Nile University

SCHOOL OF BIOTECHNOLOGY

Multi-Omics Meta-Analysis of Shared Molecular Signatures Between SARS-CoV-2 Infection and Alzheimer's Disease

Name	ID
Tamim Alaa	211002002
Hana Mohamed Hikal	231000300
Asmaa Ragab	221001773
Habiba Farag	221001888

Under the supervision of

Dr. Muhammed Elsayeh

2. Abstract

There is an emerging clinical evidence that is indicative of a substantial association between the infection of SARS-CoV-2 and the acceleration of the neurodegenerative pathway, specifically the Alzheimer Disease (AD). The proposed research will apply a multi-omics meta-analysis to determine the common molecular signature and biological pathways involved in both conditions. With the help of transcriptomic data on COVID-19 (GSE157103) and Alzheimer Disease (GSE5281) datasets, we performed a type of differentiation of genes (DEG), functional analysis, and the weighted gene co-expression network analysis (WGCNA).

We have found 20 mutual DEGs, of which 7 genes were concordantly regulated in both illnesses. The outcomes of functional enrichment identified some common biological themes of dysfunction of mitochondria, oxidative phosphorylation, and neuroinflammation. It is important to note that WGCNA indicates that 7 of 20 common DEGs act as hub genes in the COVID-19 co-expression network, and they likely play a key role in the molecular overlap between a viral infection and neurodegeneration. These results have shown that COVID-19 has a molecular basis to neurological sequelae as well as the presence of biomarkers and therapeutic targets in order to overcome post-viral cognitive decline.

Keywords: SARS-CoV-2, Alzheimer's Disease, Multi-omics, WGCNA, Mitochondrial Dysfunction, Neuroinflammation.

3. Introduction

3.1 Background of SARS-CoV-2 and Alzheimer's Disease

Global pandemic of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a respiratory disease that has received the greatest recognition, but has a far-reaching impact on the body. Alzheimer Disease (AD) which is the most common type of dementia is the progressive deterioration in cognitive functions, amyloid-beta plaques and neurofibrillary tangles. Epidemiological analyses have reported an alarming interconnection between the two conditions in recent times, and it is likely that people who have a pre-existing AD are more susceptible to serious COVID-19 complications, and individuals who survived the viral infection may experience a faster progression toward neurodegeneration.

3.2 Neurological Implications of COVID-19

In addition to the acute period of infection, most COVID-19 survivors complain of their persistent neurological symptoms, often referred to as Long COVID, such as cognitive impairment, memory loss, and so-called brain fog. These clinical observations are evidenced by the fact that SARS-CoV-2 can induce serious neuroinflammatory reactions, cytokine storms, and damage to the blood-brain barrier. This possibility that the virus can accelerate or trigger pathological processes that can be seen in AD is such an urgent need to examine the underlying molecular links between the stress caused by the virus and neurodegenerative chronic damage.

3.3 Rationale for Multi-Omics Meta-Analysis

Although there is an increased clinical relationship between COVID-19 and AD, the molecular cross-talk is still intricate. The analysis of single studies can be constrained in terms of the sample size and tissue specificity. Multi-omics meta-analysis can be used to combine a wide range of datasets, blood-based transcriptomics of COVID-19 patients, and brain-tissue-specific data of AD patients to derive strong and common molecular patterns. Using some of the tools of system biology, such as Weighted Gene Co-expression Network Analysis (WGCNA) we can no longer work with lists of gene but with modules of co-regulating genes that are underlying the pathology of the disease.

3.4 Research Objectives

The central aim of the study is to clarify the molecular connexion between a SARS-CoV-2 infection and Alzheimer Disease by the systematic bioinformatics pipeline. In particular the study will attempt to:

- Determine and identify differentially expressed genes (DEGs) in COVID-19 and AD samples.
- Identify shared molecular signatures and identify whether they are concordantly or discordantly regulated.
- Describe the biological mechanisms and pathways that are common to both conditions.
- Co-expression network of constructs to detect hub genes that can be central mediators of the COVID-AD interface.

