Transcriptomic cell type analysis of human middle temporal gyrus in the context of Alzheimer's disease

Twitter @neurojoy

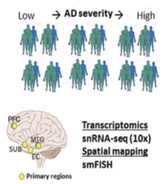
A. BUCHIN[1], R. D. HODGE[1], K. A. SMITH[1], J. GOLDY[1], N. JORSTAD[1], A. TORKELSON[1], E. S. KAPLAN[1], E. J. MELIEF[2], M. J. HAWRYLYCZ[1], D. C. KEEN[2], E. S. LEIN[1,2] and J.. A. MILLER[1]

[1]Allen Institute For Brain Science, Seattle, WA; [2]University of Washington, Seattle, WA

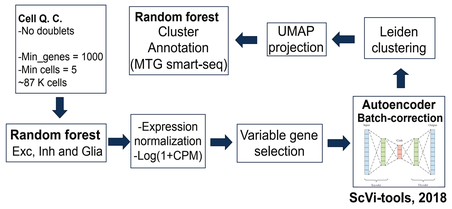
Summary

- Dementia affects nearly ~11% of the U.S. population over the age of 65, with the incidence doubling every 5 years up to 40-60% after the age of 90. Around 2/3 of these cases are correlated with Alzheimer's disease.
- **The key strategy to identify AD-related changes in cell type proportions and properties is to generate a well-annotated reference cell type classification of the neurotypical brain, and then map AD data to that reference to discern pathological changes in specific cell types during disease progression.**
- We present a workflow for constructing a neurotypical brain reference classification and strategies for mapping disease datasets for middle temporal gyrus (MTG) based on 10x Genomics single nucleus RNA sequencing data.
- Non-linear dimensionality reduction and clustering of nearly 90 thousand nuclei from 3 patients identified 29 glutamatergic neuron, 41 GABAergic neuron and 10 non-neuronal types, very similar to recently published results using SMART-seq methodologies [Hodge et al. 2019].
- We applied neural network and random forest classifiers to identify known cell types and cell classes in the tissue from aged and Alzheimer's patients.
- We found that that the new reference dataset based on 10X Genomics MTG data could be successfully used to identify known cell types in AD tissues based on patterns of gene expression.

Patient and region selection



RNA-seq clustering workflow



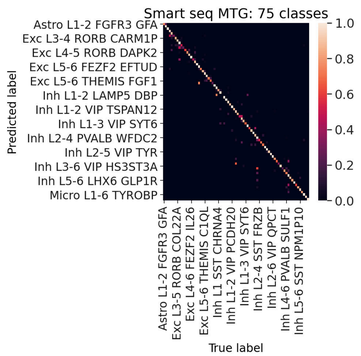
1

Transcriptomic analysis

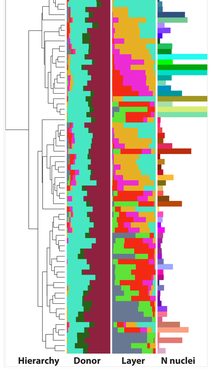
Human MTG smart-seq

~14 000 cells (smart-seq)
~90 000 cells (10x)
[Hodge et al 2019]
https://portal.brain-map.org

