Exploratory Differential Gene Expression Analysis of Alzheimer's Disease

1) INTRODUCTION:

Alzheimer's disease is a progressive neurodegenerative brain disorder and the most common cause of dementia around the world. It leads to a gradual decline in memory, thinking skills, and behaviour. This decline eventually disrupts a person's ability to carry out daily activities. The disease mainly affects older adults and poses a significant global health challenge due to its rising occurrence and the lack of a definitive cure.

At the neuropathological level, Alzheimer's disease involves the buildup of amyloid-β plaques outside cells and neurofibrillary tangles made of hyperphosphorylated tau protein inside cells. These abnormal features harm neuron structure and communication, causing neuron dysfunction and cell death, particularly in areas of the brain responsible for learning and memory, such as the prefrontal cortex and hippocampus.

However, emerging evidence indicates that the disease cannot be fully explained by amyloid and tau issues. Clinical observations show that the amount of plaques does not always match the severity of the disease, and neuron dysfunction can happen before significant plaque buildup occurs. This has led to a broader view of Alzheimer's disease as a complex condition involving long-term brain inflammation, metabolic problems, mitochondrial damage, and widespread changes in gene regulation.

Gene expression analysis provides a way to study Alzheimer's disease by capturing changes in RNA levels across thousands of genes at once. Changes in gene expression reflect stress responses, immune activation, synaptic issues, and energy metabolism problems in the affected brain tissue. Transcriptomic studies offer important insights into the disease mechanisms that go beyond single-gene mutations or specific molecular pathways.

In this context, exploring publicly available transcriptomic datasets allows researchers to identify broad gene expression patterns linked to Alzheimer's disease. These analyses uncover coordinated molecular changes that contribute to neurodegeneration and serve as a basis for forming hypotheses about how the disease progresses.

2) OBJECTIVES:

The primary objective of this study was to perform an exploratory differential gene expression analysis using publicly available transcriptomic data from human brain tissue. Specifically, the study aimed to compare gene expression patterns between Alzheimer's

disease and non-demented control prefrontal cortex samples in order to identify genes showing relative upregulation or downregulation associated with the disease state.

3) MATERIALS AND METHODS:

3.1 Dataset

Gene expression data were obtained from the Gene Expression Omnibus (GEO) database (accession: GSE44770). The dataset consists of microarray-based transcriptomic profiles generated from pooled post-mortem human prefrontal cortex tissue samples derived from individuals diagnosed with Alzheimer's disease and non-demented controls. The prefrontal cortex was selected due to its involvement in cognitive function and its relevance to Alzheimer's disease pathology.

3.2 Gene Expression Measurement

Total RNA was extracted from brain tissue samples and reverse-transcribed into complementary DNA (cDNA). The cDNA was fluorescently labeled and hybridized to microarray chips containing gene-specific DNA probes. Following hybridization, the arrays were scanned to measure fluorescence intensity, which reflects the relative abundance of RNA transcripts corresponding to each gene. These fluorescence intensity values were used as quantitative measures of gene expression.

3.3 Data Preprocessing

The gene expression dataset was analysed using Python in the Google Colab environment. Expression values were loaded into a pandas DataFrame for downstream analysis. Non-expression annotation columns were excluded, and expression values corresponding to pooled Alzheimer's disease and control samples were retained. Gene identifiers were used as row labels, with pooled samples represented as columns.

3.4 Sample Grouping

Due to the pooled nature of the samples, individual-level variability could not be assessed. Samples were therefore grouped into Alzheimer's disease and control categories based on the experimental design described in the original dataset. The analysis was treated as exploratory, with the goal of identifying broad trends in gene expression rather than performing formal statistical inference.

3.5 Differential Expression Analysis

Differential gene expression was assessed by calculating the mean expression level for each gene across Alzheimer's disease and control sample groups. Log2 fold change values were computed as the logarithm (base 2) of the ratio between Alzheimer's disease mean expression and control mean expression. Positive log2 fold change values indicate relative upregulation in Alzheimer's disease, while negative values indicate relative downregulation.

