Uncovering Common Gene Expression Signatures Between COVID-19 and Alzheimer's Disease

Final Project Report

Under supervision:

Dr. Mohamed EL-Sayeh

Name and ID:

Sara Ahmed 231001837

Youseif Hussein 221000443

Hashem Yasser 231000770

Sara Sayed Elganzory 221000509

CBIO311 / CBIO310

DEC 2026

Table of Contents

Abstract

COVID-19 and Alzheimer's disease (AD) are two complex disorders that share several biological features, including inflammation, oxidative stress, and cellular dysfunction. In this study, publicly available gene expression datasets were analyzed to identify differentially expressed genes (DEGs) associated with both conditions. Statistical analysis was performed to detect significant changes in gene expression, followed by visualization using volcano plots, heatmaps, and comparative analysis. In addition, pathway enrichment analysis was applied to explore the biological mechanisms underlying the observed gene expression patterns. The results revealed important shared molecular pathways related to mitochondrial activity, immune response, and neurodegeneration, highlighting potential links between viral infection and neurodegenerative processes.

Introduction

Alzheimer's disease is a progressive neurodegenerative disorder characterized by memory loss, cognitive impairment, and neuronal damage. Its pathology involves amyloid-beta accumulation, tau protein aggregation, chronic neuroinflammation, and mitochondrial dysfunction.

COVID-19, caused by the SARS-CoV-2 virus, primarily affects the respiratory system but has also been associated with neurological complications, including cognitive decline and inflammatory damage to the nervous system. Recent evidence suggests that systemic viral infections may influence long-term brain health.

Gene expression profiling provides a powerful approach for understanding disease mechanisms at the molecular level. By comparing diseased and healthy samples, it is possible to identify genes and pathways that are significantly altered. This study aims to investigate and compare the transcriptional changes in COVID-19 and Alzheimer's disease to identify shared biological signatures and potential molecular connections between the two conditions.

Methodology

Data Collection and Preprocessing

Gene expression datasets related to Alzheimer's disease and COVID-19 were obtained from public repositories. The datasets included patient and control samples. Data were cleaned, standardized, and filtered to ensure consistency of gene identifiers and expression values.