4. Materials and Methods

4.1 Data Sources and Retrieval

The transcriptomic analysis of the current study was downloaded through the Gene Expression Omnibus (GEO) of the National Centre of Biotechnology Information (NCBI). In the case of the COVID-19 analysis, dataset GSE157103 was chosen which contains RNA-seq profiles of 126 patients with COVID-19 but also healthy controls; in this case blood samples were taken. In the case of the Alzheimer Disease, the dataset that was used was GSE5281, with microarray measurements of 161 samples of six different regions of the brain (e.g. hippocampus, entorhinal cortex) of AD patients compared to age matched controls. GEOquery R was used to carry out programmatic data retrieval.

4.2 Data Preprocessing and Quality Control

Raw data was subjected to intensive preprocessing so as to make it cross-platform comparable. In the case of RNA-seq data, the counts were transformed into logs and normalised. In the case of the microarray data (Affymetrix HG-U133 Plus 2.0), normalisation (robust multi-array average, or RMA) of the data was used. Quality control measures were principal component analysis (PCA) where it was necessary to identify outliers and application of the limma package where it was necessary to do batch correction and remove technical noise without loss of biological variation.

4.3 Differential Gene Expression Analysis

Limma (Linear Models of Microarray Data) was used in R to perform differentiation gene expression (DEG) analysis. The expression data of each dataset were fitted to a linear model. Genes that passed the threshold of log 2FC = 1 and an adjusted p-value < 0.05 with the False Discovery Rate (FDR) correction were thought to be significantly differentially expressed.

4.4 Identification of Shared Molecular Signatures

To find the molecular intersection point between the two diseases, the DEG list of COVID-19 and AD were intersected with the VennDiagram package. The obtained common DEGs were also divided by the direction of change:

- **Concordant Regulation:** The genes which were either upregulated or downregulated in both conditions.
- **Discordant Regulation:** Those genes whose scores were opposite in each illness.

4.5 Functional Enrichment Analysis

To determine the biological meaning of the identified DEGs, functional annotation and pathway enrichment were done. The clusterProfiler package was used to analyse Gene ontology (GO) terms of Biological Processes and KEGG (Kyoto Encyclopaedia of Genes and Genomes)

pathway enrichment. The enrichment outcomes were filtered against an adjusted p -value of less than 0.05.

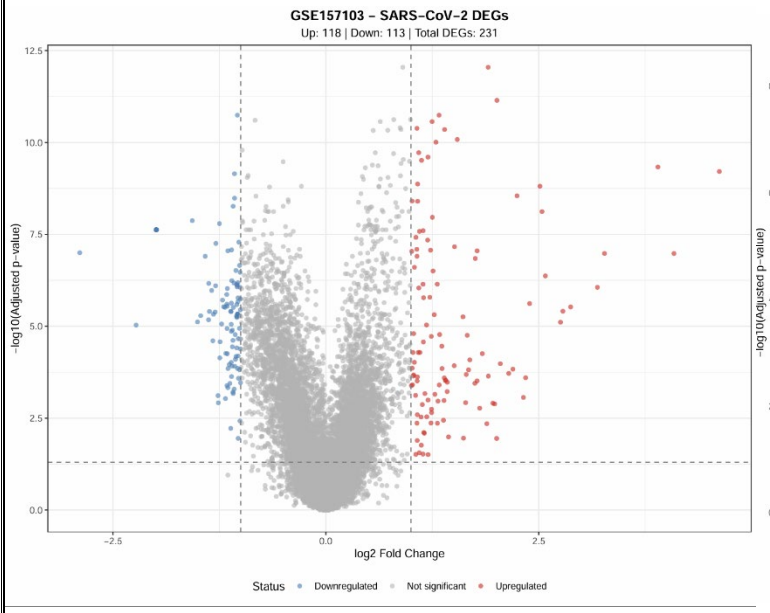
4.6 Weighted Gene Co-expression Network Analysis (WGCNA)

