Alzheimer's Disease - Single Soma Transcriptomics Analysis

Name: Nicole Ma

1. Introduction and Data Overview

In this analysis, we use single-soma transcriptomics data to explore gene expression patterns linked to Alzheimer's Disease (AD)[1]. We focused on a list of 11 genes with the most substantial effect sizes, which are likely to reveal meaningful patterns between AD and normal subjects[2]. However, of the 11 genes initially identified, only six were present in the dataset: *SLC26A3, LINGO1, PDE4DIP, LINC01609, PHYHIP, and RASGEF1B.* The missing genes are *RP11-701H24.9, AC159540.1, RP11-289H16.1, RP11-219A15.1, and RP11-745L13.2.*

The absence of certain genes in the dataset may stem from several factors. Recognizing these limitations allows for a more informed analysis. According to Prof. Fang[3], a few potential reasons may explain the absence of these genes:

- **Dataset Filtering**: Some datasets are pre-filtered to remove genes with very low expression levels or genes classified as non-coding RNAs, focusing on high-confidence data.
- **Alternate Naming Conventions**: Gene names can vary across databases. For example, a gene listed under one alias might be named differently in this dataset, particularly with Ensemble IDs or other standard identifiers.
- **Non-Coding Gene Identifier Differences**: Non-coding RNAs and pseudogenes can have unique identifiers that vary by source, so they might not always align across datasets.
- **Specific Dataset Focus**: Given that this dataset targets specific cell types or expression characteristics, it might exclude genes less directly related to its central research goals.

2. Gene Correlation Analysis

In studying potential co-expression among AD-associated genes, notable correlations included a positive relationship between LINGO1 and RASGEF1B (0.54), hinting at a possible shared role in AD progression pathways. This finding suggests that these genes might interact within similar regulatory mechanisms.

3. Key Findings Related Disease Cohort Comparison

3.1. Data Insight

- **Upregulated Genes in AD**:

- **RASGEF1B**: Exhibits the highest upregulation (+0.89) in AD subjects. This gene's involvement in neuronal signaling pathways, particularly those regulating synaptic plasticity and neurogenesis, suggests it may play a role in cellular responses to neurodegeneration, supporting attempts to adapt and protect neuronal structures against AD progression.
- **LINGO1**: Increased expression (+0.72) may inhibit neuronal repair processes, exacerbating degradation.
- **SLC26A3**: Displays a moderate increase (+0.47), likely related to ion transport, which is crucial for cellular homeostasis.

- **Downregulated Gene in AD:**

- **PDE4DIP**: The sole downregulated gene (-0.32), which is involved in maintaining cellular structure, reflecting possible cytoskeletal disruptions in Alzheimer's-affected neurons.

3.2. Detailed Analysis of PDE4DIP: Alzheimer's Downregulated Gene

- **Functional Role of PDE4DIP:**

  PDE4DIP's anchoring role in microtubule assembly supports neuronal structure, and its interaction with PDE4D, involved in cAMP signaling, is essential for memory and cognitive function. Lower PDE4DIP expression in AD could destabilize neurons, impairing signaling pathways and increasing susceptibility to neurodegenerative damage, especially in regions critical for cognitive processing. [4]

- **Industry Practical Directions:**

  Microtubule instability is a key factor in AD progression. Targeting PDE4DIP to restore its function may strengthen neuronal resilience, providing a promising therapeutic approach to support neuron structure and potentially slow disease progression.

4. Age-Related Differences within AD and Normal Groups

4.1. Data Overview: Alzheimer's Disease (AD) Group Distribution

AD counts are heavily clustered within specific age groups, while other groups include only normal subjects. This uneven distribution is somewhat unexpected, as we would typically expect a mix of AD and normal subjects across all age stages.

4.2. Detailed Analysis: Age-Related Differences within AD and Normal Groups

- **PDE4DIP**: Shows an increase in expression with age in AD subjects, while in normal subjects, expression remains nearly stable across age groups. This suggests PDE4DIP may be more actively involved in AD progression in older subjects.
- **PHYHIP**: Displays a significant increase in expression with age in AD subjects but decreases in normal subjects, indicating a potential age-dependent role specific to AD progression.

4.3. Detailed Analysis: Age-Dependent Reversal of Gene Expression Patterns

- **Younger Cohort**: Both *LINGO1* and *RASGEF1B* are significantly upregulated in AD compared to normal subjects, indicating they may play active roles in the early stages of AD pathology.
- **Older Cohort**: In older subjects, *LINGO1* and *RASGEF1B* show a shift to downregulation in AD, suggesting that as AD progresses, the functional roles or regulatory mechanisms of these genes in AD pathology may change.
- **Findings**: This finding highlights the complexity of AD, where certain genes are not consistently upregulated or downregulated throughout disease progression. Instead, these genes seem to play more active roles in the early stages of AD and experience regulatory shifts as individuals age. This age-dependent gene expression pattern could provide insights into AD progression, emphasizing the need for age-specific therapeutic interventions.

5. Gender-Based Analysis

5.1. Data Overview

Alzheimer's Prevalence by Gender: Alzheimer's cases are notably more prevalent among females than males, which aligns with broader research indicating a higher risk of Alzheimer's in women. This gender-based difference emphasizes the need to explore potential sex-specific factors, whether genetic or lifestyle-related, that could influence Alzheimer's development and progression.

5.2. Detailed Analysis

- **RASGEF1B:**

- Difference in AD and Normal (Male): RASGEF1B expression shows a distinct contrast between AD and normal male subjects, with an AD-to-normal difference of 0.884, pointing to significantly higher expression in AD males. This elevated expression suggests that RASGEF1B may play a more active role in Alzheimer's pathology among male subjects.
- AD Gender Difference: Within the AD cohort, males show a slight increase in RASGEF1B expression compared to females, with a difference of 0.138. This could imply a gender-specific function for RASGEF1B in AD, meriting further research into how this gene might uniquely impact Alzheimer's in men.

- **PDE4DIP:**

- Unlike other genes, PDE4DIP is downregulated in AD compared to normal, particularly in male subjects. This decline may hint at a possible structural or protective role for PDE4DIP that becomes compromised in AD, especially among men, where the reduction is more pronounced.

- **PHYHIP:**

- Expression patterns for PHYHIP vary more subtly across groups. Female AD subjects show slightly higher PHYHIP expression than males, though overall trends are less clear. This inconsistent pattern suggests that PHYHIP's function might be influenced by gender, but its precise role remains less defined and may require further investigation.

Reference:

[1] Otero-Garcia *et al.*, “Single-soma transcriptomics of tangle-bearing neurons in Alzheimer’s disease,” *Neuron*, [Online]. Available: https://cellxgene.cziscience.com/collections/b953c942-f5d8-434f-9da7-e726ba7c1481

[2] “Differential Expression Analysis,” *CellxGene CZ Science*, [Online]. Available: https://cellxgene.cziscience.com/differential-expression.

[3] F. Fang, "Detective Time (Solve the Missing Genes Mystery): 10 bonus points," *Reply on Canvas Discussion*, Oct. 26, 2023.

[4] J. A. Potashkin, V. Bottero, J. A. Santiago, and J. P. Quinn, "Bioinformatic Analysis Reveals Phosphodiesterase 4D-Interacting Protein as a Key Frontal Cortex Dementia Switch Gene," *Int. J. Mol. Sci.*, vol. 21, no. 11, pp. 3787, May 2020, doi: 10.3390/ijms21113787.

Code:

https://github.com/NicoleMa1220/SEM/blob/main/Load_Specific_Dataset_and_Conduct_Cohort_Analysis.ipynb