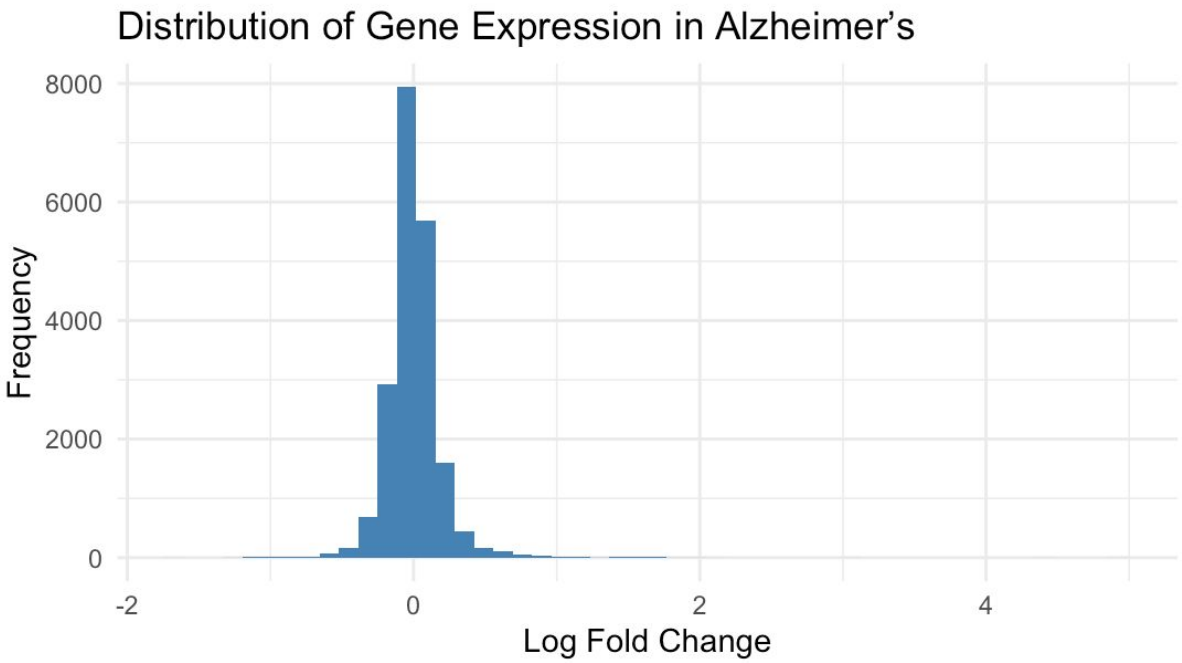

Figure 1 : Distribution of Gene Expression in Alzheimer's Disease.

This histogram shows the distribution of log fold change values for genes in the Alzheimer's dataset.

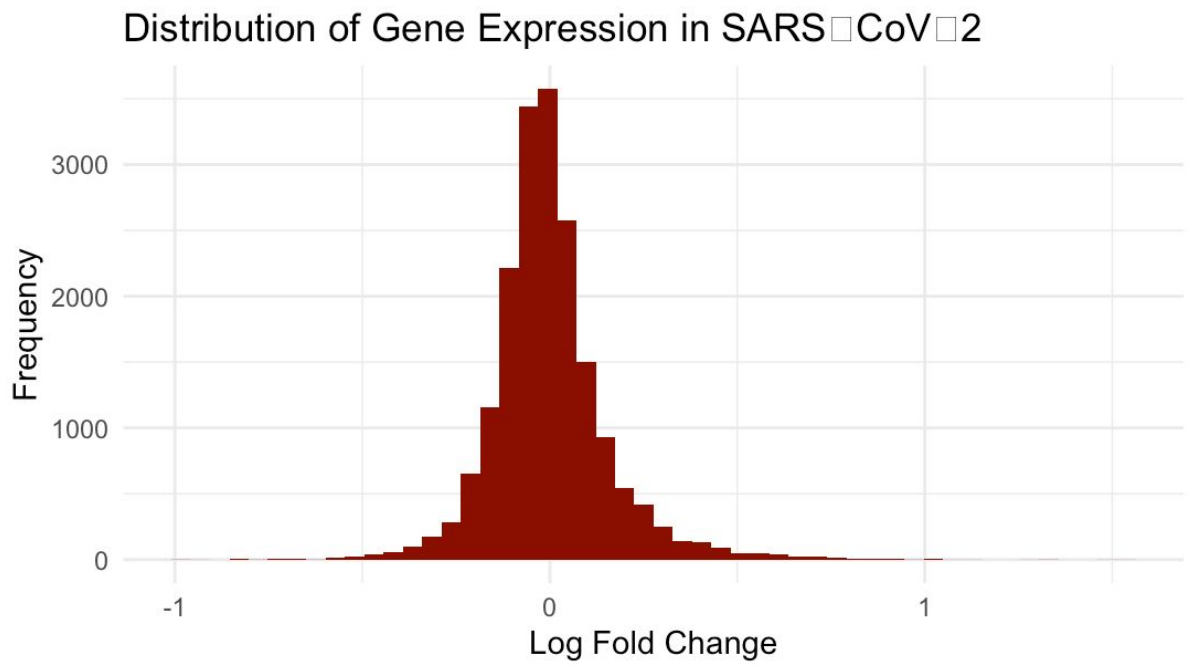

Figure2 : Distribution of Gene Expression in SARS-CoV-2

This histogram illustrates the distribution of log fold change values for genes in the COVID-19 dataset.

Differential Expression Analysis

Differential gene expression analysis was conducted using statistical thresholds of:

- Adjusted p-value < 0.05
- |Log2 Fold Change| > 0.5

Genes meeting these criteria were considered significantly upregulated or downregulated.

