

© Copyright 2025

Victoria Mallett Rachleff

Disentangling the relationship between pathology and cellular vulnerability in
Alzheimer's disease

Victoria Mallett Rachleff

A dissertation
submitted in partial fulfillment of the
requirements of the degree of

Doctor of Philosophy

University of Washington
2025

Reading Committee:

C. Dirk Keene

Ed Lein

Rebecca Hodge

Jeffrey Iliff

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

Disentangling the relationship between pathology and cellular vulnerability in Alzheimer's disease

Victoria Mallett Rachleff

Chair of Supervisory Committee

C. Dirk Keene

Department of Laboratory Medicine and Pathology

Alzheimer's Disease (AD) is the most common cause of dementia affecting approximately 57 million people worldwide. Currently, no effective preventions, treatments, or cures exist. This work ventures to study AD from a new perspective; approaching AD as a cellular disorder that preferentially affects unique subpopulations of cells over time as the disease progresses. The goal of this work is to identify cells vulnerable to dysregulation and cell death at the earliest stages of disease and investigate the role of pathogenic protein accumulation in these processes. The overarching hypothesis of this work was that **selectively vulnerable populations of cells preferentially accumulate pathologic peptides and can be identified transcriptomically based on changes in relative abundance and cell state associated with disease progression**. This dissertation leverages the massive data collection effort of the multi-center Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD) initiative, which includes single-nucleus RNA sequencing (snRNA-seq) and quantitative neuropathology data from a clinical cohort spanning a spectrum of AD pathology. Additionally, high resolution multiomic data, including spatial transcriptomics and pathologic peptide-based single cell RNA sequencing on a subset of SEA-AD donors enables direct assessment of the relationship between protein pathology and cellular vulnerability. Upon completion, the findings generated from this work will serve as a resource for mapping the cellular changes in AD and will identify opportunities for therapeutic intervention at the earliest stages of disease.

TABLE OF CONTENTS

LIST OF FIGURES.....	iii
LIST OF TABLES.....	v
ACKNOWLEDGEMENTS.....	vi
 Chapter 1. INTRODUCTION TO HUMAN BRAIN PROCESSING AND CELL TYPES.....	1
1.1 Standardization and modernization of human postmortem brain tissue processing to enable modern molecular approaches to study the brain in health and disease.....	1
1.2 Challenges with adoption and standardization of best practices for the processing of human postmortem brain tissue.....	4
1.3 Generation of transcriptomics-based human brain cell atlases.....	5
 Chapter 2. ADVANCEMENTS IN QUANTITATIVE NEUROPATHOLOGY.....	8
2.1. Introduction.....	8
2.2. Methods.....	21
2.3 Results.....	26
2.4. Discussion.....	35
 Chapter 3. LEVERAGING SINGLE CELL TRANSCRIPTOMICS TO STUDY THE RELATIONSHIP BETWEEN VULNERABLE CELLS AND PROTEIN PATHOLOGY IN AD.....	37
3.1. Introduction.....	38
3.2. Methods.....	40
3.3. Results.....	67
3.4. Discussion.....	92
 Chapter 4. DISENTANGLING THE RELATIONSHIP BETWEEN VULNERABLE CELL TYPES AND PROTEIN PATHOLOGY IN AD USING SPATIAL TRANSCRIPTOMICS.....	98
4.1. Introduction.....	98
4.2. Methods.....	100

4.3. Results.....	116
4.4. Discussion.....	124
Chapter 5. CLOSING REMARKS.....	127
5.1. Future directions for quantitative neuropathology and multiomics in AD.....	129
References.....	132

LIST OF FIGURES

Figure 1.1 Allen Institute Cellular Taxonomy.

Figure 2.1. Representative images of pTau and A β pathology in postmortem brain tissue.

Figure 2.2. Stereotyped spatiotemporal distribution of pTau NFT pathology and A β pathology in AD.

Figure 2.3. Change in cognitive score in response to Lecanemab treatment.

Figure 2.4. Schematic detailing the experimental design for the SEA-AD study.

Figure 2.5. Inverse correlation of pathology and cognitive impairment in AD.

Figure 2.6. SEA-AD cohort demographics.

Figure 2.7. Schematic depicting quantitative neuropathology workflow.

Figure 2.8. Representative cortical column visualized with immunohistochemical (IHC) staining.

Figure 2.9. Correlation heatmap of normalized quantitative neuropathological features.

Figure 2.10. Boxplots quantifying the number of pTau-bearing cells per unit area.

Figure 2.11. Barplot summarizing results from a generalized linear model predicting dementia status.

Figure 2.12. Heatmap of normalized neuropathological features.

Figure 2.13. Specific cortical layers preferentially accumulate pTau.

Figure 2.14. Pseudoprogression score as continuous metric to order donors by pathological burden.

Figure 3.1. Power calculation for pTau-seq.

Figure 3.2. Representative FACS plots from two different batches.

Figure 3.3. Representative bright field image pre and post FACS sorting.

Figure 3.4. Subclass assignments align with standard cell-type marker genes.

Figure 3.5. scVI integration of SEA-AD and pTau-seq datasets with high label transfer confidence.

Figure 3.6. Representative subclass-level QC workflow.

Figure 3.7. Subclass-level UMAP representation of integrated dataset pre-QC and post-QC.

Figure 3.8. pTau-seq UMAP representation of donor and library preparation.

Figure 3.9. Relative cellular abundances and pTau percentage by supertype and subclass.

Figure 3.10. Relative contribution of each donor to subclass proportions.

Figure 3.11. Donor-level pTau distributions in UMAP space.

Figure 3.12. Heatmap highlighting low abundance subclasses and superotypes per donor.

Figure 3.13. Integrated SEA-AD and pTau-seq dataset pseudocolored by class and subclass.

Figure 3.14. pTau-seq dataset split by pTau status pseudocolored by subclass.

Figure 3.15. pTau-positive fractions track with Braak Stage.

Figure 3.16. Comparison of pTau positivity between Otero-Garcia et al., 2022 and pTau-seq.

Figure 3.17. Known AD and senescence-related genes associated with pTau accumulation.

Figure 3.18. High fraction of pTau-positive cells in both expected and unexpected supertypes.

Figure 3.19. Vulnerable supertypes in SEA-AD track with significant pTau enrichment in pTau-seq.

Figure 3.20. Relative abundance changes in pTau-seq at the subclass-level using scCODA.

Figure 3.21. Relative abundance changes in pTau-seq at the supertype-level using scCODA.

Figure 3.22. Supertypes with top upregulated and downregulated genes across pTau status.

Figure 3.23. Heatmap of top differentially expressed genes across selected vulnerable supertypes.

Figure 3.24. Top differentially expressed genes in vulnerable excitatory populations.

Figure 3.25. Top differentially expressed genes in vulnerable inhibitory populations.

Figure 4.1. Cell segmentation boundaries following custom re-segmentation.

Figure 4.2. Consistent high-quality metrics across donors in proof-of-concept dataset.

Figure 4.3. Negative probe detection indicates low background across tissue.

Figure 4.4. Spatial assessment of low-quality cells.

Figure 4.5. Assessment of mapping confidence at the supertype level.

Figure 4.6. Ranked list of low abundance supertypes (<10 cells).

Figure 4.7. Spatial distribution of subclass-level laminar structure.

Figure 4.8. Ranked list of subclass proportions.

Figure 4.9. Spatial distribution of inhibitory interneuron subclass-level marker genes (Pvalb and Sst).

Figure 4.10. Spatial distribution of selected neuronal and non-neuronal subclasses.

Figure 4.11. UMAP representation of subclasses with mapping confidence and quality metrics.

Figure 4.12. pTau-positive cells highlighted in the context of subclass level distributions.

Figure 4.13. AD-related and DE genes visualized in spatial context

LIST OF TABLES

Table 1.1. Postmortem brain tissue quality control metrics.

Table 2.1. Complete SEA-AD cohort demographic information.

Table 2.2. Quantified features extracted from HALO image processing.

Table 3.1. pTau-seq donor cohort with associated metadata.

Table 3.2. FACS cell counts following pTau (AT8) sorting.

Table 3.3. Capture rate of single somas following 10x processing.

Table 3.4. Empirically defined subclass level QC thresholds.

ACKNOWLEDGEMENTS

This work would not have been possible without the support and guidance of so many wonderful people - family, friends, colleagues, and mentors. Many aspects of my Ph.D. were unconventional - starting the Molecular and Cellular Biology Ph.D. program at the onset of global COVID-19 pandemic, pausing my Ph.D. lab rotations to volunteer at the UW Virology testing center, followed by a formal lab switch to join Dr. Dirk Keene's UW Neuropathology group and Dr. Ed Lein's Human Cell Types program at the Allen institute, and ultimately the birth of my two incredible daughters. There were ups, downs, and pauses - and my incredible support system did not waiver once. Thank you from the bottom of my heart to everyone who helped make this lifelong dream of mine become a reality.

First, I want to thank my grandfather, Dr. Loren Corotto, who inspired my pursuit to better understand the human brain. It still strikes me - how could it be that we don't understand the one organ in our body that shapes the very essence of who we are. As a Ph.D.-trained psychologist, he opened my eyes to the awesome potential of our brain - in both health and disease. The stories he shared with me from his days working at the Napa State Hospital will always stay with me. His work, at the behavioral level, directly helped those that needed it most. His memory, presence, and values have, and always will, stay with me in my scientific career and beyond - thank you, Papa.

Next, I would like to thank my supervisors, Dr. Dirk Keene and Dr. Ed Lein, for their unwavering support throughout the past five years. I had the honor of learning from them as leaders in the field, but more importantly, as individuals who are thoughtful, ethical, caring, and collaborative - genuinely good people. The example they have set for me as a person and scientist is second to none and I will strive to maintain their level of commitment and integrity as I continue in my career. As I reflected on the process, I remembered that neither Dirk nor Ed were looking for a Ph.D. student when I approached them to inquire about joining their research programs. They agreed to jointly supervise me as a Ph.D. student anyways, and despite being extremely busy, never let me down once throughout the entire process. I have immense respect

for them both and am very thankful for the incredible mentorship they provided and the collaborative world of team science they introduced me to.

Back to family - my best friend and husband, Jake, has made so many sacrifices to support me and my goal to pursue a Ph.D. Leaving this part short and sweet, I acknowledge and appreciate every little thing that you do on an hourly, daily, weekly, monthly, yearly scale. None of it goes unnoticed, and I can confidently say that without your commitment over the past few years, I very likely would not have made it to the Ph.D. finish line. You see me for everything I am and everything I want to be. You are the best teammate (and soulmate) in the world - thank you for always helping me make my dreams come true.

To my parents - mom and dad, thank you for shaping me into the person that I am today and helping guide me without ever stepping on my toes. You let me make (many) mistakes and be better for them. I have always been able to feel your unconditional love and unwavering belief in me. You both have played a central part in making every single one of my dreams a reality, from the early days in gymnastics, through school, and my journey to become a scientist. For the Ph.D. specifically, I would be remiss not to thank you both for dropping everything and flying to Seattle on a day's notice to ensure that I wouldn't lose momentum when our childcare situation fell through. Looking back, those months that you helped with Parker and encouraged me to keep working, could have been one of those inflection points in life that could have drastically changed my life course. You both made sure that I would not lose momentum, despite being a new mom - making it to this point is a direct consequence of that support. Thank you both for every single thing you have done for me over the past 30 years - you are the best parents in the world and I strive to give my girls the same unconditional love, confidence, and support.

This brings me to Parker and Chloe - my beautiful girls. They are the brightest things in my life. They opened my world to a level of overwhelming love, joy, and meaning that I previously did not conceive possible. Parker was born in 2022, one year after I completed my qualifying exam and Chloe was eight months old when I completed my thesis defense. While I must admit that completing a Ph.D., with intensive periods of laboratory-based experimental work and two little kids, was not easy at times, I wouldn't have wanted it any other way. Parker

and Chloe, I love you both more than anything! Everything that I do now is for you both. I must also give credit to our adorable, tiny rescue dog - Olive. She has literally been by my side throughout this entire process and in many ways is my most unwavering supporter. Thank you for your unconditional love, Olive.

Another unconventional aspect of my Ph.D. was that from a scientific perspective, I did not do it alone. Different from a traditional academic lab setting, I had the privilege of working with a team of over 30 individuals that were a part of the SEA-AD study, most of which are experts in their field, who share the same goal of characterizing the early cellular changes in Alzheimer's disease. This is an extremely uncommon situation - instead of competition and isolation, I was able to learn surrounded by supportive team members with the shared goal of contributing to foundational science efforts. Collaboration is something often talked about in science but rarely practiced in its true form. Throughout my entire Ph.D., I experienced true collaboration at the highest level. Thank you to my wonderful team members at UW and the Allen Institute that played an equally important role in training me and helping me become the scientist that I am today.

Lastly, I would like to thank my committee members - Dr. Rebecca Hodge, Dr. Jessica Young, Dr. Nicole Liachko, Dr. Michael Hawrylycz, and Dr. Jeff Iliff. The feedback offered in our annual committee meetings was always extremely thoughtful and relevant. Knowing that I had each of your consistent guidance and support throughout this process was especially important to me, given that I had previously dealt with a lab switch, where my previous committee members (Thank you, Dirk!) played a crucial role in helping me navigate a complicated situation. I appreciate that I am lucky to have such sharp scientists and wonderful leaders supporting me during this process. Thank you!

Chapter 1. INTRODUCTION TO HUMAN BRAIN PROCESSING AND CELL TYPES

For over a century, human brain tissue has been collected for the explicit purpose of understanding diseases of the nervous system¹. In the United States, a systematic approach to research brain autopsies began in the 1960s, and today there are numerous brain banks worldwide that collect tissue from individuals with a wide range of neurological diseases². These efforts have advanced our understanding of the human brain significantly, influencing clinical diagnoses and revealing important clues about the pathophysiology underlying neurological diseases. However, as research studies utilizing the human brain have become increasingly complex, the gaps in the standardization and availability of tissue have become more prominent³. Modern brain atlasing efforts are particularly impacted due to the need for consistent high quality frozen brain tissue to enable molecular characterization of human cell types.

1.1 Standardization and modernization of human postmortem brain tissue processing to enable modern molecular approaches to study the brain in health and disease

Brain donations are the greatest gift one can give to science. Improved tissue preservation methods and field-wide standardization is needed to further the scientific benefit gained from this precious resource. Historically, research involving postmortem human brain tissue was forced to rely on formalin fixed tissue (and typically paraffin embedded (FFPE)) since this was the standard brain preservation across brain banks. While this approach does preserve cytoarchitecture, it cross-links proteins, making dissociation-based assays problematic^{4,5}. Additionally, it was thought that nucleic acids, such as DNA and RNA, were not preserved using such approaches, though this has been challenged recently. Recent advancements in spatial transcriptomics have demonstrated that high quality and representative RNA transcripts can be maintained in FFPE tissue blocks when freshly sectioned and assayed^{6,7}. Still, the vast majority of molecular techniques are still not compatible with formalin fixed tissue, reinforcing that updated brain collection and preservation techniques are needed for maximal impact. By employing modernized brain processing approaches, banked tissue can be used to perform both

well-established and novel molecular assays, propelling the study of postmortem human brain tissue into the modern era of molecular assays, such as single-cell transcriptomics.

Over the past decade, the work of several key groups, including the Allen Institute, have pioneered the development of improved protocols for the processing of fresh brain tissue for molecular assays^{4,7}. These methods have been refined by the UW Medicine Neuropathology group, and optimized for the processing of postmortem human brain tissue⁸. As these methods become a field standard, efforts have been made to disseminate the protocols and facilitate hands-on training in brain banks across the country.

