) Sample heatmap describing hierarchical clustering of all mRNA data samples and (B) PCA bi-plot of comparative analysis of mRNA differential expression

To overview the similarity and dissimilarity between different subsets, a sample heatmap (Fig.3A) and principal component analysis (PCA) plot are produced (Fig.3B). In the PCA plot, PC1 (x-axis) represents 42.41% of the underlying biological variation in the dataset, separating Healthy Cytokine and Healthy Treated samples from other samples. With respect to Degenerated Control, Degenerated Treated and Healthy Control samples, PC2 (y-axis) represents 10.54% of the signal between samples. Therefore, we expect a high number of differentially expressed genes to be returned by Healthy Cytokine vs Healthy Control, fewer returned by Degenerated Control vs Healthy Control, Healthy Treated vs Healthy Cytokine and Degenerated Treated vs Degenerated Control, respectively.

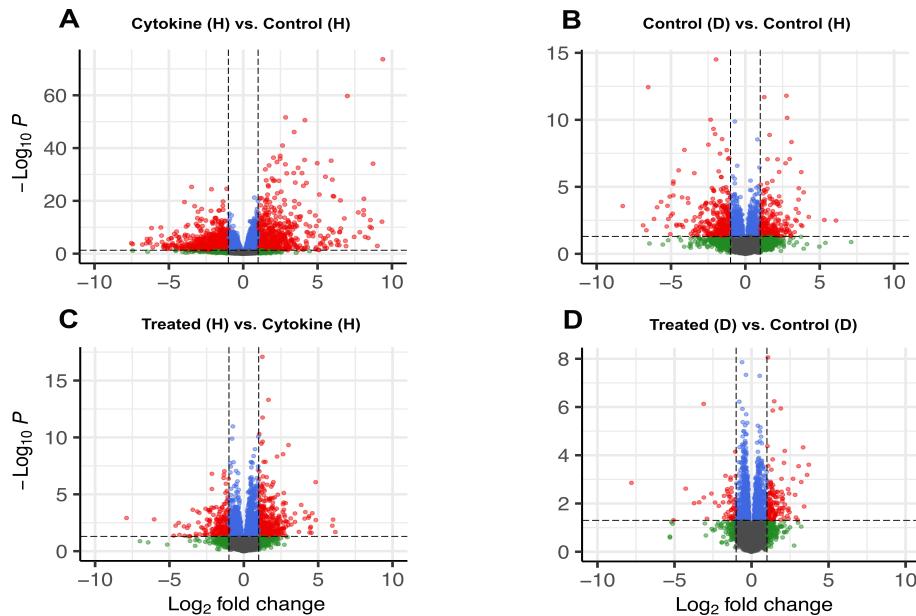


Fig. 4: Enhanced volcano plots showing differentially expressed genes between (A) Healthy Cytokine & Healthy Control, (B) Degenerated Control & Healthy Control, (C) Healthy Treated & Healthy Cytokine, (D) Degenerated Treated & Degenerated Control

3.2 Healthy Cytokine Vs Healthy Control

In the differential gene expression analysis between the healthy NP cells control group (Healthy Control) and healthy NP cells with cytokine cocktail (Healthy Cytokine) group, there are statistically significant 454 upregulated genes and 420 downregulated genes in the total of 16748 genes (Fig.4A).

In the heatmap, CXCL3, CXCL8, MARCOL, CXCL6 and IL6 are the most significant expressed genes between healthy control and healthy cytokine experiment groups (Fig.5) after applying the shrinkage estimator apeglm to reduce spurious signals with high variance (before Fig.S4, after Fig.S5).

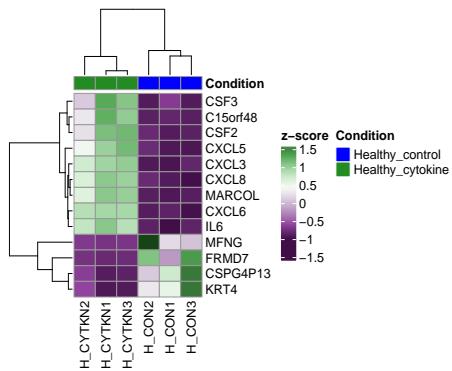


Fig. 5: Heatmap showing most differentially expressed genes using z-score between Healthy Cytokine and Healthy Control groups

In the Gene Ontology gene set enrichment analysis, GOBP Neutrophil Migration, GOBP Neutrophil Chemotaxis, GOBP Granulocyte Migration, GOMF Chemokine Activity ($p < 0.048$) serve as the most enriched gene sets in healthy NP cells with cytokine cocktail group (Healthy Cytokine) (Fig.6). There are 43 enrichment plots where the normalized enrichment score (NES) exceeds 2.2 and less than -2.2 (Fig.S6). In the heat map of specific GOBP Neutrophil Chemotaxis pathway between these two groups, CCL2, PDE4B, CXCL3, CXCL2, CXCL6, CXCL8, CXCL1, and CXCL5 are significant genes while TNFAIP6, CCL2, PDE4B, CXCL3, CXCL2, CXCL6, CXCL8, CXCL1 are also significant upregulated ones

in the Healthy Cytokine group at specific GOBP Neutrophil Migration pathway (Fig.S7). Through the application of the clusterprofiler package, generation of precursor metabolites and energy ($p < 2.37e -15$), cellular respiration($p < 8.36e -14$), ATP metabolic process ($p < 1.83e -12$) and aerobic respiration ($p < 1.73e -11$), are functionally enriched in Healthy Cytokine groups (Fig.S8). There are also 35 genes involved in neutrophil migration.

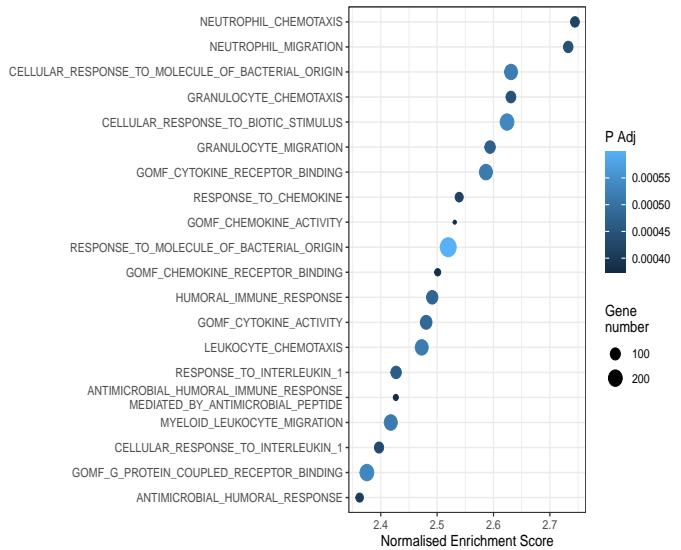


Fig. 6: Gene set enrichment analysis plot describing the most enriched Gene Ontology terms in Healthy Cytokine group

KEGG gene set enrichment analysis reported that Cytokine-Cytokine Receptor Interaction, and Chemokine Signalling pathway ($p < 0.0087$) are the most statistically represented (Fig.7), and their NES is greater than 2.2 in Healthy Cytokine group. In the heatmap of specific KEGG: Cytokine-Cytokine Receptors Interaction pathway (Fig.S9), CXCL5,

CSF2, CXCL1, CXCL2, CXCL3, CXCL6, CXCL8, CCL2, CSF1, IL6 are significantly upregulated genes in Healthy Cytokine group. At the same time, CXCL6, CCL2, CXCL2, CXCL3, CXCL8 and NFKBIA are significant upregulation in the Healthy Cytokine group at the specific pathways of KEGG: Chemokine Signalling pathway.

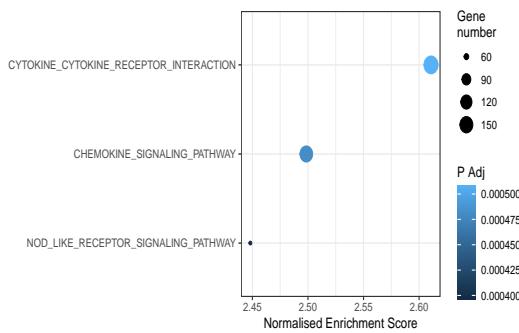


Fig. 7: Gene set enrichment analysis plot describing the most enriched KEGG terms in Healthy Cytokine group

Reactome pathway analysis affirmed that Interleukin 10 Signalling, Chemokine Receptors Bind Chemokines, and Cytokine Signaling in Immune System ($p<0.019$) are the most statistically significant gene sets in Healthy Cytokine group (Fig.8). In the specific Interleukin 10 Signalling pathway (Fig.S11), CCL2, CSF1, IL6, CXCL2, CXCL8 and CXCL1 have upregulated genes in the Healthy Cytokine group. In contrast, CXCL10, CCL7, CCL2, CXCL1, CXCL2, CXCL3, CXCL6 and CXCL8 are significantly upregulated in the Healthy Cytokine group through specific Reactome: Chemokine Receptors Bind Chemokine pathway. These Reactome terms are functionally enriched, such as the citric acid (TCA) cycle and respiratory electron transport ($p<3.64e-12$), signalling by interleukins ($p<2.66e-10$) and respiratory electron transport ATP synthesis by chemiosmotic coupling ($p<1.74e-10$) in Healthy Cytokine group via clusterProfiler package (Fig.S12).

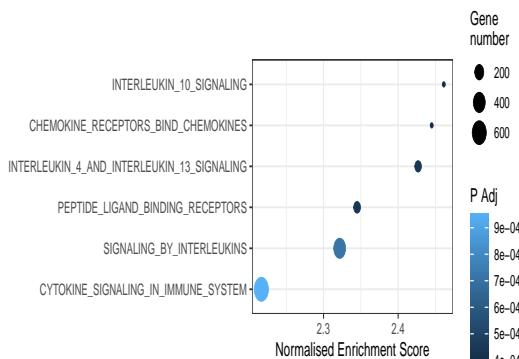


Fig. 8: Gene set enrichment analysis plot describing the most enriched Reactome terms in Healthy Cytokine group

3.3 Degenerated Control Vs Healthy Control

The difference in NP cells between these two groups is that one is isolated from non-degenerated IVD (Healthy Control), and the remaining one is isolated from healthy cells of degenerated cells (Degenerated Control) and, we explored to find the difference of biological gene set between these groups. In the differential gene expression analysis between healthy NP cells from healthy IVD control (Healthy Control) and healthy NP cells from degenerative IVD control groups (Degenerated Control), there are statistically significant 36 upregulated and 26 downregulated genes in the

total of 16919 genes (Fig.4B). Through the fgsea package, there is no enrichment pathway in Gene Ontology, KEGG and Reactome enrichment analysis. Using clusterProfiler package, collagen-containing extracellular matrix ($p<3.81e-07$) and regulation of ion transmembrane transport ($p<1.06e-03$) are the functionally negatively enriched Gene Ontology pathways in Degenerated control. At the same time, KEGG Ribosome ($p<3.54e-11$) and Reactome ECM Organisation ($p<0.0001$) are the top functionally downregulated enriched pathways in the Degenerated control group. The LAMC2 gene is differentially expressed in specific Reactome ECM organization and Reactome assembly of collagen fibrils and other multimeric structures ($p<0.0108$).

3.4 Healthy Treated Vs Healthy Cytokine

In the differential gene expression analysis between the healthy NP cells with cytokine cocktail group (Healthy Cytokine) and the mixture of healthy NP cells and cytokine cocktail treated with 3Fax-Peracetyl Neu5Ac group (Healthy Treated), there are 27 upregulated and 58 downregulated genes in a total of 15155 genes (Fig.4C).

Regarding Gene Ontology enrichment analysis via the fgsea package, the GOMF Structural Constituent of Ribosome ($p<0.017$) is the most enriched pathway in the Healthy Treated group. In this specific pathway heatmap, RPL35A, RPL22, RPS25 and RPL23 genes are downregulated in the Healthy Treated group. Moreover, these four pathways, such as GOCC Cytosolic Ribosome, GOCC Ribosomal Subunit, GOBP Cytoplasmic Translation and GOCC Ribosome pathways ($p<0.017$), are enriched in the Healthy Treated group via fgsea and clusterProfiler packages (Fig.9, S13). There are 37 genes related to these pathways. In specific GOCC Cytosolic Ribosome pathways, RPL27A, RPL39, RPLP1 and RPS13 genes are downregulated in the Healthy Treated group. In addition, focal adhesion ($p<4.5e-15$), cell-substrate junction ($p<8.8e-15$) and transmembrane receptor protein serine/threonine kinase signalling pathways ($p<3.8e-03$) are also functionally enriched through the analysis of clusterProfiler tools (Fig.S13).

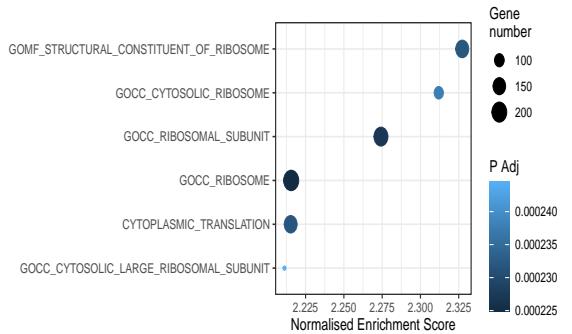


Fig. 9: Gene set enrichment analysis plot describing the most enriched Gene Ontology terms in Healthy Treated group

The KEGG enrichment analysis enriched the KEGG Ribosome pathway ($p<0.011$), and functional analysis validated TGF beta signalling pathway ($p<1.04e-02$) in the Healthy Treated group (Fig.S15). In this specific KEGG Ribosome pathway, RPL27, RPS10, RPL13A, RPS3A, RPL38, and RPS25 genes are downregulated in the Healthy Treated group (Fig.S14).

Reactome pathway analysis through fgsea and clusterProfiler packages disclosed that Eukaryotic Translation Elongation, SRP Dependent Cotranslational Protein Targeting to Membrane, Response of EIF2AK4 GCN2 to Amino Acid Deficiency and Selenoamino Acid Metabolism pathways ($p<0.011$) are also the most represented pathways in Healthy

Treated group (Fig.10). In the specific Eukaryotic Translation Elongation pathway, these genes, including RPL27, RPS10, RPL38, RPL41, RPL35A, RPL22 and EEF1B2, are also downregulated in the Healthy Treated group (Fig.S16). In addition, RPL35A, RPL22, RPS25, RPL23, RPS17 and RPS3 are downregulated genes of the Healthy Treated group in the SRP Dependent Cotranslational Protein Targeting pathway. Cellular Response to starvation pathway ($p < 9.3e -23$) is also functionally enriched in the group (Fig.S17).

