Molecular Subtypes of Amyotrophic Lateral Sclerosis: A Gene Expression Analysis

Caitlyn Wong

Abstract

Amyotrophic Lateral Sclerosis (ALS) is a fatal neurodegenerative disorder characterized by progressive paralysis, with most patients surviving only 2-5 years after diagnosis. While the exact cause is unknown, TDP-43 protein dysfunction and sporadic gene mutations are common. Recently, researchers have categorized ALS into three distinct molecular subtypes: ALS-Glia (microglial activation and neuroinflammation), ALS-TE (TDP-43 pathology and transposable element silencing), and ALS-Ox (mitochondrial stress and oxidative phosphorylation). Using a dataset of post-mortem samples of control and ALS patients from the NYGC ALS Consortium, this study analyzed gene expression across these three subtypes. We identified subtype-specific genes linked to metabolic and immune pathways using differential expression and biological pathway analysis. These insights contribute to a deeper understanding of the molecular basis of ALS and can ultimately inform the development of more targeted, individualized therapeutic approaches.

Introduction

Amyotrophic Lateral Sclerosis (ALS) is a neurodegenerative disease characterized by the breakdown of motor neurons, with a complex combination of genetic and environmental factors contributing to its onset and progression. The disease begins by affecting the upper and lower motor neurons, which control voluntary and involuntary motion. This manifests as clinical symptoms, often beginning with weakness in one limb. Eventually, over a few weeks to a few months, weakness will occur in another limb. The disease will then progress until the patient experiences complete paralysis of voluntary muscles. ALS patients often die within 2 to 5 years after showing early symptoms, but many can live for up to 15 years after. A majority of ALS patients have no familial history of the disease, however, some have familial ALS. Despite years of research since ALS's discovery, little is known about the causes of the disease, and current therapeutic measures remain limited.

Clinical features and gene expression are often incredibly heterogeneous among ALS patients. Genetic factors associated with ALS include 25 gene variants with strong evidence of association and more than 120 gene variants with weaker evidence. Key genetic mutations linked to ALS include repeat expansions in the C9orf72 gene and mutations in the SOD1 and TARDBP genes, although only 2% of patients display these (Oliver et al., 2019). In contrast to rare and poorly associated genetic mutations, cytoplasmic aggregates of mislocalized TDP-43 are a hallmark in nearly all ALS cases, seen in nearly 97% of ALS patients (Sar et al., 2024). The disease involves various mechanisms, such as disturbances in protein quality control, hyper-activated microglia, decreased energy supply from oligodendrocytes, glutamate excitotoxicity, RNA metabolism disturbances, and cytoskeletal defects (Swati Dhasmana et al., 2023). Additionally, dysfunction in signaling between motor neurons and glia, as well as mutations in genes related to autophagy and the endolysosomal system, are implicated in ALS (Silvia Giovedì et al., 2020). While the majority of ALS cases are sporadic with no known genetic

mutation or family history, large-scale sequencing studies continue to uncover new genetic associations.

A recent advancement in our knowledge of the disease, though, has been the definition of three molecular subtypes of ALS: ALS-Glia, ALS-TE, and ALS-Ox. A deep-layer ALS neural classifier was able to concretely assign subtypes to a large sample of patients. By identifying these subtypes, researchers can better study and understand the underlying gene expression and clinical features of each. The ALS-Glia subtype is categorized by microglial activation and neuroinflammation. ALS-TE samples show dense TDP-43 pathology and associated transposable element silencing. Finally, ALS-Ox displays signatures of mitochondrial stress and oxidative phosphorylation (Kathryn O'Neill et al., 2024). This new classification of ALS allows us to better understand the disease as well as the possibility of more targeted treatment methods. This study aims to investigate the gene expression patterns across ALS subtypes using RNA sequencing data to identify potential molecular pathways and therapeutic targets. In this paper, we will focus on breaking down gene expression within these subtypes, identifying differences and similarities across subtypes, allowing us to better understand the driving forces behind the clinical features of ALS.

Materials and Methods

Data Selection

The data used in this project was taken from the NYGC ALS Consortium. This data consists of publicly available RNA sequencing, post-mortem samples taken from healthy control brains, ALS patients, Frontal Temporal Dementia patients, and other neurological disorders. After looking through the superseries GSE137810, we decided to focus on the subseries GSE124439 because of its usage in a study on ALS subtypes as well as its lower sample number compared to other series.