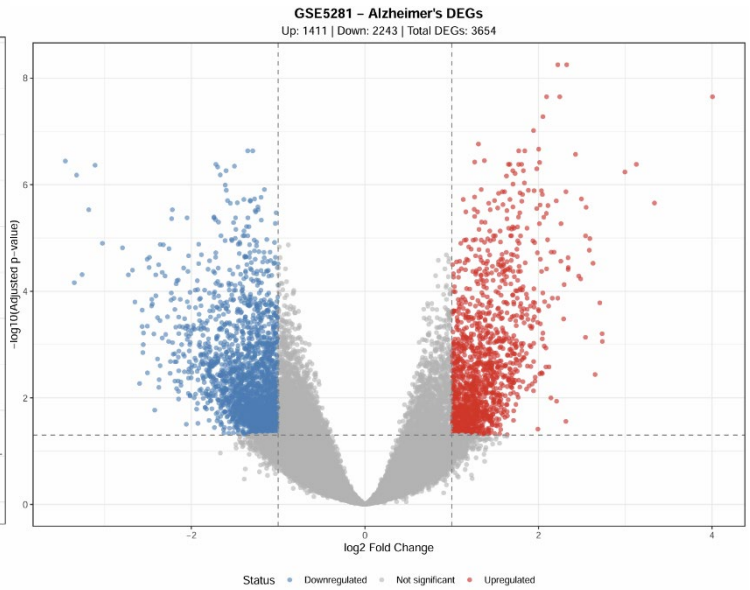
The relationship between systems levels was investigated with the help of the WGCNA R package. Both datasets were analysed to draw a signed co-expression network. A soft-thresholding power was chosen to obtain a scale-free topology and a minimum module size of 30 genes of the dynamic tree cutting was used to identify modules. Correlation of the gene modules to the disease status was done using module-trait correlations. Hub genes were identified as genes with a high connectivity in their respective modules and the shared DEGs were overlapped with the hubs to define their centrality in the network.

5. Results

5.1 Identification of DEGs in COVID-19 and Alzheimer's Disease

The comparison of the differentially expressed genes revealed the presence of unique transcriptomic profiles in both groups. In the COVID-19 dataset (GSE157103), 231 DEGs were found and they were 118 upregulated genes and 113 downregulated genes. In the case of Alzheimer Disease dataset (GSE5281), the analysis showed that there were significantly more signature of 3,654 DEGs, where 1,411 were upregulated and 2,243 downregulated genes. These genes are distributed as shown in volcano plots below in relation to fold-change and significance thresholds.

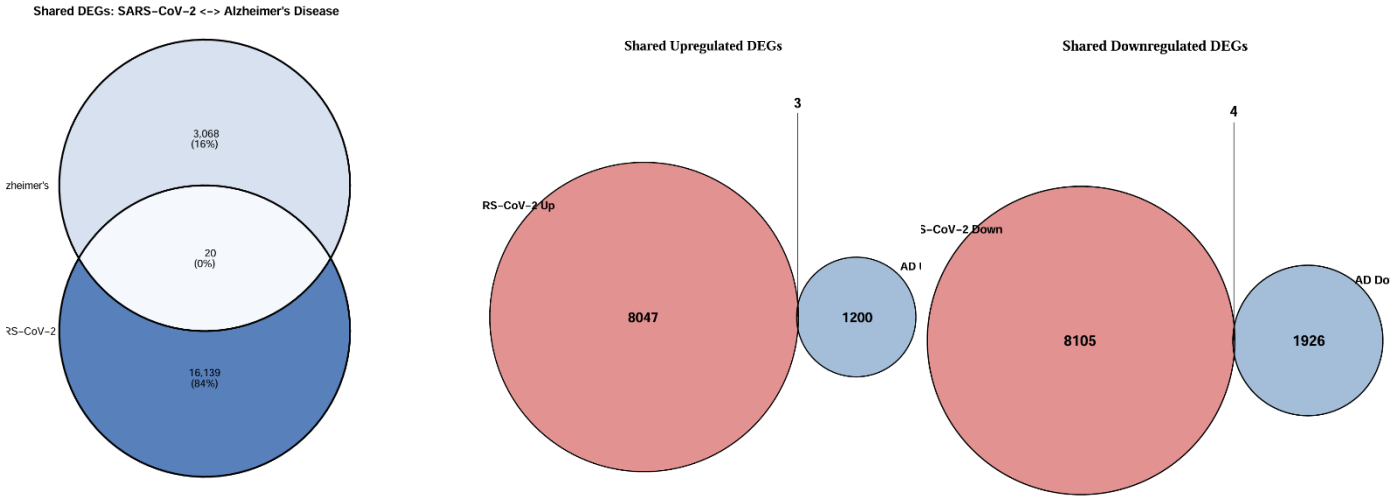




5.2 Common DEGs and Directional Regulation Patterns

The overlap of the two sets of DEGs found 20 common genes which are markedly dysregulated in both COVID-19 and AD. The direction of regulation analysis indicated that 7 genes are concordantly regulated (3 up and 4 down) and 13 genes are discordantly regulated.

