Noise Induced Hearing Loss and Gene Expression Analysis using the R language

Introduction

Exposure to loud noise can cause hearing sensitivity or hearing loss, the intensity and duration of the exposure determines whether the hearing loss is temporary or permanent. Despite these clear physiological definitions, what's lacking is a comprehensive understanding of the complex molecular mechanisms responsible for noise-induced threshold elevation and the processes responsible for hearing threshold recovery.

The aim of the chosen research [1] was to obtain a deep understanding of the mechanisms contributing to NIHL, and in particular cochlear synaptopathy, since this could accelerate the ability to develop strategies for prevention and treatment. Multiple experiments were performed using the cochlea of mice that were exposed to different sound pressure levels (SPL), e.g. performing hearing tests to measure the recovery over time, comparing the cochlear proteome to spot changes in proteins, physical inspections and more. The cochlea is the part of the inner ear that converts mechanical vibrations into nerve impulses and transmits them to the brain, it houses a lot of the important central units contributing to hearing.

The goal of this project is to try to replicate one of the experiments that was done as part of the chosen research; the RNA-sequencing analysis of gene expression in cochlea.

Materials and Methods

60-day old awake mice (Mus Musculus) were exposed to 70, 94 and 105dB SPL for 30 minutes, immediately after noise exposure the cochlea were removed, RNA was harvested with Trizol reagent and the RNA libraries were sequenced with the Illumina NovaSeq 6000. The resulting count data has 3 groups each with 4 replicates and is publicly available as a single file in the NCBI Gene Expression Omnibus [2].

To perform the statistical analysis and find differentially expressed genes (DEGs), the R programming language was used in RStudio to create an R Markdown file. First the quality and structure of the data was assessed by plotting a few key visuals, after which was concluded that all samples were sufficient and nothing had to be directly done to them. Genes with low counts were filtered and using the Bioconductor package edgeR the actual analysis was performed.

The data was normalized using edgeR's TMM (Trimmed Mean of M-values) normalization method, resulting P-values were adjusted using Benjamini and Hochberg's approach. Genes with an adjusted P value < 0.05 and fold change > 1.5 found by edgeR were assigned as differentially expressed genes.