Differential Expression Analysis

Our main method of analysis was the usage of differential expression analysis on several different groups found in the study. We did this using the DESeq2 package within RStudio. The package tests for differentially expressed genes through the use of negative binomial generalized linear models. When reviewing data and results, we limited the results to p-adjusted values less than 0.1 and $\log_2$Fold Changes greater than 2 or -2. Positive fold changes meant the genes were upregulated in one group compared to another, while negative fold changes meant they were downregulated.

Pathway Analysis

We used the g:Profiler [https://biit.cs.ut.ee/gprofiler/page/docs] on the genes that met the conditions set in the previous section. This showed which pathways held most of the up or down-regulated genes, allowing us to draw connections between clinical features and biological pathways.

Results

In these results, we have summarized the most prominent differentially expressed genes and biological pathways between each ALS subtype. These differences have been highlighted with volcano plots for differentially expressed genes and heat maps for biological pathways.

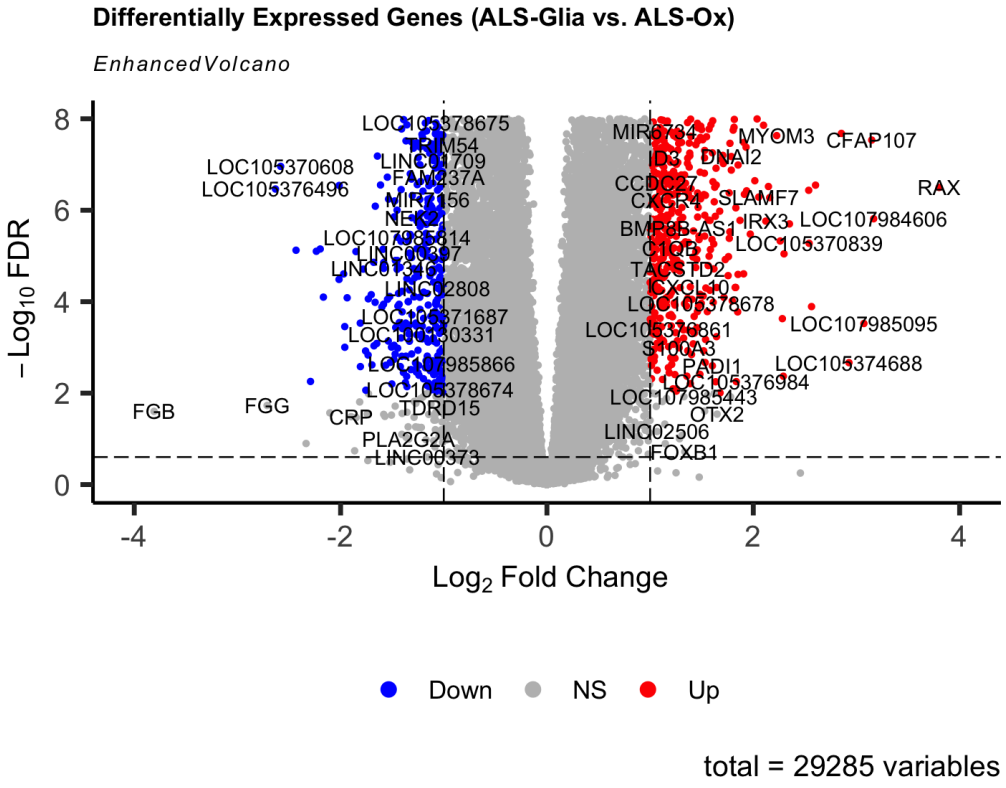


Figure 1.1. ALS-Glia vs ALS-Ox. A volcano plot of upregulated (red) and downregulated (blue) genes in ALS-Glia compared to ALS-Ox (reference) samples. These were determined by differences in Log_2Fold change of gene expression on the x-axis, as well as the p-adjusted values on the y-axis. For thresholds regarding determining regulations, see “Thresholds” in the Methods section. This plot visualizes which genes were used in the g:Profiler to obtain biological pathway information. NS = not significant.