Visualization

Several visualization techniques were applied, including:

- Volcano plots to display significance and fold change
- Heatmaps to show expression patterns and clustering
- Venn diagrams to identify shared DEGs

- Correlation plots to compare gene behavior between diseases

Pathway Enrichment Analysis

Gene Set Enrichment Analysis (GSEA) was performed using the KEGG database to identify significantly enriched biological pathways associated with the identified DEGs.

Results

Differentially Expressed Genes

The analysis revealed many significantly altered genes in both COVID-19 and Alzheimer's datasets. These changes indicate extensive molecular remodeling in response to disease conditions.

Volcano plots demonstrated clear separation between significant and non-significant genes, highlighting strong upregulation and downregulation patterns.

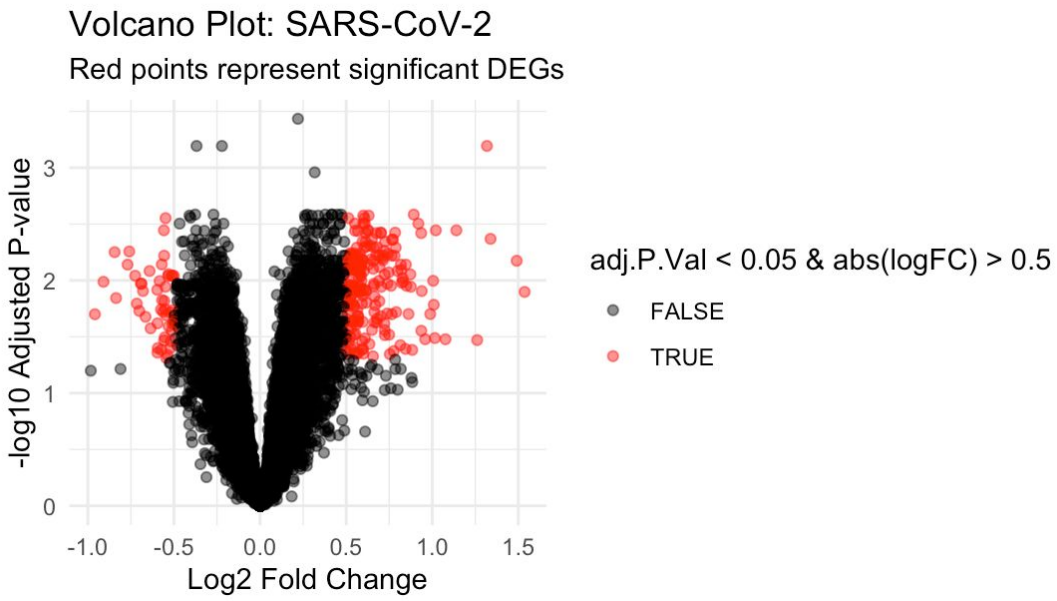

Figure 3: Volcano Plot of Differentially Expressed Genes in SARS-CoV-2

This volcano plot highlights significantly upregulated and downregulated genes based on adjusted p-value and log fold change thresholds.

Shared Gene Analysis

Comparison between the two datasets identified a subset of genes that were commonly dysregulated in both diseases, suggesting overlapping molecular mechanisms. However, heatmap and correlation analyses showed consistent expression trends for many of these shared genes.