Random forest Classifier
Precision: 90%

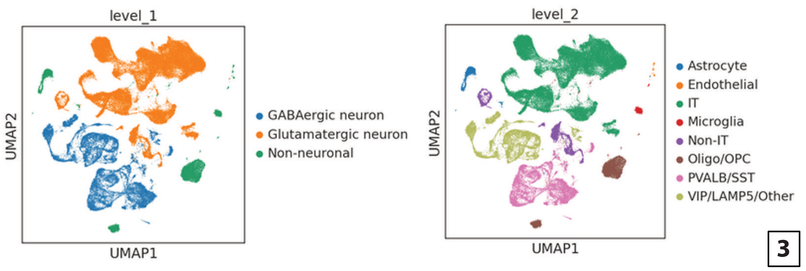


Cells info



2

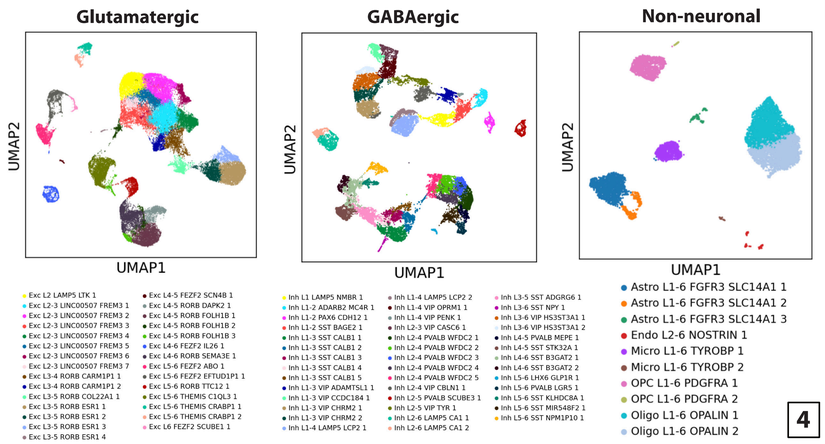
Initial identification of cells major classes in 10x data



3

De novo clustering of 10x MTG data

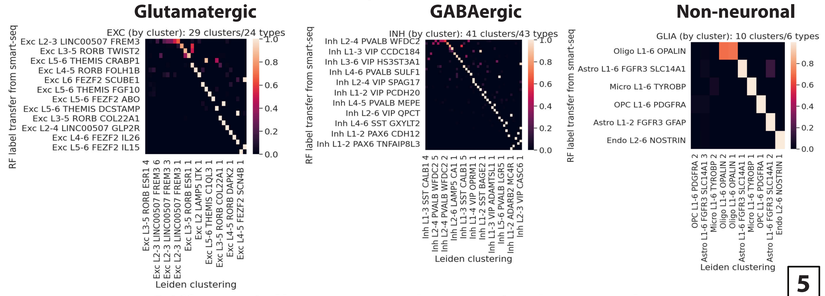
-Cluster names are assigned based on [Hodge et al. 2019] using the Random Forest classifier
-The clusters are named after the majority class in every cluster



4

Comparisons of Hodge et al 2019 and *de novo* clustering

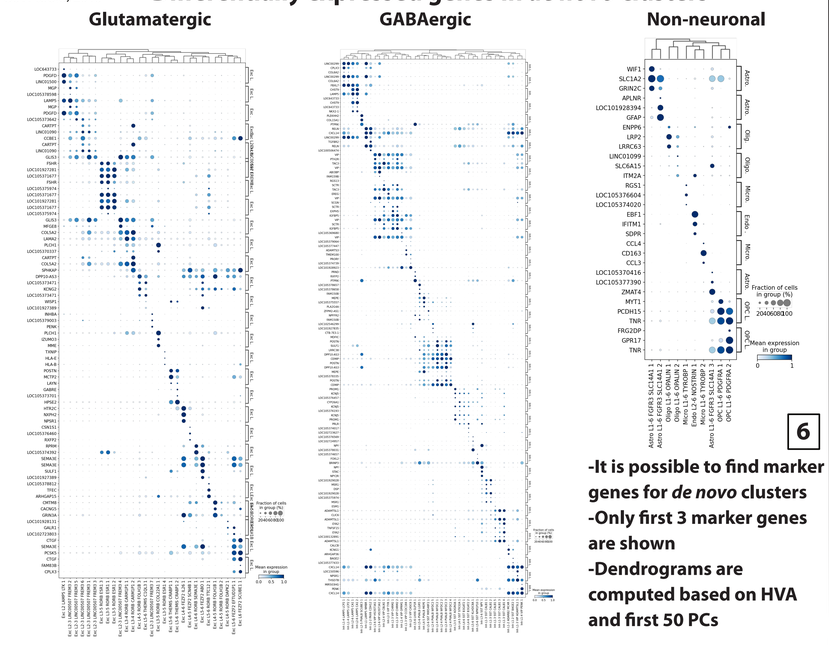
-The novel clusters are similar to smart-seq Random forest prediction