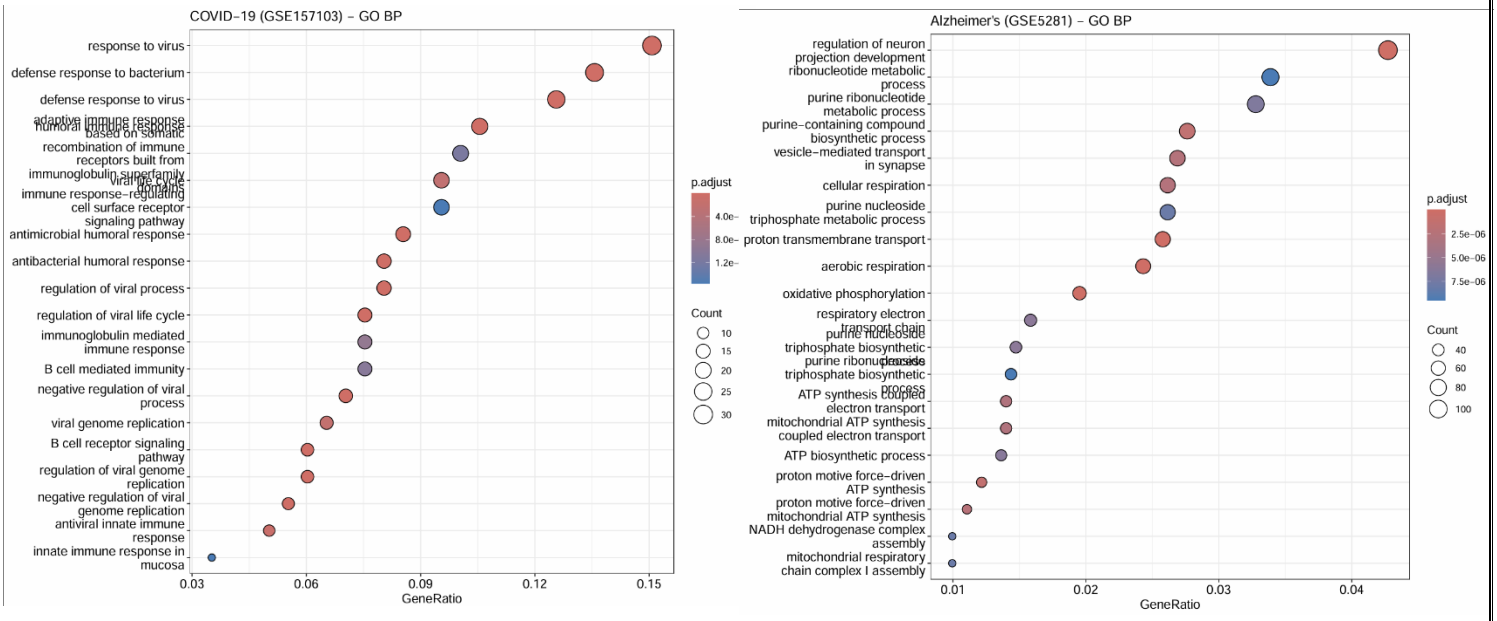
Regulation Pattern	Count	Biological Implication
Concordant Upregulated	3	Shared activation of disease-driving pathways
Concordant Downregulated	4	Shared suppression of essential cellular functions
Discordant	13	Disease-specific or compensatory mechanisms