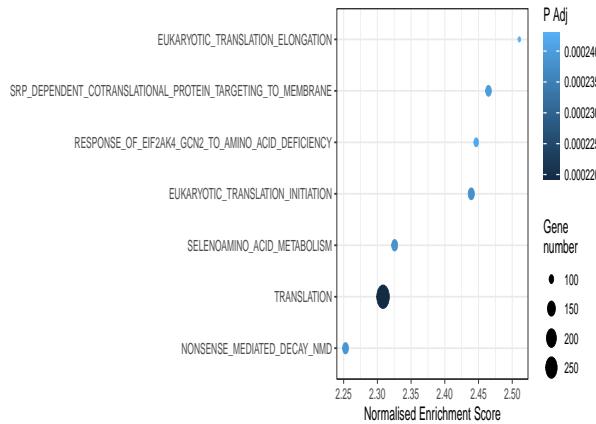


Fig. 10: Gene set enrichment analysis plot describing the most enriched Reactome terms in Healthy Treated group

3.5 Degenerated Treated Vs Degenerated Control

In the differential gene expression analysis between the Degenerated Control group and degenerated NP cells with 3Fax-Paracetamol Neu5Ac treatment (Degenerated Treated) group, there are statistically significant four downregulated genes in the 14714 genes. In the heatmap, HYDIN2, SEMA3A, IL31RA, and CLDN1 are the most represented expressed genes between these groups (Fig.4D).

Gene Ontology enrichment analysis via fgsea package revealed that GOCC Mitochondrial Protein Containing Complex ($p < 0.048$), GOMF Extracellular Matrix Structural Constituent, GOBP ATP Synthesis Coupled Electron Transport ($p < 0.039$), GOBP Oxidative Phosphorylation ($p < 0.041$) and GOBP Mitochondrial Respiratory Chain Complex Assembly ($p < 0.039$) are the enriched pathways in the Degenerated Control group when GOMF Extracellular Matrix Structural Constituent, HP Osteoarthritis and GOBP External Encapsulating Structure Organisation ($p < 0.026$) are the enriched pathways in Degenerated Treated groups (Fig.11). Through the analysis of the clusterProfiler package, Mitochondrial Inner Membrane ($p < 1.24e -07$) and ATP Metabolic Process pathways ($p < 8.44e -06$) are also functionally enriched (Fig.S18). Nine genes are also involved in ATP Synthesis Coupled Electron Transport and Oxidative Phosphorylation (Fig.12). In the heatmap of specific GOCC Inner Mitochondrial Membrane Protein Complex pathway, NDUFAB1, NDUFA3, NDUFA8, NDUFC1, NDUFB10 and MICOS13 genes are upregulated in the Degenerated Treated group. In contrast, NDUFA8, UQCC3, ATP5PD, NDUFA3, NDUFB10, SDHC and STOML2 genes are upregulated in the Degenerated Treated group via a specific GOBP Oxidative Phosphorylation pathway. In the specific ATP Synthesis Coupled Electron Transport pathway, the Degenerated Treated group also upregulated UQCC3, NDUFB3, NDUFB10, NDUFA3, NDUFA8 and NDUFAB1 genes.

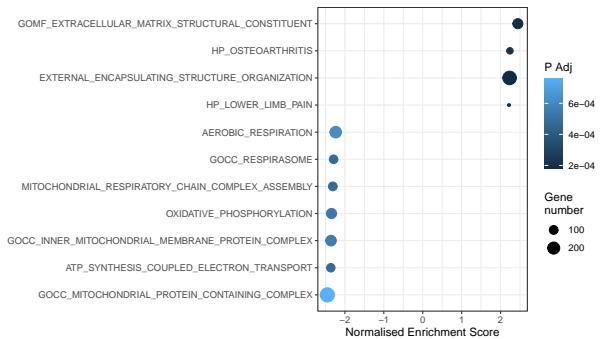


Fig. 11: Gene set enrichment analysis plot describing the most enriched Gene Ontology terms (positive NES) in Degenerated Treated group and the most enriched Gene Ontology terms (negative NES) in Degenerated Control group

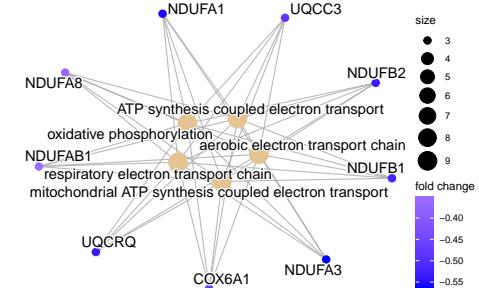


Fig. 12: Category plot of the genes relating with specific gene ontology: ATP synthesis coupled electron transport and oxidative phosphorylation pathways between Degenerated Treated and Degenerated Control groups

KEGG enrichment analysis, not only the fgsea package and clusterProfiler package, proved that KEGG Oxidative Phosphorylation ($p < 0.027$) pathway is enriched (Fig.S19). In this specific pathway, NDUFB3, NDUFB5, NDUFB10, NDUFAB1, NDUFA3, SDHC and ATP5F1C genes are differentially expressed between these groups.

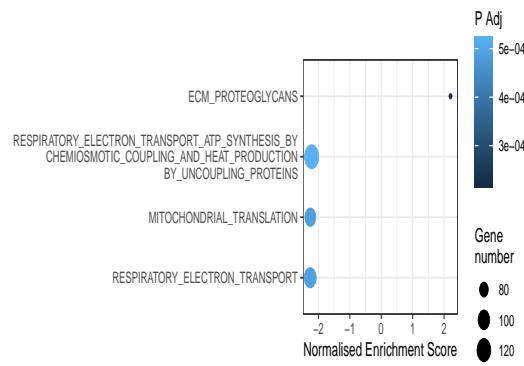


Fig. 13: Gene set enrichment analysis plot describing the most enriched Reactome terms (positive NES) in Degenerated Treated group and the most enriched Reactome terms (negative NES) in Degenerated Control group

In Reactome enrichment analysis, Mitochondrial Translation ($p < 0.032$) and Respiratory Electron Transport ($p < 0.032$) pathways are enriched in the Degenerated Control group, and ECM Proteoglycans ($p < 0.026$) are enriched in the Degenerated Treated group via fgsea package (Fig.13). In Respiratory Electron Transport pathways, NDUFA8, UQCC3, ATP5PD, NDUFA3, NDUFB10, SDHC and STOML2 genes are upregulated in the Degenerated Treated group via a specific GOBP Oxidative Phosphorylation pathway. In the specific ATP Synthesis Coupled Electron Transport pathway, the Degenerated Treated group also upregulated UQCC3, NDUFB3, NDUFB10, NDUFA3, NDUFA8 and NDUFAB1 genes.

NDUFB10, NDUFAB1, and SDHC genes are upregulated in the Degenerated Treated group (Fig.S20). In specific Mitochondrial Translation pathway, MRPL18, MRPL41, MRPL43, MRPS15, MRPS18A and TUFM genes are highly expressed in the Degenerated Treated group. Moreover, Complex I biogenesis ($p < 7.87e-06$) and the Citric acid (TCA) cycle and respiratory electron transport ($p < 2.77e-05$) are functionally enriched.

4 Discussion

Although the current treatment of IVDD, such as analgesics, rest and surgery, mostly emphasises reducing the symptoms, it does not provide to regenerate the normal anatomy and physiology of IVD. Regarding focusing on the regeneration of degenerative IVD, there is still left to understand the cellular and genetic influence on the pathophysiology of IVDD. Developing bioinformatics tools helps to understand more about various human diseases' cellular and molecular levels. In this research, we explored how Neu5Ac-inhib affects glycosylation inhibition through different gene set enrichment analyses of transcriptomic data.

The statistically differentially expressed genes between the Healthy Control and Healthy Cytokine group are the largest in the number of 454 upregulated genes and 420 downregulated genes compared with other groups' differentiation. Pro-inflammatory cytokine produced by IVD cells is considered as a mediator of IVDD (Risbud and Shapiro, 2014). These inflammatory cells can stimulate autophagy, and apoptosis of IVD cells (Purmessur et al., 2013; Shen et al., 2011). Finally, the metabolic imbalance results in the degeneration of IVDD (Risbud and Shapiro, 2014). Through the differential gene expression analysis of Healthy Control and Healthy Cytokine experimental groups, gene ontology terms such as neutrophil migration, neutrophil chemotaxis, granulocyte chemotaxis, cytokine receptor binding, and chemokine activity pathways (Fig.6) are enriched at the Healthy Cytokine groups in this present study. In addition, the most significant differentially expressed gene is chemokines C-C motif ligand 6, CXCL6, (Fig.5), which is increased in the serum levels of disc degeneration patients (Grad et al., 2016). One of the most noticeable differentially expressed genes in this comparison, IL6, performs an important factor in IVDD (Deng et al., 2016). CCL20, which is highly expressed in human degenerative cells of this study, and the CCR6 system will involve the smuggling of IL-17 secreting T-cell subset (Th 17) to degenerative cells (Zhang et al., 2013). Increased cytokine levels and chemokines recruit inflammatory mediators, MMPs and enhancement of matrix degradation (Gabr et al., 2011; Smith et al., 2011). In our study, KEGG terms such as cytokine-cytokine receptor interaction and nod-like receptor signalling pathways are also enriched in the Healthy Cytokine group (Fig.7). Metabolic imbalance reduces the ECM's quality, quantity and functional properties (Zhou et al., 2021). IL1 β stimulated MMPs, and upregulated MMPs gene expression results in IVDD (Fan et al., 2022). This study highly expresses MMP-12, 3, 1, 10 in the Healthy Cytokine group. Moreover, Reactome terms such as cytokine signalling in the immune system, chemokine receptor bind chemokines and signalling by interleukins (Fig.8) are enriched in this study and are compatible with the pro-inflammatory cytokines process. Through the stimulation of cytokine cocktail (IL1 β , IL-6 and TNF α) in healthy NP cells, this gene set enrichment and differential gene analysis will describe neutrophil migration, cytokine activity and metabolic impairment, which help to understand more about the pathophysiological process of IVDD.

Although there are 36 upregulated and 26 downregulated genes between Healthy Control and Degenerated Control experimental groups, there is no specific enriched gene set of Gene Ontology, KEGG and Reactome pathway between these two groups. The reason is that when the cells are cultured under the same conditions, they may perform

similarly, although some are received from healthy patients, and some are received from degenerated disc cells. The degenerated disc cells may perform as healthy cells because they are grown in a healthy environment with oxygen and nutrients. However, WIF1 and LAMC2 genes are the groups' most significant differentially expressed genes. The reduction of the Wnt signalling pathway during NP chondrofication can be interpreted by increasing WIF1 and Frizzled-related protein (Wnt inhibitors) levels in chondrodystrophic dogs (Iwata et al., 2015).

Ribosomes are essential to maintain mRNA-guided protein homeostasis (Zhao et al., 2022), most of which are located in the cytoplasm. The ribosome is composed of ribosomal RNA (rRNA), ribosomal proteins (RPs) and small nucleolar RNAs (snoRNAs). The large eukaryotic subunit of ribosomes consists of 5S rRNA, 5.8S rRNA and 47 RPs and its function are to catalyse peptide bond formation with the peptidyl transferase centre (PTC). Ribosomes are a significant administrator in the precise regulation process of translation, in which initiation, prolongation, termination and ribosome recycling has happened (Schuller and Green, 2018). This study shows 27 upregulated and 58 downregulated genes in the differential gene expression analysis between Healthy Treated and Healthy Cytokine experiment groups. In the mixture of healthy NP cells and cytokine cocktail treated with Neu5Ac-inhib group (Healthy Treated), GO terms such as cytosolic ribosome and large ribosomal subunit are enriched (Fig.9) while Reactome terms such as eukaryotic translation initiation and elongation are also enriched (Fig.10). In these specific pathways, RPL27A, RPL35A, RPL39, RPLP1, RPS25 and RPS26 are downregulated in the Healthy Treated groups. Selenium is one of the essential trace elements of humans. Selenoamino acid is formed when selenium is substituted for sulphur. In the ocular inflammation, lipopolysaccharides induced inflammation, suppressed the selenoamino acid metabolism to decline antioxidant activity, and finally created oxidative stress (Chu et al., 2022). In the synovial fluid metabolomic analysis, selenoamino acid metabolism is downregulated in the early osteoarthritis phenotypes (Carlson et al., 2019). Our study enriched the selenoamino acid metabolism pathway in the healthy NP cells and cytokine cocktail treated with Neu5Ac-inhib (Healthy Treated). The role of signal recognition particle (SRP) serves as cotranslational protein conveyance for translocation into the endoplasmic reticulum (ER). Without SRP, normally translocated to the ER was mistargeted into mitochondria where defects result, and SRP is proved to maintain the efficiency and precision of protein targeting (Costa et al., 2018). After the protein translocation process into the ER lumen, glycosylation, the most common post-translational modification of the protein occurs. It also maintains folding, localisation and protein functions (Hirata et al., 2022; Schjoldager et al., 2020). Moreover, these SRP-dependent cotranslational protein targeting and selenoamino acid metabolism pathways are also enriched in the Healthy Treated group of this study (Fig.10). Neu5Ac-inhib may prevent ribosomal dysfunction through the large ribosomal subunit and cytosolic translational pathways and support antioxidant activity and protein targeting specificity via selenoamino acid pathway and SRP dependent pathway respectively.

Apoptosis is one of the essential factors in the amelioration of degeneration (Ding et al., 2013). Due to the NP cells' apoptosis, the disc cannot afford to produce enough cells for ECM. Significant oxidative stress is one of the causative agents of apoptosis (Chen et al., 2014; Yang et al., 2015). Aerobic respiration and electron transport chain synthesise this oxidative stress and reactive oxygen species (ROS) in the mitochondria (Wang et al., 2022). The initial duty of mitochondria is to supply energy to the cell by ATP production (Orsini et al., 2004). Mitochondria in disc cells are the primary attack of ROS from the excessive collection of ROS. Due to oxidative stress, dysfunctional mitochondria produce imbalanced mitochondria dynamics and irregular mitophagy (Ansari et al., 2018; Zhang et al., 2018). Through the differential gene analysis between Degenerated Control and Degenerated Treated experimental groups,

oxidative phosphorylation, aerobic respiration, mitochondria respiratory chain complex assembly and ATP synthesis coupled electron transport pathway are enriched in Degenerated Control groups (Fig.11). The straight connection between sialic acid and oxidative stress is not fully explained, however, it says that sialic acids serve as the scavenger of H₂O₂ radicals and oxidative sialic acid residues will be the target of ROS (Cho et al., 2017). One of the interesting things is that Neu5Ac-inhib may reduce the electron transport chain, which will prevent oxidative stress and ROS in the degenerated NP cells. NDUFA8, NDUFA3, UQCRC3, ATP5PD and SDHC genes are upregulated in the Degenerated Treated group through the specific GO oxidative phosphorylation pathway and ATP synthesis coupled electron transport pathway. Moreover, the metabolism of ECM is balanced in healthy IVD cells (Kepler et al., 2013). When different stimuli destroy the balance of ECM, IVDD will occur. In these cases, the number of proteoglycans and collagen type II decline enormously. In the present study, gene ontology molecular function extracellular matrix structural constituent and reactome ECM proteoglycans are enriched in the Degenerated Treated group, which shows the maintenance of ECM by Neu5Ac-inhib in degenerated NP cells (Fig.13). This result shows that not only Neu5Ac-inhib will prevent apoptosis, oxidative stress and ROS in NP cells and reverse mitochondrial dysfunction by reducing the pathways of oxidative phosphorylation, and ATP synthesis coupled electron transport pathway but also balance the metabolism of ECM by ECM structural constituent and ECM proteoglycans pathways.