In this comparison, we saw which genes in ALS-Glia were upregulated or downregulated compared to ALS-Ox (Figure 1.1). The genes upregulated in ALS-Glia compared to ALS-Ox were found to be linked with metabolic and catabolic processes, especially chitin (Figure 1.2). Commonly found in fungi, chitin is a fiber that has the ability to cause allergy-like immune responses. According to Dr. Brian Doctrow, “the immune response to chitin has benefits for metabolic health” (Do-Hyun Kim et al., 2023). Because ALS-Glia is linked to microglial activation, i.e., an immune response, it makes for an interesting connection to see upregulation in a metabolic process possibly related to immune response. However, this must be interpreted with the caveat that the correlation is not strong.

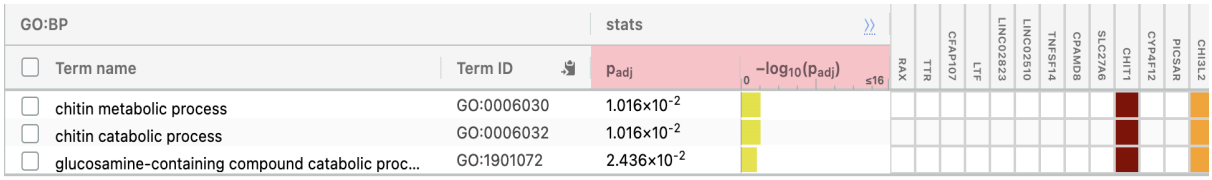


Figure 1.2 A heatmap of upregulated genes in ALS-Glia compared to ALS-Ox, demonstrating which biological pathways these genes are linked to through the use of g:Profiler.

Along with upregulated genes, we also explored genes downregulated in ALS-Glia, which refers to the genes more prevalent in ALS-Ox compared to ALS-Glia. ALS-Ox is associated with mitochondrial dysfunction and oxidative stress. The genes more prominent in ALS-Ox are related to many functions (See Figure 1.3). Here, we considered processes with an adjusted p-value less than 0.01. A summary of the following processes is a combination of blood coagulation and regulatory processes of transport and release of signaling molecules.

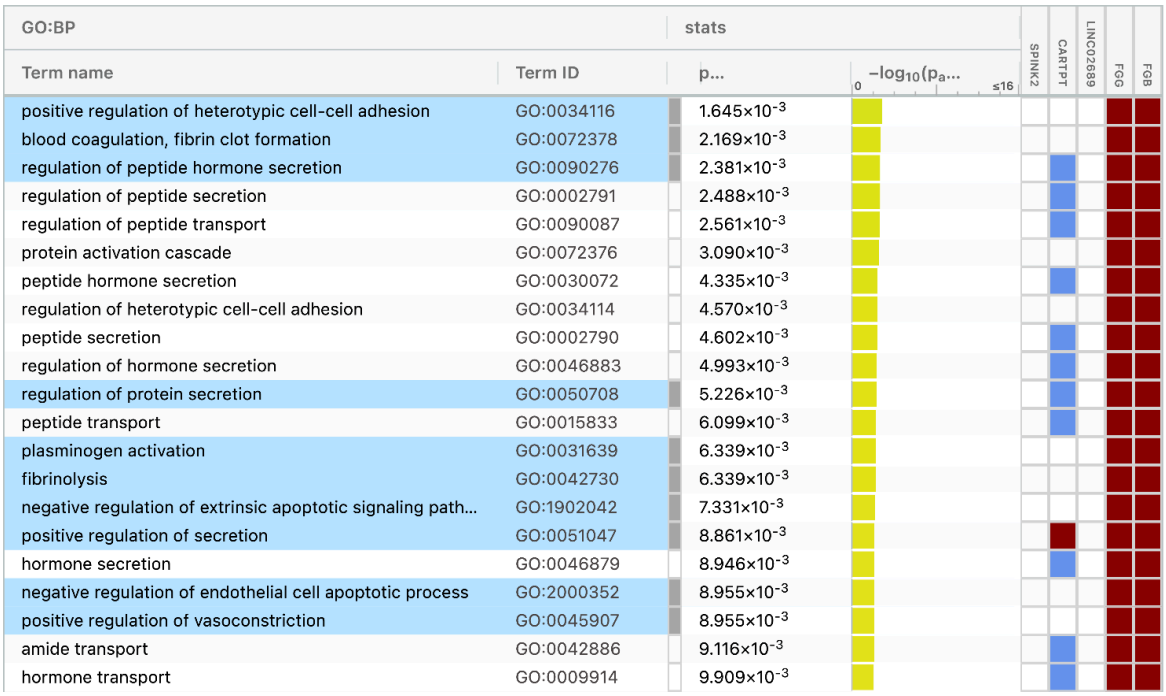


Figure 1.3 A heatmap of downregulated genes in ALS-Glia compared to ALS-Ox and which biological pathways these genes are linked to through the use of g:Profiler. In our analysis, we only viewed the pathways with a p-adjusted value less than 0.01.