Founded in 2013, National Institutes of Health (NIH)-funded NeuroBioBank (NBB) has been contracted by the federal government to develop a federated resource of human postmortem brain tissue distribution for research across the United States⁹. The NBB has been an important source of human postmortem brain tissue for the study of the nervous system and has developed practices that outline enrollment/informed consent, brain collection, tissue processing and preparation, neuropathology, serology, toxicology, and the distribution for brain biospecimens and tissues¹⁰. Protocols developed at the Allen Institute and UW Medicine BioRepository and Integrated Neuropathology (BRaIN) laboratory for use in BRAIN Initiative and other state-of-the-art multimodal molecular omics projects have been incorporated into the NBB's best practices for their Human Brain Cell Atlas Collection (HBCAC), for which the UW protocol for processing of fresh postmortem brain tissue for molecular assays is the centerpiece.

This improved brain processing protocol involves several critical steps to enable the spatial and molecular study of the postmortem human brain, rapid brain procurement, fresh coronal sectioning, and fast controlled freezing of sliced tissue. Rapid brain procurement requires a complex logistical setup to ensure that the time between death and brain removal, referred to as the postmortem interval (PMI), is as short as possible. If the PMI is less than 12 hours, a rapid autopsy procedure is performed, and the tissue is deemed of the highest quality and amenable to a wide range of molecular assays. The target PMI is typically 7 hours, which is typically associated with a brain pH of approximately 7 and a RNA integrity number (RIN) of 8-9.

The SEA-AD consortium is a multi-institutional effort between the University of Washington (UW) School of Medicine, the Allen Institute for Brain Science (AIBS), and Kaiser Permanente Washington Health Research Institute (KPWHRI). The study aims to characterize donors spanning the complete spectrum of Alzheimer’s disease neuropathologic change (ADNC) using single-nucleus RNA sequencing (snRNA-seq), single-nucleus Assay for Transposase-Accessible Chromatin sequencing (snATAC-seq), and spatial transcriptomics across 84 brain donors processed by rapid autopsy at UW Medicine Neuropathology¹¹. The postmortem brain tissue included in the study is consistently high quality across the cohort, with an average PMI of 6.9 hours, brain pH of 6.8, and RIN of 8.6 (Table 1.1).

Donor ID (AIBS)	PMI	Brain pH	RIN
H19.33.004	8.1	7	8.33
H20.33.001	7.7	6.8	8.6
H20.33.002	4.3	7.3	7.87
H20.33.004	8.8	6.7	7.83
H20.33.005	7.6	6.8	8.4
H20.33.008	7	6.4	8.68
H20.33.011	6	7	9.07
H20.33.012	7.7	6.6	8.75
H20.33.013	4.5	7.6	9.08
H20.33.014	7.4	6.8	8.67
H20.33.015	11.4	6.8	8.57
H20.33.016	7.7	6.8	8.82
H20.33.017	5.5	7.2	8.82
H20.33.018	4.5	6.8	9.55
H20.33.019	11.1	7	8.65
H20.33.020	6.6	6.8	8.98
H20.33.024	3.5	7.2	9.5
H20.33.025	4.5	7.4	9.43
H20.33.026	7	7.2	9.53
H20.33.027	9.2	6.7	8.87
H20.33.028	7.4	6.6	7.9
H20.33.029	6.1	6.5	9.43
H20.33.030	6	6.7	8.65
H20.33.031	7.9	6.4	8.85
H20.33.032	6.9	6.8	9.45
H20.33.033	6.5	6.8	7.93

H20.33.034	10.0	6.8	8.98
H20.33.035	4.6	6.6	9.04
H20.33.036	10.6	6.4	8.63
H20.33.037	10	6.4	8.63
H20.33.038	6.2	6	6.3
H20.33.039	10.3	6.8	9.48
H20.33.040	5.03	6.4	9.17
H20.33.041	5.9	6.8	8.43
H20.33.043	4.5	7	9.37
H20.33.044	4.7	7.6	9.2
H20.33.045	3.3	7.3	8.15
H20.33.046	8	6.2	5.67
H21.33.001	9.6	7.2	8.1
H21.33.002	3.2	6.8	9.18

Table 1.1 Postmortem brain tissue quality control metrics. Donors listed by their Allen Institute ID. Brain quality metrics include postmortem interval (PMI), brain pH, and RNA integrity number (RIN).

1.2 Challenges with adoption and standardization of best practices for the processing of human postmortem brain tissue

Despite the increased prevalence of tissue freezing at most modern brain banks, issues of standardization remain. Unlike historical methods of tissue preservation, novel approaches optimized to preserve high quality genetic material are substantially more complex leading to increased variability across sites. Numerous tissue freezing methods exist, each with strengths and weaknesses. Such protocols include freezing entire, intact hemispheres or coronal slabs at -80°C , or dissecting select regions and snap freezing in liquid nitrogen⁴. Such methods of snap freezing often lead to widespread tissue damage and cell rupture^{12,13}. As described above, the Allen Institute, UW Medicine Neuropathology, and others have recently demonstrated that using a dry ice isopentane slurry at -40°C , provides a controlled approach for freezing the tissue in a manner that preserves both the cytoarchitecture and genomic material¹⁴. Studies are needed to conclusively compare each protocol regarding its efficacy at preserving tissue specifically for modern molecular approaches.

Establishing the infrastructure to enable rapid autopsies requires substantial funding resources and oversight to create and maintain an on-call team of specialized technicians to perform brain procurement, processing, and preservation 24/7 to enable brain collection with the shortest possible PMI¹⁵. Larger, more well-resourced biorepositories are better positioned to implement such programs, though this leaves substantial geographical and demographic gaps in available donors on a national scale. Efforts for standardization are ongoing and are becoming more widespread thanks to initiatives such as BICCN/BICAN, where practices for processing postmortem human brain tissue have been incorporated into national scale research projects.

The current human brain atlasing efforts face several additional challenges. The first is the lack of precision and consistency in anatomical sampling. Brain sampling requires an expertly trained technician or neuropathologist to visually locate and dissect the precise brain region that is intended to be studied. If photodocumentation is not used, information regarding the physical location of the sampled tissue is dependent on the quality of notes retained. Unfortunately, many biorepositories and studies do not appropriately describe the precise region that was sampled (e.g., not distinguishing which part of the frontal cortex was sampled), causing issues both within and across institutes. Beyond photodocumentation, efforts are being made to register images of the coronal brain tissue slabs to a standardized reference atlas, such that any brain sample can be mapped to a consistent anatomical framework and aid in delineating the exact substructure that was sampled. Until this modernized practice becomes more widespread, very careful attention must be made to photodocumentation and manual annotation of anatomy. In the human brain, there are many very subtle, yet topographically and functionally distinct brain regions, and the ability to discern those differences is essential. Improved mapping to standard anatomical coordinates should improve reproducibility across all studies using human brain tissue.

1.3 Generation of transcriptomics-based human brain cell atlases

Modern single-cell atlases play a crucial role in defining the cells that make up our brain. In 2025, the complete complement of cell types that make up the human brain, their distinct roles, and their relative proportions in healthy and diseased brain are incompletely understood. A

primary goal of these atlasing efforts is to develop a resource to enable the precise study of human cell types and understand their role in a healthy, non-disease context to then better understand changes in aging, injury, and disease. One of the most impactful early contributions to this effort was published by the Allen Institute and used snRNA-seq to profile the human Middle Temporal Gyrus (MTG) of eight neurotypical brain donors, defining a cellular taxonomy of 69 neuronal and 9 non-neuronal cell types¹⁶. This work was the first published example of human brain research using tissue collected with the updated brain processing methods described above. The NIH BRAIN Initiative Brain Cell Atlas Network (BICAN), of which the Allen Institute is a key contributor, continues to devote substantial resources to the effort to identify and catalogue the cell types that make up the human brain in health and disease.

Recent studies, including SEA-AD and work performed as part of this dissertation, have continued to successfully leverage human postmortem brain tissue in a range of modern molecular assays due to the high quality of preserved genomic material. Several other projects including a collaborative effort between the Allen Institute and Karolinska Institute generated a first draft of a human whole brain atlas which profiled over three million nuclei from 100 subsampled brain regions across three adult donors processed and preserved using the methods described above¹⁷. Other studies have combined multiple modalities in Alzheimer's disease such as snRNA-seq and snATACseq, as well as methods such as VINE-seq to profile vascular cell types¹⁸. More recently, an effort to aggregate over 70 human studies comprising over 26 million cells has been published¹⁹.

Currently, the human brain is thought to have three major classes of cells: excitatory (Glutamatergic), inhibitory (GABAergic), and non-neuronal. Cortical neurons are typically split by their layer (Layer 1 - 6) and projection pattern (intratelencephalic, extratelencephalic, and corticothalamic). Inhibitory neurons are more broadly dispersed throughout the brain and are typically named based on marker genes (e.g., *Sst*, *Pvalb*). Non-neuronal subclasses include glia (astrocytes, oligodendrocytes, and microglia) and other cell types such as endothelial cells. These designations result in a three-level hierarchy of cell types, with a middle (subclass) level of approximately 20 unique cell subclasses in the cortex. Finally, to establish more granular definitions of cell types, a combination of morphological, electrophysiological, and

transcriptomic information has been used to distinguish more than 130 individual cortical cell types, referred to as supertypes (Figure 1.1).

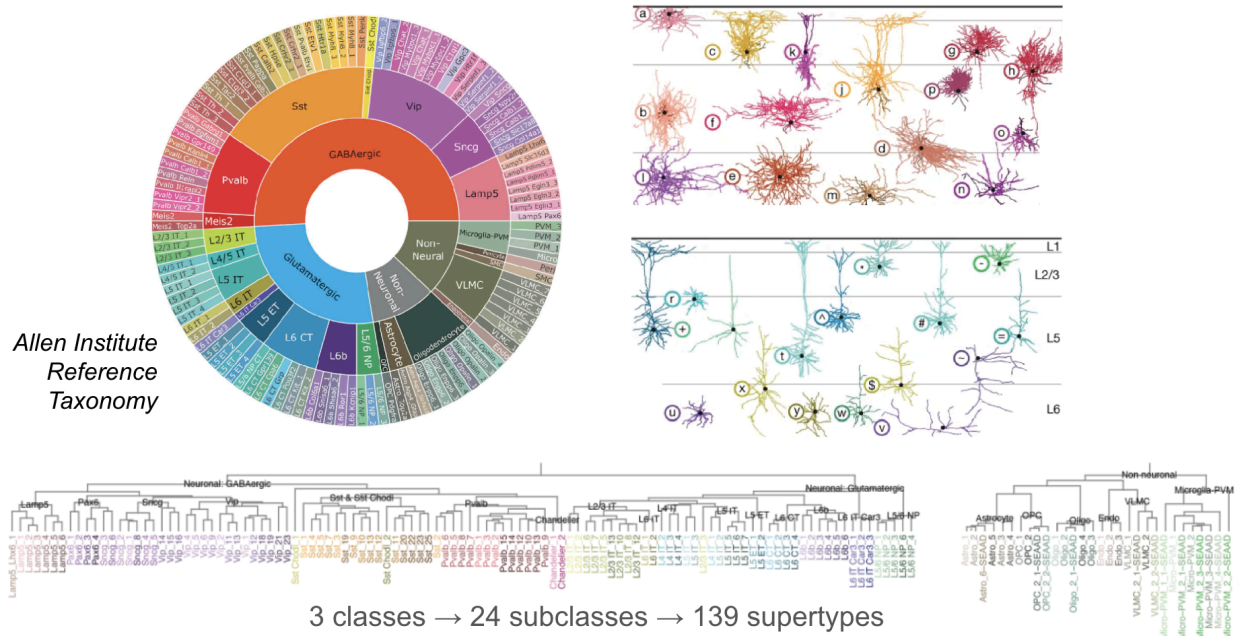


Figure 1.1 Allen Institute Cellular Taxonomy. Adapted from the Allen Institute (BICCN 2021) showing a ‘cell types wheel’ of an Allen Institute reference taxonomy. To the right morphological features of key cell types and below a tree diagram showing up to 139 distinct transcriptomically-defined supertypes.

One challenge facing studies using human brain tissue for the purpose of mapping cell types, have not reached consensus on a standardized cellular taxonomy. Lack of a consistent naming scheme for human brain cell types has posed substantial hurdles to the productivity of brain atlasing efforts at the cell type level. Unsurprisingly, this negatively impacts reproducibility and leads to confusion in the field. The Allen Institute has devoted substantial efforts to this cause and have made numerous meticulously curated cellular taxonomies from key brain regions publicly available. Further, tools, such as MapMyCells, have been developed for easier adoption of standard taxonomies by lowering the requirement of computational expertise of the end user²⁰. Together, improved sample tracking and registration with consensus of cell type taxonomies used will lead to increased progression in this important area.

Chapter 2. ADVANCEMENTS IN QUANTITATIVE NEUROPATHOLOGY

Components of this chapter are adapted from the following work:

Gabitto, M. I.*, Travaglini, K. J.*, **Rachleff, V. M.**, Kaplan, E. S., Long, B., Ariza, J., ... & Lein, E. S. (2024). Integrated multimodal cell atlas of Alzheimer's disease. *Nature Neuroscience*, 27(12), 2366-2383.

Introduction

The field of neuropathology has remained largely unchanged for decades, relying on histological stains to visualize, diagnose, and qualitatively assess neurological diseases. The expanding subdiscipline of research neuropathology has begun to embrace more modern and higher throughput approaches to studying the brain, opening the door to entirely new areas of discovery. Two advances have been critical in enabling this rapid progress. First, replacing single or duplex histological stains with highly multiplexed proteomic readouts, has enabled the study of multiple disease features, such as cell types or pathogenic proteins, simultaneously in the same piece of tissue. Second, advances in image analysis now enable rapid and accurate quantification of entire sections of tissue (whole slide images (WSIs)), leading to more quantitative datasets that cover multiple functional and anatomical regions. This is a substantial improvement over random region of interest (ROI) selection, which involves manually counting a small region that may or may not represent the larger tissue context. Together, integrating multiplexed immunohistochemical approaches with advanced image processing strategies is changing the field and much of this work has been developed by the UW Medicine Neuropathology group.

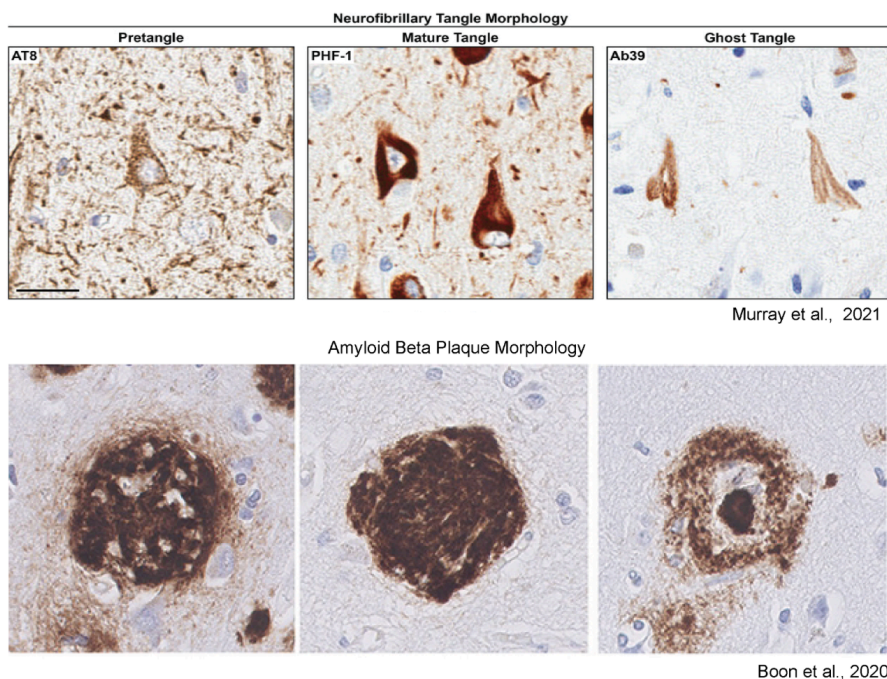
Neuropathological introduction to Alzheimer's disease

AD is the most common cause of dementia affecting approximately 57 million people worldwide¹. Age is the primary risk factor for the disease; given the aging global population, it is essential that progress be made in understanding this complex and devastating disease. Currently, no effective preventions, treatments, or cures exist. Decades of research has failed to translate to

the clinic, in part due to the limited translatability of animal models and the absence of technology enabling precise and high-resolution study of the human brain. Further, efforts to characterize the distribution of pathological peptides, without accounting for the cell types in which they accumulate, left the field with more questions than answers as to why certain cells are selectively vulnerable. To understand the complete picture of the disease, both the proteins and the cell types they directly affect must be studied jointly at high resolution.