Overall, the data analysis between Healthy Control and Healthy Cytokine groups explains the inflammatory process in IVDD. The analysis from the Healthy Treated group shows that Neu5Ac-inhib will limit dysfunctional ribosomes through the ribosomal translational pathway and produce anti-oxidation and specific and efficient protein targeting through selenoamino acid metabolism and SRP-dependent cotranslational protein targeting to membrane. Moreover, when Neu5Ac-inhib apply to degenerative NP cells, it reduces oxidative stress through oxidative phosphorylation and aerobic respiration, and it also maintains ECM homeostasis by the ECM proteoglycans pathway. Neu5Ac-inhib is assumed to be beneficial in the regeneration process of IVDD because of the mentioned facts.

However, poor differentiation between Healthy and Degenerated Control groups through the PCA plot may show that the samples exhibit a similar transcriptional profile. In the future, sequencing additional samples to increase the sample size (power) could supply more differentially expressed genes. This transcriptomic data analysis focused on the gene set enrichment analysis, not specific gene findings. In future work, specific gene-finding research about the sialylation inhibition in the IVDD and proteomics analysis about the ribosome involvement will need to be carried out. Moreover, the role of sialylation inhibition in the selenoamino acid metabolism and SRP-dependent cotranslational protein targeting to membrane might be interesting areas to investigate more in the future.

Acknowledgements

I want to thank my supervisors, Professor Abhay Pandit, Dr Kieran Joyce and Barry Digby, for the guidance and precious feedback on the project this year. I also want to acknowledge other CÚRAM lab members and my classmates, especially Liam Porter for their advice. Lastly, I extend my heartfelt appreciation to my family and fiancé, May, for their unwavering support and motivation during the year.

Funding

This publication has emanated from research conducted with the financial support of the College of Medicine, Nursing and Health Sciences, NUI

Galway, Science Foundation Ireland (SFI) and the European Regional Development Fund (Grant Number 13/RC/2073 P2)

References

- Andersson, G. B. (1999). Epidemiological features of chronic low-back pain. *The lancet*, 354(9178):581–585.
- Andrews, S. (2010). Fastqc: a quality control tool for high throughput sequence data.
- Ansari, M. Y., Khan, N. M., Ahmad, I., and Haqqi, T. M. (2018). Parkin clearance of dysfunctional mitochondria regulates ros levels and increases survival of human chondrocytes. *Osteoarthritis and cartilage*, 26(8):1087–1097.
- Blighe, K. and Lun, A. (2022). *PCAtools: PCAtools: Everything Principal Components Analysis*. R package version 2.8.0.
- Buckwalter, J. A. (1995). Aging and degeneration of the human intervertebral disc. *Spine*, 20(11):1307–1314.
- Carlson, A. K., Rawle, R. A., Wallace, C. W., Brooks, E. G., Adams, E., Greenwood, M. C., Olmer, M., Lotz, M. K., Bothner, B., and June, R. K. (2019). Characterization of synovial fluid metabolomic phenotypes of cartilage morphological changes associated with osteoarthritis. *Osteoarthritis and cartilage*, 27(8):1174–1184.
- Chen, J.-W., Ni, B.-B., Li, B., Yang, Y.-H., Jiang, S.-D., and Jiang, L.-S. (2014). The responses of autophagy and apoptosis to oxidative stress in nucleus pulposus cells: implications for disc degeneration. *Cellular Physiology and Biochemistry*, 34(4):1175–1189.
- Chen, X. and Varki, A. (2010). Advances in the biology and chemistry of sialic acids. *ACS chemical biology*, 5(2):163–176.
- Cheung, K. M., Karppinen, J., Chan, D., Ho, D. W., Song, Y.-Q., Sham, P., Cheah, K. S., Leong, J. C., and Luk, K. D. (2009). Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. *Spine*, 34(9):934–940.
- Cho, A., Christine, M., Malicdan, V., Miyakawa, M., Nonaka, I., Nishino, I., and Noguchi, S. (2017). Sialic acid deficiency is associated with oxidative stress leading to muscle atrophy and weakness in gne myopathy. *Human molecular genetics*, 26(16):3081–3093.
- Chu, K. O., Chan, K. P., Yip, Y. W. Y., Chu, W. K., Wang, C. C., and Pang, C. P. (2022). Systemic and ocular anti-inflammatory mechanisms of green tea extract on endotoxin-induced ocular inflammation. *Frontiers in Endocrinology*, 13:899271.
- Clouet, J., Vinatier, C., Merceron, C., Pot-Vaucel, M., Hamel, O., Weiss, P., Grimandi, G., and Guicheux, J. (2009). The intervertebral disc: from pathophysiology to tissue engineering. *Joint Bone Spine*, 76(6):614–618.
- Cloyd, J. M., Malhotra, N. R., Weng, L., Chen, W., Mauck, R. L., and Elliott, D. M. (2007). Material properties in unconfined compression of human nucleus pulposus, injectable hyaluronic acid-based hydrogels and tissue engineering scaffolds. *European spine journal*, 16:1892–1898.
- Collin, E., Kilcoyne, M., White, S., Grad, S., Alini, M., Joshi, L., and Pandit, A. (2016). Unique glycosignature for intervertebral disc and articular cartilage cells and tissues in immaturity and maturity. *Scientific reports*, 6(1):23062.
- Costa, E. A., Subramanian, K., Nunnari, J., and Weissman, J. S. (2018). Defining the physiological role of srp in protein-targeting efficiency and specificity. *Science*, 359(6376):689–692.
- da Silva Correia, J. and Ulevitch, R. J. (2002). Md-2 and tlr4 n-linked glycosylations are important for a functional lipopolysaccharide receptor. *Journal of Biological Chemistry*, 277(3):1845–1854.
- De Geer, C. M. (2018). Intervertebral disk nutrients and transport mechanisms in relation to disk degeneration: a narrative literature review. *Journal of chiropractic medicine*, 17(2):97–105.
- Deng, X., Zhao, F., Kang, B., and Zhang, X. (2016). Elevated interleukin-6 expression levels are associated with intervertebral disc degeneration. *Experimental and Therapeutic Medicine*, 11(4):1425–1432.
- Ding, F., Shao, Z.-w., and Xiong, L.-m. (2013). Cell death in intervertebral disc degeneration. *Apoptosis*, 18:777–785.
- Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). Biomart and biocconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*, 21:3439–3440.
- Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). Multiqc: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19):3047–3048.
- Fan, N., Yuan, S., Hai, Y., Du, P., Li, J., Kong, X., Zhu, W., Liu, Y., and Zang, L. (2022). Identifying the potential role of il-1 β in the molecular mechanisms of disc degeneration using gene expression profiling and bioinformatics analysis. *Journal of Orthopaedic Surgery*, 30(1):23094990211068203.
- Gabr, M. A., Jing, L., Helbling, A. R., Sinclair, S. M., Allen, K. D., Shamji, M. F., Richardson, W. J., Fitch, R. D., Setton, L. A., and Chen, J. (2011). Interleukin-17 synergizes with ifn γ or tnf α to promote inflammatory mediator release and

- intercellular adhesion molecule-1 (icam-1) expression in human intervertebral disc cells. *Journal of Orthopaedic Research*, 29(1):1–7.
- Gerardy-Schahn, R., Delannoy, P., von Itzstein, M., Hinderlich, S., Weidemann, W., Yardeni, T., Horstkorte, R., and Huizing, M. (2015). Udp-glcnaac 2-epimerase/mannac kinase (gne): a master regulator of sialic acid synthesis. *SialoGlyco Chemistry and Biology I: Biosynthesis, structural diversity and sialoglycopathologies*, pages 97–137.
- Grad, S., Bow, C., Karppinen, J., Luk, K., Cheung, K., Alini, M., and Samartzis, D. (2016). Systemic blood plasma ccl5 and cxcl6: potential biomarkers for human lumbar disc degeneration. *Eur Cell Mater*, 31:1–10.
- Hadjipavlou, A., Tzermiadis, M., Bogduk, N., and Zindrick, M. (2008). The pathophysiology of disc degeneration: a critical review. *The Journal of Bone & Joint Surgery British Volume*, 90(10):1261–1270.
- Hirata, T., Yang, J., Tomida, S., Tokoro, Y., Kinoshita, T., Fujita, M., and Kizuka, Y. (2022). Er entry pathway and glycosylation of gpi-anchored proteins are determined by n-terminal signal sequence and c-terminal gpi-attachment sequence. *Journal of Biological Chemistry*, 298(10).
- Ignatiadis, N. and Huber, W. (2017). Covariate-powered weighted multiple testing with false discovery rate control. *arXiv*.
- Iwata, M., Aikawa, T., Hakozaki, T., Arai, K., Ochi, H., Haro, H., Tagawa, M., Asou, Y., and Hara, Y. (2015). Enhancement of runx2 expression is potentially linked to β -catenin accumulation in canine intervertebral disc degeneration. *Journal of cellular physiology*, 230(1):180–190.
- Johnson, Z. I., Schoepflin, Z. R., Choi, H., Shapiro, I. M., and Risbud, M. V. (2015). Disc in flames: roles of tnf- α and il-1 β in intervertebral disc degeneration. *European cells & materials*, 30:104.
- Joyce, K., Isa, I., Fahey, R., Creemers, L., Devitt, A., and Pandit, A. (2018). The glycomic profile of the intervertebral disc in health and degeneration for biomaterial functionalization. In *Orthopaedic Proceedings*, volume 100, pages 117–117. Bone & Joint.
- Joyce, K., Mohd Isa, I., Krouwels, A., Creemers, L., Devitt, A., and Pandit, A. (2021). The role of altered glycosylation in human nucleus pulposus cells in inflammation and degeneration. *Eur Cell Mater*, 41:401–420.
- Katz, J., Lipson, S., Larson, M., McInnes, J., Fossel, A., and Liang, M. (2006). The journal of bone & joint surgery. *J Bone Joint Surg Am*, 88(Suppl 2):21–24.
- Kazezian, Z., Li, Z., Alini, M., Grad, S., and Pandit, A. (2017). Injectable hyaluronic acid down-regulates interferon signaling molecules, igfbp3 and ifit3 in the bovine intervertebral disc. *Acta biomaterialia*, 52:118–129.
- Keppler, C. K., Ponnappan, R. K., Tannoury, C. A., Risbud, M. V., and Anderson, D. G. (2013). The molecular basis of intervertebral disc degeneration. *The Spine Journal*, 13(3):318–330.
- Kim, Y. J. and Varki, A. (1997). Perspectives on the significance of altered glycosylation of glycoproteins in cancer. *Glycoconjugate journal*, 14:569–576.
- Kolde, R. (2019). pheatmap: Pretty Heatmaps. R package version 1.0.12.
- Komekado, H., Yamamoto, H., Chiba, T., and Kikuchi, A. (2007). Glycosylation and palmitoylation of wnt-3a are coupled to produce an active form of wnt-3a. *Genes & Cells*, 12(4):521–534.
- Korotkevich, G., Sukhov, V., and Sergushichev, A. (2019). Fast gene set enrichment analysis. *bioRxiv*.
- Li, Y. and Chen, X. (2012). Sialic acid metabolism and sialyltransferases: natural functions and applications. *Applied microbiology and biotechnology*, 94:887–905.
- Liu, Y.-C., Yen, H.-Y., Chen, C.-Y., Chen, C.-H., Cheng, P.-F., Juan, Y.-H., Chen, C.-H., Khoo, K.-H., Yu, C.-J., Yang, P.-C., et al. (2011). Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. *Proceedings of the National Academy of Sciences*, 108(28):11332–11337.
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biology*, 15:550.
- Matta, A. and Erwin, W. M. (2020). Injectable biologics for the treatment of degenerative disc disease. *Current reviews in musculoskeletal medicine*, 13:680–687.
- Mohd Isa, I. L., Abbah, S. A., Kilcoyne, M., Sakai, D., Dockery, P., Finn, D. P., and Pandit, A. (2018). Implantation of hyaluronic acid hydrogel prevents the pain phenotype in a rat model of intervertebral disc injury. *Science Advances*, 4(4):eaaoq0597.
- Orsini, F., Migliaccio, E., Moroni, M., Contursi, C., Raker, V. A., Piccini, D., Martin-Padura, I., Pelliccia, G., Trinei, M., Bono, M., et al. (2004). The life span determinant p66shc localizes to mitochondria where it associates with mitochondrial heat shock protein 70 and regulates trans-membrane potential. *Journal of Biological Chemistry*, 279(24):25689–25695.
- Purmessur, D., Walter, B., Roughley, P., Laudier, D., Hecht, A., and Iatridis, J. (2013). A role for tnf α in intervertebral disc degeneration: a non-recoverable catabolic shift. *Biochemical and biophysical research communications*, 433(1):151–156.
- Rillahan, C. D., Antonopoulos, A., Lefort, C. T., Sonon, R., Azadi, P., Ley, K., Dell, A., Haslam, S. M., and Paulson, J. C. (2012). Global metabolic inhibitors of sialyl-and fucosyltransferases remodel the glycome. *Nature chemical biology*, 8(7):661–668.
- Risbud, M. V. and Shapiro, I. M. (2014). Role of cytokines in intervertebral disc degeneration: pain and disc content. *Nature Reviews Rheumatology*, 10(1):44–56.
- Roberts, S., Caterson, B., Menage, J., Evans, E. H., Jaffray, D. C., and Eisenstein, S. M. (2000). Matrix metalloproteinases and aggrecanase: their role in disorders of the human intervertebral disc. *Spine*, 25(23):3005–3013.
- Schjoldager, K. T., Narimatsu, Y., Joshi, H. J., and Clausen, H. (2020). Global view of human protein glycosylation pathways and functions. *Nature reviews Molecular cell biology*, 21(12):729–749.
- Schuller, A. P. and Green, R. (2018). Roadblocks and resolutions in eukaryotic translation. *Nature Reviews Molecular Cell Biology*, 19(8):526–541.
- Shen, C., Yan, J., Jiang, L.-S., and Dai, L.-Y. (2011). Autophagy in rat annulus fibrosus cells: evidence and possible implications. *Arthritis Research & Therapy*, 13(4):1–10.
- Smith, L. J., Chiaro, J. A., Nerurkar, N. L., Cortes, D. H., Horava, S. D., Hebela, N. M., Mauck, R. L., Dodge, G. R., and Elliott, D. M. (2011). Nucleus pulposus cells synthesize a functional extracellular matrix and respond to inflammatory cytokine challenge following long term agarose culture. *European cells & materials*, 22:291.
- Soneson, C., Love, M. I., and Robinson, M. D. (2015). Differential analyses for rna-seq: transcript-level estimates improve gene-level inferences. *F1000Research*, 4.
- Teraguchi, M., Yoshimura, N., Hashizume, H., Muraki, S., Yamada, H., Minamide, A., Oka, H., Ishimoto, Y., Nagata, K., Kagotani, R., et al. (2014). Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: the wakayama spine study. *Osteoarthritis and cartilage*, 22(1):104–110.
- Urban, J. and Roberts, S. (2003). Arthritis research & therapy. *Arthritis Res Ther*, 5(3):120–130.
- Varki, A. (2007). Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature*, 446(7139):1023–1029.
- Vo, N. V., Hartman, R. A., Yurube, T., Jacobs, L. J., Sowa, G. A., and Kang, J. D. (2013). Expression and regulation of metalloproteinases and their inhibitors in intervertebral disc aging and degeneration. *The Spine Journal*, 13(3):331–341.
- Wang, D.-K., Zheng, H.-L., Zhou, W.-S., Duan, Z.-W., Jiang, S.-D., Li, B., Zheng, X.-F., and Jiang, L.-S. (2022). Mitochondrial dysfunction in oxidative stress-mediated intervertebral disc degeneration. *Orthopaedic surgery*, 14(8):1569–1582.
- Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., xiaochong Fu, Liu, S., Bo, X., and Yu, G. (2021). clusterprofiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, 2(3):100141.
- Wuertz, K. and Haglund, L. (2013). Inflammatory mediators in intervertebral disk degeneration and discogenic pain. *Global spine journal*, 3(3):175–184.
- Yang, L., Rong, Z., Zeng, M., Cao, Y., Gong, X., Lin, L., Chen, Y., Cao, W., Zhu, L., and Dong, W. (2015). Pyrroloquinoline quinone protects nucleus pulposus cells from hydrogen peroxide-induced apoptosis by inhibiting the mitochondria-mediated pathway. *European Spine Journal*, 24:1702–1710.
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterprofiler: an r package for comparing biological themes among gene clusters. *OMICS: A Journal of Integrative Biology*, 16(5):284–287.
- Zhang, W., Nie, L., Wang, Y., Wang, X.-p., Zhao, H., Dongol, S., Maharjan, S., and Cheng, L. (2013). Ccl20 secretion from the nucleus pulposus improves the recruitment of crf6-expressing th17 cells to degenerated ivd tissues. *PLoS one*, 8(6):e66286.
- Zhang, Z., Xu, T., Chen, J., Shao, Z., Wang, K., Yan, Y., Wu, C., Lin, J., Wang, H., Gao, W., et al. (2018). Parkin-mediated mitophagy as a potential therapeutic target for intervertebral disc degeneration. *Cell death & disease*, 9(10):980.
- Zhao, P.-y., Yao, R.-q., Zhang, Z.-c., Zhu, S.-y., Li, Y.-x., Ren, C., Du, X.-h., and Yao, Y.-m. (2022). Eukaryotic ribosome quality control system: a potential therapeutic target for human diseases. *International Journal of Biological Sciences*, 18(6):2497.
- Zheng, W., Chung, L. M., and Zhao, H. (2011). Bias detection and correction in rna-sequencing data. *BMC bioinformatics*, 12:1–14.
- Zhou, Z., Cui, S., Du, J., Richards, R. G., Alini, M., Grad, S., and Li, Z. (2021). One strike loading organ culture model to investigate the post-traumatic disc degenerative condition. *Journal of orthopaedic translation*, 26:141–150.
- Zhu, A., Ibrahim, J. G., and Love, M. I. (2018). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics*.