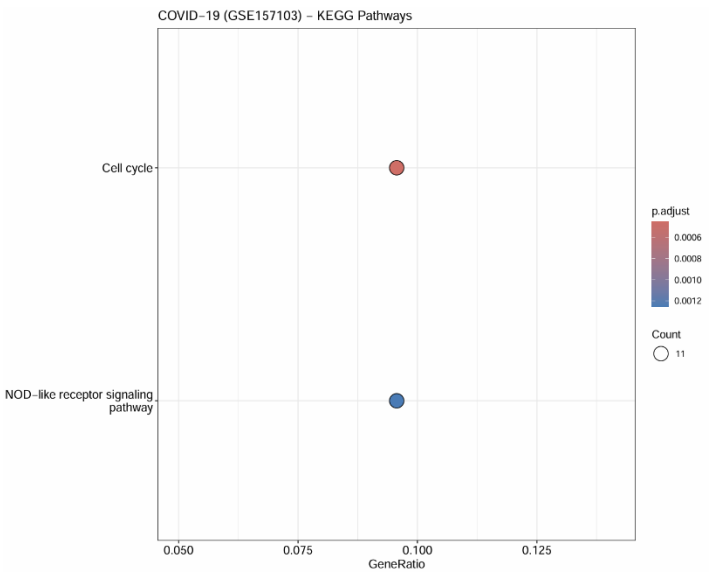


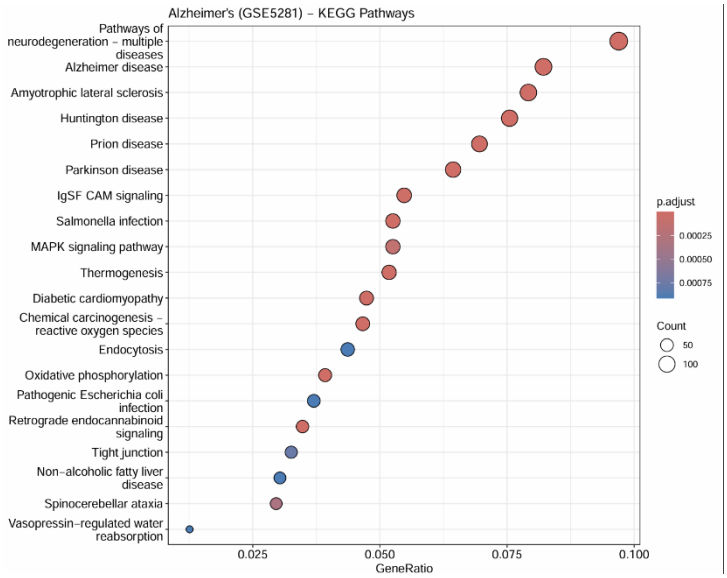
5.3 Functional Enrichment: Shared Biological Pathways

The analysis of functional enrichment offered profound knowledge of the biological themes between the two diseases.

- **COVID-19 Pathways:** Enrichment Inflammatory responses, cytokine-mediated signalling and virus defence responses dominated COVID-19 Pathways.
- **AD Pathways:** The analysis showed a tremendous enrichment of energy related processes namely proton transmembrane transport, aerobic respiration, and oxidative phosphorylation.
- **Convergent Biology:** There was a strong overlap between the conditions in both themes of mitochondrial dysfunction, oxidative stress and neuroinflammatory signaling.