To receive a neuropathological diagnosis of ADNC, the presence of two pathologic peptides, hyperphosphorylated tau (pTau) in the form of neurofibrillary tangles (NFTs) and extracellular amyloid beta (A β) in the form of A β plaques, are assessed across several affected brain regions^{2,3,4} (Figure 2.1). The primary diagnostic assessments are Braak staging (pTau-positive NFTs), Thal phasing (A β plaques), and Consortium to Establish a Registry for Alzheimer's Disease (CERAD) scoring (neuritic plaques)^{5,6}.

Alzheimer's disease accounts for 60-70% of dementia cases. Presents as changes in thinking and memory and is characterized by abnormal deposits of amyloid plaques and tau tangles.



Other Dementias

Vascular Dementia

Prevalence: 20%
Conditions, such as blood clots, disrupt blood flow in the brain

Lewy Body Dementia

Prevalence: 10-15%
Abnormal deposits of alpha-synuclein proteins, called "Lewy bodies"

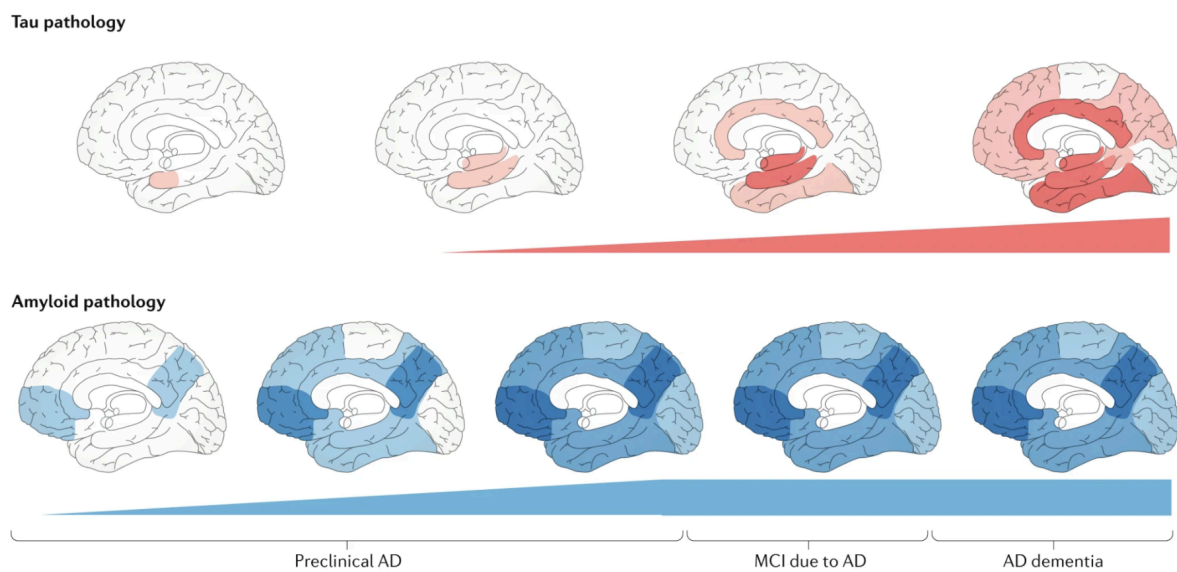
Frontotemporal dementia

Prevalence: 2%
Abnormal tau and TDP-43 protein in the frontal and temporal lobes

nia.nih.gov

Figure 2.1. Representative images of pTau and A β pathology in postmortem brain tissue. Adapted from Murray et al., 2021, Boon et al., 2020 and [nih.gov](https://www.nih.gov). Summary Alzheimer's disease and other dementias. Bright field histochemical stains of neurofibrillary tangles (pretangle, mature tangle, and ghost tangle) and amyloid beta plaque morphology (diffuse and dense core).

Broadly speaking, both plaques and NFTs follow a stereotyped pattern of deposition across the brain as the disease progresses⁷. The spatiotemporal spread of pTau pathology tracks closely with the region-specific nature of cognitive changes. Braak staging maps the spatiotemporal deposition pTau pathology which begins in the entorhinal cortex (Stage I), then the hippocampal CA1 (Stage II), followed by temporal lobe structures (Stage III and IV), and finally broad neocortical distribution (Stage V and VI). Thal Phasing maps the spread of A β pathology which follows the opposite pattern beginning cortically and moving subcortically at later stages of the disease (Figure 2.2).



Kant et al., 2020

Figure 2.2. Stereotyped spatiotemporal distribution of pTau NFT pathology and A β pathology in AD. Adapted from Kant et al., 2020. Visualizing progression of tau pathology (red) from subcortical, to temporal lobe, then broadly throughout the neocortex at later stages of disease. Contrasting amyloid spread (blue) which begins in the frontal neocortex and only in the latest stages of disease reaches subcortical structures.

Decades of research on the brain-wide accumulation, distribution, and clearance of these peptides, particularly A β , has yielded minimal clinical success⁸ (Figure 2.3). Recently, several groups have ventured to study AD from a new perspective redefining it as a cellular disorder that preferentially affects unique subpopulations of cells over time as the disease progresses. The unprecedented cellular resolution afforded by single-cell RNA sequencing-based approaches has enabled the identification of highly granular cell types that are vulnerable at distinct stages of

disease progression¹¹. However, to date, these atlases lack the pathological protein context required to address the fundamental question of why certain cell types are vulnerable.

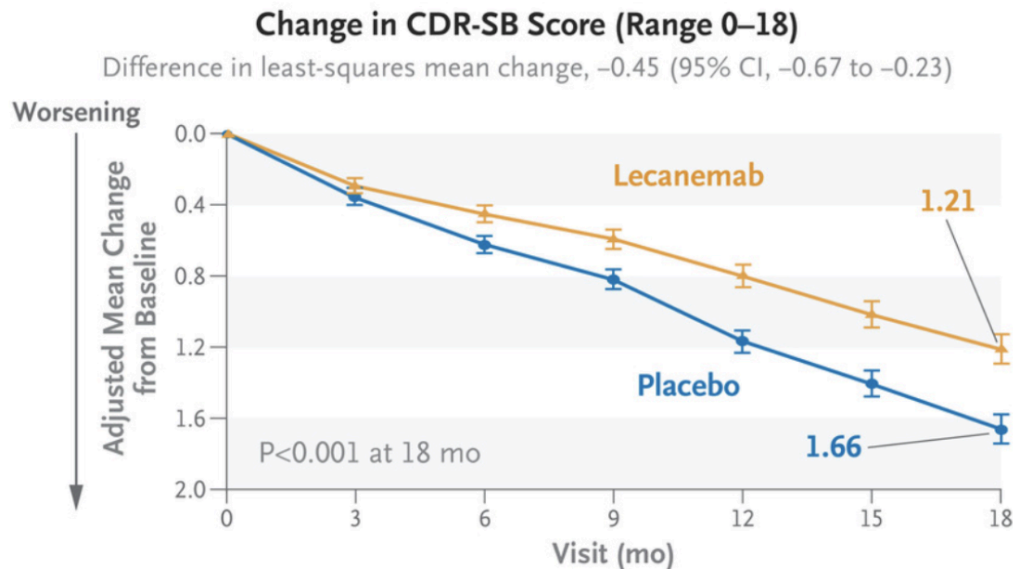


Figure 2.3. Change in cognitive score in response to Lecanemab treatment. Adapted from Katayama et al., 2023. The adapted figure shows the new Alzheimer’s drug Lecanemab (monoclonal antibody targeting amyloid beta) and the marginal impact it has on adjusted mean change from baseline (1.21) relative to placebo control (1.66).

Background in traditional neuropathological approaches and hurdles to overcome

Several biological, technological, and logistical factors have prolonged the field of neuropathology from entering the modern quantitative era. Most notably, the field has been limited by the possible research applications that are amenable to formalin fixed tissues, as described in Chapter 1. Next, it has been a challenge to move beyond traditional low plex or bulk methods in human postmortem tissue, in part due to tissue preservation methods, and in part due to technological limitations (e.g., microfluidics and imaging). Further, the standard histochemical workflows are time-intensive, making it difficult to increase throughput.

Standard human postmortem FFPE blocks are sectioned at 5µm and are typically stained using histochemical methods such as hematoxylin and eosin (H&E), cresyl violet (Nissl), various

silver-based stains, and Luxol fast blue, among many others⁹. These histochemical stains allow for simple and consistent diagnostic assessment of pathology but have limited research application, because only 1-2 stains can be assessed at a time and the assessments are typically subjective or ordinal. As a slightly better alternative, immunohistochemical approaches allow for antibody antigen interactions to more precisely target proteins of interest, while also enabling slightly higher plex. This immunological reaction can be paired with either a histochemical readout, as described above, or a fluorescent dye which enables slightly more throughout (up to four proteins per sample).

While fluorescent stains are not standard in diagnostic neuropathology, they are an important tool for research neuropathology and neuroscience more broadly. However, fluorescent approaches pose additional methodological hurdles when applied to human brain tissue, such as the challenge of high background autofluorescence that is particularly pronounced in aged brain tissues. Standard neuroscience antibodies include NeuN (neurons), GFAP (glia), IBA1 (microglia), S100 (astrocytes), AT8 (phosphorylated Tau at Ser202/Thr205), 6e10 (A β), pTDP-43 (phosphorylated TDP-43), and α -synuclein (Lewy body pathology). Together, the ability to combine up to four markers in one tissue section enables modest improvement for understanding static cellular and protein relationships, though they are still a far cry from painting a comprehensive picture of disease.

While the limitation in readout plex is a contributing factor, the time-intensive nature of scaling up the traditional histology pipeline has also contributed to neuropathology's slow evolution to becoming highly quantitative. Histology has a long history in the sciences and involves three major steps: tissue staining, slide imaging, and image analysis. Histological staining is time-consuming. Fully automated staining systems, such as IntelliPATH BioCare, have helped to overcome this potential bottleneck by accommodating 50 slides per run and five different protocols within the same run. The batch processing time is only 4-6 hours for a monoplex stain. Not only do automated approaches increase productivity, but they also greatly improve consistency and reproducibility which is essential for both diagnostics and research.

The second time-intensive step is image acquisition which is typically carried out on a standard or confocal microscope. It requires the individual to manually select, label, focus, and set thresholds for each feature to be imaged. Because of the manual nature of the process, typically only a subset of the tissue is imaged with this approach. Automated systems have become more common with current systems capable of scanning 400+ slides per run, including the Leica Aperio AT2, which dramatically accelerate the imaging process, while capturing high-resolution WSIs.

The third and final hurdle in the process is image processing and analysis. Fortunately, like the prior steps, automated approaches for image processing and analysis, such as Indica Labs HALO, Imaris, and QuPath, are now widely available. These solutions offer a variety of features such as tissue and region segmentation, cell detection and phenotyping, and area-based quantifications. While the staining and image acquisition steps are essential, it is this final image processing step that brings traditional neuropathology into the highly quantitative realm. Of note, standardized image processing and analysis pipelines are also beginning to play a crucial role in the improved reproducibility within and across studies.

The current work performed by UW Neuropathology, utilizes one of the field's most advanced, high-throughput quantitative neuropathology workflows. First, slides are duplex stained using the IntelliPATH BioCare automated staining robots and covered using a robotic coverslipper. Then they are imaged using the Leica Aperio AT2 or Leica Aperio GT450 and uploaded to cloud storage (Amazon AWS). Finally, they are processed and analyzed using Indica Lab HALO software to set tissue boundaries and quantify relevant features. The extracted measures are analyzed using custom Python scripts.

Introduction to the Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD) consortium

The SEA-AD consortium leverages advances in quantitative neuropathology, single cell multiomics, and spatial transcriptomics to produce the highest-resolution, multimodal, brain-wide cell atlas of AD and related dementias. The primary goals of SEA-AD were to (1)

generate a high-quality donor cohort spanning the full spectrum of ADNC (instead of a case-control design), recruited from longitudinal prospective cohort studies with well-characterized participants; (2) use improved tissue preparation methods for single-nucleus and spatial genomics; (3) develop deep donor characterization with all analytical methods applied to the same donors; (4) perform sufficient sampling to analyze the full diversity of cell types; (5) map profiled cells to the highly granular BICCN reference; and (6) validate cellular phenotypes across cortical areas, orthogonal modalities and independent datasets (Figure 2.4).

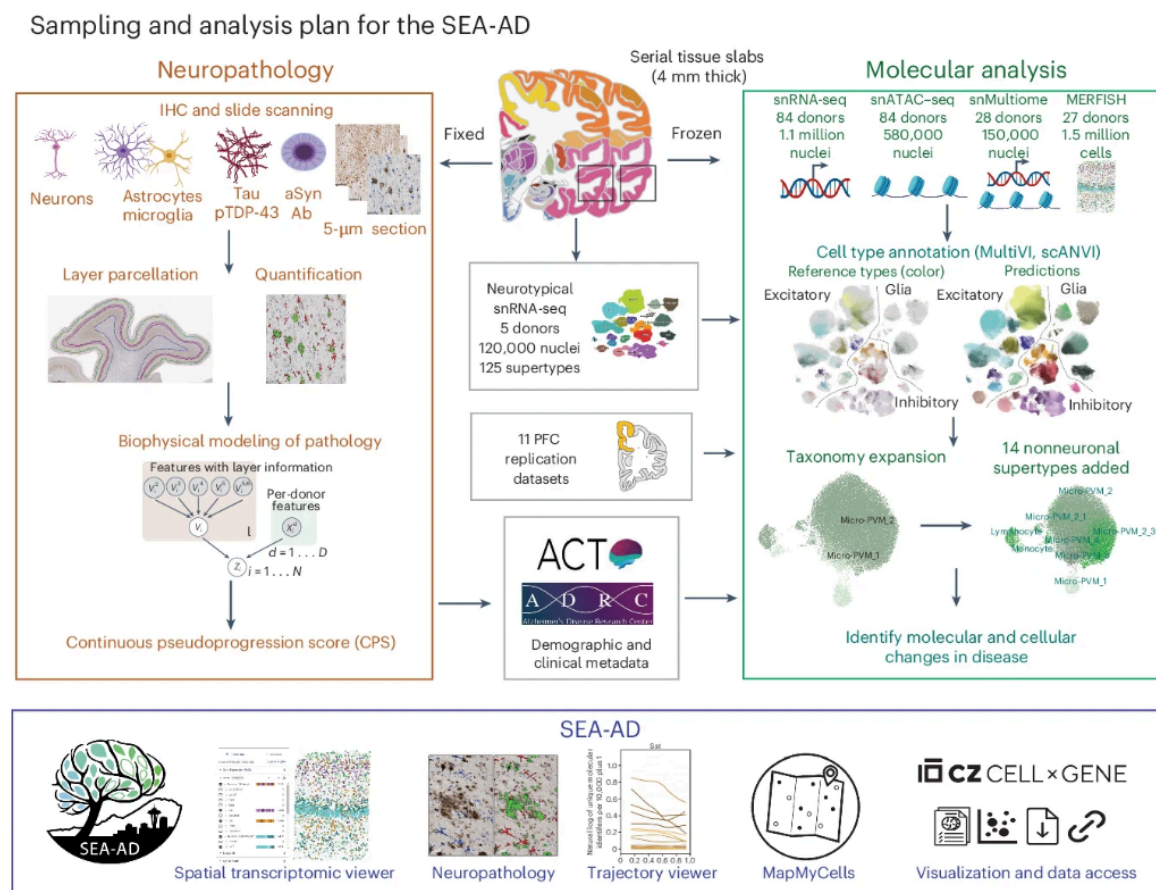


Figure 2.4. Schematic detailing the experimental design for the SEA-AD study. Quantitative neuropathology (left panel) serves as the basis to inform multiomic analyses, including snRNA-seq, snATAC-seq, snMultiome, and spatial transcriptomics.