5 Supplementary Materials

5.1 Supplementary Table

Table S1. Multiqc report of all raw files after trimming with trim-galore tools (DCON: Degenerated Control, DTREAT: Degenerated Treated, HCON: Healthy Control, HCYTKN: Healthy Cytokine, HTREAT: Healthy Treated)

Sample Name	Trimmed %	GC Content %	Total Sequences(millions)
DCON1	1.6	50	75.4
DCON2	1.6	50	72.8
DCON3	1.6	50	78
DTREAT1	1.6	51	86.6
DTREAT2	1.4	50	85
DTREAT3	1.1	50	68.8
HCON1	1.4	50	84.2
HCON2	1.6	50	69.4
HCON3	1.6	50	69.8
HCYTKN1	1.6	50	85.4
HCYTKN2	1.4	50	65.4
HCYTKN3	1.6	50.5	77.8
HTREAT1	1.6	50	67.2
HTREAT2	1.6	50	79.4
HTREAT3	1.6	50	69.2

5.2 Supplementary Figures

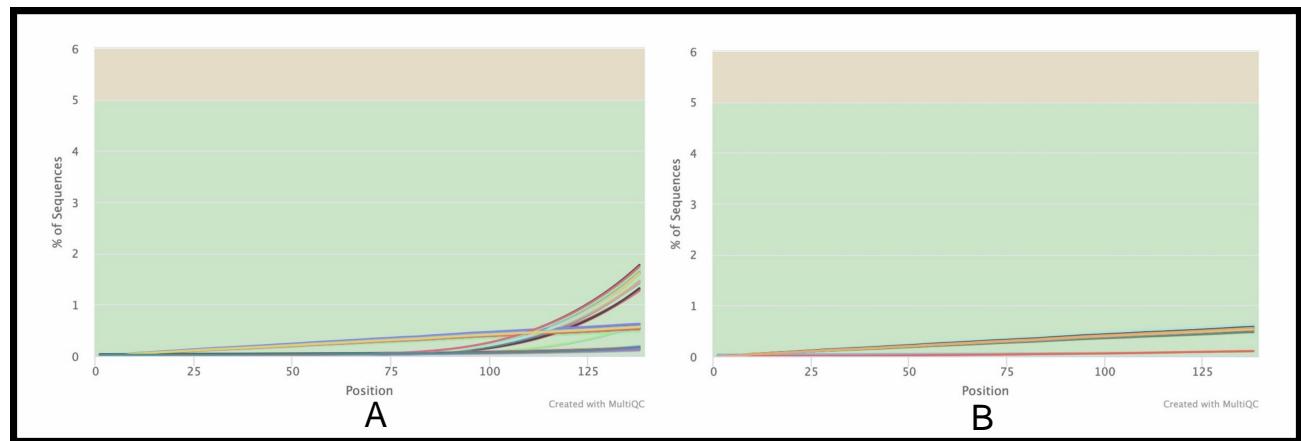


Fig. S1: Adapter content plot of RNA-seq data (A) before the application of trim-galore tools and (B) after the application of trim-galore tools



Fig. S2: Sequence Quality Histogram in the multiqc report after trimming with trim-galore

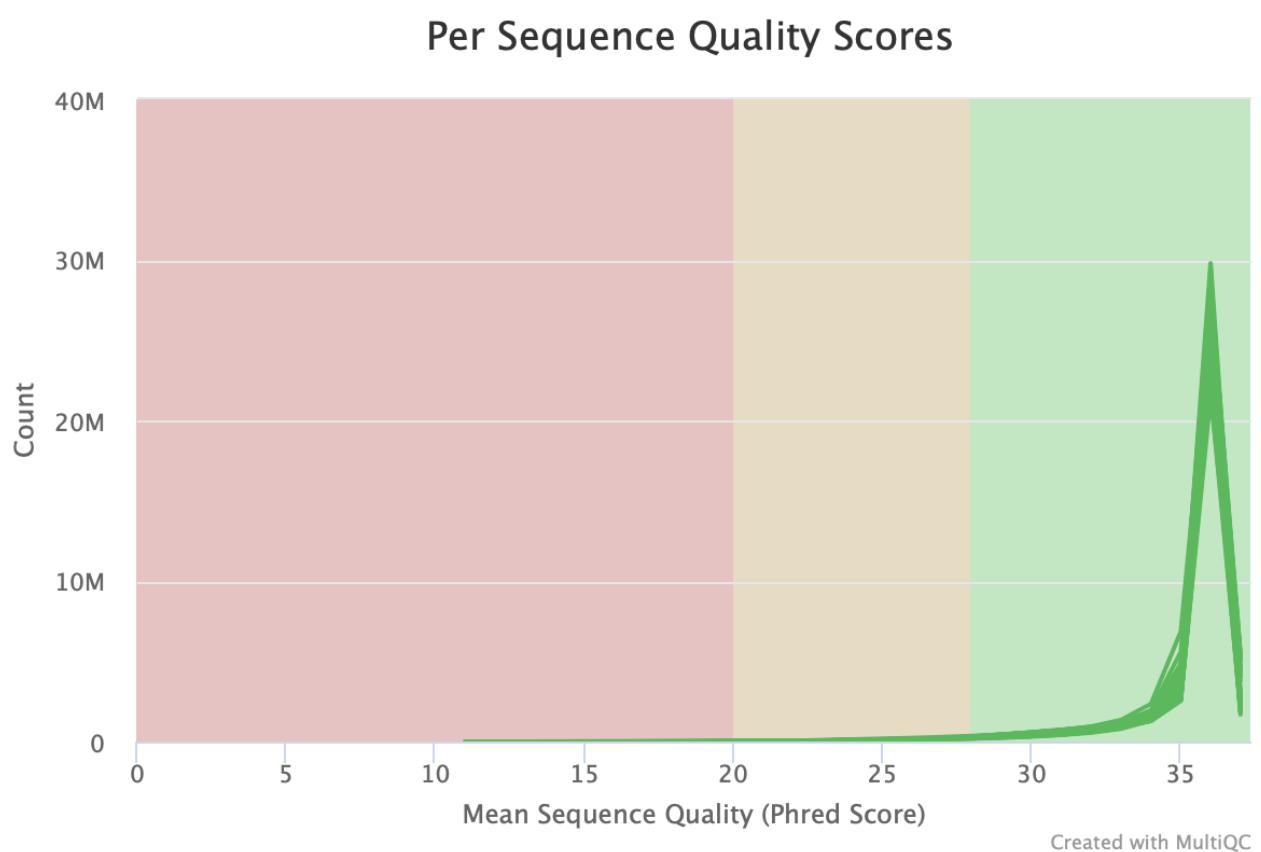


Fig. S3: Per Sequence Quality Scores plot in the multiqc report after trimming with trim-galore

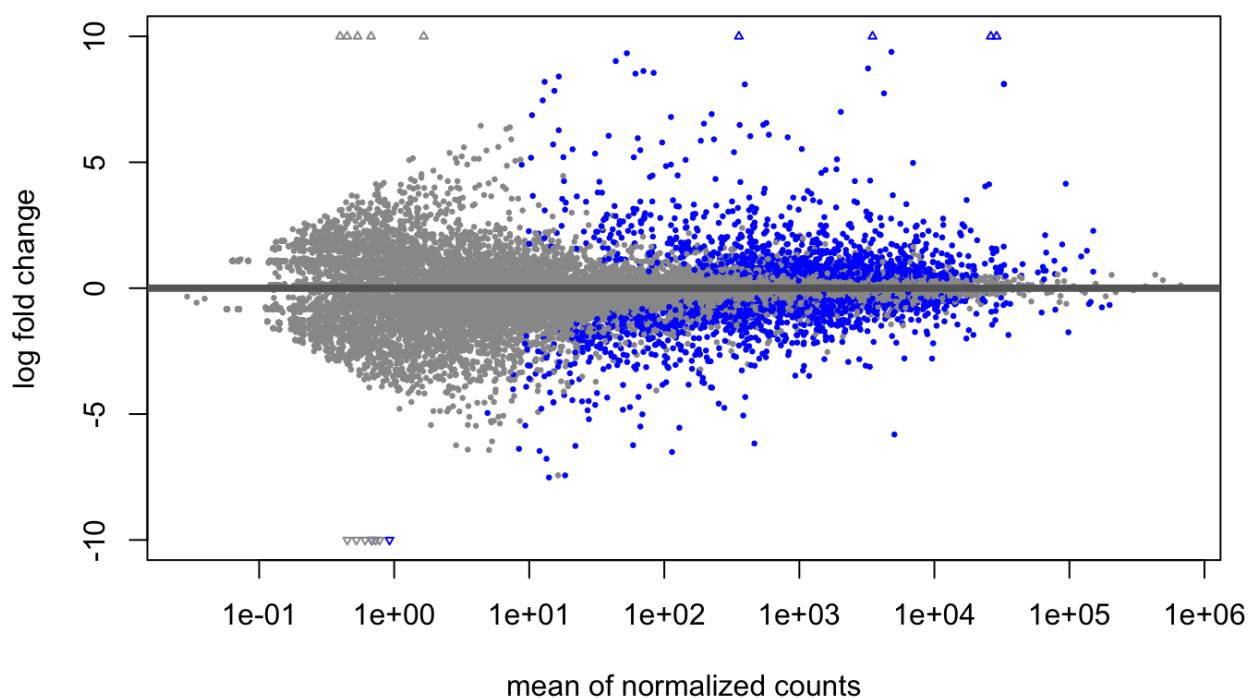


Fig. S4: Differential gene expression data between Healthy Cytokine and Healthy Control groups before the application of shrinkage estimator, apeglm

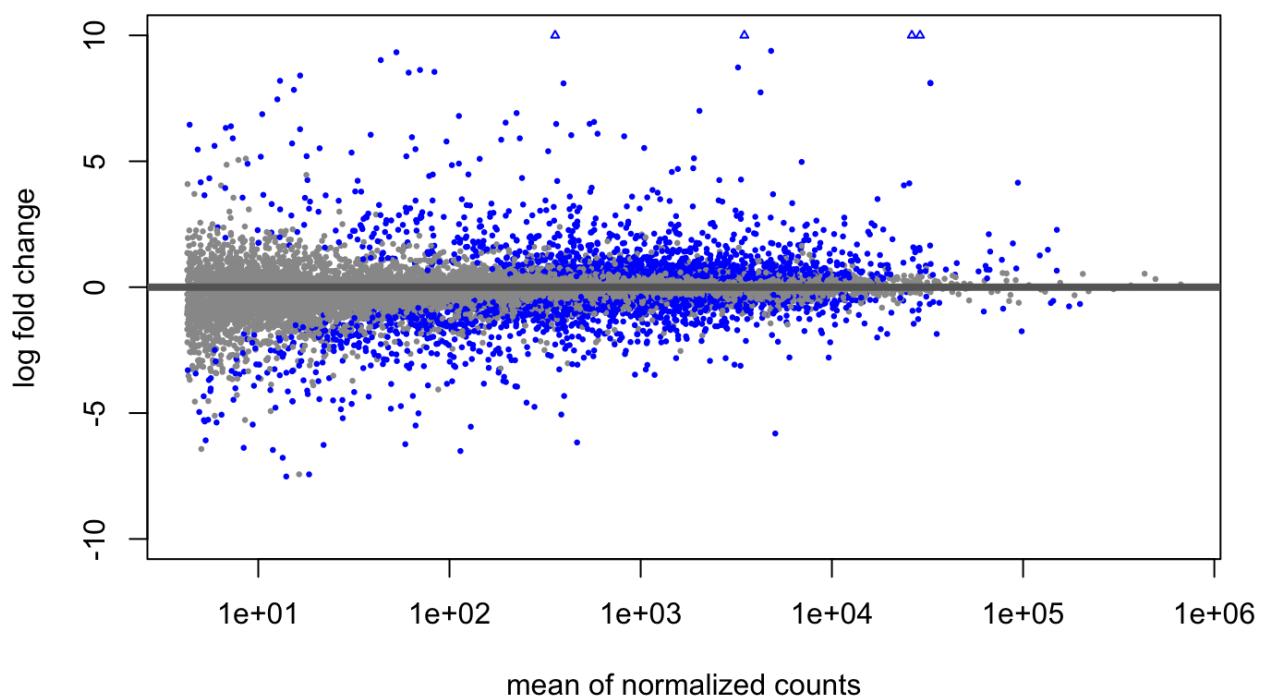


Fig. S5: Differential gene expression data between Healthy Cytokine and Healthy Control groups after the application of shrinkage estimator, apeglm

