Single-nuclei epigenomics and transcriptomics to uncover the gene regulatory landscape in Alzheimer's disease

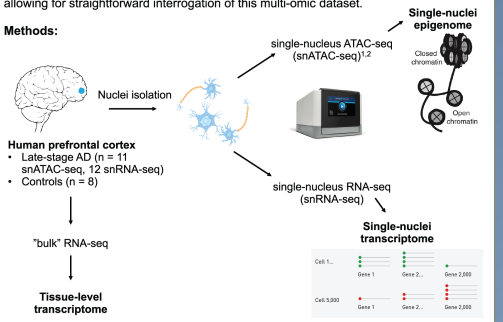
Samuel Morabito[1,3], Emily Miyoshi[2,3], Neethu Michael[2,3], Saba Shahin[2,3], Vivek Swarup[2,3]

1. Mathematical, Computational and Systems Biology (MCSB) Program, University of California, Irvine, CA 92697, USA
2. Department of Neurobiology and Behavior, University of California, Irvine, CA 92697, USA
3. Institute for Memory Impairments and Neurological Disorders (MIND), University of California, Irvine, CA 92697, USA

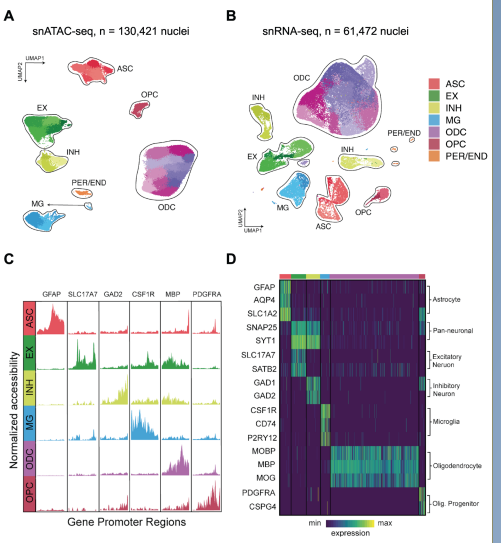
Abstract

The gene regulatory landscape of the brain is highly dynamic in health and disease, coordinating a menagerie of biological processes across distinct cell-types. Fully contextualizing abnormal molecular signatures in disease with respect to specific cell-types requires a holistic multi-layered experimental and analytical approach. Single-cell transcriptomics has been used extensively in human disease systems; however, very few single-cell epigenomic studies have been carried out in primary disease samples. Here, we present a multi-omic single-cell study of 191,897 nuclei in late-stage Alzheimer's Disease (AD), in which we profiled and analyzed chromatin accessibility and gene expression in the same biological samples, uncovering vast glial heterogeneity in late-stage AD. We describe cis-regulatory relationships in specific cell-types at AD risk loci, defined by genome wide association studies (GWAS), demonstrating the utility of this multi-omic single-cell framework for uncovering disease and cell-type-specific regulatory mechanisms. In addition, we characterize transcription factor regulatory patterns in the transition between healthy and diseased states through trajectory analysis of glial populations. We also introduce scWGCNA, a co-expression network analysis strategy, robust to the sparsity of single-cell data, to perform a systems-level meta-analysis of AD transcriptomics. Finally, this work is highly accessible through our intuitive web-portal, allowing for straightforward interrogation of this multi-omic dataset.

Methods:

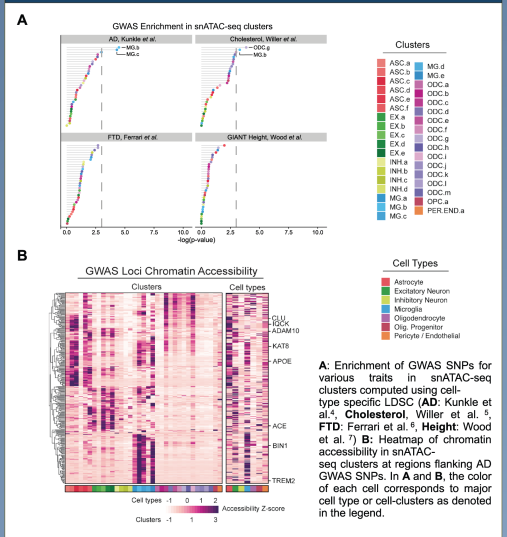


snATAC-seq and snRNA-seq to describe cellular heterogeneity in the aged human cortex