The current work focuses on a cohort of 84 well-characterized individuals who have donated their brains for research and comprise the SEA-AD consortium's first study, which profiled the middle temporal gyrus (MTG). Quantitative neuropathology provides the foundation

for the study and serves as the basis for all downstream omics-based analyses. The steps involved in processing and analyzing these data is the basis of this chapter and my contribution to the larger consortium publication.

The MTG has been extensively characterized in both neurotypical and AD tissue. Functionally, the MTG is an area involved in language and semantic memory processing and higher-order visual processing. Clinically, the MTG is thought to be a transition zone between aging-related or preclinical AD-related medial temporal lobe pTau and more advanced stages of AD, where pTau-related neurofibrillary degeneration extends across the brain and is strongly correlated with dementia. AD is thought to develop slowly over decades with a gradual accumulation of pathology that corresponds to a gradual decline in cognition, until the intersection point at which cognitive decline becomes more rapid and interventions may have limited utility^{10,11} (Figure 2.5).

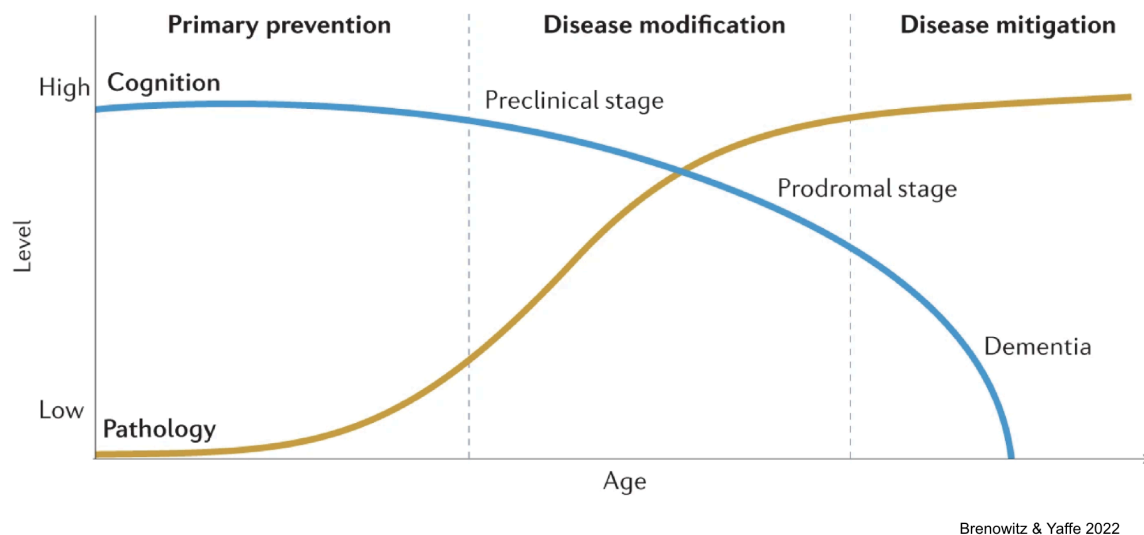


Figure 2.5. Inverse correlation of pathology and cognitive impairment in AD. Adapted from Brenowitz and Yaffe 2022. Inverse relationship between cognition and pathology. Illustrates the preclinical stage at which disease modifying interventions would have highest therapeutic benefit.

The temporal lobe is one of the earliest locations in the brain where both pTau and A β intersect. By combining temporal modeling of disease severity based on quantitative neuropathology with single-nucleus genomics and spatial analyses, the SEA-AD study provides a comprehensive understanding of the specific, highly granular cell types affected over the

course of disease, where those affected cells are located in tissue microarchitecture and when they are affected as disease progresses. All data associated with the study are freely available on the SEA-AD website (sea-ad.org).

SEA-AD cohort demographics

The SEA-AD cohort was derived from longitudinally characterized research brain donors from the community-based Adult Changes in Thought (ACT) study and the University of Washington (UW) Alzheimer's Disease Research Center (ADRC). Brains were collected using highly optimized brain preparation methods (mean PMI = 7 hours) that enable exceptionally high-quality genomics data. Donors were included if death occurred within the specific time of data collection. SEA-AD includes donors across the range of AD neuropathological change (ADNC) (9, no AD; 12 low; 21 intermediate; 42 high ADNC) who were all aged (minimum age at death = 65, mean = 88) (Figure 2.6).

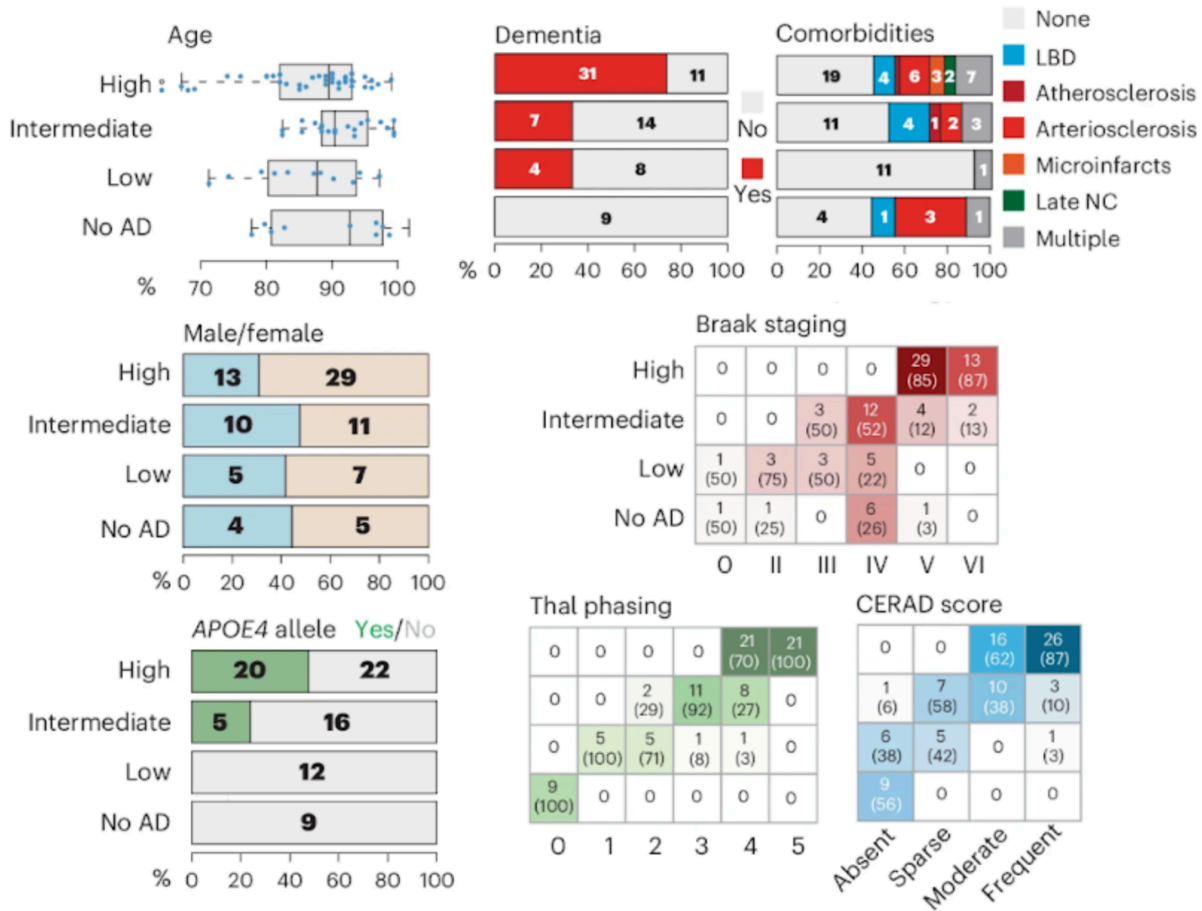


Figure 2.6. SEA-AD cohort demographics. Summarizing age at death, biological sex and APOE4 allele, stratified according to ADNC score. Age at death is represented by box-and-whisker plots with the box representing the interquartile range (IQR) and the whiskers representing 1.5 times the IQR. The solid line indicates the median. SEA-AD cohort composition stratified according to ADNC versus Braak stage Thal phase (left), and CERAD score as heatmaps, with dementia or comorbidities as bar plots. The number of donors in each box and the fraction are shown in parentheses.

Female donors outnumbered male donors (51 females, 33 males), particularly in those with high ADNC (29 females, 13 males), which is consistent with the known prevalence of AD in females. Donors with an APOE4 allele included nearly half (20 of 42) of high ADNC cases, a quarter (five of 21) of intermediate cases and no low ADNC or no AD cases. Braak stage (tangles), Thal phase (plaques) and Consortium to Establish a Registry for Alzheimer's Disease (CERAD) (neuritic plaques) increased as expected with ADNC. Nearly three-quarters (31 of 42) of high ADNC cases had dementia before death, versus a third in intermediate (seven of 21) and low (four of 12) ADNC cases, and none in no AD cases. Donors with any level of Lewy body dementia (LBD), vascular pathology or limbic-predominant age-related TDP-43 encephalopathy

(LATE) were included because these conditions are common comorbidities in AD¹². Roughly half (42 of 84) had one or more severe co-pathologies (Table 2.1).

Donor ID	Age at Death	Sex	Years of education	APOE Genotype	Cognitive Status	ADNC	Thal	Braak	CERAD score	Overall CAA Score	LATE
H19.33.004	80	F	17	3/3	No dementia	Not AD	Thal 0	Braak IV	Absent	Not identified	Not Identified
H20.33.001	82	M	16	3/3	No dementia	Low	Thal 2	Braak IV	Sparse	Not identified	LATE Stage 2
H20.33.002	97	F	12	2/3	No dementia	Not AD	Thal 0	Braak IV	Absent	Not identified	LATE Stage 2
H20.33.004	86	M	15	3/4	Dementia	High	Thal 5	Braak V	Frequent	Moderate	LATE Stage 1
H20.33.005	99	F	12	2/3	No dementia	Intermediate	Thal 3	Braak IV	Moderate	Moderate	LATE Stage 1
H20.33.008	92	F	18	3/4	No dementia	High	Thal 4	Braak V	Moderate	Moderate	Not Identified
H20.33.011	93	F	16	3/4	Dementia	High	Thal 5	Braak V	Moderate	Moderate	LATE Stage 3
H20.33.012	91	F	18	3/3	No dementia	Low	Thal 1	Braak III	Absent	Not identified	Not Identified
H20.33.013	94	M	14	3/3	No dementia	Intermediate	Thal 3	Braak IV	Moderate	Mild	Not Identified
H20.33.014	82	F	16	3/3	No dementia	Intermediate	Thal 3	Braak IV	Sparse	Mild	Not Identified
H20.33.015	88	M	18	3/3	Dementia	Intermediate	Thal 3	Braak V	Moderate	Not identified	LATE Stage 3
H20.33.016	93	F	14	2/3	Dementia	Intermediate	Thal 4	Braak IV	Moderate	Moderate	Not Identified
H20.33.017	69	M	14	3/3	Dementia	High	Thal 4	Braak VI	Frequent	Moderate	Unclassifiable
H20.33.018	81	F	21	3/4	Dementia	High	Thal 5	Braak VI	Frequent	Not identified	LATE Stage 2
H20.33.019	87	F	15	3/4	No dementia	Intermediate	Thal 4	Braak III	Sparse	Mild	Not Identified
H20.33.020	81	M	12	4/4	Dementia	High	Thal 5	Braak VI	Frequent	Moderate	LATE Stage 2
H20.33.024	90	M	19	3/3	No dementia	Intermediate	Thal 4	Braak V	Sparse	Not identified	LATE Stage 2
H20.33.025	94	M	12	3/3	No dementia	High	Thal 4	Braak V	Moderate	Not identified	LATE Stage 2
H20.33.026	75	F	12	4/4	Dementia	High	Thal 4	Braak VI	Frequent	Moderate	LATE Stage 2
H20.33.027	99	F	17	3/3	No dementia	Intermediate	Thal 3	Braak IV	Moderate	Moderate	LATE Stage 1
H20.33.028	94	F	12	3/3	Dementia	High	Thal 4	Braak V	Moderate	Moderate	LATE Stage 2
H20.33.029	91	F	13	3/3	Dementia	High	Thal 4	Braak V	Moderate	Not identified	LATE Stage 2

H20.33.030	86	F	16	3/4	No dementia	High	Thal 4	Braak V	Frequent	Not identified	Not Identified
H20.33.031	87	F	12	3/3	Dementia	High	Thal 4	Braak VI	Frequent	Mild	LATE Stage 2
H20.33.032	98	M	17	3/3	No dementia	High	Thal 5	Braak V	Moderate	Moderate	LATE Stage 1
H20.33.033	68	M	18	3/3	Dementia	High	Thal 5	Braak VI	Frequent	Moderate	Not Identified
H20.33.034	85	F	16	2/2	No dementia	Intermediate	Thal 3	Braak IV	Absent	Moderate	Not Identified
H20.33.035	99	F	16	3/3	No dementia	Not AD	Thal 0	Braak IV	Absent	Not identified	LATE Stage 2
H20.33.036	100	F	15	2/3	No dementia	High	Thal 5	Braak V	Moderate	Mild	LATE Stage 2
H20.33.037	96	F	20	3/3	Dementia	High	Thal 5	Braak V	Frequent	Moderate	LATE Stage 2
H20.33.038	90	F	18	3/3	Dementia	High	Thal 4	Braak V	Frequent	Mild	LATE Stage 2
H20.33.039	96	F	14	3/3	No dementia	High	Thal 4	Braak V	Frequent	Moderate	LATE Stage 2
H20.33.040	98	M	13	3/3	Dementia	Intermediate	Thal 4	Braak IV	Frequent	Not identified	Not Identified
H20.33.041	91	F	18	3/3	Dementia	High	Thal 4	Braak V	Moderate	Not identified	LATE Stage 2
H20.33.043	85	M	20	4/4	No dementia	Intermediate	Thal 4	Braak IV	Sparse	Moderate	LATE Stage 2
H20.33.044	81	M	19	2/3	No dementia	Not AD	Thal 0	Braak V	Absent	Not identified	LATE Stage 1
H20.33.045	77	F	18	4/4	Dementia	High	Thal 5	Braak VI	Frequent	Not identified	LATE Stage 2
H20.33.046	94	M	21	3/3	Dementia	High	Thal 5	Braak VI	Frequent	Moderate	LATE Stage 2
H21.33.001	80	M	21	2/3	Dementia	Low	Thal 2	Braak III	Absent	Not identified	Not Identified
H21.33.002	70	F	18	3/4	Dementia	High	Thal 5	Braak VI	Frequent	Moderate	LATE Stage 2
H21.33.003	78	M	16	3/3	No dementia	Not AD	Thal 0	Braak 0	Absent	Not identified	Not Identified
H21.33.004	93	M	18	2/3	No dementia	Not AD	Thal 0	Braak II	Absent	Not identified	Not Identified
H21.33.005	95	M	16	3/3	Dementia	Intermediate	Thal 3	Braak V	Moderate	Not identified	LATE Stage 3
H21.33.006	97	M	16	3/4	No dementia	Intermediate	Thal 4	Braak IV	Sparse	Mild	Not Identified
H21.33.007	86	F	21	3/3	Dementia	High	Thal 4	Braak V	Frequent	Not identified	LATE Stage 2
H21.33.008	91	F	16	3/3	Dementia	High	Thal 4	Braak V	Frequent	Severe	Not Identified
H21.33.009	65	F	16	4/4	Dementia	High	Thal 5	Braak VI	Frequent	Moderate	Not Identified
H21.33.010	93	F	13	3/3	Dementia	High	Thal 5	Braak VI	Frequent	Severe	Not Identified