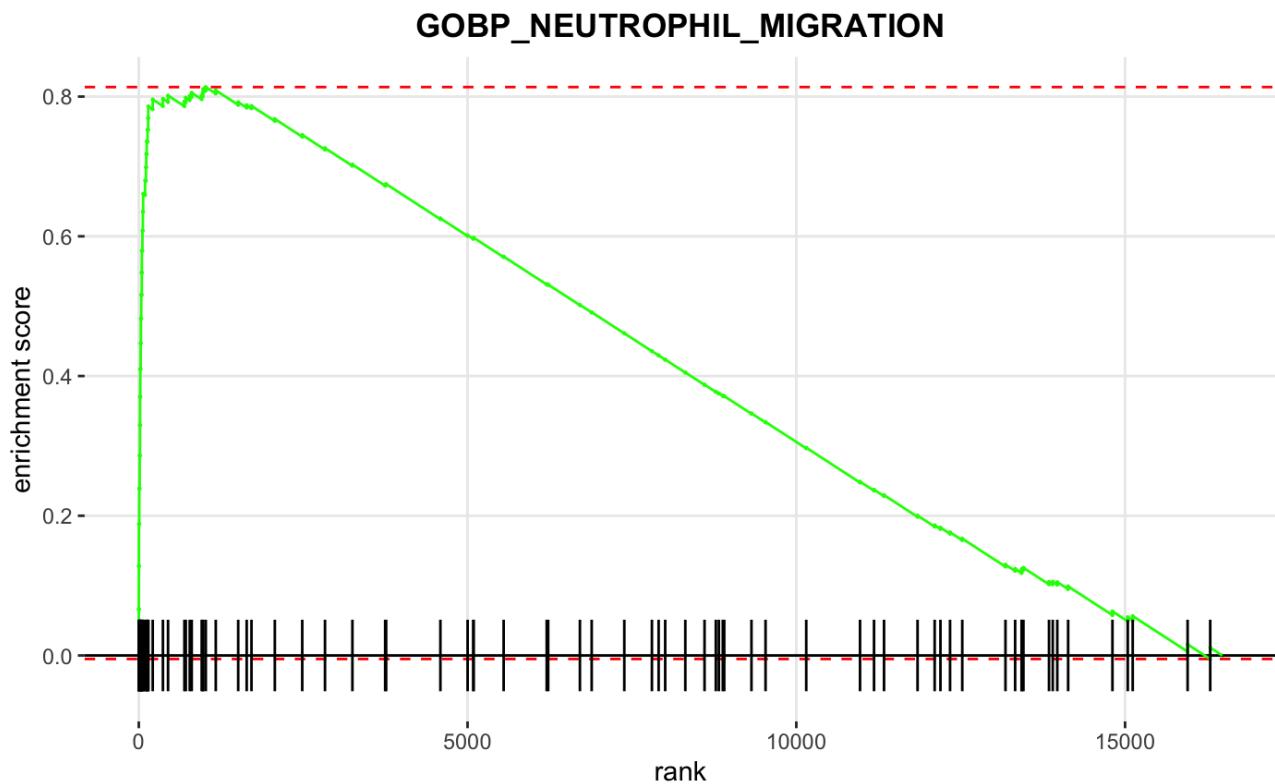


Fig. S6: Enrichment plot of GOBP: Neutrophil migration pathway in Healthy Cytokine group, compared with Healthy control group

GOBP_NEUTROPHIL_MIGRATION

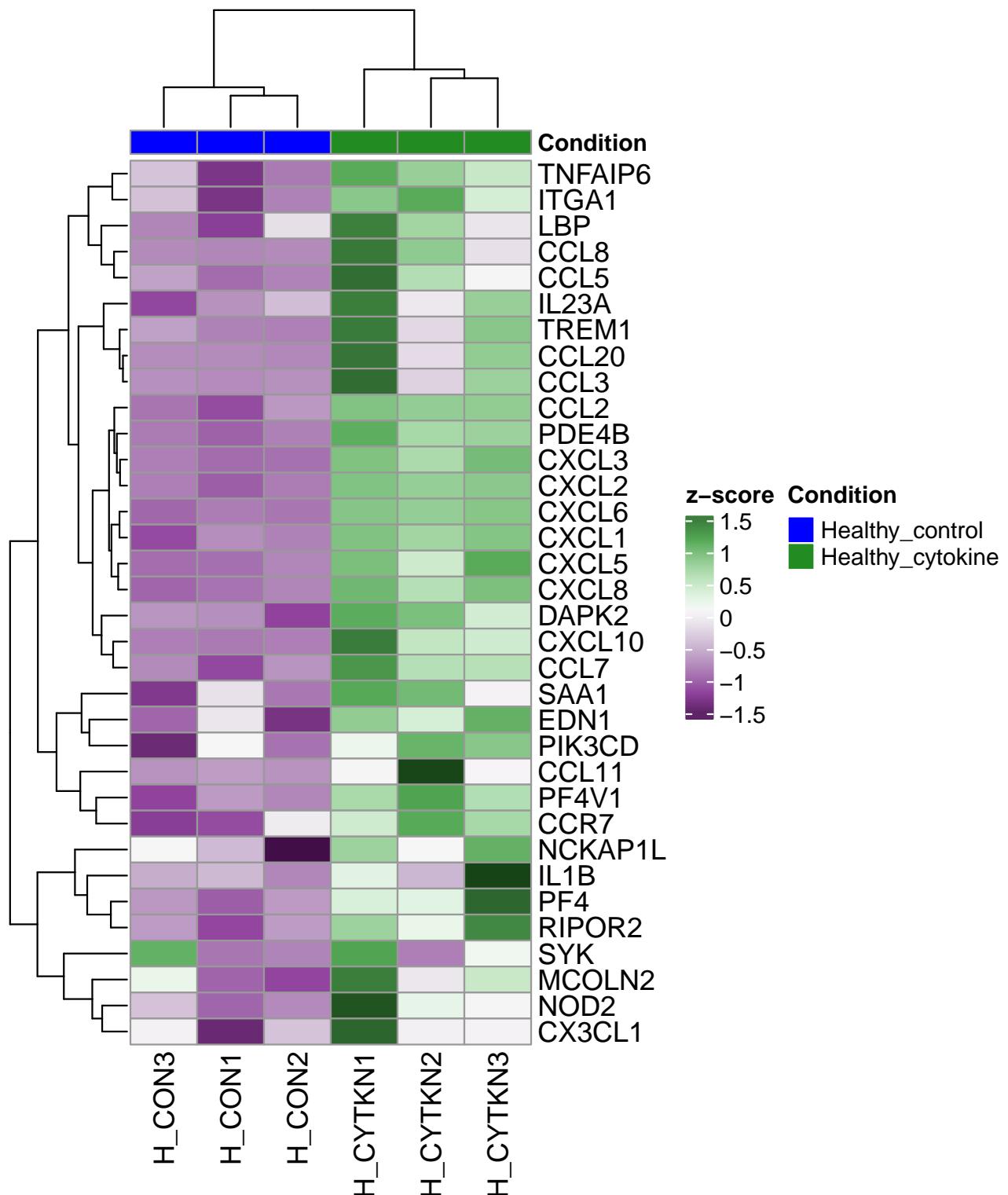


Fig. S7: Heatmap of differentially expressed genes between Healthy Cytokine and Healthy Control in specific GOBP: Neutrophil Migration pathway

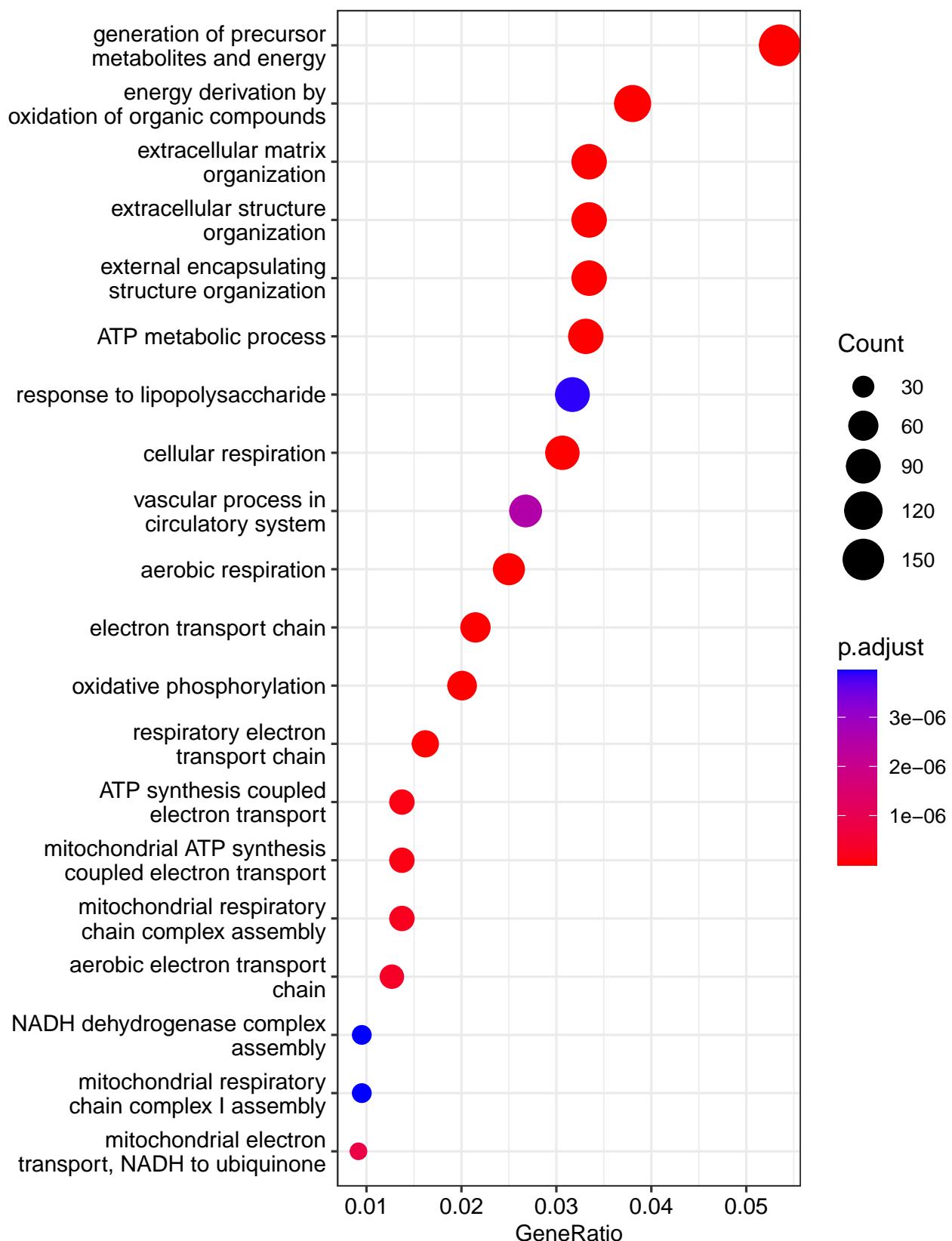


Fig. S8: Most functionally enriched Gene Ontology terms in Healthy Cytokine group via clusterProfiler package

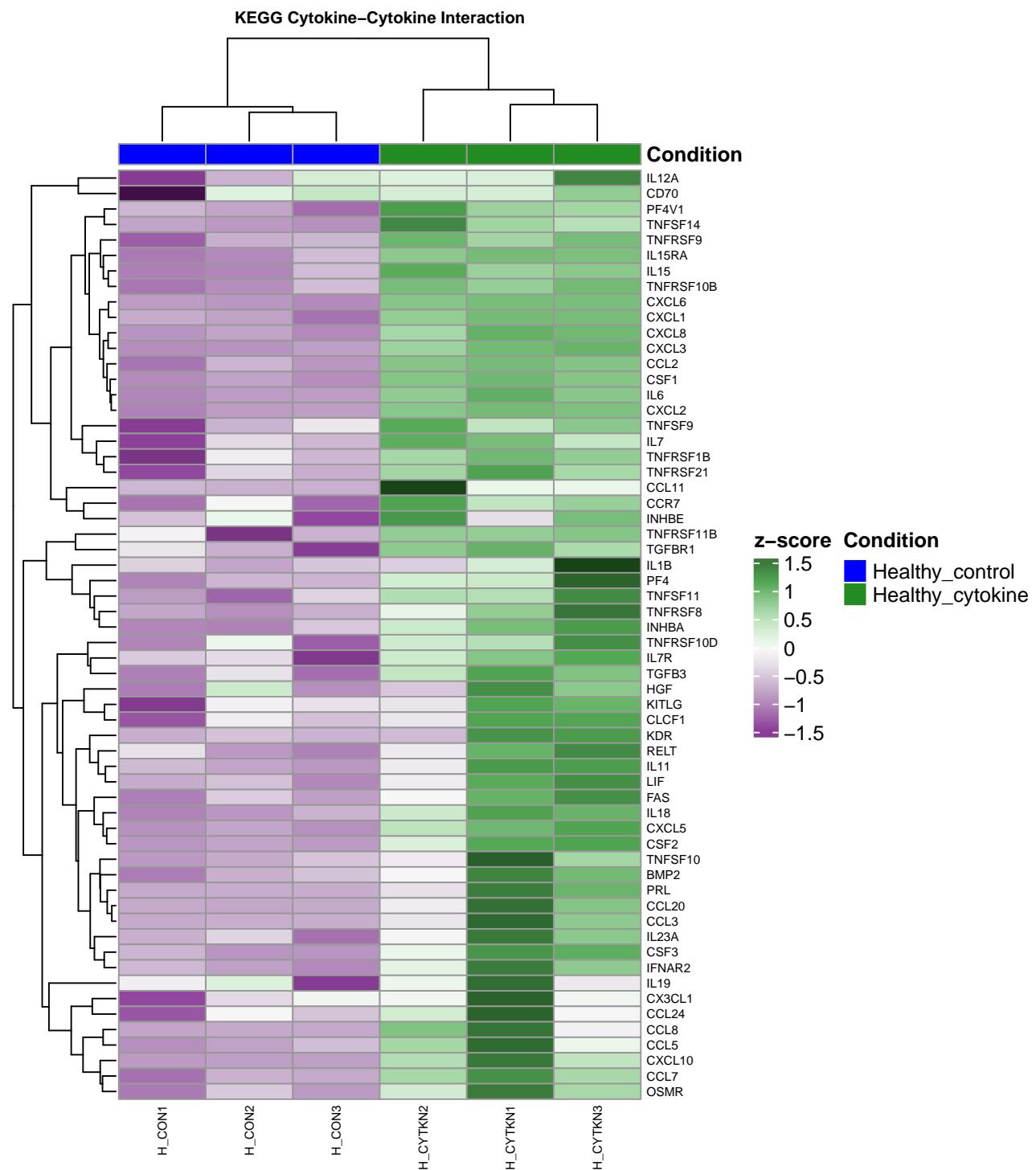


Fig. S9: Heatmap of differentially expressed genes between Healthy Cytokine and Healthy Control in specific KEGG Cytokine-Cytokine Interaction pathway