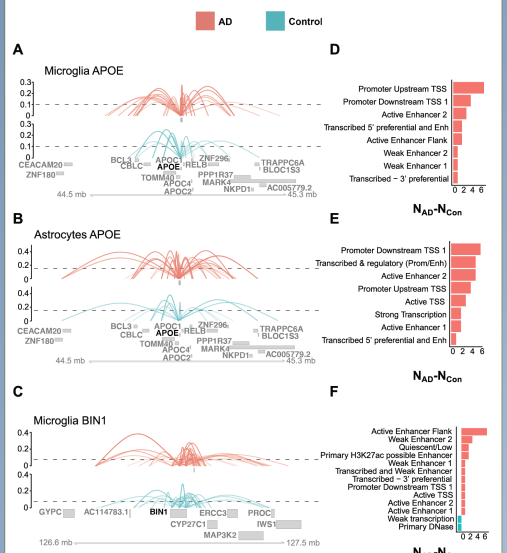
A-B: Uniform Manifold Approximation and Projection[3] (UMAP) visualizations of 130,425 nuclei profiled with snATAC-seq **(A)** and 61,472 nuclei profiled with snRNA-seq **(B)**, colored by major cell-type as denoted in the legend. **C:** Pseudo-bulk chromatin accessibility plots for canonical cell-type marker genes. **D:** Heatmap of normalized gene expression of cell-type marker genes. Genes shown in **C** and **D** as well as other cell type marker genes were used in order to determine the cell-type of each major cluster in snATAC-seq and snRNA-seq.

Cell-type specific gene regulation of AD GWAS loci



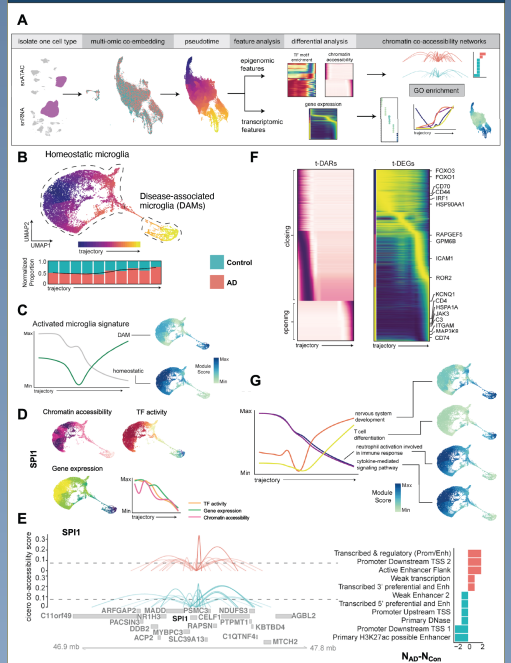
A: Enrichment of GWAS SNPs for various traits in snATAC-seq clusters computed using cell-type specific LDSC (**AD**: Kunkle et al.[4], **Cholesterol**, Willer et al. [5], **FTD**: Ferrari et al. [6], **Height**: Wood et al. [7]) **B:** Heatmap of chromatin accessibility in snATAC-seq clusters at regions flanking AD GWAS SNPs. In **A** and **B**, the color of each cell corresponds to major cell type or cell-clusters as denoted in the legend.

Cell-type and disease specific cis-regulatory connections at AD GWAS loci



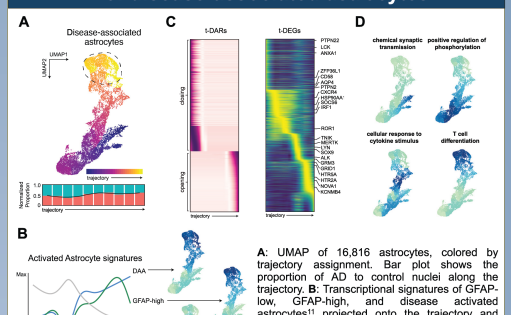
A, B, C: Co-accessible regulatory links from cell-type specific cis-co-accessibility network analysis (CCANs) using the R package cicero[8] at the APOE GWAS loci in microglia **(A)** and astrocytes **(B)** and at the BIN1 **loci in microglia (C)**. **D, E, F:** Bar plots are the number of annotated connections in AD minus control for **(A), (B)** and **(C)**, respectively.