5

ScVi-tools, 2018

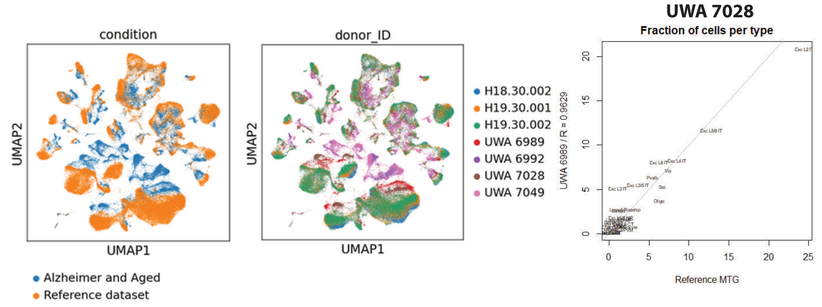
Differentially expressed genes in *de novo* clusters



6

-It is possible to find marker genes for *de novo* clusters
-Only first 3 marker genes are shown
-Dendrograms are computed based on HVA and first 50 PCs

Donor analysis and cell type proportions in Alzheimer disease data

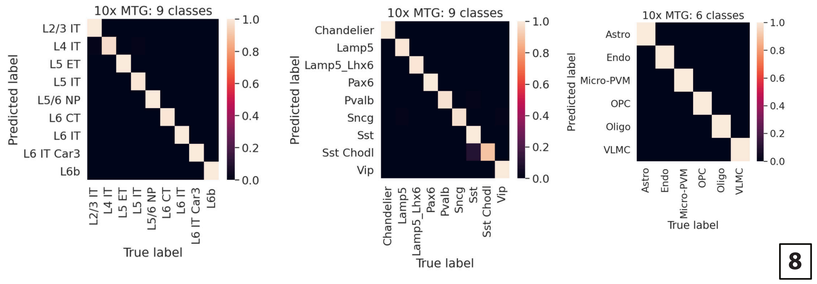


-There are expected disease-specific differences between Alzheimer and MTG reference data
-There is minor patient-specific clustering, but the majority of clusters are overlapped
-The number of cells in diseased and MTG reference is linearly correlated

7

Training Random Forest classifier based on10x MTG reference on the subclass level

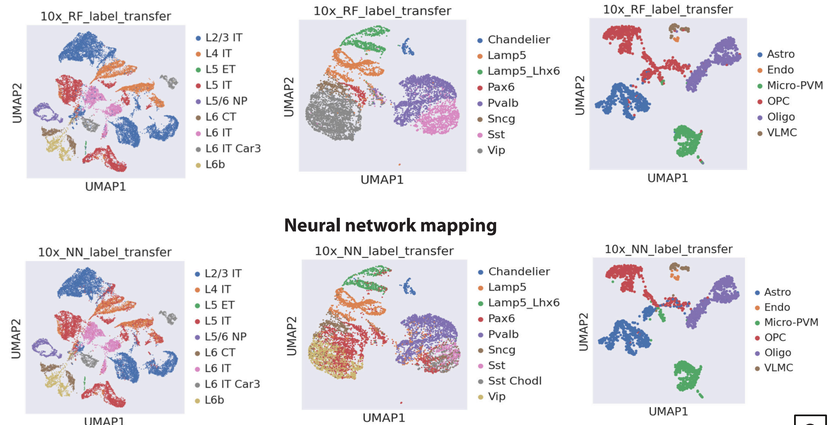
-Performance of the Random Forest classifiers trained on the subclass level (test set)
-There is high precision of the classification for all subclasses



8

Mapping of the Alzheimers disease data to MTG reference

-Random Forest and Neural Network classifiers provide consistent mapping of the Alzheimer data to MTG reference on the subclass level
-The inhibitory neurons are not consistently mapped between two classifiers



9

We wish to thank the Allen Institute for Brain Science founder, Paul G. Allen, for his vision, encouragement and support

brainscience.alleninstitute.org
brain-map.org