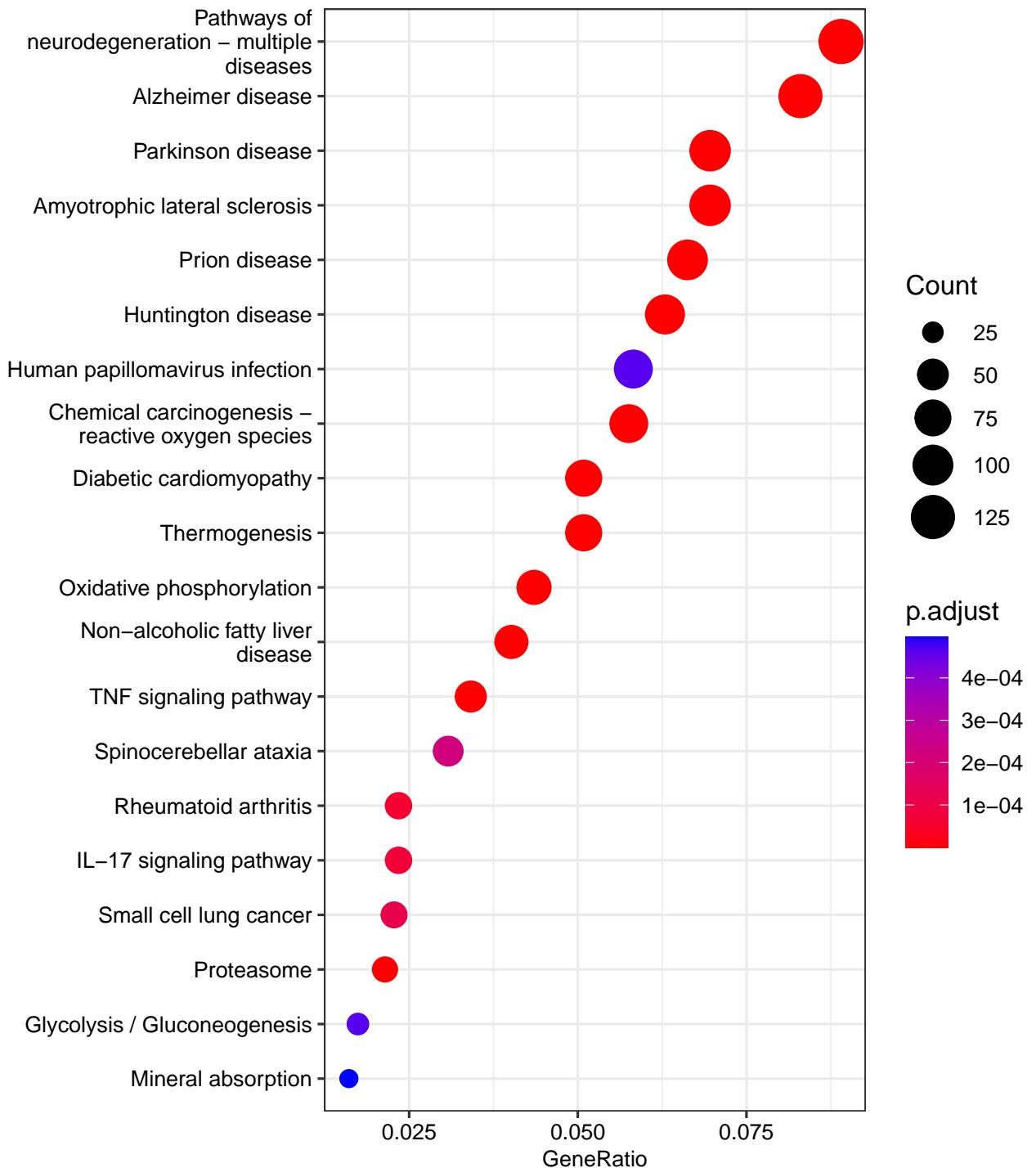


Fig. S10: Most functionally enriched KEGG terms in Healthy Cytokine group via clusterProfiler package

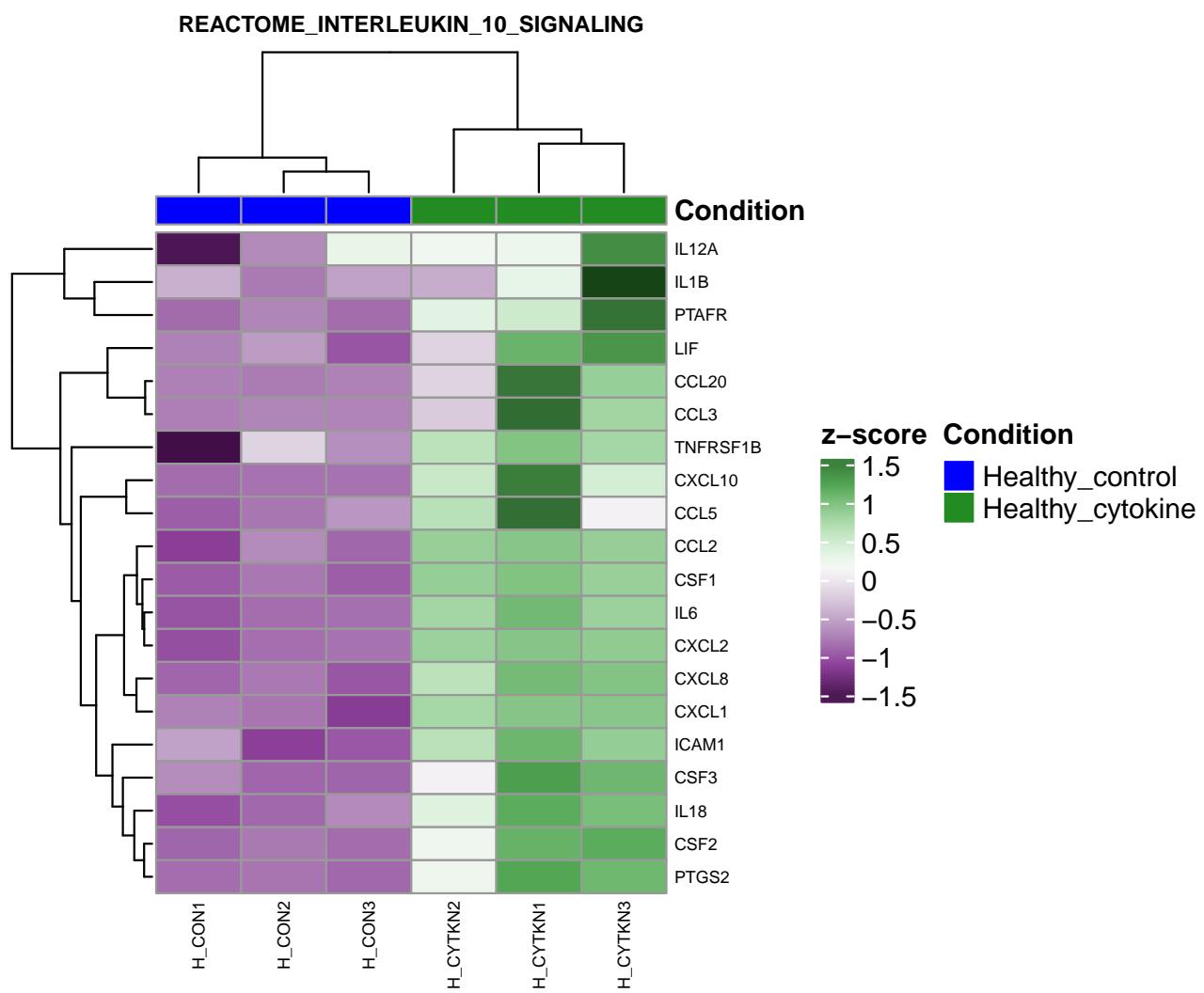


Fig. S11: Heatmap of differentially expressed genes between Healthy Cytokine and Healthy Control in specific Reactome Interleukin 10 Signalling pathway

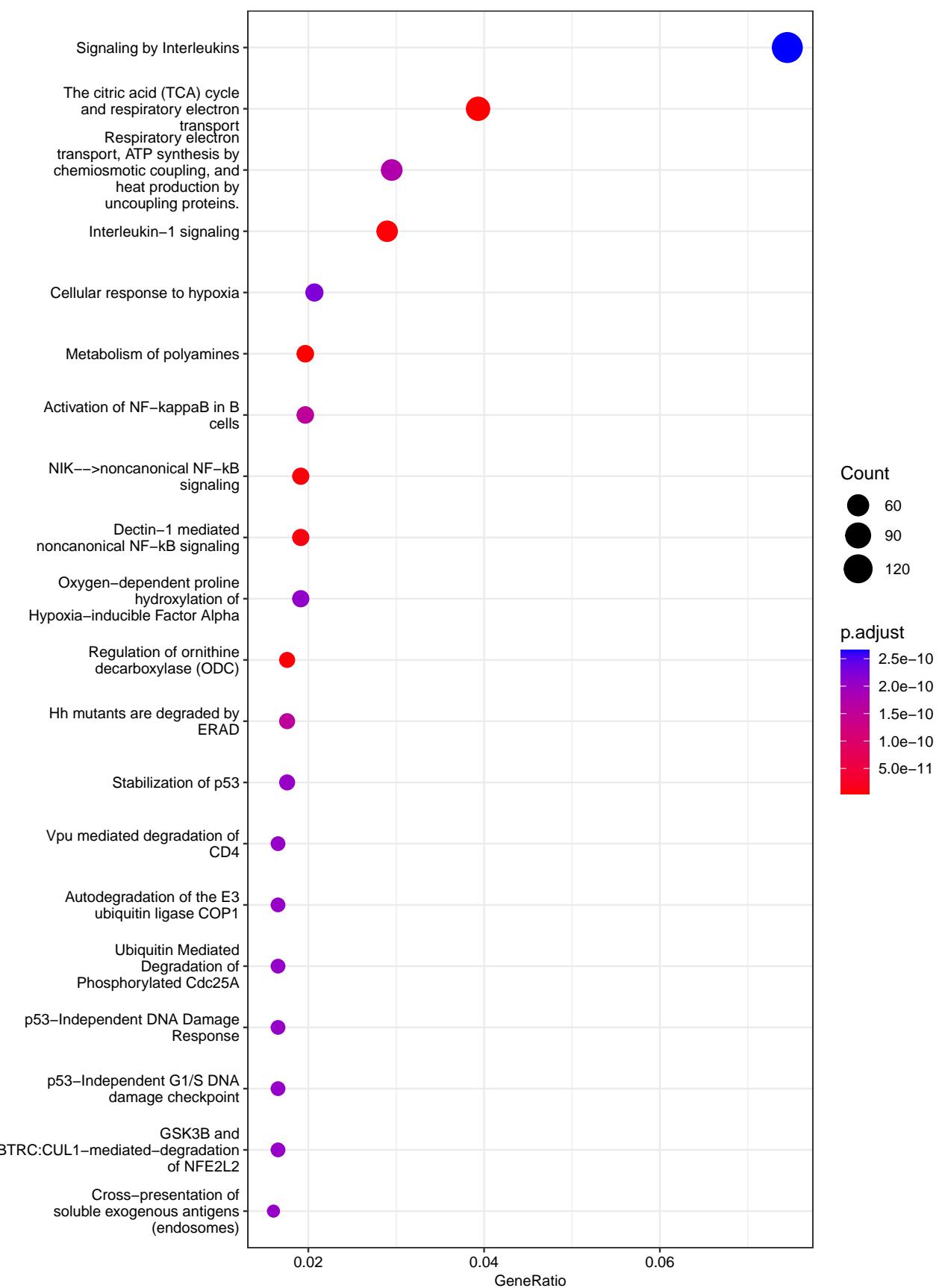


Fig. S12: Most functionally enriched Reactome terms in Healthy Cytokine group via clusterProfiler package

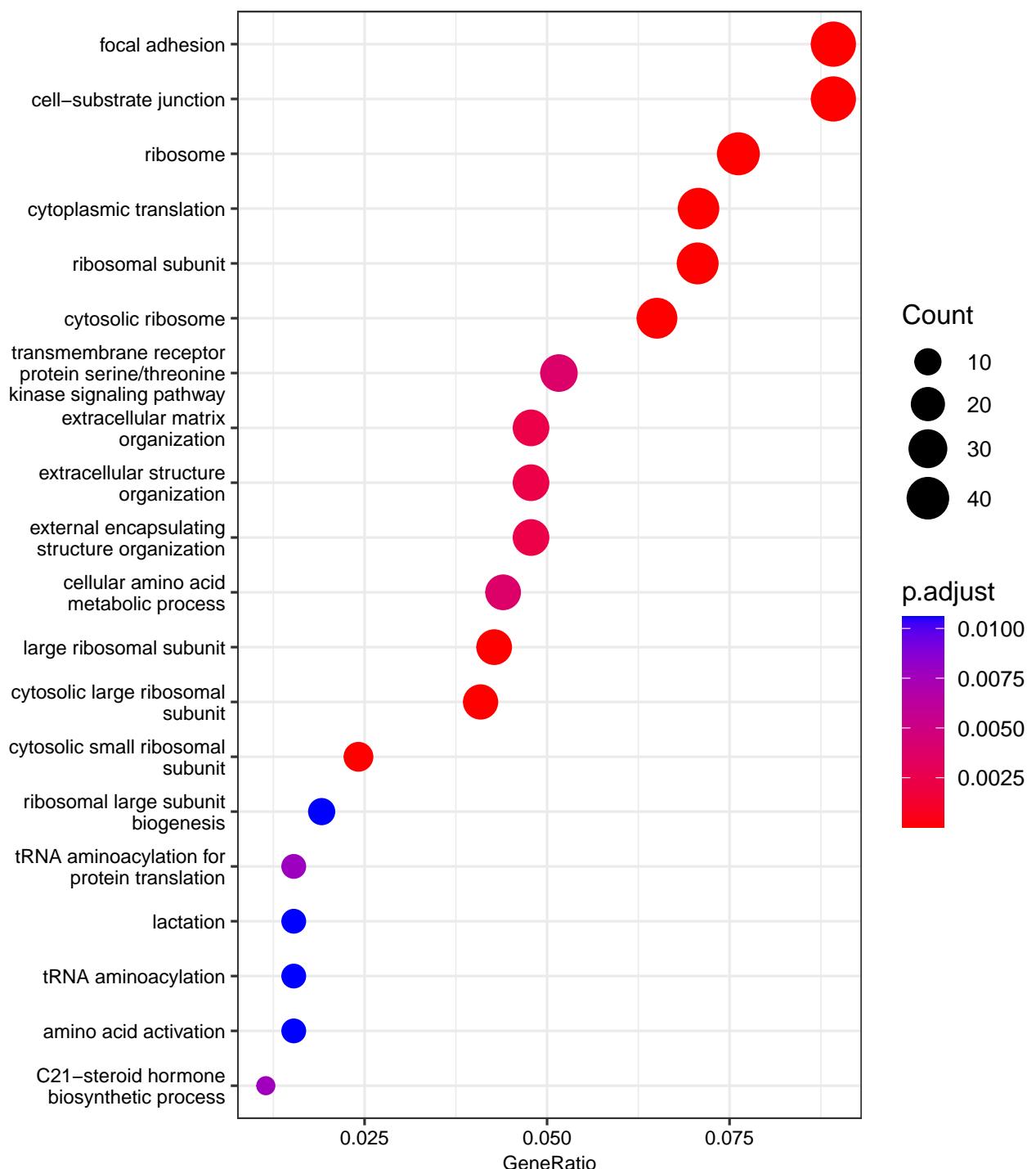


Fig. S13: Most functionally enriched Gene Ontology terms in Healthy Treated group via clusterProfiler package