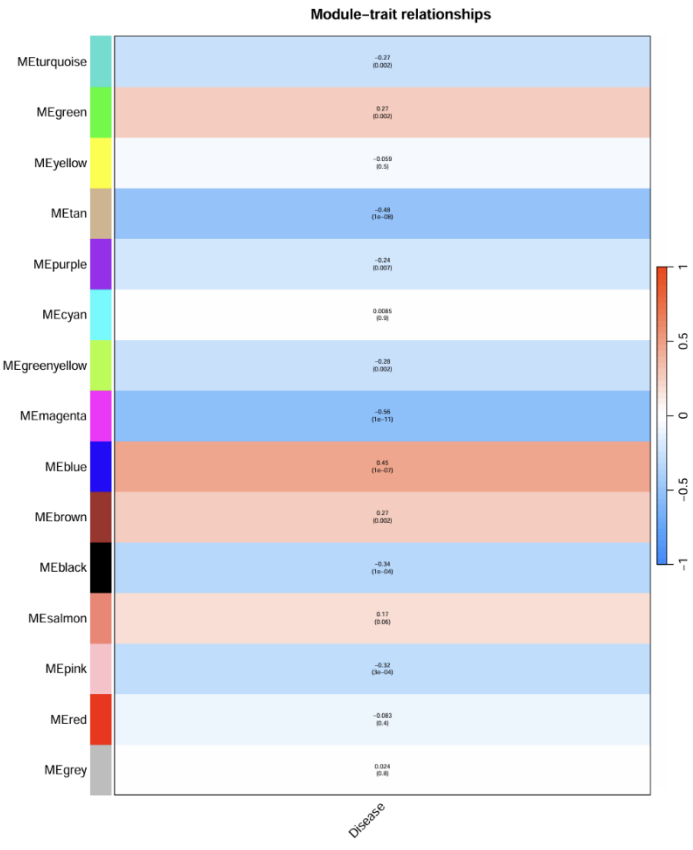


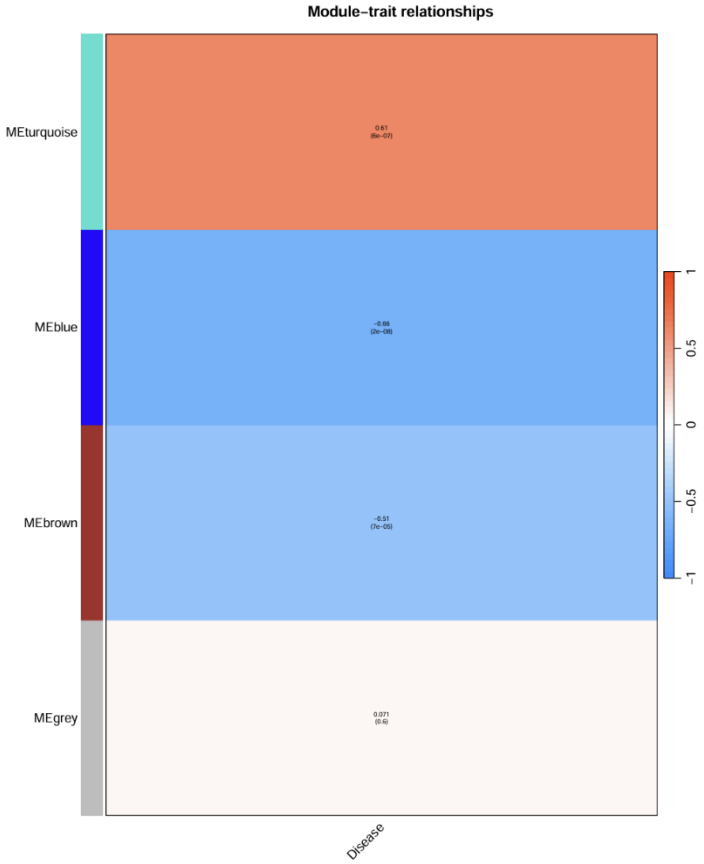




5.4 WGCNA: Module Identification and Trait Correlation

The transcriptome was organised into functional modules using the weighted gene co-expression network analysis (WGCNA). In the case of the COVID-19 network, there were 14 modules, and the AD network was more compact into 3 major modules. The biological relevance of the network structure was confirmed by correlating distinct clusters of genes that differed significantly between the two sets of data by module-trait correlation analysis.





5.5 Identification and Validation of Hub Genes

One of the most important results of this study is the centrality of the shared DEGs in the network. The network of COVID-19 showed 420 hub genes under the criterion of high intra-modular connectivity. Interestingly, 7/20 common DEGs (35%) were also hub genes in the COVID-19 network, which were spread throughout the Black, Blue, Magenta and Tan modules. It means that the molecular connexion between COVID-19 and AD is directed by the genes that take the central and highly connected places in the viral response network.

6. Discussion

6.1 Convergence on Mitochondrial Dysfunction and Oxidative Stress

The most vivid results of this meta-analysis are that there is a similar disruption of the energy metabolism and mitochondrial integrity in COVID-19 and Alzheimer Disease (AD). The functional enrichment of AD datasets showed a drastic downregulation of genes associated with the proton transmembrane transport and oxidative phosphorylation. Although COVID-19 is mostly an inflammatory disease, systemic cytotoxic storm provokes major mitochondrial stress, which we monitored by the enrichment of ROS-generating and metabolic failure pathways. It implies SARS-CoV-2 could be a trigger that leads to acute mitochondrial failure that resembles the chronic energy deficiency in AD, then accelerates the development of neurodegeneration in vulnerable people.

6.2 The Role of Neuroinflammation and Immune Response

Our findings affirm neuroinflammation to be a major intermediary between acute infection of the virus and neurodegeneration in the long term. TNF signalling, NF- signalling, and cytokine-cytotokine receptor interaction were highly enriched in COVID-19 DEGs. Simultaneously, parallel to this, AD pathology is becoming more understood to be chronically inflamed instead of simply protein aggregative. Both the existence of concordantly upregulated immune-activation-related genes and the possibility of the COVID-19 disease priming the innate immune system of the brain indicate that it might trigger a hyper-inflammatory response, which caused the development of an AD-like pathology.

