Gene expression analysis of Alzheimer's Disease on single-nucleus level

Nikolay Makarenko, Rucsanda Juncu, Mieke Nicolaï
Supervisors: Darina Abaffyová, Alexandra Pančíková

KU LEUVEN

Introduction

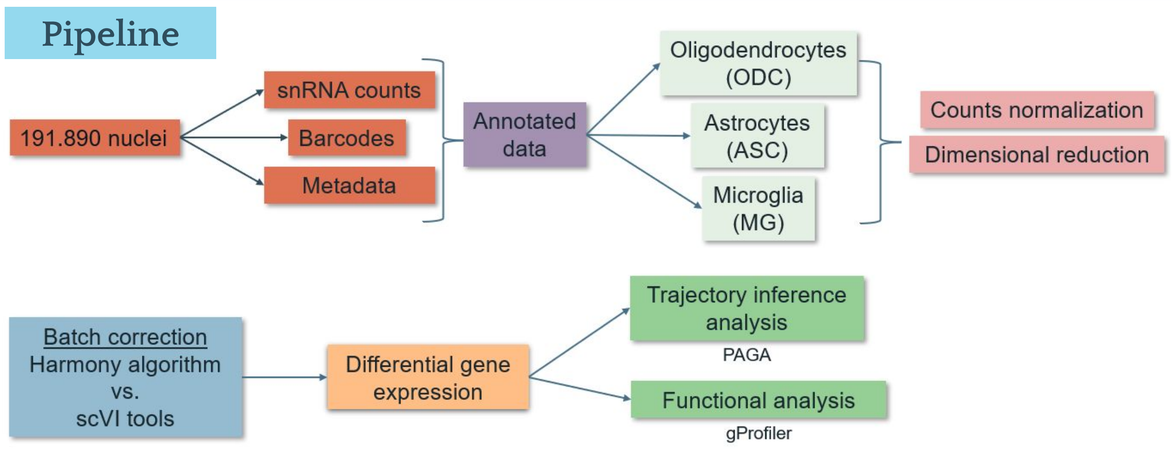
Alzheimer's Disease (AD) is a common neurodegenerative disease with high heritability. It leads to many psychological symptoms, such as memory loss, confusion, and depression. The mechanism of disease progression is unknown and is difficult to elucidate due to heterogeneous cell expression and regulation. AD is characterized by the accumulation of β-amyloid (Aβ) plaques in the neocortex, neurofibril formation, and inflammation.

Objectives

- Compare different methods to correct for batch effect
- Explore differential gene expression in AD for 3 subtypes using PAGA
- Perform functional analysis on the differentially expressed genes

Methodology

Pipeline



Data availability:

Code Availability:

Results

Batch correction

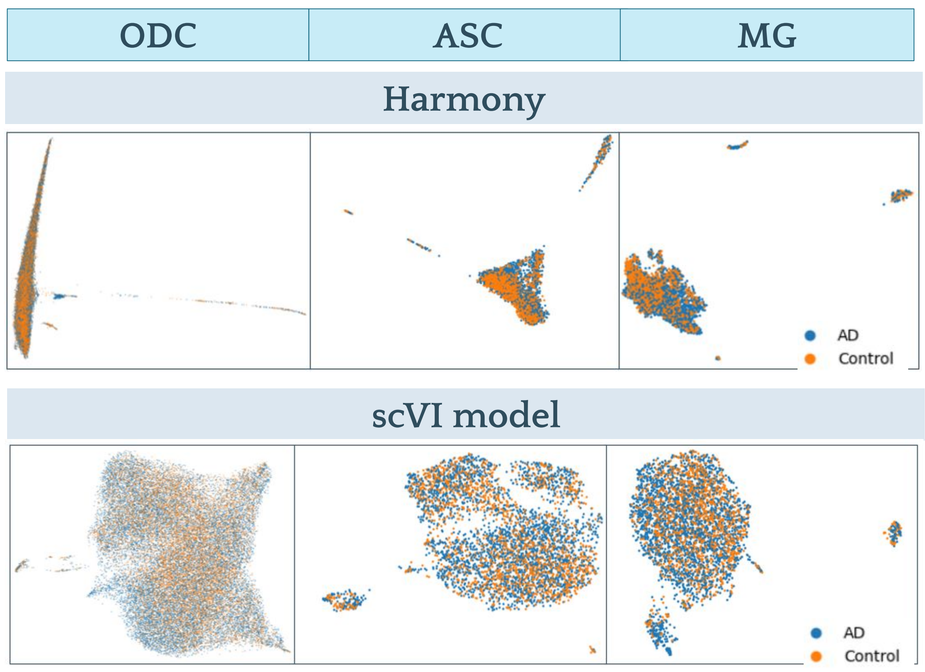


Figure 1. Comparison of the use Harmony algorithm and scVI model for batch correction. UMAPs of the data per cell type corrected with respectively Harmony and scVI model.

Correcting the batch effects in the data with the scVI model results in better clusters with less outliers than when corrected with the harmony algorithm.

Gene changes along PAGA paths

ODC

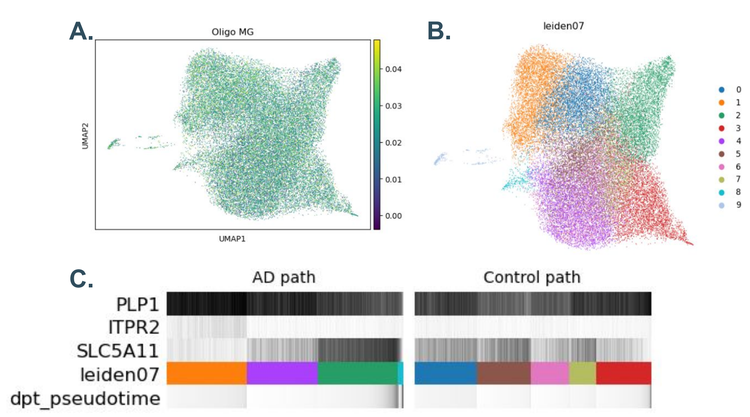


Figure 2. Gene changes in ODCs in AD. (a) UMAP showing the presence of ODC marker genes in all the clusters. (b) UMAP showing the clusters generated with the Leiden algorithm. (c) PAGA paths showing gene changes per cluster.

Visualization of ODC marker genes in UMAPs show similar expression in all clusters. No cluster filtering is necessary here. In ODCs, *PLP1* and *ITPR2* are upregulated at the onset of AD, but decreases during disease development. The opposite is observed in *SLC5A11*.

ASC

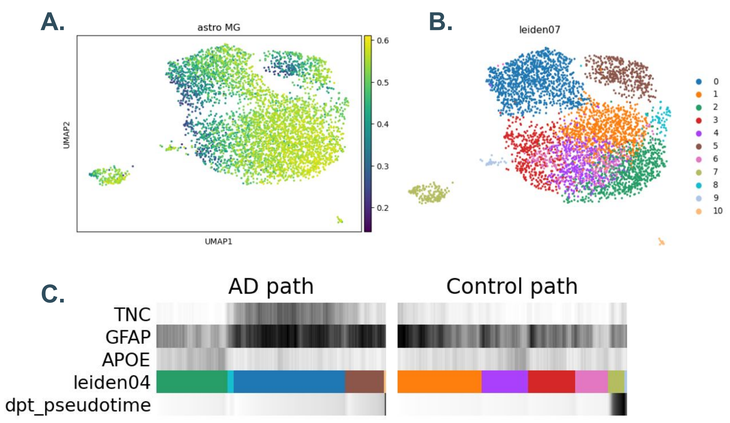


Figure 3. Gene changes in ASCs in AD. (a) UMAP showing the presence of ASC marker genes in all the clusters. (b) UMAP showing the clusters generated with the Leiden algorithm. (c) PAGA paths showing gene changes per cluster.

Visualization of ASC marker genes in UMAPs show sufficiently high expression in all clusters. Thus, again no cluster filtering is necessary here. *TNC* and *GFAP* are upregulated in mature AD ASCs, matching results from the literature. *APOE* shows downregulation in AD development.