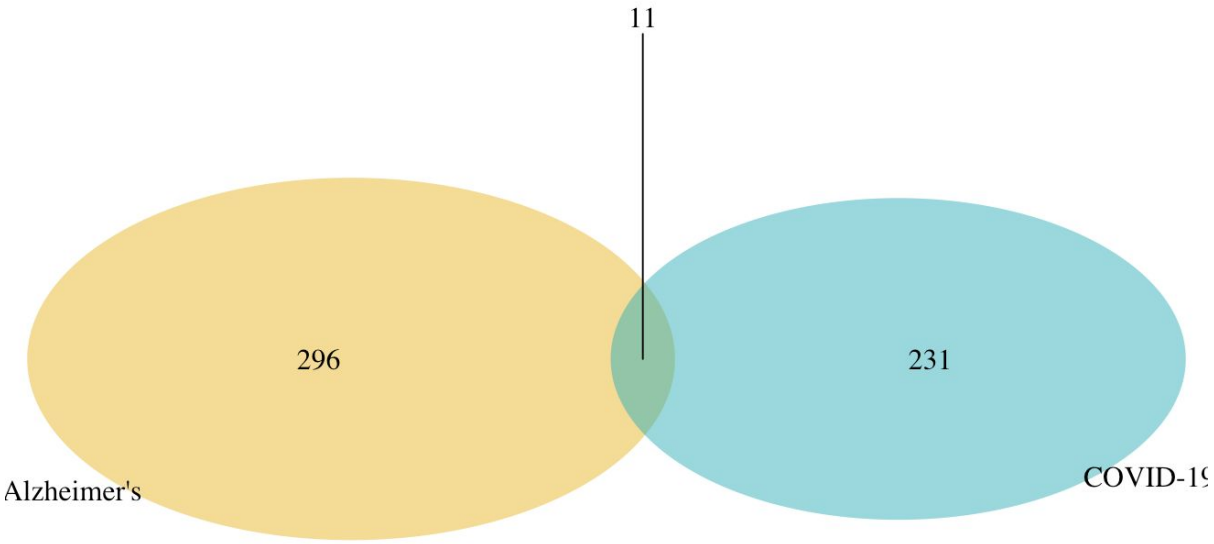

Figure 4: Venn Diagram of Differentially Expressed Genes This diagram shows the number of shared and unique differentially expressed genes between COVID-19 and Alzheimer's disease

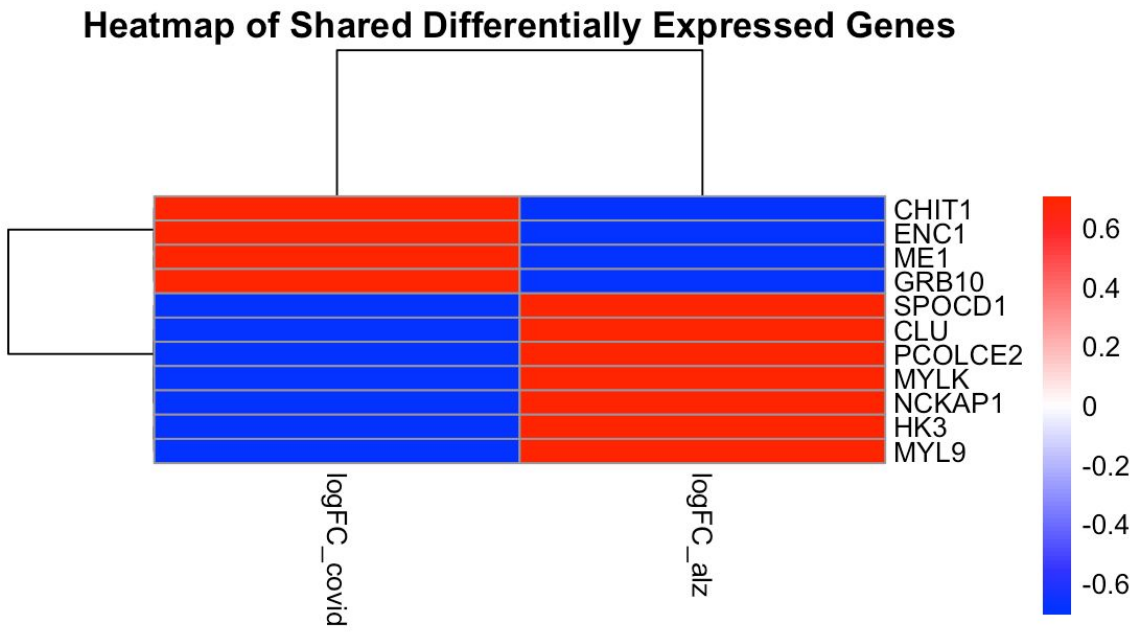

Figure 5: Heatmap of Shared Differentially Expressed Genes

This heatmap represents the expression patterns and clustering of genes that are commonly dysregulated in both diseases.