H21.33.011	83	F	17	3/3	No dementia	Not AD	Thal 0	Braak IV	Absent	Not identified	LATE Stage 2
H21.33.012	93	F	21	2/4	Dementia	Intermediate	Thal 3	Braak IV	Sparse	Mild	LATE Stage 2
H21.33.013	94	F	21	3/4	Dementia	High	Thal 4	Braak V	Moderate	Not identified	LATE Stage 2
H21.33.014	92	M	15	2/3	No dementia	Intermediate	Thal 4	Braak IV	Sparse	Not identified	LATE Stage 2
H21.33.015	98	M	12	3/3	No dementia	Low	Thal 2	Braak IV	Sparse	Mild	LATE Stage 2
H21.33.016	94	F	21	3/3	Dementia	Low	Thal 1	Braak IV	Absent	Not identified	LATE Stage 2
H21.33.017	92	F	16	3/3	Dementia	High	Thal 5	Braak V	Frequent	Moderate	Not Identified
H21.33.018	89	F	20	3/3	Dementia	Low	Thal 3	Braak II	Frequent	Not identified	Not Identified
H21.33.019	75	M	15	2/3	No dementia	Low	Thal 1	Braak 0	Sparse	Not identified	Not Identified
H21.33.020	82	M	21	3/3	Dementia	High	Thal 4	Braak V	Frequent	Mild	LATE Stage 2
H21.33.021	99	M	14	3/3	Dementia	Intermediate	Thal 4	Braak III	Moderate	Not identified	Not Identified
H21.33.022	82	F	12	3/3	No dementia	Intermediate	Thal 2	Braak VI	Moderate	Mild	LATE Stage 1
H21.33.023	102	M	12	3/3	No dementia	Not AD	Thal 0	Braak IV	Absent	Not identified	LATE Stage 1
H21.33.025	88	F	21	3/3	No dementia	Intermediate	Thal 3	Braak IV	Moderate	Moderate	Not Identified
H21.33.026	90	F	16	3/4	No dementia	High	Thal 4	Braak V	Frequent	Moderate	Not Identified
H21.33.027	92	M	18	3/4	Dementia	High	Thal 5	Braak V	Moderate	Moderate	LATE Stage 1
H21.33.028	72	M	16	3/3	No dementia	Low	Thal 1	Braak II	Absent	Moderate	Not Identified
H21.33.029	89	M	18	2/4	Dementia	High	Thal 5	Braak V	Frequent	Moderate	LATE Stage 1
H21.33.030	89	M	17	3/4	No dementia	Intermediate	Thal 3	Braak III	Moderate	Mild	LATE Stage 1
H21.33.031	84	M	18	3/4	Dementia	High	Thal 5	Braak V	Moderate	Moderate	LATE Stage 2
H21.33.032	98	F	16	3/3	No dementia	Low	Thal 2	Braak IV	Sparse	Moderate	LATE Stage 2
H21.33.033	83	F	14	3/4	No dementia	High	Thal 5	Braak V	Moderate	Moderate	LATE Stage 1
H21.33.034	90	F	16	3/4	Dementia	High	Thal 5	Braak VI	Frequent	Moderate	LATE Stage 2
H21.33.035	97	F	14	3/3	No dementia	High	Thal 5	Braak V	Moderate	Moderate	LATE Stage 2
H21.33.036	93	F	14	3/3	No dementia	High	Thal 4	Braak V	Frequent	Mild	LATE Stage 2
H21.33.037	88	F	18	2/3	No dementia	Low	Thal 2	Braak IV	Absent	Moderate	LATE Stage 1

H21.33.038	84	F	15	3/3	No dementia	Low	Thal 1	Braak III	Absent	Not identified	Not Identified
H21.33.039	88	F	13	3/3	Dementia	High	Thal 4	Braak V	Moderate	Not identified	LATE Stage 3
H21.33.040	83	M	17	3/4	No dementia	High	Thal 4	Braak V	Frequent	Moderate	LATE Stage 1
H21.33.041	98	F	12	2/3	No dementia	Not AD	Thal 0	Braak IV	Absent	Not identified	Not Identified
H21.33.042	91	F	15	4/4	Dementia	High	Thal 5	Braak V	Moderate	Moderate	LATE Stage 2
H21.33.043	95	F	16	3/3	Dementia	Low	Thal 4	Braak II	Sparse	Not identified	LATE Stage 1
H21.33.044	88	F	15	3/3	Dementia	Intermediate	Thal 3	Braak VI	Frequent	Moderate	LATE Stage 1
H21.33.045	94	F	12	3/4	Dementia	High	Thal 4	Braak VI	Frequent	Moderate	LATE Stage 3
H21.33.046	97	M	17	3/3	Dementia	High	Thal 4	Braak V	Moderate	Moderate	LATE Stage 2
H21.33.047	90	M	21	3/3	No dementia	Intermediate	Thal 2	Braak V	Frequent	Not identified	LATE Stage 2

Table 2.1. Complete SEA-AD cohort demographic information.

Methods

Rapid brain processing

In a rapid autopsy, the brain is hemisected by performing a midsagittal cut through the cerebrum, cerebellum, and brainstem. One hemisphere is embedded in dental alginate and transferred to a slicing sled for uniform 4mm coronal sectioning anterior to posterior (first slice is the frontal pole). Each slab is photographed. Alternating slabs are either fixed in 10% neutral buffered formalin (NBF) or frozen in a dry ice isopentane slurry. The formalin fixed slabs are then subsampled into blocks to be paraffin embedded (FFPE) for 5um sectioning and histological staining. Slabs are frozen in a controlled manner which is essential for maintaining the integrity of cell membranes and tissue cytoarchitecture. The frozen slabs are immediately vacuum sealed in pre-labeled bags and transferred to -80C storage where they remain until they are sampled for nuclei isolation or other molecular assays. The other hemisphere of the brain is placed in a container of 10% NBF, medial side facing down for a minimum of two weeks prior to gross examination and ex vivo MRI.

Single and duplex IHC for quantitative neuropathology

STG-MTG tissue blocks were sectioned at 5 μ m, deparaffinized by immersion in xylene for 3 min three times, rehydrated in graded ethanol series (100%, 3 $\times$, 96%, 70% and 50% for 3 min each), and washed with Tris-buffered saline with 0.25% Tween-20 (TBST) twice for 3 min. The slides were immersed in Diva Decloaker 10X solution (cat. no. DV2004, Biocare Medical) for heat-induced epitope retrieval using the Decloaking Chamber at 110 °C for 15 min for most of the antibodies. To detect the α -synuclein protein, enzymatic antigen retrieval with protein kinase (cat. no. AR551, Leica Biosystems) was used. After heat-induced epitope retrieval was completed, the slides were cooled for 20 min at room temperature. Afterward, slides were washed with TBST for 5 min twice.

Chromogenic staining was performed using the fully automated IntelliPATH system (Biocare Medical). Blocking with 3% hydrogen peroxide (cat. no. PX968, Biocare Medical), Bloxall (Vector Laboratories), Background Punisher (Biocare Medical) and levamisole (Vector Laboratories) was performed to avoid any cross-reactivity and background. The following primary antibodies were used for the first target protein at the dilutions indicated: NeuN (1:500, clone A60, mouse, MAB5374, Merck Millipore), pTDP-43 (1:1,000, Ser409/410, clone ID3, rat, cat. no. 829901, BioLegend), β -amyloid (1:1,000, clone 6E10, mouse, cat. no. 803003, BioLegend), α -synuclein (1:200, clone LB509, mouse, cat. no. 180215, Invitrogen) and GFAP (1:1,000, rabbit, cat. no. Z033401-2, DAKO). After incubation with primary antibodies, sections were washed four times for 2 min with TBST and stained with species-appropriate secondary probe or antibody with a polymer horseradish peroxidase (HRP) (manufacturer's proprietary dilution, MACH 3 mouse (cat. no. M3M530) and MACH 3 rabbit (cat. no. M3R531), Biocare Medical; manufacturer's proprietary dilution, ImmPRESS goat anti-rat IgG (cat. no. MP-7444), Vector Laboratories). Sections were washed two times for 2 min with TBST; the antibody complex was visualized after 3–7 min by HRP-mediated oxidation of 3,3'-diaminobenzidine (DAB) (IntelliPATH, cat. no. IPK5010) by HRP (brown precipitate). Counterstaining was done with hematoxylin after the DAB reaction.

In duplex IHC (6E10/IBA1 and AT8/pTDP-43), slides were washed for 22 min in TBST and then incubated with primary antibodies at the dilutions indicated after the DAB reaction: IBA1 (1:1,000, rabbit, cat. no. 019-19741, Wako) and pTau (1:1,000, clone AT8, mouse, cat. no. MN1020, Thermo Fisher Scientific). They were washed as above and stained with species-appropriate secondary polymers conjugated to an alkaline phosphatase (MACH 3 mouse (cat. no. M3R532), MACH 3 rabbit (cat. no. M3R533), Biocare Medical). The complex was then visualized with the IntelliPATH Ferengi Blue Chromogen Kit (cat. no. IPK5027, Biocare Medical; blue precipitate). Once staining was completed, the slides were removed from the automated stainer and immersed in TBST for 3 min, then dehydrated in graded ethanol (70%, 96%, 100%, 2×) for 3 min and xylene (or xylene substitute in the case of duplex IHC), three times each for 3 min. Finally, coverslipping was carried out with a Tissue-Tek automated cover slipper (Sakura Finetek) using Ecomount medium (cat. no. EM897L, Biocare Medical).

Acquisition of whole slide images

Slides were scanned on the Aperio AT2 digital scanner (Leica, software version 102.0.7.5), which captures sequential images of a 20x field of view, using slide settings optimized for our IHC protocols which are subsequently assembled or stitched into whole slide images (WSIs) to be exact replicas of the glass slides. All images are scanned at 20x magnification and using the same gain, brightness and exposure times to avoid image to image variations.

Processing of whole slide images

The quantitative pathological assessment for the WSIs obtained from the MTG region were analyzed using the HALO® v.3.4.2986 (Indica labs, Albuquerque, New Mexico, USA). First, DenseNet2, a deep learning convolutional neural network was trained to segment the MTG cortical ribbon into cortical layers. The DenseNet network is a minimally pretrained classifier developed to recognize patterns in the tissue structure provided by Halo. Training data was created by manually annotating cortical layers labelled with NeuN in 10 cases. Based on the cellular architecture and the relative position within the cortical ribbon, the following layers were

annotated: Layer1 (molecular layer), layer 2 (external granular layer), layer 3 (external pyramidal layer), layer 4 (Internal granular layer) and layers 5-6 (internal pyramidal and multiform layers) (Figure 1). Then the trained classifier was applied to the NeuN-labelled sides from all donors. All results of the automatic segmentation were examined by a scientist trained in cortical neuroanatomy and adjusted when necessary. Manual adjustment of the annotations also included removal of staining artifacts and non-parenchymal structures, such as large blood vessels by drawing exclusion areas around them.

Second, using the Serial Section registration tool, all 5 WSIs belonging to the same case (labelled with NeuN, GFAP, α -Syn, A β combined with Iba1, and pTau combined with pTDP-43) were registered to each other in order to establish anatomical correspondence between tissue sections, and the cortical annotations from the NeuN-labelled slide were copied to the other 4 IHC stained slides (noted above). We then applied different algorithms and approaches to obtain stain-specific metrics from all the slides for each cortical layer. The Area Quantification module was used for determining the area of positive staining for all proteins of interest (NeuN, GFAP, Iba1, A β , pTau, α -Syn, and pTDP-43). The Multiplex IHC module was used to determine the number of cells displaying positive labelling for NeuN, pTau, α -Syn, and pTDP-43).

For double labelled slides, the Multiplex IHC module was used to estimate the area of co-localization of pTau with pTDP-43, and A β with Iba1. The Microglia Activation module was used to determine the number of cells positive for Iba1, measure the cell process area and length, and classify the cells according to the activation state (activated vs not, based on the process thickness). The same module was adapted to estimate the process length and process area for cells positive for GFAP. In the slides double-labelled for A β and Iba1 the Object Colocalization module was used to determine the number of A β -positive objects (amyloid plaques), the average object area, median object diameter, and the number of objects that were double-positive for A β and Iba1. The final dataset includes raw measurements (absolute values) and metrics normalized to the unit area.

Feature quantification and quality control measures

Given that generating large quantitative datasets is not yet standard in the field of neuropathology, extensive quality control and standardization was performed to generate a reproducible pipeline that could be used on the initial dataset (MTG) and across brain regions. To this end, a custom script was generated in Python to enforce consistency with column naming (rename dictionary), assert if values were logical (e.g., no percentages above 100%), and align dataset structure. With few exceptions, the quantitative neuropathology datasets generated for the SEA-AD study contain a set of core features for each stain which are quantified (Table 2.2). Some features such as A β pathology (6e10) and IBA1 (microglia), performed as a duplex stain, were amenable to a larger list of robust and biologically relevant measurements. Whereas the monoplex GFAP stain needed to be constrained to fewer standard measurements, due in part to high background in certain areas.

AT8 & pTDP-43	6e10 & IBA1	GFAP	NeuN	aSyn
Total area analyzed	Total area analyzed	Total area analyzed	Total area analyzed	Total area analyzed
Total [x] positive area	Total [x] positive area	Total [x] positive area	Total [x] positive area	Total [x] positive area
Percent [x] positive area	Percent [x] positive area	Percent [x] positive area	Percent [x] positive area	Percent [x] positive area
Number of [x] positive cells	Total [IBA1] positive process length; Total [IBA1] positive process area		Number of [x] positive cells	Number of [x] positive cells
Number of [x] positive cells per area	Number of [6e10] positive objects; Number of [6e10] positive objects per area		Number of [x] positive cells per area	Number of [x] positive cells per area
Average [x] positive cell area	Average [6e10] positive object area; Average [6e10] positive object diameter		Average [x] positive cell area	Average [x] positive cell area
	Number of [IBA1] positive cells; Number of [IBA1] positive cells per area			
	Average [IBA1] positive process area per cell; Average [IBA1] positive process length per cell			
	Number of [6e10] positive object co-localized with [IBA1] positive cells; Number of [6e10] positive object co-localized with [IBA1] positive cells per area			
	Percent [6e10] positive object co-localized with [IBA1] positive cells; Percent [6e10] dense core plaque area; Percent [6e10] diffuse plaque area; Percent [6e10] fibrillar plaque area			

Table 2.2. Quantified features extracted from HALO image processing. Most features consist of a core set of measurements. Some categories such as the amyloid beta (6e10) and microglia (IBA1) duplex stain were amenable to more specialized measurements (e.g., average process length).