In this comparison, we saw which genes in ALS-Ox were upregulated or downregulated compared to ALS-TE (Figure 2.1). The upregulated genes in ALS-Ox compared to ALS-TE showed no biological pathway connections through the g:Profiler. Although it may look strange because of the large number of upregulated genes compared to downregulated on the volcano plot, the difference in the results of the g:Profiler was mostly because we filtered the genes that had greater than 2^2 or 2^{-2} fold changes.

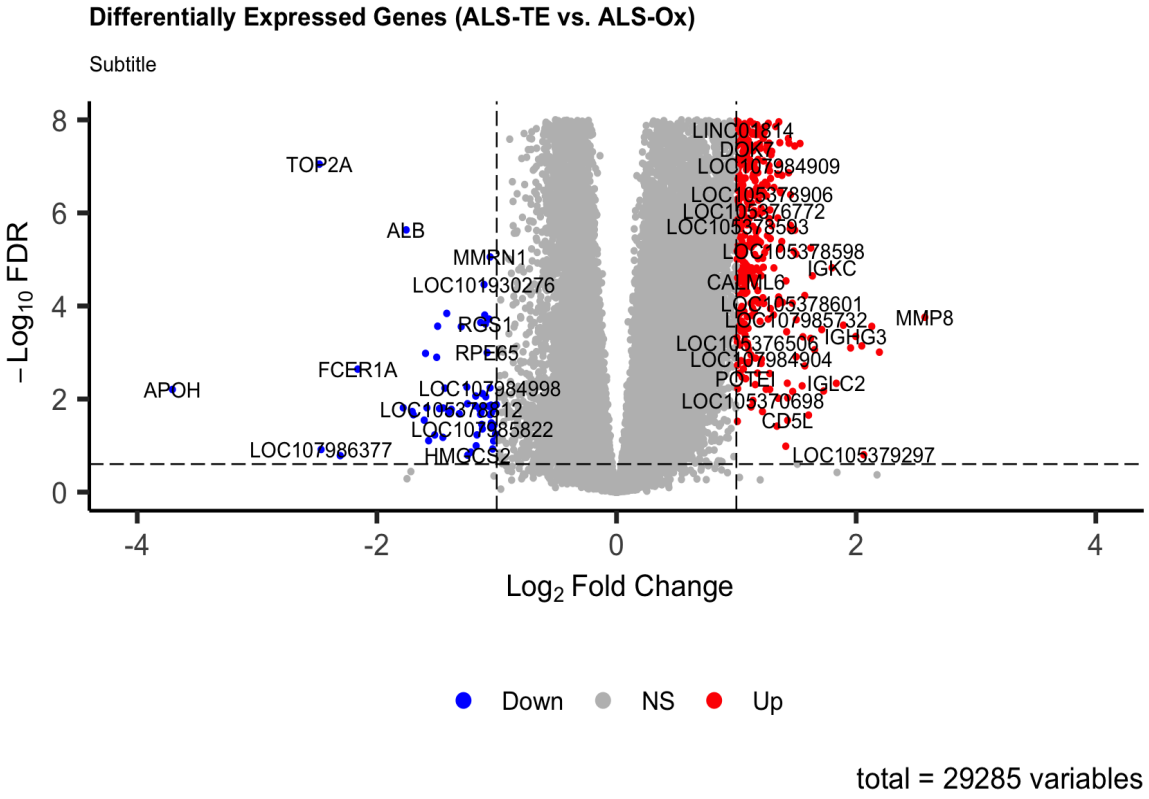
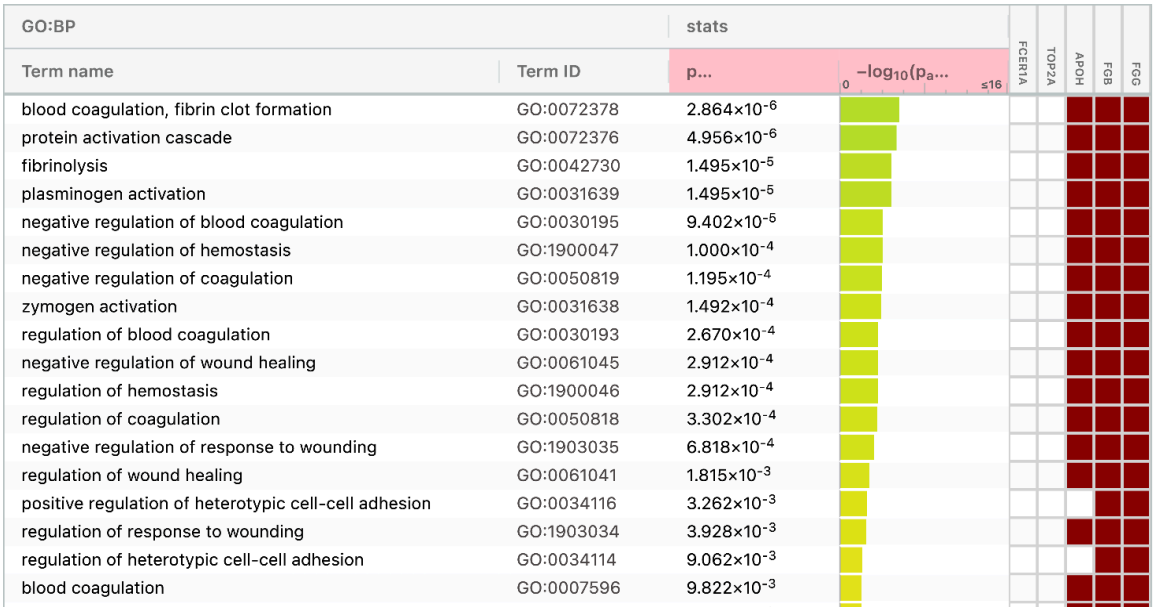