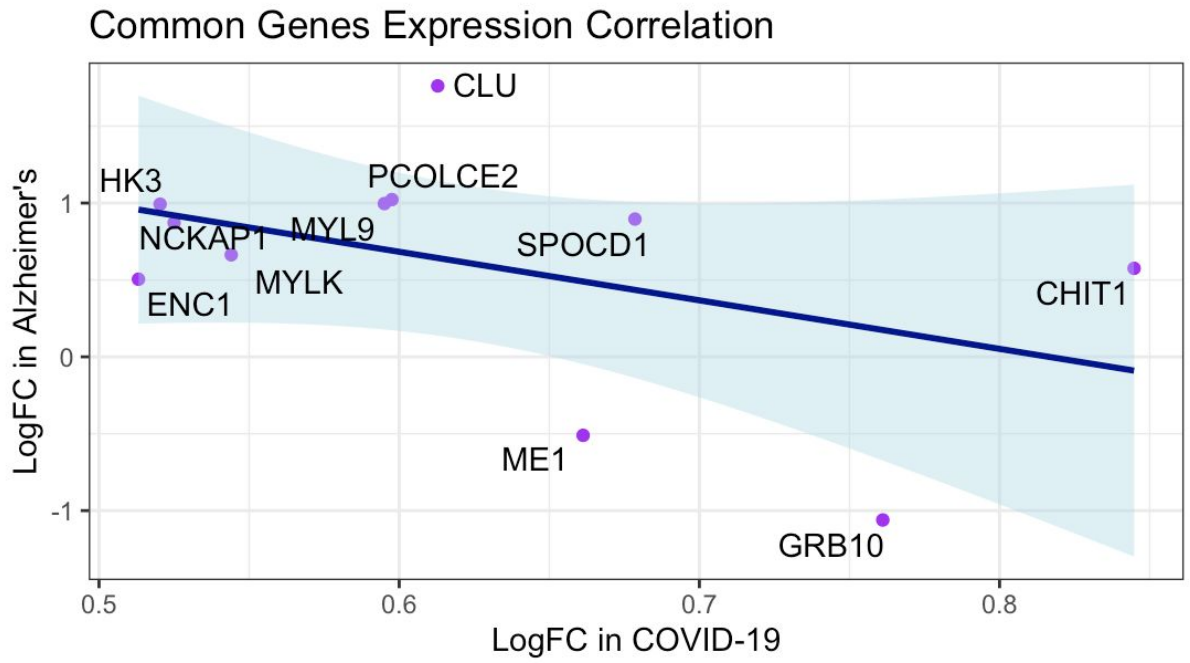

Figure 6: Correlation of Common Gene Expression Between Diseases

This scatter plot illustrates the relationship between log fold change values of shared genes in COVID-19 and Alzheimer's disease.