Results

Quantitative neuropathological assessment of pathological peptides and cell types

The dataset includes whole slide imaging and quantitative neuropathology data generated from 84 donor brains sampled for the Middle Temporal Gyrus (MTG). To summarize the tissue methods, following paraffin embedding and sectioning at 5 μ m, each set of donor slides are subjected to seven immunohistological stains including: duplex AT8 and TDP-43, duplex 6e10 and Iba1, GFAP, NeuN, and hematoxylin (Figure 2.7). AT8 is a marker of both early pTau aggregates and mature NFTs. AT8 and pTDP-43 are often comorbid pathologies, so they were stained in duplex to enable the assessment of the potential synergy between the two, where elevated levels of TDP-43 are associated with increases in pTau². Similarly, 6e10 labels A β and is performed as a duplex stain with IBA1 to enable the assessment of microglial elevation in response to plaque accumulation³. GFAP labels astrocytes which both respond to AD pathology and are thought to be altered by the course of disease⁴. NeuN is a marker of mature neurons, used to establish a baseline number of neurons per tissue area. Pathological α -synuclein (the basis of Lewy bodies) is also often comorbid with AD pathology and was included to assess potential relationships with other neuropathological features, though like pTDP-43 in in this cohort, it was present in only a small subset of donors, so was not sufficiently powered for quantitative analysis.

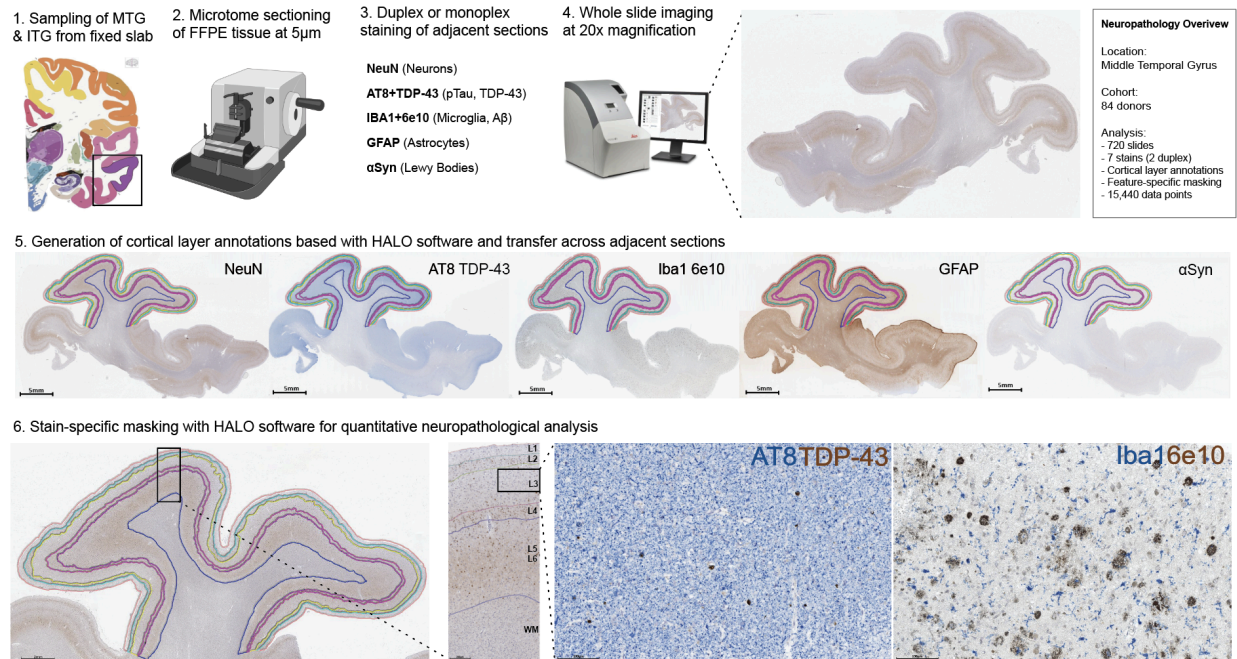


Figure 2.7. Schematic depicting quantitative neuropathology workflow. The top panel summarizes tissue sampling, sectioning, multiplex or duplex staining, followed by WSI. The middle panel shows representative transfer of cortical layer annotations using the HALO ML-based image processing. The bottom panel highlights cytoarchitectural features of the cortical layers, with an inset of pTau (AT8) and Aβ (6e10) staining.

Quantitative analysis was performed using HALO software's ML-based platform as described above. Cortical layer annotations and masks to identify relevant pathological features based on immunohistochemical staining were applied to all 720 whole slide images, resulting in over 15,000 data points. All annotations and masking were reviewed by an expert neuropathologist and determined to be highly consistent and accurate (Figure 2.8).

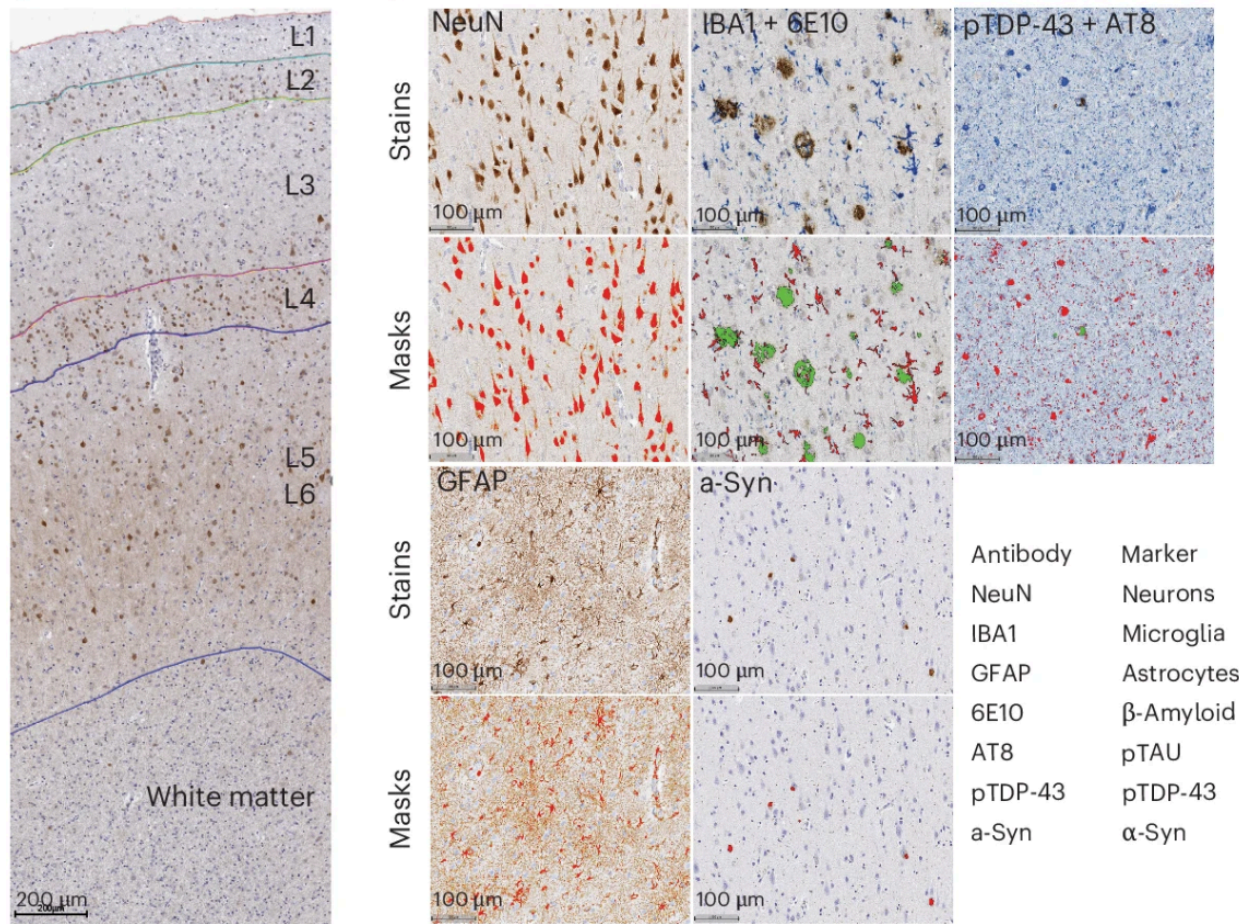


Figure 2.8. Representative cortical column visualized with immunohistochemical (IHC) staining. Cortical layers (L1–L6) and white matter are indicated. Immunostaining was performed in the entire SEA-AD cohort (n = 84). On the right-hand side, higher-powered micrographs showing IHC staining for protein aggregates and cellular populations. Bottom, masks showing positive voxels generated by HALO in red for single staining and both red and green for duplex staining. Immunostaining was performed in the entire SEA-AD cohort (n = 84).

Understanding the dynamics of each quantified feature was especially important for this dataset given the unprecedented scale and level of complexity of the data. To assess feature dynamics and relative contributions each normalized quantification was correlated against one another and visualized in a heatmap. The correlation heatmap captures all features quantified and groups them into correlation clusters (1-8) along both axes (Figure 2.9). Interestingly, several off-diagonal clusters emerge which indicate sets of features from different stains that correlate with one another. The red box shows an area of distinct features, NeuN immunoreactivity [66] that the immunoreactivity of mature neurons to the NeuN antibody is decreasing with increasing pathology. The blue box indicates an area of positive correlation between pTau pathology (AT8

antibody) and pTDP-43 pathology, which is a known relationship, as increased TDP-43 pathology has been shown to be associated with increased pTau pathology in affected brain regions.

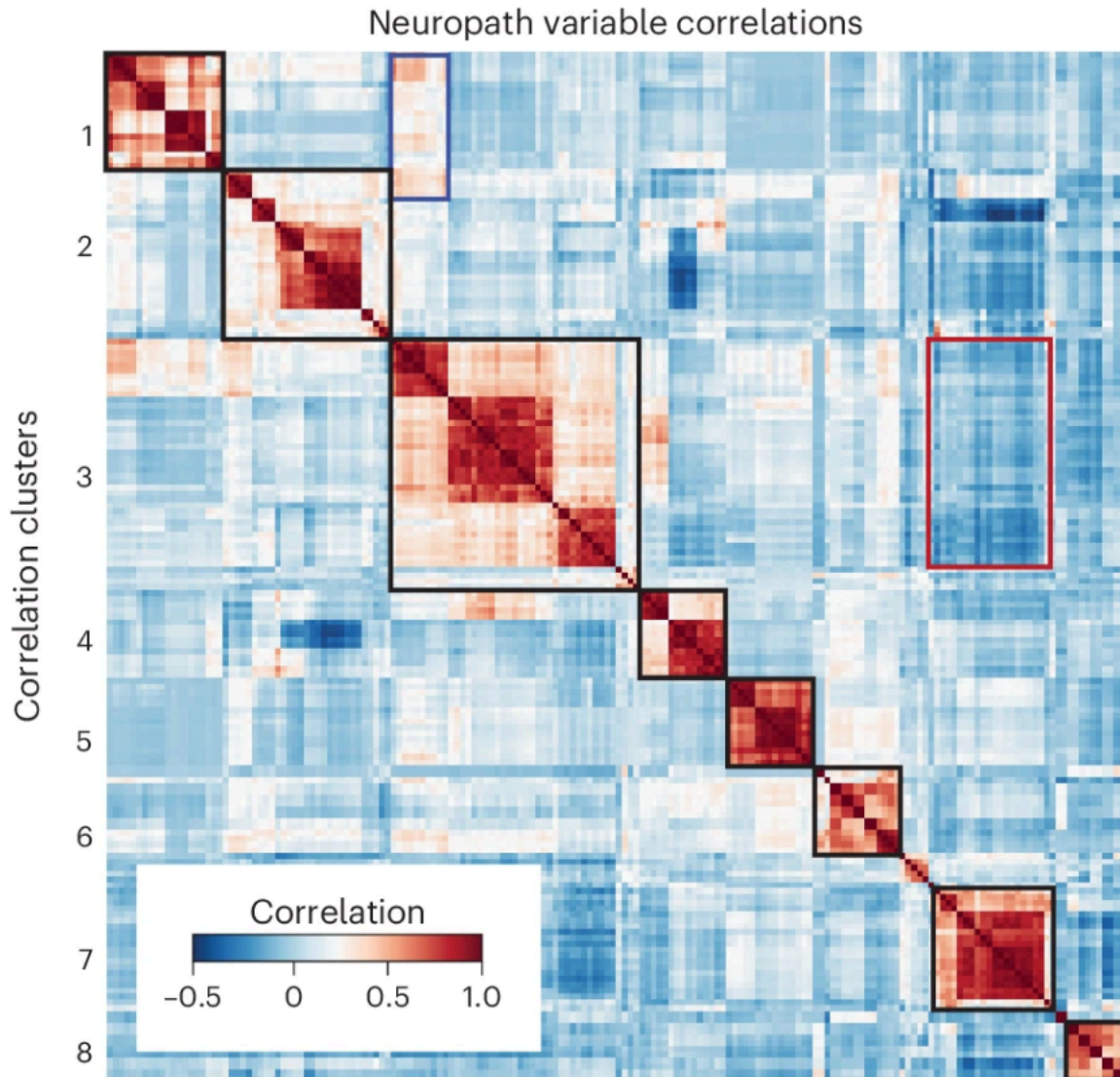


Figure 2.9. Correlation heatmap of normalized quantitative neuropathological features. Highlighting a hierarchically organized co-correlation matrix of quantitative neuropathology variables. The black boxes on the diagonal indicate eight correlated clusters. The red box indicates the anti-correlation representing AD protein pathologies and NeuN immunoreactivity (NeuN-ir), respectively. The blue box indicates the correlation between variables related to NFTs and pTDP-43 variables.

Quantitative measures of pathological proteins track with traditional clinical staging metrics

Traditional clinical staging for Alzheimer’s disease (AD), including Braak Staging and Thal Phasing, relies on semi-quantitative brain-wide assessments of key pathological proteins to definitively diagnose AD. To confirm that the quantitative neuropathological data capture these discrete semi-quantitative diagnostic categories, normalized layer-specific count values for pTau and A β (number of positive objects divided by area analyzed) were visualized as a function of Braak and Thal Phase, respectively. While the quantitative neuropathology values do correlate closely with traditional staging metrics, this approach revealed substantial variation of pathological protein levels within each discrete stage (Figure 2.10). This underscores the limitations of semi-quantitative scores and highlights the importance of quantitative measures in neuropathological data. This means that donors can be further subgrouped into high or low pathological groups within a given stage and more importantly, opens the opportunity to rank donors continuously related to their actual pathological burden rather than discretized stages.

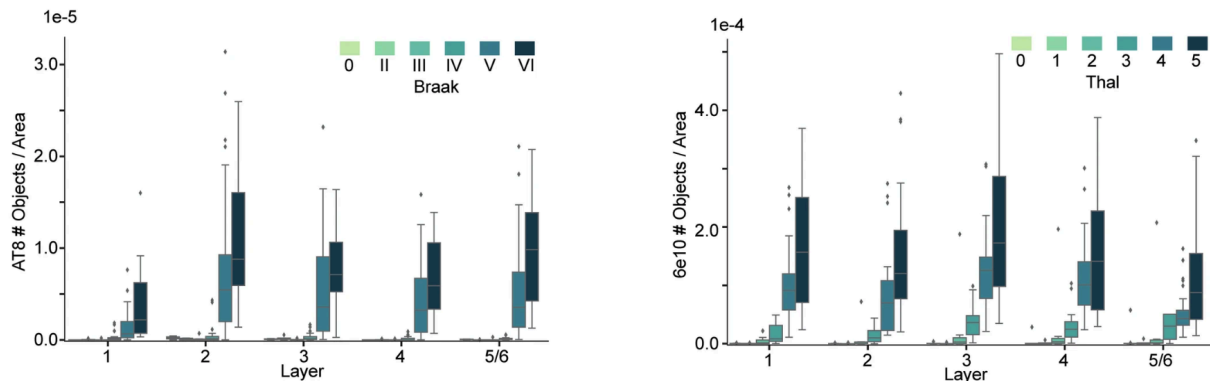


Figure 2.10. Boxplots quantifying the number of pTau-bearing cells per unit area. Organized by Braak stage (left) and number of A β plaques per unit area organized by Thal phase (right) across donors. Note, in later stages there is considerable variability in plaque and tangle number, underscoring limitations in classical staging.