MG

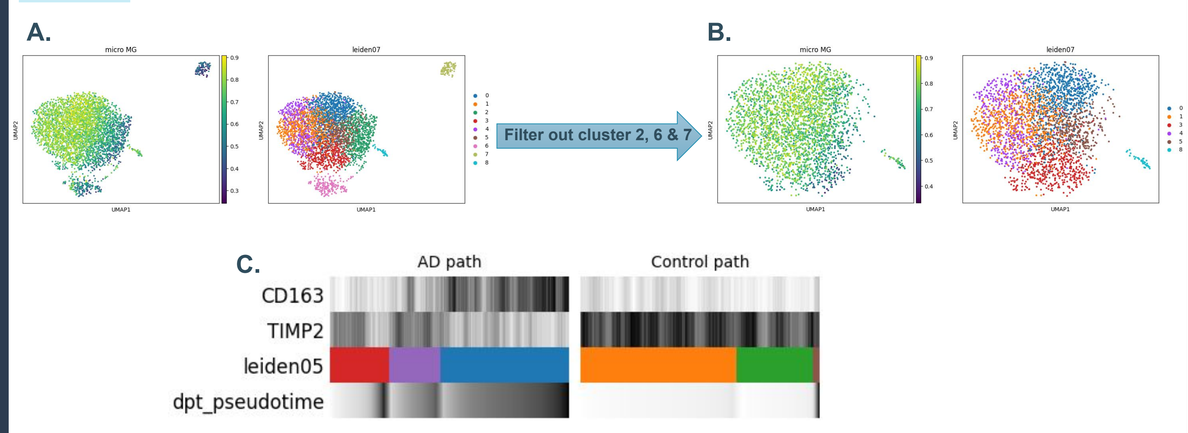


Figure 4. Gene changes in MGs in AD. (a) UMAP showing the presence of MG marker genes in all the clusters and a UMAP showing the clusters generated with the Leiden algorithm. (b) UMAPs with clusters showing low marker gene expression filtered out. (c) PAGA paths showing gene changes per cluster.

Gene marker exploration in the MG data showed that cluster 2, 6 and 7 did not contain high level expression of MG marker genes. Therefore these clusters were filtered out for further analysis. MG support previous results by showing upregulation of *CD163* in AD cells. A downregulation of *TIMP2* can be seen in AD cells and an overall lower expression than in control cells.

Functional analysis

	AD		CTRL	
	GO term	P-value	GO term	P-value
ODC	Protein folding Biological regulation	1.1×10^{-8} 5.6×10^{-11}	Protein binding Neurogenesis	4.2×10^{-8} 1.5×10^{-12}
ASC	Protein folding Response to stimulus	6.9×10^{-15} 4.5×10^{-24}	Catalytic activity Generation of neurons	2.1×10^{-4} 4.8×10^{-5}
MG	Protein folding Regulation of molecular function	2.9×10^{-7} 9.8×10^{-12}	Protein binding Nervous system development	2.1×10^{-4} 1.3×10^{-5}

Table 1. Functional analysis with g:Profiler. Associated GO terms with their p-value per cell type per condition.

Discussion

Batch correction using scVI gave more homogenous cluster dispersion and less outliers than with Harmony. Previously identified DEGs (Morabito, S., 2021) are shown in PAGA paths for each cell type in control and diseased states. *PLP1* and *ITPR2* are good markers for newly differentiated, myelin forming ODCs (Morabito et al.), while *SLC5A11* and *PDE1A* are upregulated in mature ODCs. Upregulation of structural and filament proteins *GFAP* and *TNC* is observed for ASCs, while *APOE* is downregulated. In MGs, anti-inflammatory marker *CD163* is upregulated in AD cells, while *TIMP2* is downregulated. Functional analysis using g-Profiler confirmed association of AD DEGs with protein folding, which should be explored as a connection to neurofibrillary tangle formation.

References

Gatz, M. et al. Role of genes and environments for explaining Alzheimer disease. Arch. Gen. *Psychiatry 63*, 168–174 (2006).
Brown, M. R., Radford, S. E., & Hewitt, E. W. (2020). Modulation of β-Amyloid Fibril Formation in Alzheimer's Disease by Microglia and Infection. *Frontiers in molecular neuroscience*, 13, 609073.
Morabito, S., Miyoshi, E., Michael, N., Shahin, S., Martini, A. C., Head, E., Silva, J., Leavy, K., Perez-Rosendahl, M., & Swarup, V. (2021). Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. Nature genetics, 53(8), 1143–1155.