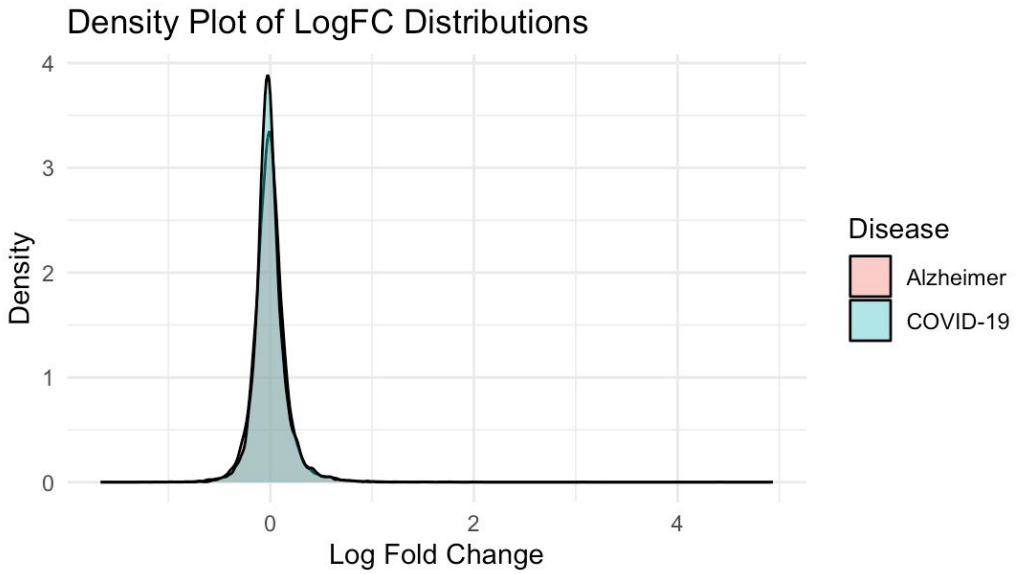

Figure 7: Density Plot of LogFC Distributions

This figure compares the overall gene expression distributions between the two conditions.

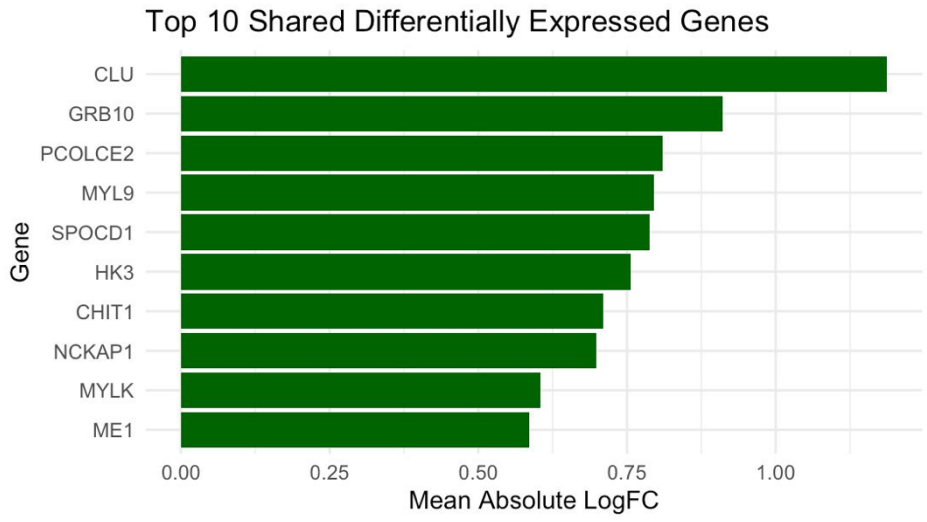

Figure 8 : Top 10 Shared Differentially Expressed Genes

This bar plot displays the most strongly dysregulated genes shared between COVID-19 and Alzheimer's disease.

Pathway Enrichment

Enrichment analysis demonstrated that the shared DEGs were mainly involved in:

- Oxidative phosphorylation
- Mitochondrial dysfunction
- Immune and inflammatory signaling
- Neurodegenerative disease pathways

These findings indicate that energy metabolism and cellular stress responses play key roles in both COVID-19 and Alzheimer's disease.

Discussion

The results support the hypothesis that COVID-19 and Alzheimer's disease share important molecular features. The involvement of mitochondrial and inflammatory pathways suggests that viral infection may exacerbate neurodegenerative processes through systemic stress and immune activation.

The overlap in neurodegenerative pathways also indicates that chronic inflammation and metabolic imbalance may represent common drivers of neuronal damage.

These findings emphasize the importance of transcriptomic and pathway-based approaches for understanding complex disease interactions.

Conclusion

This study used gene expression and pathway analysis to explore the molecular relationship between COVID-19 and Alzheimer's disease. Significant shared genes and pathways were identified, particularly those related to mitochondrial activity, oxidative stress, and neurodegeneration. The integration of statistical analysis and visualization provided a comprehensive view of disease-associated molecular alterations and may support future biomarker discovery and therapeutic research.

References

1. National Center for Biotechnology Information (NCBI). (2023). GSE222333: Gene expression profiling of Alzheimer's disease samples [Data set]. Gene Expression Omnibus (GEO). https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE222333
2. National Center for Biotechnology Information (NCBI). (2024). GSE300129: Gene expression analysis of SARS-CoV-2 infected patients [Data set]. Gene Expression Omnibus (GEO). https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300129