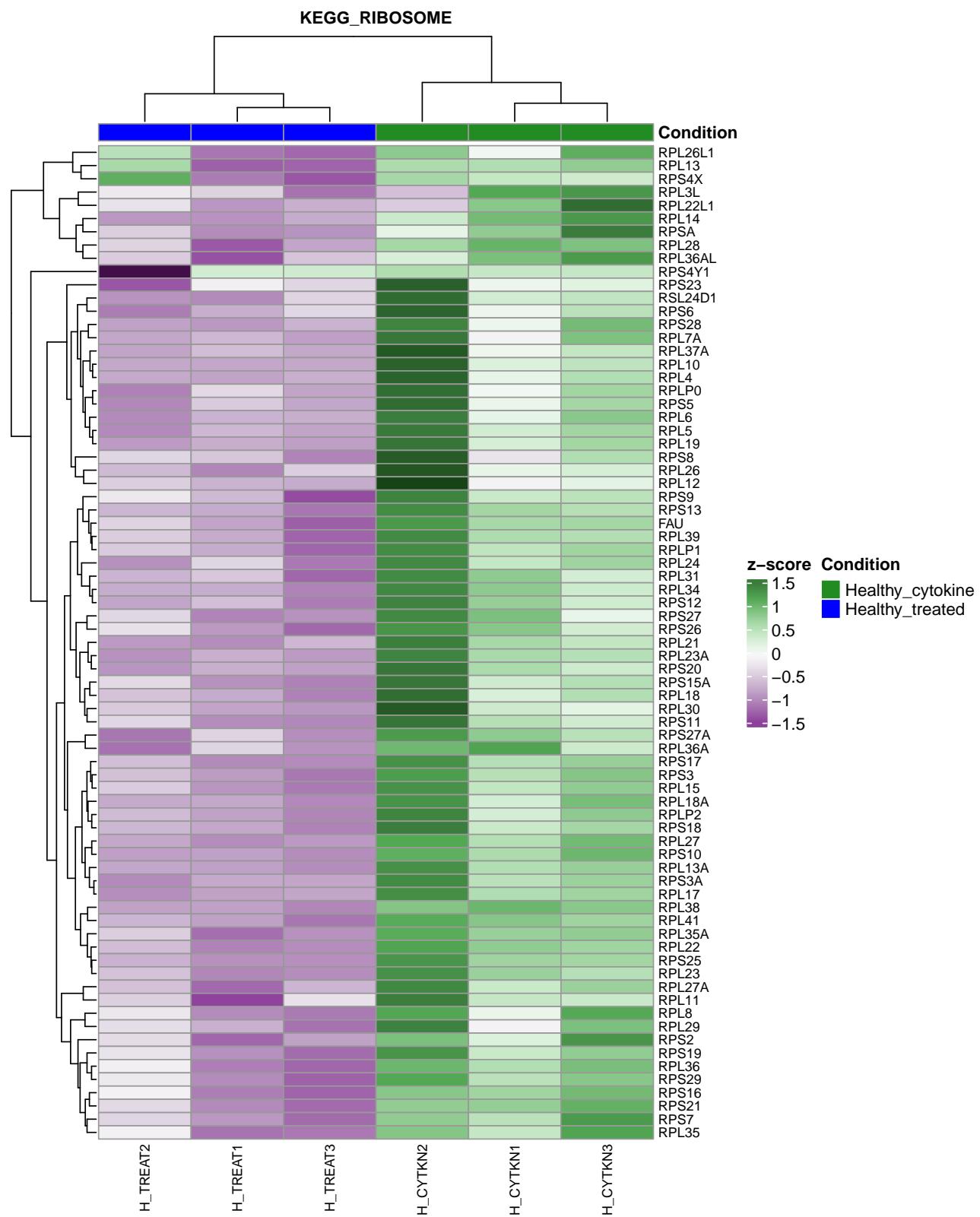


Fig. S14: Heatmap of differentially expressed genes between Healthy Treated and Healthy Cytokine in specific KEGG Ribosome pathway

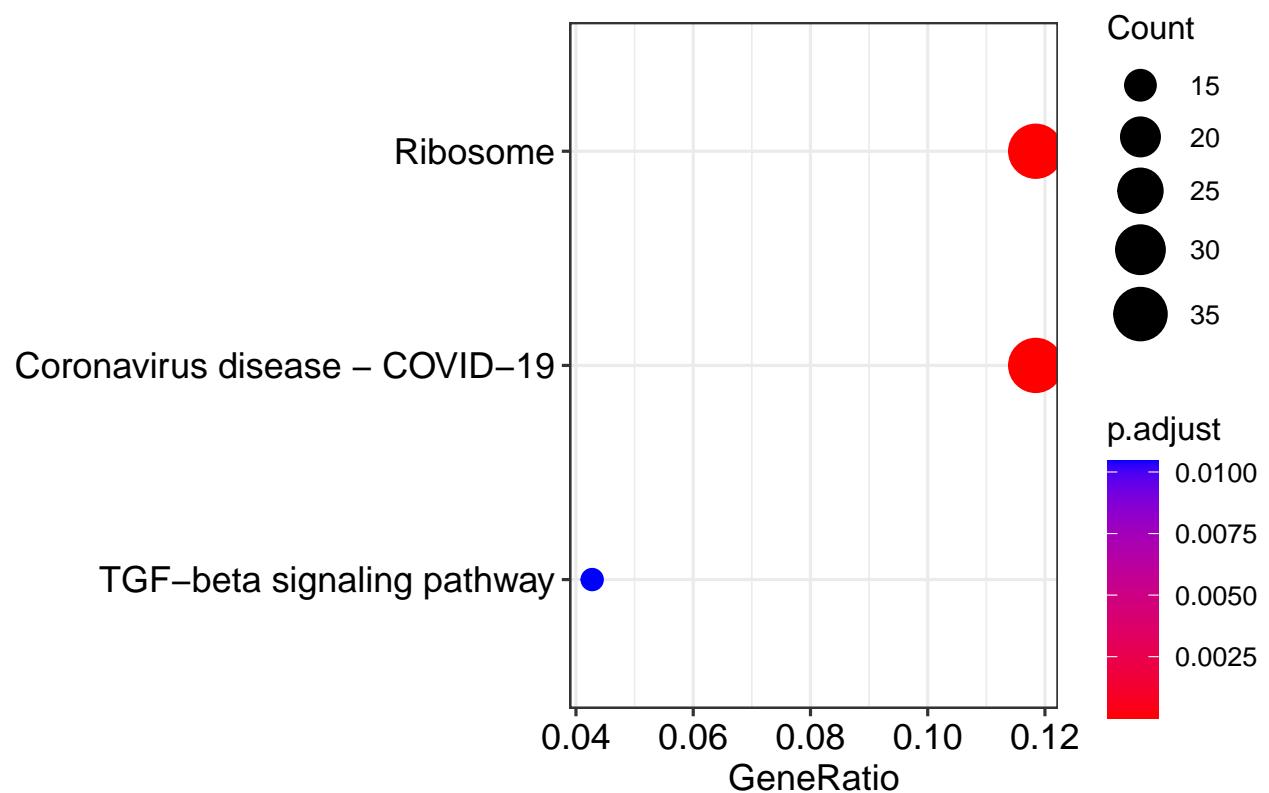


Fig. S15: Most functionally enriched KEGG terms in Healthy Treated group via clusterProfiler package

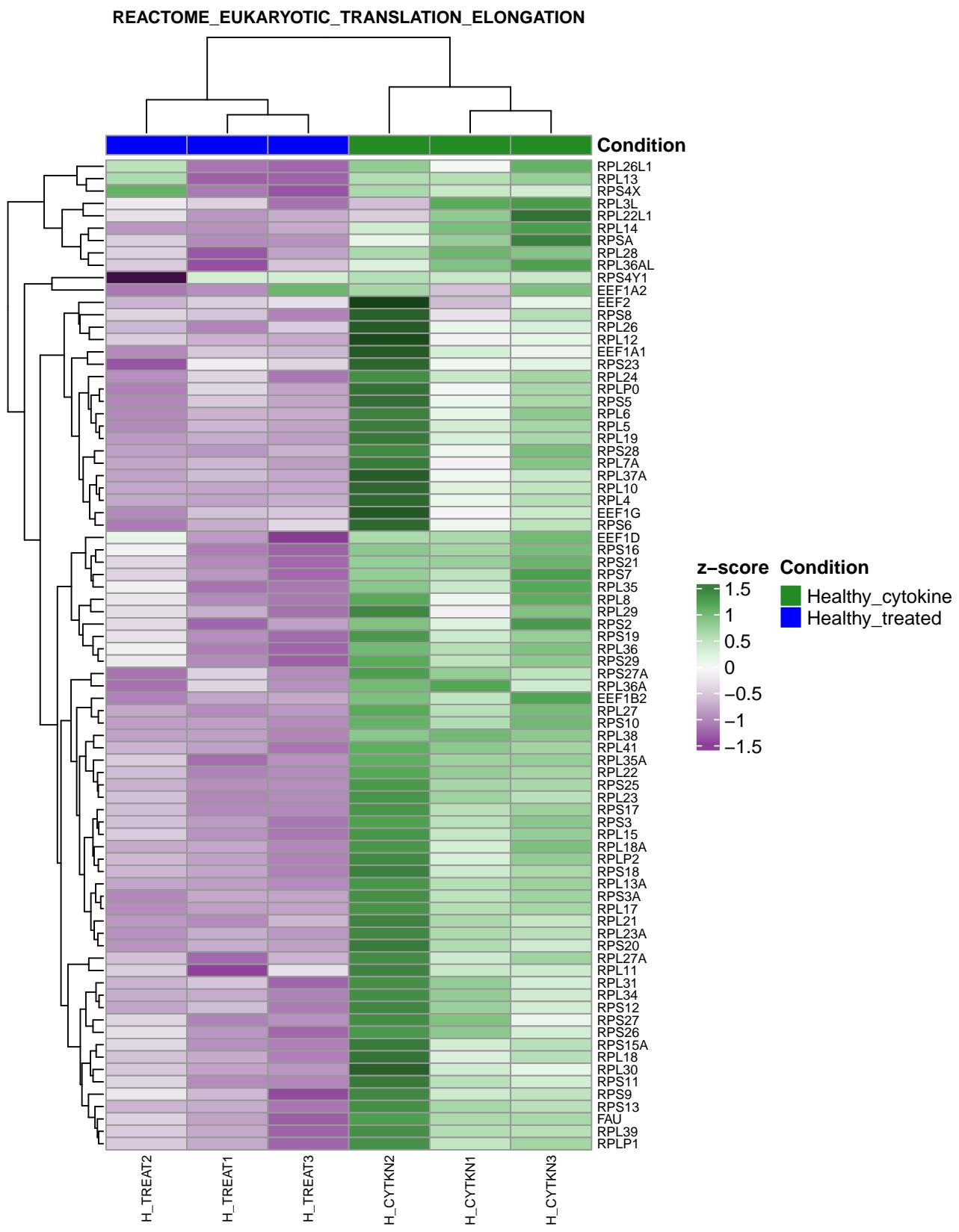


Fig. S16: Heatmap of differentially expressed genes between Healthy Treated and Healthy Cyotkine in specific Reactome Eukaryotic Translation Elongation pathway

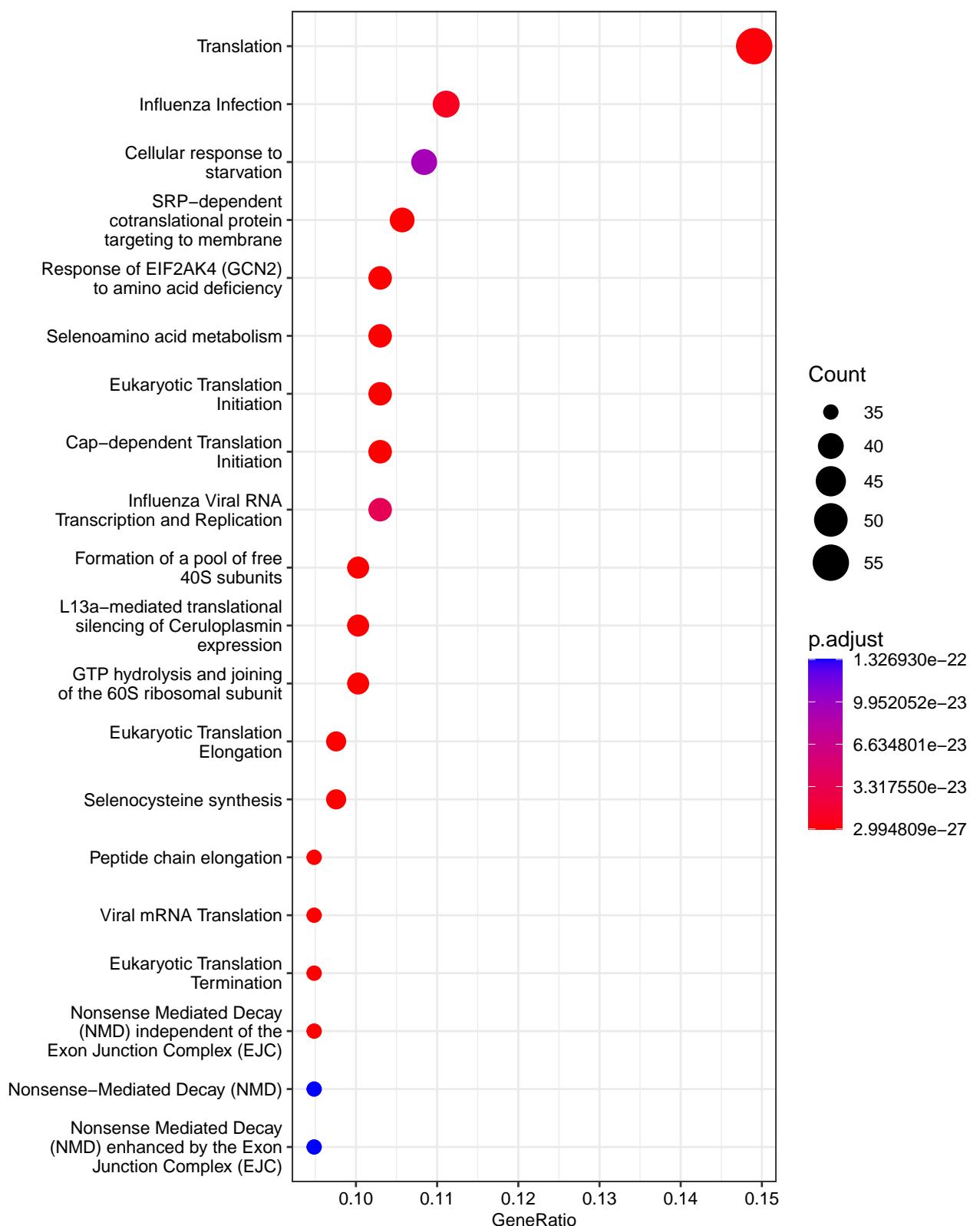


Fig. S17: Most functionally enriched Reactome terms in Healthy Treated group via clusterProfiler package