Quantitative measures of pTau and A β in MTG predict cognitive outcomes comparable to brain wide metrics (Braak Staging and Thal Phasing, respectively). To determine whether features of interest, such as pTau and A β , can predict relevant clinical outcomes, a generalized linear model was employed where dementia status was the binary outcome, and either quantitative neuropathological metric or traditional staging metric was used as the predictor (accounting for known covariates of sex and age). By comparing standardized beta coefficients, it was apparent that the quantitative assessment of pTau in the MTG was a better predictor of the outcome of dementia status than Braak Staging. Interestingly, the quantitative assessment of A β

in the MTG cannot predict the outcome of dementia status as well as Thal Phasing (Figure 2.11). Both the Receiver Operating Characteristic (ROC)–based area-under-the-curve (AUC) and the pseudo- R^2 are measures of model fit. In this case, values below 1 indicate that while the models capture some of the variance in dementia status, they do not provide perfect classification. This is expected in a complex, multifactorial condition like Alzheimer’s disease, where neuropathological metrics are important but not the sole determinants of clinical outcome.

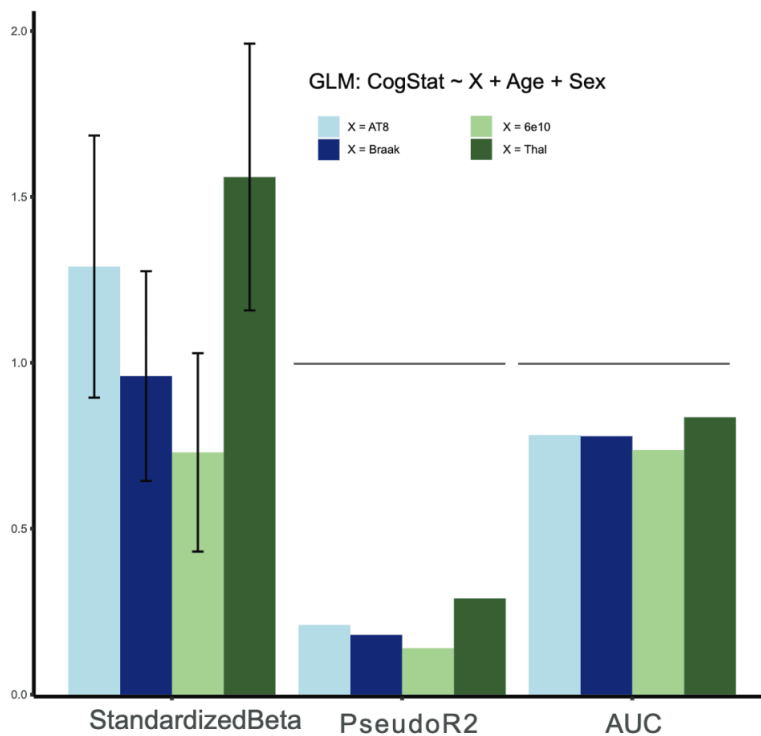


Figure 2.11. Barplot summarizing results from a generalized linear model predicting dementia status. Model (CogStat) as the outcome and either traditional staging (Braak stage or Thal phase) or quantified protein pathology (pTau or amyloid beta) as the predictor. Sex and age were accounted for as covariates. The single region (MTG) quantifications of pathology predict dementia status slightly better than the brain wide non-quantitative staging metrics.

PCA reveals key neuropathological features across disease progression

To determine which quantitative neuropathological features capture the most variance within the dataset, we performed PCA on all normalized features. The top features along PC1 are visualized as a heatmap and capture key AD-relevant categories, including increasing $A\beta$,

increasing microglia co-localizing with A β , increasing pTau, and decreases in nucleus size and NeuN immunoreactivity (Figure 2.12). The Alzheimer's Disease Neuropathological Change (ADNC) score of each donor is shown at the top of the heatmap and demonstrates that ADNC score tracks with the disease-associated quantitative neuropathology metrics captured by PC1. This approach captures an inherent ordering of the donors by ADNC and highlights key disease-relevant categories within our dataset.

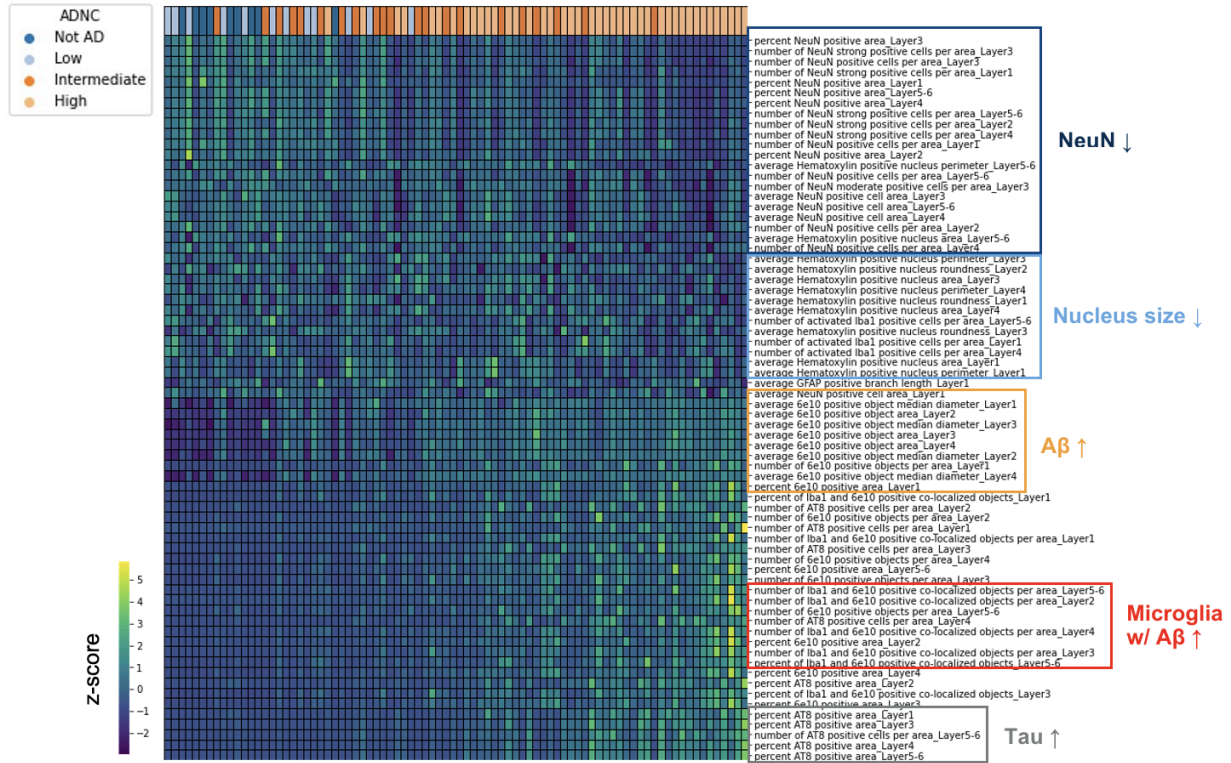


Figure 2.12. Heatmap of normalized neuropathological features. The y-axis includes the top features along PC1 and donors are along the x-axis. The ADNC score for each donor is color coded along the top of the heatmap and shows a clear ordering of donors from low to high disease burden (Not AD to High ADNC).

Specific cortical layers are preferentially affected in disease

Using the ML-based platform, HALO, cortical layers were assigned to each whole slide image. Retaining cortical layer information in the analysis of neuropathological specimens is a significant advantage. Braak & Braak's 1991 observations highlight that pTau accumulates in the cortex in a highly layer-specific manner. These qualitative observations have never been validated in a large quantitative dataset. As such, it was of interest to determine whether

quantitative neuropathology data was able to capture these subtle qualitative observations. Interestingly, layer-specific pTau patterns were captured in the quantitative neuropathology data; however, significant layer-specific enrichment was only captured with some pTau-related metrics, but not others. More specifically, the percent pTau-positive area, which captures a wider range of positive features, from neuritic pTau in the neuronal processes to extracellular pTau in neuritic plaques, does not show layer-specificity. However, when using the normalized number of pTau positive cells, which captures primarily NFTs in neurons, a clear layer-specific pattern emerges (Figure 2.13). There was not a discernable layer-specific pattern for A β .

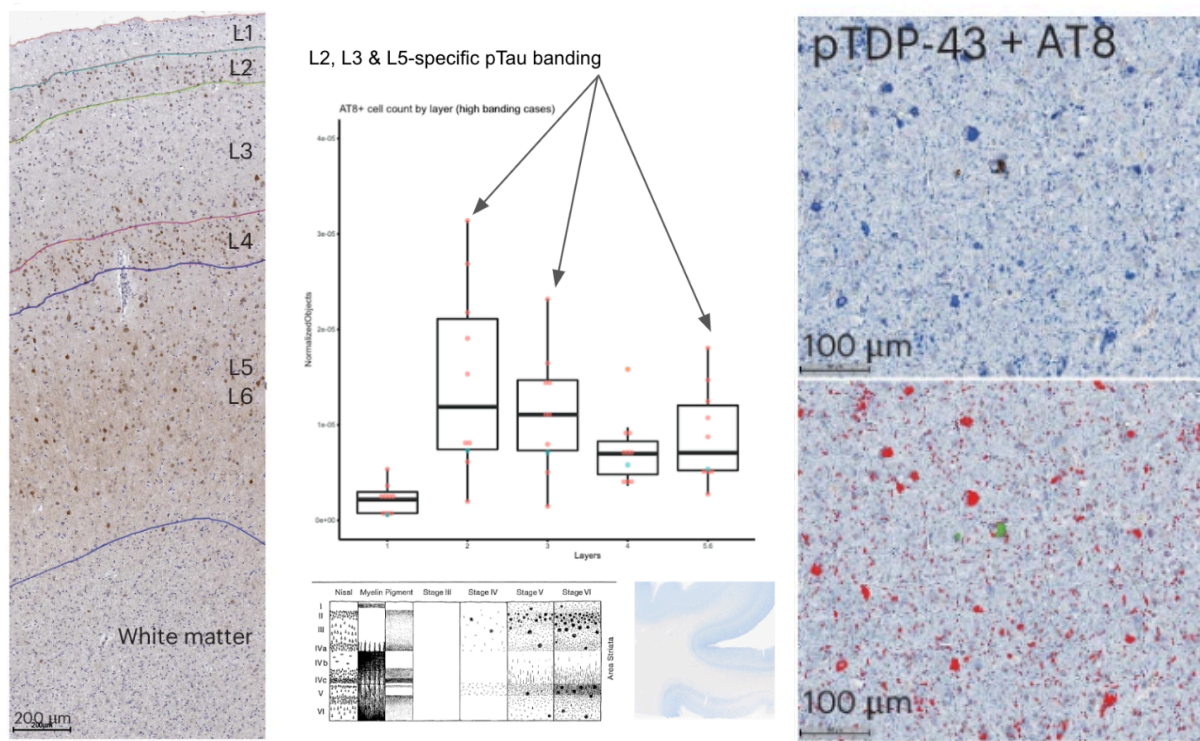


Figure 2.13. Specific cortical layers preferentially accumulate pTau. Boxplots quantifying AT8 levels across the cortical layers (Layer 1 - 5/6), showing enrichment in layers 2/3 and layers 5/6, as expected from qualitative assessment of the tissue that is referred to the stereotype Alzheimer's tau 'banding' pattern in the cortex.

Leveraging quantitative neuropathology to generate a continuous pathology score

At the onset of the SEA-AD study, there were no established frameworks for incorporating quantitative neuropathology data into genomics analyses. Additionally, a major

limitation in AD research is the reliance on ordinal staging systems (e.g., Braak Staging) to characterize disease progression. The quantitative neuropathology component of SEA-AD, presented a powerful new approach to address lack of granularity and loss of statistical power associated with binning donors into several neuropathological stages. Traditionally, the field has relied on Braak staging, a semi-quantitative system that assigns donors to one of six stages based on the distribution of pTau pathology across several stereotyped brain regions. However, as highlighted earlier in this chapter, quantitative neuropathology measurements of pTau revealed substantial variability within each Braak stage. Similar patterns of high variability were observed for Thal phases, which assesses distribution of A β plaque pathology. More integrative classification systems that incorporate both pTau and A β pathologies, such as Alzheimer's disease neuropathologic change (ADNC), still reduce donors into discrete bins, limiting resolution and statistical power.

As a part of the collaborative effort, the co-first authors of the SEA-AD MTG study, Mariano Gabitto and Kyle Travaglini, developed a continuous metric of pathology called the continuous pseudoprogression score (CPS). This data-driven pathology metric takes normalized quantitative neuropathology data as input, excluding sparse metrics (such as pTDP-43 and α -synuclein). As an output, each donor is assigned an ordered position on a continuous scale from 0 to 1, with zero being the least aggregate pathological state and one being the highest. Ordering donors along CPS correlates closely independent donor-level clinical data (Braak Stage, Thal Phase, and ADNC) despite those metrics not being included in the model development (Figure 2.14). This approach increased the power of key downstream analyses of the SEA-AD data, including relative abundance analysis (using scCODA) and differential expression analysis (using Nebula), resulting in the discovery of disease-related cellular changes even in the sparsest, and typically overlooked cellular populations in the context of AD.

