Understanding Microglial Phenotype Switching Through Dimensionality Reduction and Clustering Analysis

Anthony Wolfe[+], Kamila Larripa [×]

[+] indicates Student Presenter, [×] indicates Faculty Advisor

Research Objectives

We aim to answer the following questions utilizing machine learning and single cell analysis:

- Do distinct microglia clusters in our dataset correspond to known biological phenotypes?
- How do these phenotypic clusters change over time under neurodegenerative conditions?
- Are there transcriptional states that do not align with canonical phenotypes?
- How well do literature-derived gene signatures align with data-driven cluster identities?

Introduction

Microglia are resident immune cells of the central nervous system that regulate neural homeostasis through dynamic functional states, or phenotypes. Dysregulation of these states is implicated in neurodegenerative diseases such as Alzheimer's disease (AD). Advances in single-cell RNA sequencing (scRNA-seq) enable characterization of cellular heterogeneity, allowing phenotypes to be defined by gene expression patterns. In this study, we analyze hippocampal microglia from transgenic mice across four time points over a six-week period to investigate temporal changes in transcriptional states and their relationship to known and potentially novel microglial phenotypes and biological signatures (Mathys et al., 2017).

Data & Methods

- GEO scRNA seq data from hippocampal microglia in transgenic mice, modeling AD-like neurodegeneration, collected at four time points over six weeks to capture temporal progression
- Data processed in Python using Scanpy; gene annotation via AnnData & MyGene, phenotype definitions derived from literature
- Quality control and gene signature scoring to quantify M1/M2-associated expression patterns
- PCA to assess variance structure, followed by UMAP for visualization and Leiden clustering to identify cell populations
- Cluster relationships evaluated using differential expression, dot plots with hierarchical clustering, and temporal dynamics assessed across time

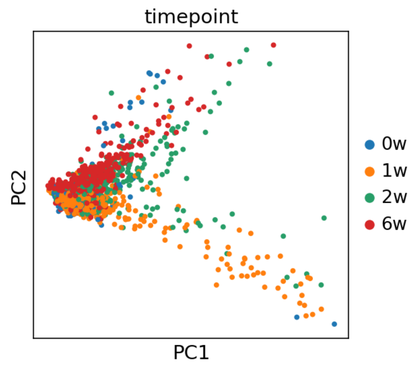
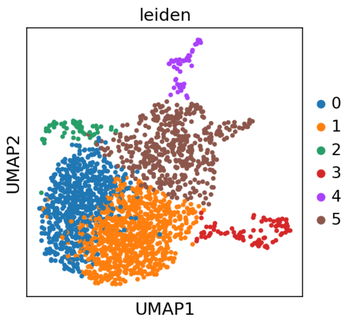
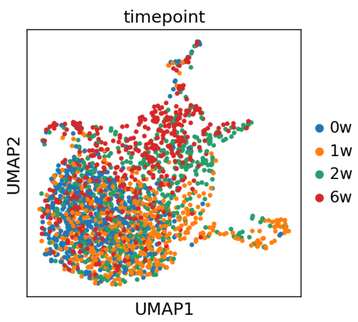


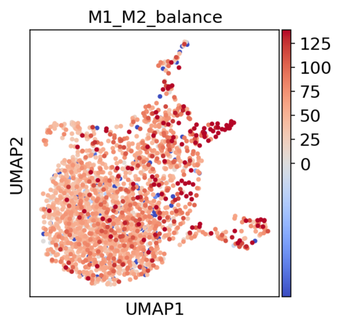
Figure: PCA reveals time-dependent shifts in gene expression



(a) UMAP leiden clusters separate into distinct transcriptional populations.



(b) UMAP's temporal structure emerges across time points.



(c) UMAP colored by gene expression shows phenotype gradients.

Analysis & Results

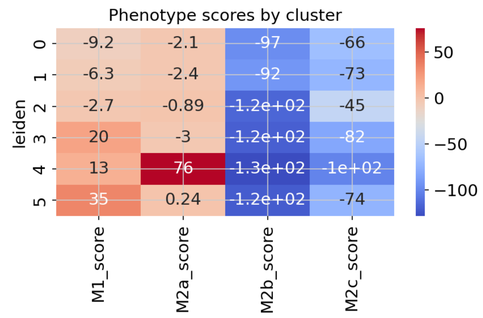


Figure: Cluster-specific gene expression patterns show partial alignment with canonical phenotypes, with M1-like signatures in select clusters and limited representation of M2 subtypes.

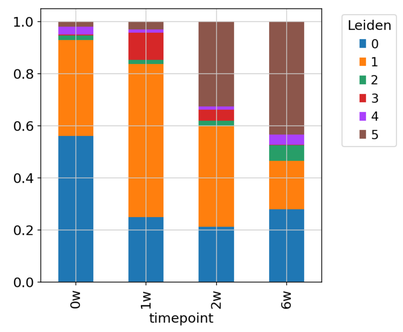


Figure: Cluster proportions shift over time, with certain clusters emerging and declining across time points, notably clusters 5 and 3.

- Temporal shifts in cluster composition reveal dynamic microglial state transitions, with expansion of late-stage clusters and reduction of early-state populations.
- Leiden clustering identifies 6 distinct transcriptional populations with clear time-dependent structure.
- Cluster-specific expression partially aligns with canonical M1/M2 phenotypes, with strong M1-like signatures and limited M2 composition.
- Several clusters exhibit mixed or weak phenotype scores, suggesting transcriptional heterogeneity and potentially uncharacterized microglial states.

Conclusion & Significance

- Microglial states evolve dynamically over time and form distinct transcriptional clusters.
- Canonical M1/M2 phenotypes partially explain the dataset's structure, indicating a spectrum of cell states revealed by our single-cell analysis.

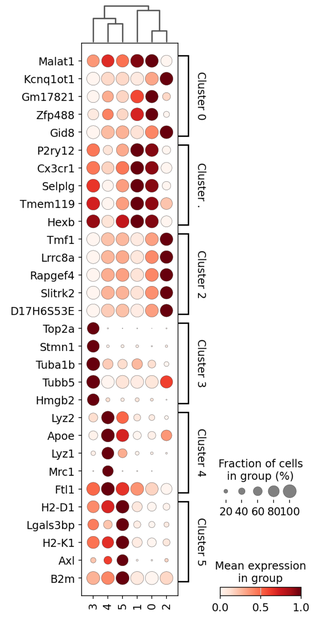


Figure: Dotplot of top marker genes shows clusters defined by activation (Top2a, Stmn1), immune-associated (Lyz2, Axl), and homeostatic (Cx3cr1, P2ry12) gene expression programs.

Supplemental Information

Acknowledgements & Data Source

Supported by NSF grant 2245839 through the Division of Mathematical Sciences Mathematical Biology program and the Systems and Synthetic Biology Program.

Data source: Mathys et al. (2017), hippocampal microglia single-cell RNA-seq data from transgenic mice over six weeks.