Figure 2.1. A volcano plot of upregulated (red) and downregulated (blue) genes in ALS-Glia compared to ALS-Ox. These were determined by differences in Log_2Fold change on the x-axis, as well as the p-adjusted values on the y-axis. This plot visualizes which genes were used in the g:Profiler to obtain biological pathway information. NS = not significant.

The downregulated genes in ALS-Ox compared to ALS-TE, which are genes more expressed in ALS-TE, were shown to have strong correlations with multiple pathways similar to those observed in the Ox vs Glia results. These pathways mainly relate to blood, for example, blood coagulation, hematopoiesis, and other blood clotting pathways (Figure 2.2). In general, though, compared to the other comparisons between subtypes, ALS-Ox and ALS-TE had the lowest number of genes that met the threshold of a p-value lower than 0.1 as well as log_2Fold Changes greater than 2 or less than -2.



GO:BP		stats						
Term name	Term ID	p...	$-\log_{10}(p_a...$ 0 ≤16	FGER1A	TOP2A	APOH	FGB	FGG
blood coagulation, fibrin clot formation	GO:0072378	2.864×10^{-6}						
protein activation cascade	GO:0072376	4.956×10^{-6}						
fibrinolysis	GO:0042730	1.495×10^{-5}						
plasminogen activation	GO:0031639	1.495×10^{-5}						
negative regulation of blood coagulation	GO:0030195	9.402×10^{-5}						
negative regulation of hemostasis	GO:1900047	1.000×10^{-4}						
negative regulation of coagulation	GO:0050819	1.195×10^{-4}						
zymogen activation	GO:0031638	1.492×10^{-4}						
regulation of blood coagulation	GO:0030193	2.670×10^{-4}						
negative regulation of wound healing	GO:0061045	2.912×10^{-4}						
regulation of hemostasis	GO:1900046	2.912×10^{-4}						
regulation of coagulation	GO:0050818	3.302×10^{-4}						
negative regulation of response to wounding	GO:1903035	6.818×10^{-4}						
regulation of wound healing	GO:0061041	1.815×10^{-3}						
positive regulation of heterotypic cell-cell adhesion	GO:0034116	3.262×10^{-3}						
regulation of response to wounding	GO:1903034	3.928×10^{-3}						
regulation of heterotypic cell-cell adhesion	GO:0034114	9.062×10^{-3}						
blood coagulation	GO:0007596	9.822×10^{-3}						

Figure 2.2 A heatmap of downregulated genes in ALS-Ox compared to ALS-TE and which biological pathways these genes are linked to through the use of g:Profiler. In our analysis, we only viewed the pathways with a p-adjusted value less than 0.01.