Results

Using the edgeR library, comparisons between sample groups and statistical tests were done, resulting in data sets with o.a. p-values and fold change (FC) values for the genes. The resulting data sets were visualized in volcano plots, highlighting the genes considered DEGs by implementing the p-value and FC thresholds (Figure 1).
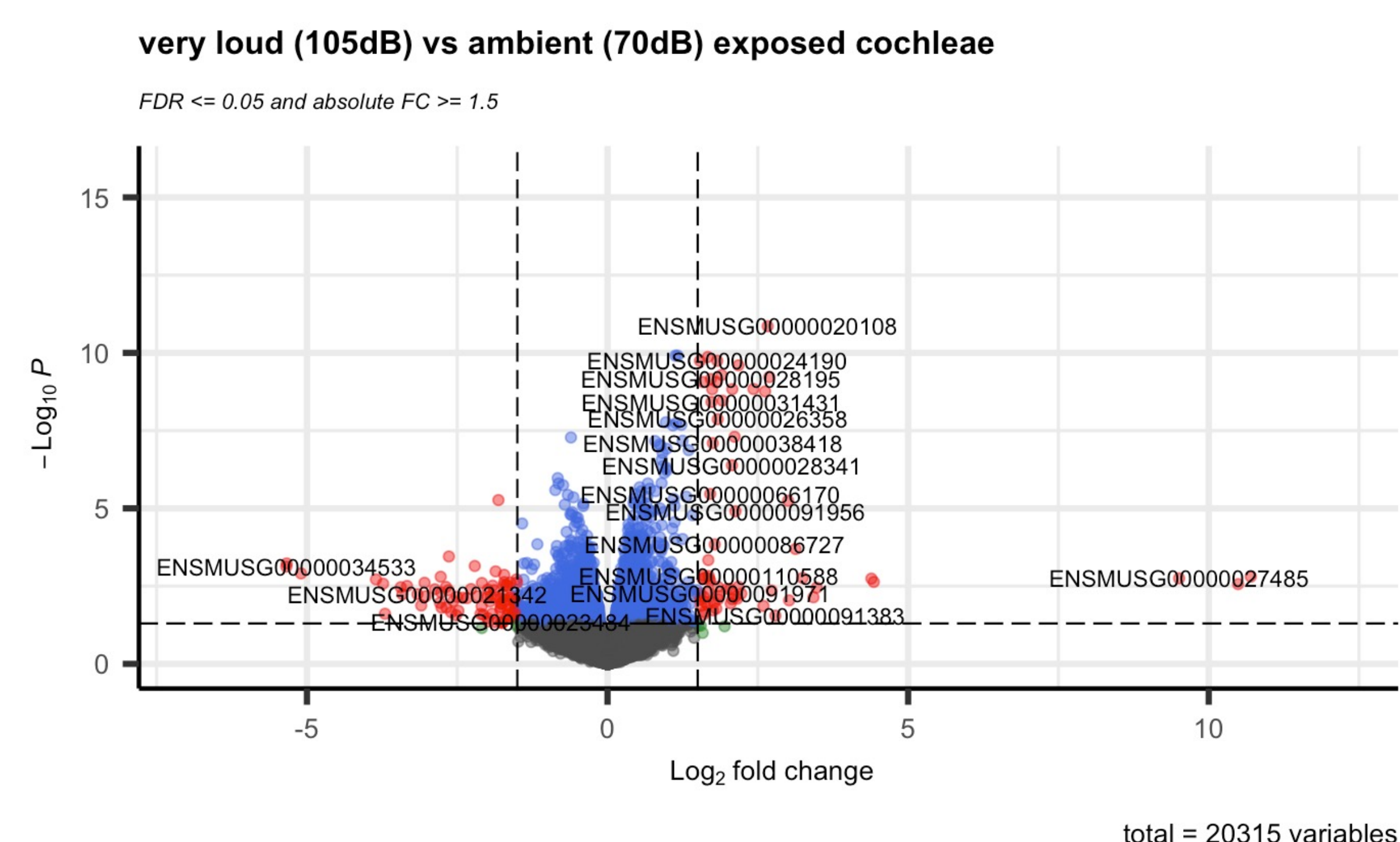
Figure 1: Volcano Plot plotting all genes with their $-\log_{10}$ p-value against $\log_2$ fold change.
Blue dots are statistically significant but don't have a high enough fold change. Gray dots have neither significance nor a high enough fold change. Red dots have both significance and a high enough fold change and are thus considered DEGs.