Epigenetic and transcriptional characterization of disease-associated microglia



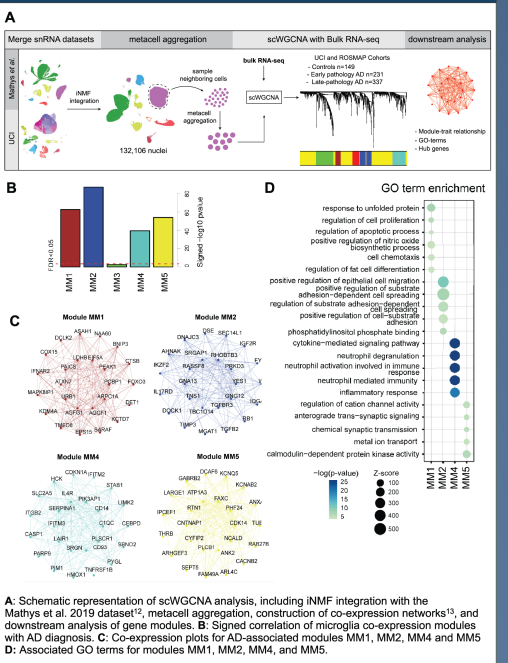
A: Schematic representation of cell-type specific multi-omic trajectory analysis using monocle[9]. **B**: UMAP of 14,887 microglia, colored by trajectory assignment. Bar plot shows the proportion of AD to control nuclei along the trajectory. **C**: Transcriptional signatures of homeostatic and disease associated microglia[10] projected onto the trajectory and shown as UMAP feature plots. **D**: UMAP feature plots and trajectories of chromatin accessibility, motif enrichment, and gene expression for SPI1. **E**: CCANs for SPI1. Bar plot is the number of annotated connections in AD minus control. **F**: Heatmaps of chromatin accessibility for 2391 t-DARs and of gene expression for 2138 t-DEGs. **G**: Select GO terms for microglial t-DEGs projected onto the microglia trajectory and shown as UMAP feature plots.

Epigenetic and transcriptional characterization of disease-associated astrocytes



A: UMAP of 16,816 astrocytes, colored by trajectory assignment. Bar plot shows the proportion of AD to control nuclei along the trajectory. **B**: Transcriptional signatures of GFAP-low, GFAP-high, and disease activated astrocytes[11] projected onto the trajectory and shown as UMAP feature plots. **C**: Heatmaps of chromatin accessibility for 928 t-DARs and of gene expression for 1797 t-DEGs. **D:** Select GO terms for astrocyte t-DEGs shown as UMAP feature plots.

Single-cell gene co-expression networks using scWGCNA



A: Schematic representation of scWGCNA analysis, including iNMF integration with the Mathys et al. 2019 dataset[12], metacell aggregation, construction of co-expression networks[13], and downstream analysis of gene modules. **B**: Signed correlation of microglia co-expression modules with AD diagnosis. **C**: Co-expression plots for AD-associated modules MM1, MM2, MM4 and MM5 **D:** Associated GO terms for modules MM1, MM2, MM4, and MM5.

Conclusion

We present a rigorous multi-omic analysis of 191,897 nuclei to interrogate cell-type specific epigenomic and transcriptomic changes occurring in late-stage human AD. Network analysis of the chromatin accessibility landscape identified thousands of cis co-accessibility networks that have altered topology in disease and that are cell-type specific, highlighting the vast systems level perturbations occurring in AD. In addition, modern data integration techniques allowed us to simultaneously analyze single-nuclei open-chromatin profiles and transcriptomes, thus we examined the cellular heterogeneity in glia, identifying the spectrum of epigenomic and transcriptomic changes involved in the glial immune response in AD. Further, our data serves as an excellent public data resource for other researchers to ask specific questions about cell-type specific changes in gene expression and gene regulation using our data, which we may not have looked at in our analysis.

In the future, we wish to extend our work by studying more regions of the brain using the same sequencing techniques, as well as using new techniques to spatially resolve gene expression in a tissue of interest.

References

1. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10, 1213–1218 (2013).
2. Buenrostro, J. D. *et al.* Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 523, 486–90 (2015).
3. Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* 37, 38–44 (2018).
4. Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat Genet* 51, 414–430 (2019).
5. Consortium, G. L. G. *et al.* Discovery and refinement of loci associated with lipid levels. *Nat Genet* 45, 1274–1283 (2013).
6. Ferrari, R. *et al.* Frontotemporal dementia and its subtypes: a genome-wide association study. *Lancet Neurology* 13, 686–699 (2014).
7. Consortium, T. E. M. R. and G. (eMERGE) *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* 46, 1173–1186 (2014).
8. Pliner, H. A. *et al.* Cicero Predicts cis-Regulatory DNA Interactions from Single-Cell Chromatin Accessibility Data. *Mol Cell* 71, 858-871.e8 (2018).
9. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 32, 381–6 (2014).
10. Keren-Shaul, H. *et al.* A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell* 169, 1276-1290.e17 (2017).
11. Habib, N. *et al.* Disease-associated astrocytes in Alzheimer's disease and aging. *Nat Neurosci* 1–6 (2020) doi:10.1038/s41593-020-0624-8.
12. Mathys, H. *et al.* Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* 570, 332–337 (2019).
13. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *Bmc Bioinformatics* 9, 559 (2008).