In this comparison, we saw which genes in ALS-TE were upregulated or downregulated compared to ALS-Glia (Figure 3.1). In this instance, the g:Profiler was unable to determine any biological pathways in both the upregulated and downregulated genes.

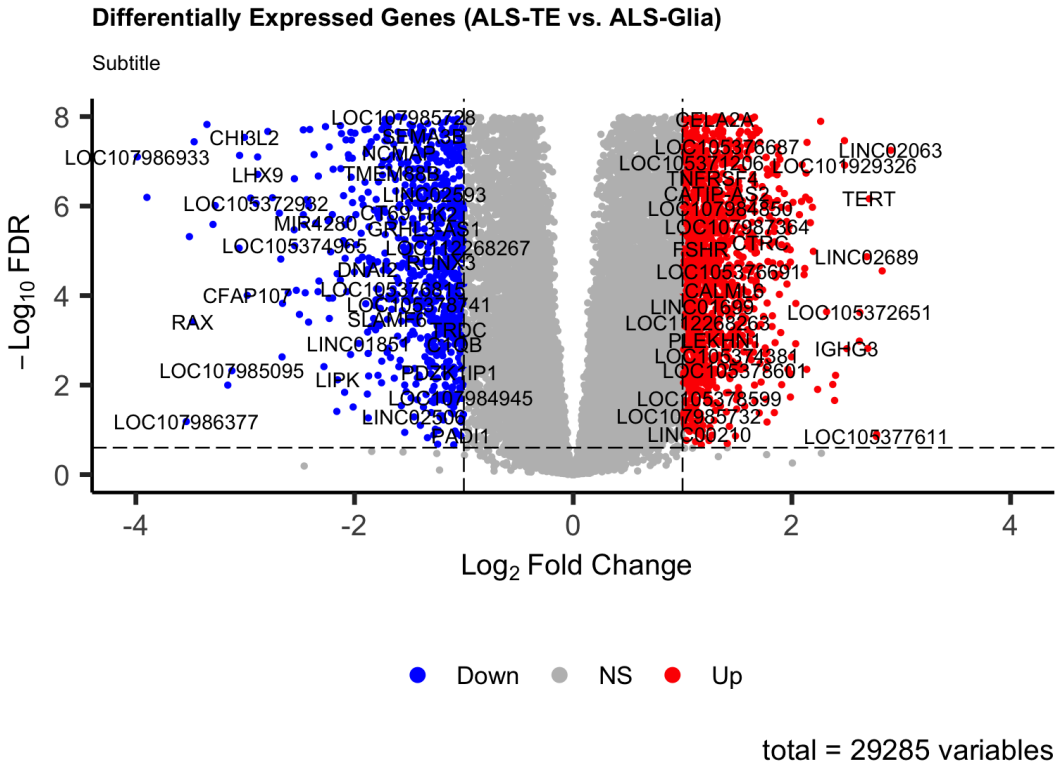


Figure 3.1. A volcano plot of upregulated and downregulated genes in ALS-Glia compared to ALS-TE. These were determined by differences in Log_2Fold change on the x-axis, as well as the p-adjusted values on the y-axis. This plot visualizes which genes were actually used in the g:Profiler in order to obtain biological pathway information. NS = not significant.

Discussion

Based on the comparisons made between different ALS subtypes in this study, we see how different biological pathways relate to each subtype. ALS-Ox genes, for example, are more upregulated in pathways that relate to vascular processes. When compared to ALS-TE, blood clotting pathways are also upregulated. ALS-Glia showed upregulation in the immune system, and we were able to connect this to chitin metabolic and catabolic processes, which in recent studies have been shown to relate to improved immune systems. Unfortunately, we were unable to locate any upregulated pathways in ALS-TE, as the g:Profiler found no commonality in the genes highlighted by the DESeq data for this analysis.