To show the amount of genes that are considered differentially expressed in both comparisons against the control group, a Venn diagram was made (Figure 2).
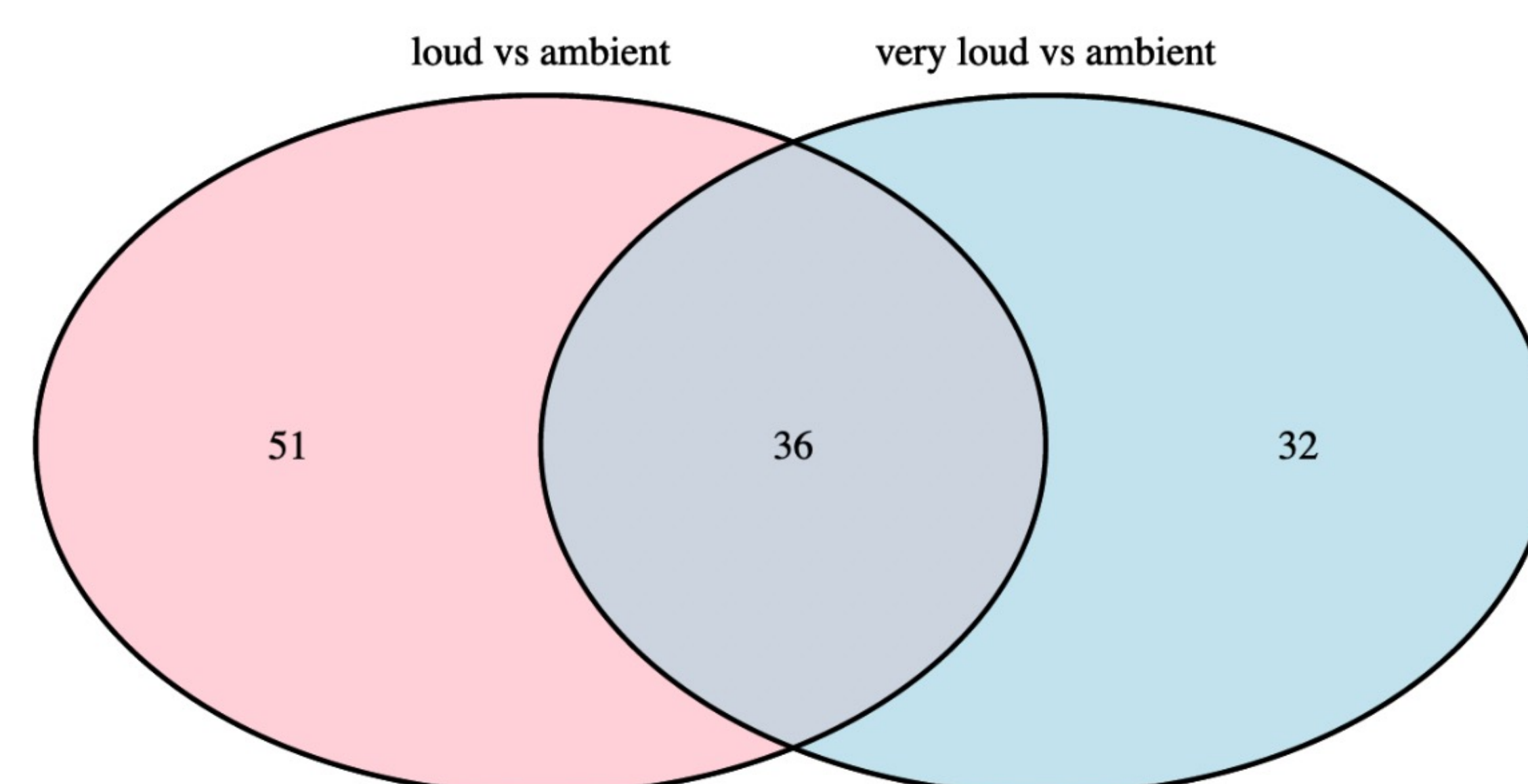
Figure 2: Venn Diagram showing the amount of shared DEGs between group comparisons.

In the end a total of 119 DEGs were found, of which 36 are shared between the comparisons and thus these are the most important.

Conclusion / Discussion

The majority of DEGs that were found are up-regulated, this was sort of surprising since it's hearing loss, and you'd think that there is mainly down-regulation because of a loss of function. This could possibly be explaining there are a lot of genes that were expressed more to try and fix damage caused by the noise insults.

When performing Multi-Dimensional scaling replicates clustered nicely but the very loud group was found to be more similar to the ambient group. This is also in line with that there were more DEGs found in the loud group (87) than in the very loud group (68).

Something extra that could have been done is to perform the analysis using different normalization methods or packages, this could lead to different results or strengthen the already achieved results. The next step for after this experiment would be to look at the discovered DEGs in more detail and find out what they do exactly.

Hanzehogeschool Groningen
University of Applied Sciences

Vincent Talen
Bioinformatics – BFV2
Institute of Life Sciene & Tech
v.k.talen@st.hanze.nl

References

[1] Nopporn Jongkamonwiwat, Miguel A. Ramirez, Seby Edassery, Ann C.Y. Wong, Jintao Yu, Tirzah Abbott, Kwang Pak, Allen F. Ryan, and Jeffrey N. Savas, *"Noise Exposures Causing Hearing Loss Generate Proteotoxic Stress and Activate the Proteostasis Network"*, Nov-2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7722268/. [Accessed: 11-Mar-2022].

[2] "NCBI *GEO Accession viewer"*, Nov-2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7722268/. [Accessed: 11-Mar-2022].