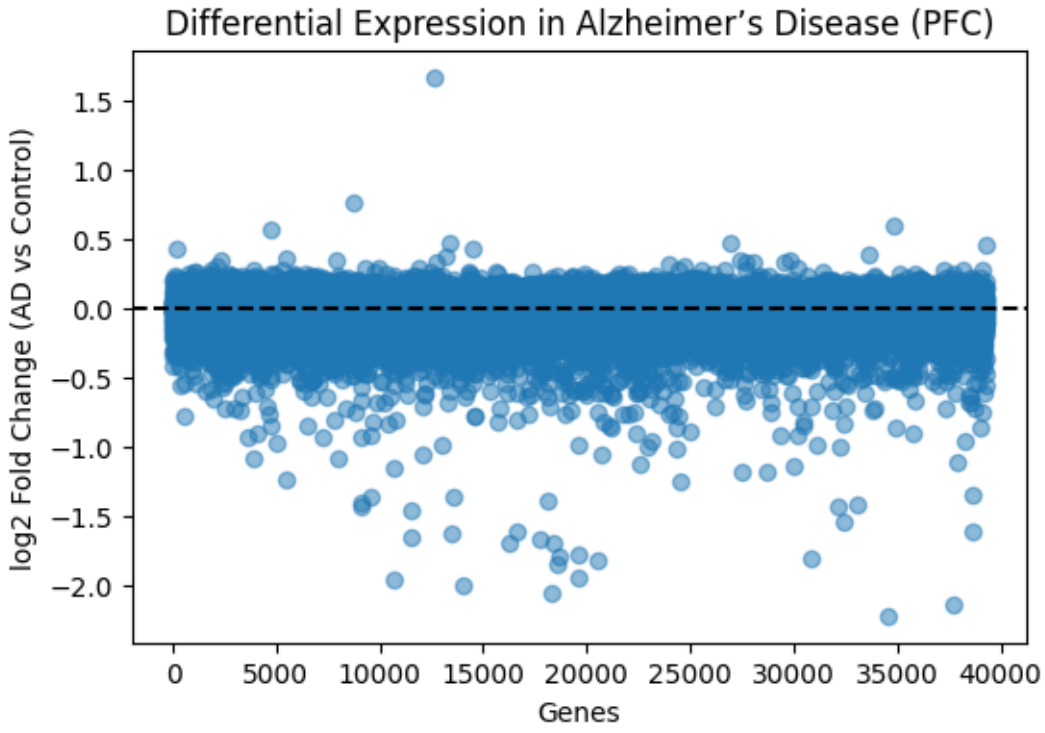
4) Results

Exploratory differential gene expression analysis was performed by comparing mean RNA expression levels between Alzheimer's disease and control prefrontal cortex samples. Log2 fold change values were calculated for each gene to quantify relative differences in expression between the two conditions.

The majority of genes exhibited log2 fold change values close to zero, indicating minimal differences in expression between Alzheimer's disease and control samples. This observation is expected for complex biological tissues such as the brain, where only a subset of genes is altered in disease states.

A subset of genes showed notable positive log2 fold change values, indicating relative upregulation in Alzheimer's disease samples. These genes included several associated with immune response, inflammatory signalling, and cellular stress. Conversely, a subset of genes exhibited strongly negative log2 fold change values, reflecting downregulation in Alzheimer's disease samples. Many of these genes were related to mitochondrial function, oxidative phosphorylation, and neuronal metabolism.