While we purposefully selected a smaller sample group, this means the data collected may not fully encompass the scope of ALS patients. Although it doesn't discount these findings, it does mean we can't be sure the pathways are found in every patient with a certain subtype. We were unable to find upregulated pathways in ALS-TE. This, however, does not mean that the upregulated genes found through the analysis are unrelated to one another. There is a possibility that the g:Profiler was simply unable to find pathways because of a lack of sufficient data.

I hope that with the analysis in this project, fellow ALS researchers can expand our analysis to larger datasets to see if there is replication of these pathways among larger groups. I also hope that the pathways discovered can lead to targeted therapeutics for patients. For example, knowing the upregulation of blood coagulation/clotting in ALS-Ox could aid in developing treatments that target that pathway or mitigate clotting. By linking specific pathways to the clinical manifestations of ALS subtypes, we can move closer to personalized treatments that address the unique needs of patients, enhancing both their quality of life and long-term prognosis.

References

Dhasmana, Swati, Anupam Dhasmana, Sudhir Kotnala, Varsha Mangtani, Acharan S. Narula, Shafiul Haque, Meena Jaggi, Murali M. Yallapu, and Subhash C. Chauhan. "Boosting Mitochondrial Potential: An Imperative Therapeutic Intervention in Amyotrophic Lateral Sclerosis." *Current Neuropharmacology* 21, no. 5 (2023): 1117-38. https://doi.org/10.2174/1570159x20666220915092703.

Eshima, Jarrett, Samantha A. O'Connor, Ethan Marschall, Robert Bowser, Christopher L. Plaisier, and Barbara S. Smith. "Molecular Subtypes of ALS Are Associated with Differences in Patient Prognosis." *Nature Communications* 14, no. 1 (2023). https://doi.org/10.1038/s41467-022-35494-w.

Giovedì, Silvia, Margherita Maria Ravanelli, Barbara Parisi, Barbara Bettegazzi, and Fabrizia Claudia Guarnieri. "Dysfunctional Autophagy and Endolysosomal System in Neurodegenerative Diseases: Relevance and Therapeutic Options." *Frontiers in Cellular Neuroscience* 14 (December 17, 2020). https://doi.org/10.3389/fncel.2020.602116.

Kim, Do-Hyun, Yilin Wang, Haerin Jung, Rachael L. Field, Xinya Zhang, Ta-Chiang Liu, Changqing Ma, James S. Fraser, Jonathan R. Brestoff, and Steven J. Van Dyken. "A Type 2 Immune Circuit in the Stomach Controls Mammalian Adaptation to Dietary Chitin." *Science* 381, no. 6662 (2023): 1092-98. https://doi.org/10.1126/science.add5649.

O'Neill, Kathryn, Regina Shaw, Isobel Bolger, Oliver Tam, Hemali Phatnani, and Molly Gale Hammell. "ALS Molecular Subtypes Are a Combination of Cellular, Genetic, and Pathological Features Learned by Deep Multiomics Classifiers." *bioRvix*, July 20, 2024. https://doi.org/10.1101/2024.07.19.603731.

Sar, Sonya. "Analysis of the Therapeutic Strategies Used to Treat ALS Caused by TDP-43 Aggregation." *Research Archive of Rising Scholars*, January 20, 2024. https://doi.org/10.58445/rars.884.

Tam, Oliver H., Nikolay V. Rozhkov, Regina Shaw, Duyang Kim, Isabel Hubbard, Samantha Fennessey, Nadia Propp, Delphine Fagegaltier, Brent T. Harris, Lyle W. Ostrow, Hemali Phatnani, John Ravits, Josh Dubnau, Molly Gale Hammell, Hemali Phatnani, Justin Kwan, Dhruv Sareen, James R. Broach, Zachary Simmons, Ximena Arcila-Londono, Edward B. Lee, Vivianna M. Van Deerlin, Neil A. Shneider, Ernest Fraenkel, Lyle W. Ostrow, Frank Baas, Noah Zaitlen, James D. Berry, and Andrea Malaspina. "Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia." *Cell Reports* 29, no. 5 (2019): 1164-1177.e5. https://doi.org/10.1016/j.celrep.2019.09.066.