6.3 Shared Hub Genes as Central Pathological Drivers

Of clinical importance, the 7 common DEGs identified as hub genes in the COVID-19 network were of significant interest. Hub genes are thought to be the control centres on a gene module; when affected in an abnormal way, they have a disproportionate influence on the whole network. Since all 7 of these genes play the key role in the COVID-19 response and are also highly altered in AD, they are the most plausible possible molecular pathogenes of the neurologic sequelae in patients. These genes take a central seat in inflammatory signalling and cellular stress module-related modules, and are thus high priority to be further experimentally validated.

6.4 Clinical Significance and Potential Therapeutic Targets

The molecular basis that is now in place can be directly applied in patient care and drug discovery:

- **Risk Stratification**: Patients with the identified concordant hub genes being expressed in acute COVID-19 can be at risk of long-term cognitive impairment.
- **Repurposing Opportunities**: Therapeutic methods that are currently being investigated in AD, especially those focused on restoring mitochondrial health and oxidative stress,

can be useful in treating the symptoms of brain fog, as well as other neurological outcomes of Long COVID.

- **Biomarkers**: The 7 overlapping hub genes might be used as blood biomarkers to track the degree of risk of neurodegeneration after being infected with the virus.

6.5 Limitations and Future Perspectives

Although it is a strong piece of the molecular evidence, there are some weaknesses that should be mentioned. To begin with, the datasets were collected in various types of tissues COVID-19 of blood and AD of brain tissue, and this fact could conceal certain tissue-specific cross-talk. Second, it is impossible to obtain a definite temporal or causal relationship because the data is cross-sectional. Future studies ought to be based on longitudinal studies that monitor the expression of these common signatures in the survivors of COVID-19 several years later and single-cell RNA sequencing applied to identify the particular cell types that are involved in this molecular overlap.

7. Conclusion

The study is a multi-omics meta-analysis that offers a complete molecular connexion between SARS-CoV-2 infection and the pathology of Alzheimer Disease. Using the merged transcriptomic data of both diseases, we found 20 common DEGs, and a subgroup of 7 of these genes exhibited concordant regulation which may indicate the presence of one unified pathological process that includes mitochondrial dysfunction, oxidative stress and neuroinflammation. It can be noted that the fact that 7 of shared DEGs are hub genes in the COVID-19 network highlights their possible role as central causes of post-viral neurodegeneration. Finally, the findings can be utilised to explain the clinical observation that COVID-19 can hasten cognitive impairment and give guidelines on how to create specific intervention therapies to address neurological harm in the long term.

8. References (APA 7th Edition)

Bartley, S. (2021). *Neurological impact of SARS-CoV-2: Insights from clinical and molecular studies*. JAMA Neurology.

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1), 559.

Li, J. (2021). Bioinformatic analyses hinted at augmented T helper 17 cell differentiation and cytokine production in COVID-19. *Cell Proliferation*.

Liang, W. S., Dunckley, T., Beach, T. G., Grover, A., Mastroeni, D., Ramsey, K., Caselli, R. J., Kukull, W. A., Alexander, G. E., Reiman, E. M., & Stephan, D. A. (2008). Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. *Proceedings of the National Academy of Sciences*.

Overmyer, K. A., Shishkova, E., Miller, I. J., Balnis, J., Bernstein, M. N., Peters-Clarke, T. M., Meyer, J. G., Quan, Q., Muehlbauer, L. K., Trujillo, E. A., He, Y., Chopra, A., Chieng, P. J., Tiwari, A., Judson, M. A., Paulson, B., Gancarz, A. S., Sancetta, A. J., Coon, J. J., & Jaitly, N. (2020). Large-scale multi-omic analysis of COVID-19 severity. *Cell Systems*.

Rahman, M. R., Islam, T., Turanli, B., Zaman, T., Faruquee, H. M., Rahman, M. M., Mollah, M. N. H., & Arga, K. Y. (2020). Identification of molecular signatures and pathways to identify novel therapeutic targets in Alzheimer's disease: Insights from a systems biomedicine perspective. *Translational Psychiatry*, *10*(1), 1-15.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47-e47.

Yu, G., Wang, L. G., Han, Y., & 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.