The distribution of log2 fold change values across all genes is shown in **Figure 1**, illustrating both upregulated and downregulated expression trends associated with Alzheimer's disease.



	gene	log2FC
12611	MUC17,MUC3	1.671804
8737	DEFA1,MRS,DEF1,HP-1,DEFA2,HNP-1	0.766013
34835	HSS00085715	0.593411
4720	DEFA3,DEF3,HNP3,HP-3,HNP-3	0.568008
26986	hCT1818463	0.471754
13420	CACNG6	0.469365
39265	XIST,XCE,XIC,SXI1,swd66,DXS1089,DXS399E	0.457537
14525	CHI3L2,YKL39,YKL-39	0.432889
162	SERPINA3,ACT,AACT,MGC88254	0.428833
33595	NaN	0.393146
13220	XIST,XCE,XIC,SXI1,swd66,DXS1089,DXS399E	0.374803
5443	LAMB3,LAMNB1	0.363588

UPREGULATED GENES

Several of the most upregulated genes in Alzheimer's disease samples were related to immune and inflammatory responses. These included defensin family genes (such as DEFA1 and DEFA3), SERPINA3, and CHI3L2, which are commonly associated with activation of immune cells and inflammatory signalling. The increased expression of these genes suggests the presence of chronic inflammation in the Alzheimer's disease brain.

In addition, some upregulated genes were linked to cellular stress responses and structural components of tissue, indicating ongoing stress and remodelling within affected brain regions. The presence of regulatory

and non-coding transcripts among the upregulated genes further suggests widespread changes in gene regulation associated with Alzheimer's disease.

DOWNREGULATED GENES

```
downregulated_genes
```

	gene	log2FC
34521	FLJ25778	-2.225381
37685	CYTB,MTCYB	-2.152332
18338	ND3,MTND3	-2.057674
14006	COX2,MTCO2	-2.007301
10662	LOC440552	-1.961902
19635	COX2,MTCO2	-1.951621
18624	GFRA3	-1.848609
20509	COX2,MTCO2	-1.826349
30827	HSS00141642	-1.808065
18715	ND2,MTND2	-1.802870
19607	ND1,MTND1	-1.783856

Many of the most strongly downregulated genes in Alzheimer's disease samples were related to mitochondrial function and energy metabolism. These included mitochondrial genes such as **CYTB**, **ND1**, **ND2**, **ND3**, and **COX2**, which are essential components of the oxidative phosphorylation pathway responsible for ATP production.

The reduced expression of these genes suggests impaired mitochondrial activity in the Alzheimer's disease brain. Since neurons require large amounts of energy to maintain synaptic signalling and cellular function, decreased mitochondrial gene expression may contribute to neuronal dysfunction and degeneration observed in Alzheimer's disease.

Additionally, the downregulation of genes involved in neuronal growth and survival, such as **GFRA3**, further supports the presence of compromised neuronal health in affected brain tissue.

6) Limitations

This study has several limitations that should be considered when interpreting the results. First, the analysis was based on pooled brain tissue samples, which prevents assessment of individual-level variability and limits the use of formal statistical testing. As a result, the findings represent broad expression trends rather than statistically validated differential expression.

Second, the dataset was generated using microarray technology, which has a limited dynamic range compared to RNA sequencing and depends on predefined probes. This may restrict the detection of low-abundance or novel transcripts.

Finally, changes in gene expression observed in bulk brain tissue may reflect shifts in cell-type composition, such as increased immune cell activity or neuronal loss, rather than gene regulation changes within a single cell type. Despite these limitations, the analysis provides valuable insights into large-scale transcriptomic patterns associated with Alzheimer's disease

7) Conclusion

In this study, an exploratory transcriptomic analysis was performed to investigate gene expression changes associated with Alzheimer's disease in the human prefrontal cortex. By comparing mean RNA expression levels between Alzheimer's disease and control samples, patterns of gene upregulation and downregulation were identified.

The results revealed increased expression of genes related to immune and inflammatory processes, supporting the role of chronic neuroinflammation in Alzheimer's disease pathology. In contrast, genes involved in mitochondrial function and energy metabolism were consistently downregulated, suggesting impaired neuronal metabolic capacity in the diseased brain.

Together, these findings highlight Alzheimer's disease as a complex, multi-system disorder involving both immune activation and metabolic dysfunction rather than disruption of a single molecular pathway. This complexity may help explain why effective disease-modifying treatments remain limited. Exploratory transcriptomic analyses such as this provide an important foundation for future studies aimed at understanding disease mechanisms and identifying potential therapeutic targets