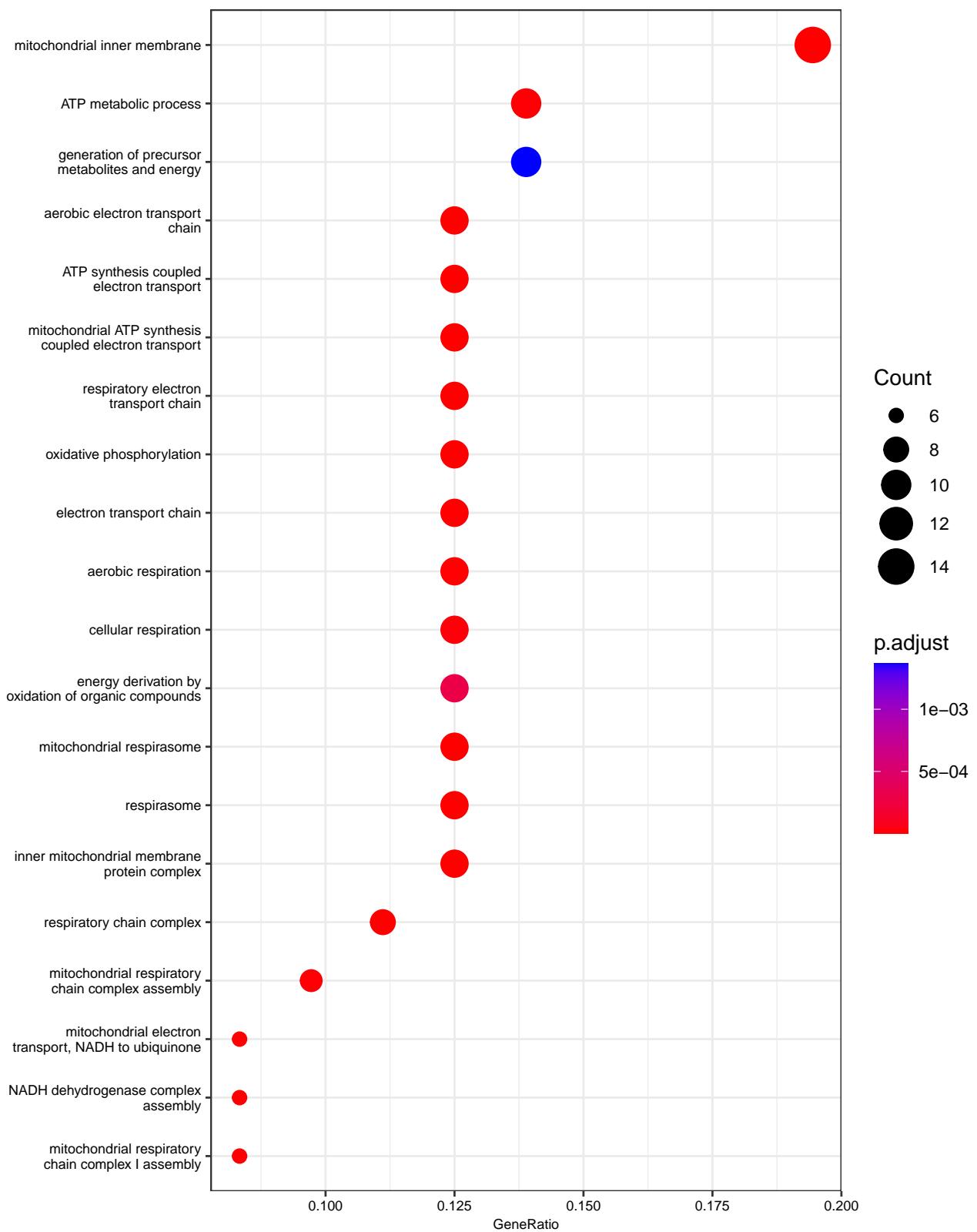


Fig. S18: Most functionally enriched Gene Ontology terms in Degenerated Treated group via clusterProfiler package

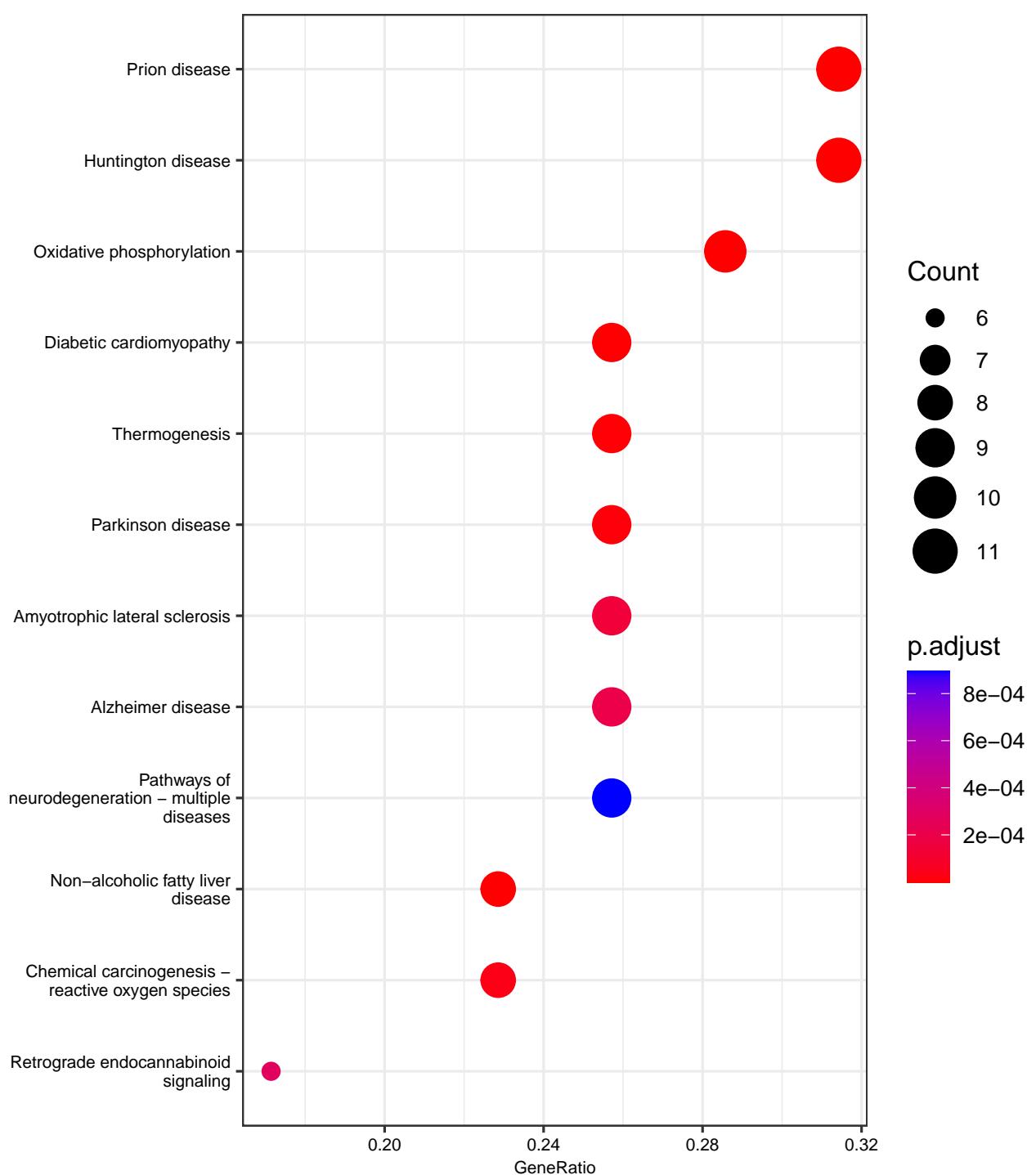


Fig. S19: Most functionally enriched KEGG terms in Degenerated Treated group via clusterProfiler package

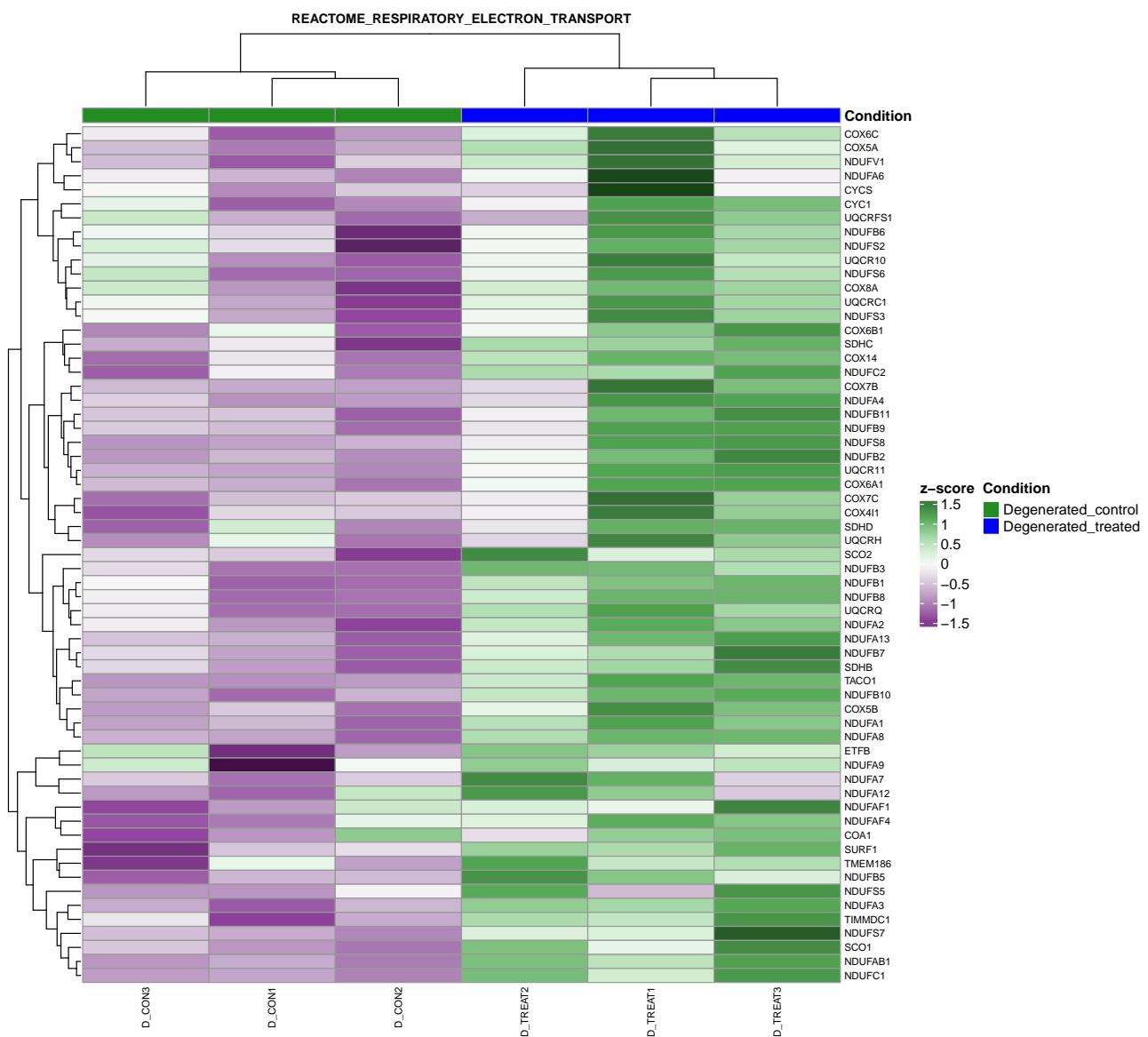


Fig. S20: Heatmap of differentially expressed genes between Degenerated Treated and Degenerated Control in specific Reactome Respiratory Electron Transport pathway

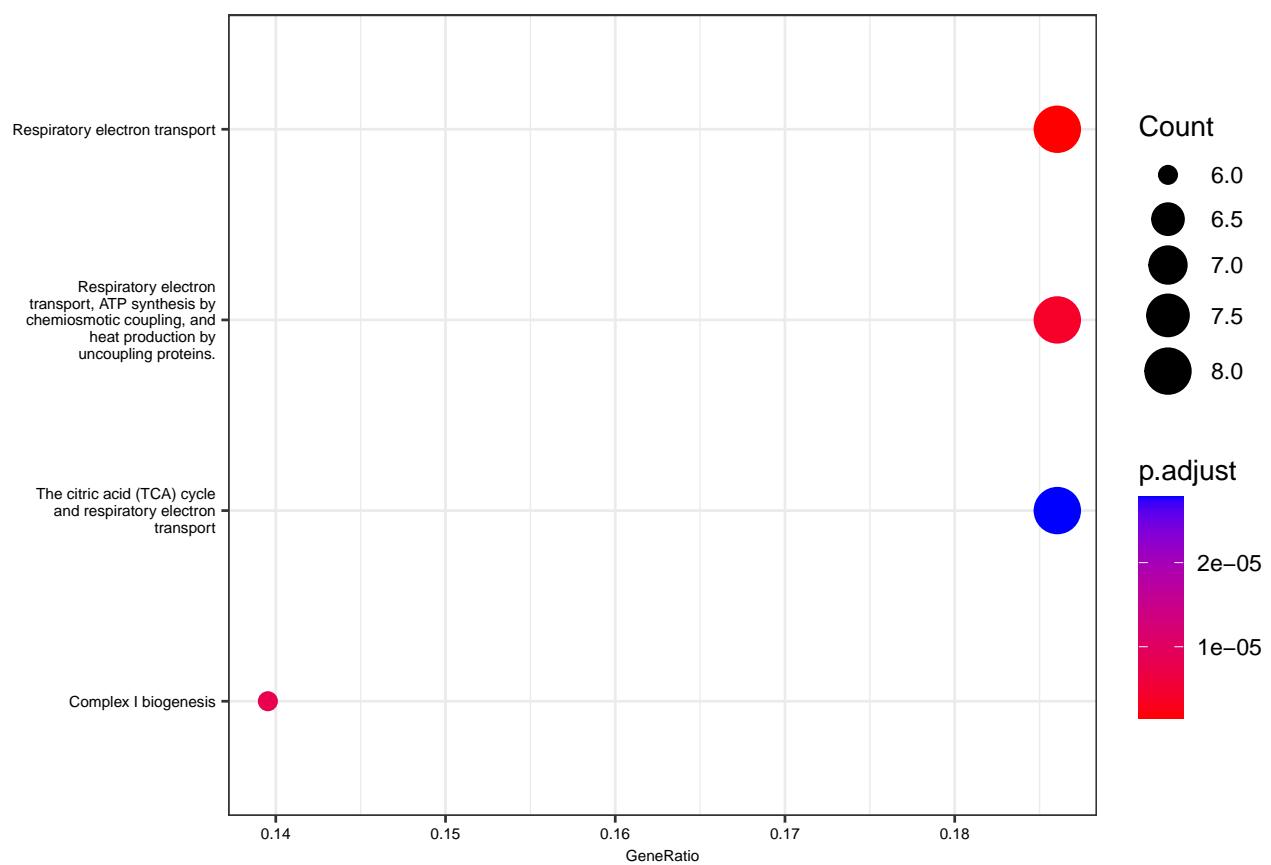


Fig. S21: Most functionally enriched Reactome terms in Degenerated Treated group via clusterProfiler package