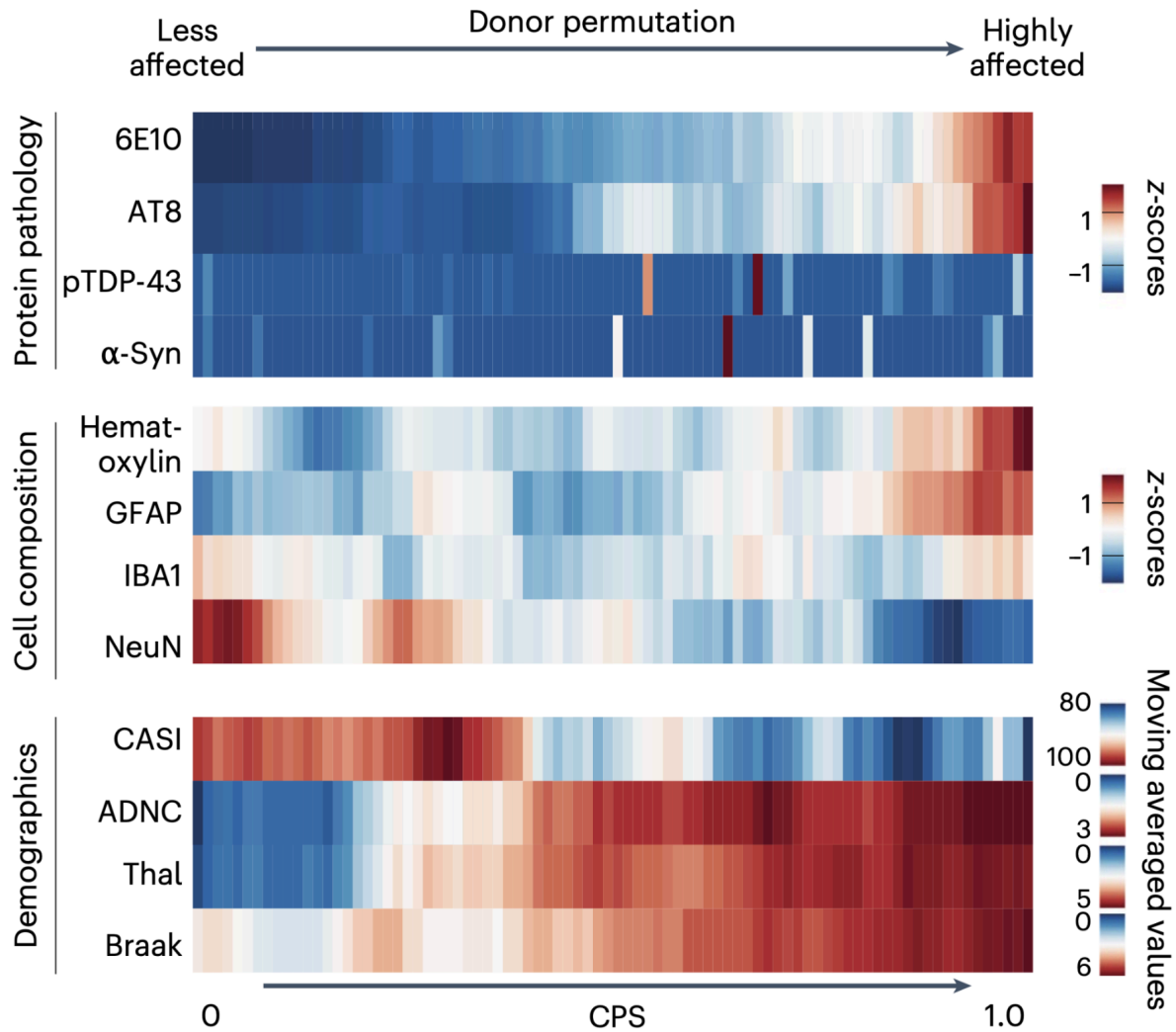


Figure 2.14. Pseudoprogression score as continuous metric to order donors by pathological burden. Heatmap (top) showing the number of pathological protein objects detected per unit area across all cortical layers in each donor, ordered along a CPS. All values were converted to z-scores and adjusted according to a moving average. Heatmap (middle) showing the number of cellular objects detected per unit area, ordered along the CPS. Hem, hematoxylin+ nuclei; GFAP, IBA1 and NeuN indicates the number of positive cells. All values were converted to z-scores and adjusted according to a moving average. Heatmap (bottom) showing the cognitive scores at the last visit (CASI) and AD pathology stage (ADNC, Thal, Braak), ordered along the CPS. All values were adjusted according to a moving average.

Discussion

Modernizing neuropathological disease staging with quantitative neuropathology

The SEA-AD study demonstrates how large-scale quantitative neuropathology data can be used alongside traditional diagnostic assessments. SEA-AD data was generated using standard methodologies employed in most neuropathology cores – (FFPE tissue stained histochemically for AT8 and 6e10), making it an ideal testing ground to explore how quantitative data can complement standard neuropathological assessments. Further, it can serve as the basis for generating a continuous score of donor pathological state to increase the statistical power of complementary genomics analyses from the same donors, as in the case of SEA-AD. By applying consistent and scalable methods to quantify these histological stains, SEA-AD demonstrates how these approaches can be combined to increase resolution and statistical power of analyses aimed at pinpointing the cellular underpinnings of AD.

This approach can be applied to any neurodegenerative disease, provided that the relevant cell types and protein pathologies are being stained and quantified. Combining quantitative neuropathology with complementary data on highly granular transcriptomically-defined cell types from the same donors will enable unprecedented insight into the cellular specificity of disease vulnerability - bringing the field closer to a targeted understanding of neurodegenerative progression at the cellular level.

Challenges with new technologies: standardization and reproducibility

Quantitative neuropathology offers a transformative addition - and potentially an alternative - to traditional neuropathological staging frameworks. However, several barriers currently limit its widespread adoption. First, although digitization of diagnostic slides using WSI is becoming more common in neuropathology cores, it has still not become standard practice across the field. While digitization poses an initial barrier, a perhaps larger hurdle is access to the software tools and technical expertise required to extract high quality data from the scanned slides. Until these tools become more affordable, standardized, and user-friendly, broader implementation of quantitative neuropathology in clinical diagnostics is likely to remain limited.

Important progress can be made in the interim. First, before neuropathologists can meaningfully and safely incorporate a new data modality into their standardized diagnostic frameworks, rigorous analysis is needed to systematically compare traditional and novel quantitative approaches. While the introduction of more quantitative, high-resolution data is generally assumed to improve diagnostic accuracy and predictive power, it also carries the risk of introducing variability or noise that could undermine the robustness and reproducibility of current simpler staging systems. It will be essential to carefully evaluate how quantitative measures align with, complement, or potentially conflict with existing criteria.

Ongoing research efforts, particularly those described as a part of SEA-AD, play a critical role in laying this groundwork. By validating quantitative metrics against established neuropathological classifications, and by demonstrating their utility to enhance disease understanding and modeling, the field can move toward broader adoption. Ultimately, thoughtful integration of quantitative neuropathology has the potential to refine diagnostic precision and bridge the gap between clinical neuropathology and a mechanistic understanding of disease.

Chapter 3. LEVERAGING SINGLE CELL TRANSCRIPTOMICS TO STUDY THE RELATIONSHIP BETWEEN VULNERABLE CELLS AND PROTEIN PATHOLOGY IN AD

Introduction

This chapter aims to delineate precisely which cell types in the human MTG are selectively vulnerable to the accumulation of intracellular pTau. This extends upon SEA-AD's effort to study AD as a cellular disorder that preferentially affects unique subpopulations of cells over time as the disease progresses but adds an important layer of resolution - single cell protein pathology. Linking protein pathology to cellular vulnerability is a critical gap in our understanding of disease progression and will help inform development of pTau-based therapeutics, which unlike therapeutics for extracellular proteins, must carefully consider cell-type specificity.

Selective vulnerability is a common feature of neurodegenerative diseases, which can refer to the selective accumulation of pathologic peptides and/or selective vulnerability to cell death^{1,2}. The existing picture of selective vulnerability in AD encompasses several brain regions and broad cell classes, including noradrenergic neurons in the locus coeruleus, cholinergic neurons in the basal forebrain, stellate cells in the entorhinal cortex, and superficial glutamatergic neurons in the neocortex. Beyond these broad designations, the study of cellular vulnerability in neurodegenerative diseases has remained relatively static, based on traditional histological stains of fixed tissue. With the newfound ability to resolve highly specific cell types in human tissue, it is becoming clear that the diversity of cell types which are selectively vulnerable to pTau pathology is more intricate than previously appreciated.

Ongoing atlasing efforts, such as SEA-AD, have provided a unified picture of the distinct cell-type specific molecular changes that occur across AD progression. In the MTG, these AD-associated cellular changes can be categorized in an early and late phase based on overall neuropathological burden. The early phase includes microglial and astrocyte activation, as well as selective loss of Sst interneurons in superficial cortical layers, loss of myelinating

oligodendrocytes, and up-regulation of a re-myelination program by oligodendrocyte precursor cells. The later phase involves loss of excitatory neurons, such as superficial intratelencephalic projecting glutamatergic neurons, as well as superficial Pvalb and Vip inhibitory neuron subtypes. Taken together, this presents an intriguing picture where early loss of SST interneurons in superficial cortical layers may contribute to local circuit dysregulation, perhaps pushing excitatory neurons in the same layer towards a similar fate in the later stages of disease. While a massive step forward in our understanding of selective vulnerability in AD, these results do not consider the role of pathological protein accumulation, a hallmark of the disease.

To disentangle the relationship between pathology and cellular vulnerability, both protein and RNA must be assessed in the same cell. Recent advances have made this possible and scalable. A new methodological approach for isolating pTau-positive somas (i.e., neuronal cell bodies) from postmortem tissue for single-cell RNA sequencing was recently published³. This method enables high throughput assessment of the affected cell types and molecular changes associated with pTau pathology in human brain tissue. In this chapter, this method was used to isolate single somas with and without pTau pathology (referred to as pTau-seq), to enable the study of pTau-driven molecular changes in the context of highly resolved transcriptomic cell types.

pTau-seq, a variation on the standard scRNA-seq workflow, carefully retains the cell membrane while delicately removing cellular processes that would otherwise clog the microfluidics required for droplet-based single cell RNA sequencing workflows. This is a critical distinction as it allows for the retention of intracellular pTau which resides within the soma, rather than the nucleus. The process involves a gentle tissue dissociation without enzymes or detergents, followed by antibody-based sorting of the target protein (e.g., AT8 monoclonal antibody, recognizing pTau at two sites, serine 202 and threonine 205), followed by the standard 10x Genomics single cell RNA sequencing workflow.

This yields high coverage and depth of the complete transcriptome in every single cell, making precise cell type labeling and robust analysis of gene expression changes across conditions possible, while retaining protein level information. The resulting data can then be

used to precisely delineate which cells are vulnerable or resistant to pTau pathology, assess how pTau status affects gene expression dynamics, and compare pTau-vulnerable cells to cells which are vulnerable or resistant to cell death (cellular abundance changes).

In this chapter, the RNA content of each cell will be utilized in two primary ways to extract meaning from the dataset. First, we and others have concretely demonstrated that highly resolved cell types can be defined transcriptomically. While other cellular features, such as morphology and electrophysiology, are still critical to refining a complete cellular taxonomy, the unique composition and relative levels of RNA serve as the basis of this classification system. Of note, the field has not settled on a singular approach for cell type mapping or a singular taxonomy, as discussed in Chapter 1. Three main approaches for cell type mapping have become common - correlation-based, hierarchical, and deep learning (of which scVI is an example). These approaches use different mathematical strategies to estimate the most probable cell type label given that cell's RNA content, relative to other cells in the dataset. The pTau-seq dataset was mapped to the standardized SEA-AD taxonomy using deep-learning based approaches (iterative_scANVI described in methods) to maintain resolution and consistency of cell type labels with existing datasets.

Second, the RNA composition with a single cell can also be leveraged to understand how cell state is altered in different contexts, such as development or disease. This is also central to this work, as it enables the assessment of the pathological protein (pTau) state of individual cells in parallel with associated transcriptomic changes in individual cells. After mapping all cell types, comparison of the composition and levels of RNA in pathological (pTau-positive) versus non-pathological (pTau-negative) cells through approaches such as differential expression analysis, exposes how the transcriptomic state of each cell type is changed in the context of pTau pathology. Finally, performing Gene Set Enrichment Analysis (GSEA) with standardized Gene Ontology (GO) terms on these differentially expressed genes, allows for a simple and reproducible way to assess how the RNA-based dynamics may relate to functional changes. Results must be followed up with experimental data to confirm cause and effect, but serve as a starting point for mechanistic studies.

Methods

Donor selection

This study was structured as a two-by-two within and across subject design. For each donor, both pTau-positive and pTau-negative samples were collected, such that within each donor we could compare pTau associated changes within a given cell type, in addition to broad changes associated with a given pTau status that was retained across multiple cell types. Each donor was processed identically to enable comparisons of cell types and associated pTau-related changes that were retained across donors.

We included three groups based on Alzheimer's disease associated pTau pathology, Braak Stage IV (Alzheimer's Disease Associated Change (ADNC) Intermediate), Braak V (ADNC Intermediate or High), and Braak VI (ADNC High). We eliminated donors that were deemed clinical or neuropathologically complex or had specific comorbidities, such as Lewy Body Disease. Donors with more than two microinfarcts and age outliers (younger than 70 years old) were removed. Additionally, since these donors were assessed as a part of the larger SEA-AD study, donors with identified technical or biological outliers from other data modalities were not considered. Within each pathology group, best efforts were made to balance for sex and age, such that these potential confounding factors were also matched across groups. Importantly, a donor's dementia status, APOE genotype, and LATE (Limbic-predominant Age-related TDP-43 Encephalopathy) Stage were not considered in the selection process, though are noted. In total, twelve donors were selected, with four donors in each subgroup. Due to technical error in the steps leading up to sequencing, three donors ultimately needed to be removed from the analysis (Table 3.1).

AIBS ID	UWA ID	Age	Sex	PMI	RIN	pH	Cognitive status (Braak)	Additional notes
H20.33.018	7045	81	F	4.5	9.55	6.8	Dementia (Braak VI)	LATE Stage 2, APOE4
H20.33.024	6999	90+	M	3.9	9.5	7.2	No dementia (Braak V)	LATE Stage 2
H20.33.013 (!)	7041	90+	M	4.5	9.08	7.6	No dementia (Braak IV)	

H20.33.046	6966	90+	M	8	5.67	6.2	Dementia (Braak VI)	LATE Stage 2, SA
H20.33.031	7117	87	F	7.9	8.85	6.4	Dementia (Braak VI)	LATE Stage 2
H20.33.029	7101	90+	F	6.1	9.43	6.5	Dementia (Braak V)	LATE Stage 2
H20.33.015	7043	88	M	11.4	8.57	6.8	Dementia (Braak V)	LATE Stage 3
H20.33.016 (!)	7146	90+	F	7.7	8.82	6.8	Dementia (Braak IV)	
H20.33.040	7179	90+	M	5.03	9.17	6.4	Dementia (Braak IV)	
H20.33.008	7048	90+	F	7	8.68	6.4	No dementia (Braak V)	APOE4
H21.33.039 (!)	7370	88	F	6.5	7.2	6.5	Dementia (Braak V)	LATE Stage 3, SA
H20.33.028	7067	90+	F	7.4	7.9	6.6	Dementia (Braak V)	Late Stage 2, SA

Table 3.1. pTau-seq donor cohort with associated metadata. Note three donors that failed QC before library preparation (H20.33.016, 7146), (H21.33.039, 7370), and (H20.33.013, 7041). SA indicates donors that were designated as Severely Affected based on transcriptional changes described in the SEA-AD publication. Donors with (!) failed QC upstream of sequencing.

Power analysis

Power analysis was performed before onset of the pTau-seq study. At the time, the primary goal was to achieve sufficient supertype representation. The strategy employed was modeled after the SEA-AD study where effect sizes were expected to range from 0.6 to 2.0 - where the measure was whether a cell type had significantly changed (adjusted p-value < 0.05) relative abundance along the continuous pseudoprogession score (CPS). For low abundance supertypes (e.g., Sst), it was determined that >80 donors were required to detect real changes in abundance if the effect sizes were more subtle (0.6 to 1.4). This approach suggests that in the highest abundance cell types (e.g., L2/3 IT) only ten donors would be needed to detect real changes in abundance, down to an effect size of 1.0; however to detect effect sizes at or below 0.6, more donors may be needed. The focus of the study ultimately became centered around whether pTau-positive populations (binary) were significantly increased or decreased at the supertype level (within type comparisons), particularly in novel AD-vulnerable types (namely, Sst interneurons). The initial assumption was that the biological baseline would be close to zero (and modeled as 0). If this were the case, simple power analysis for a paired t-test where baseline

= 0.0, power_target = 0.8, alpha = 0.05, and min_cells = 30 suggests that 12 donors would be sufficient to detect even the smallest effect sizes in low abundance types. This initial assumption that the null expectation (pTau baseline) would be zero was not entirely correct, as the majority of cell types express moderate levels of pTau, according to the pTau-seq and Otero-Garcia data. Changing the null value however, dramatically changes the power calculation. This further underscores the importance of understanding what the true null value should be for basal pTau levels in non-affected cell types.

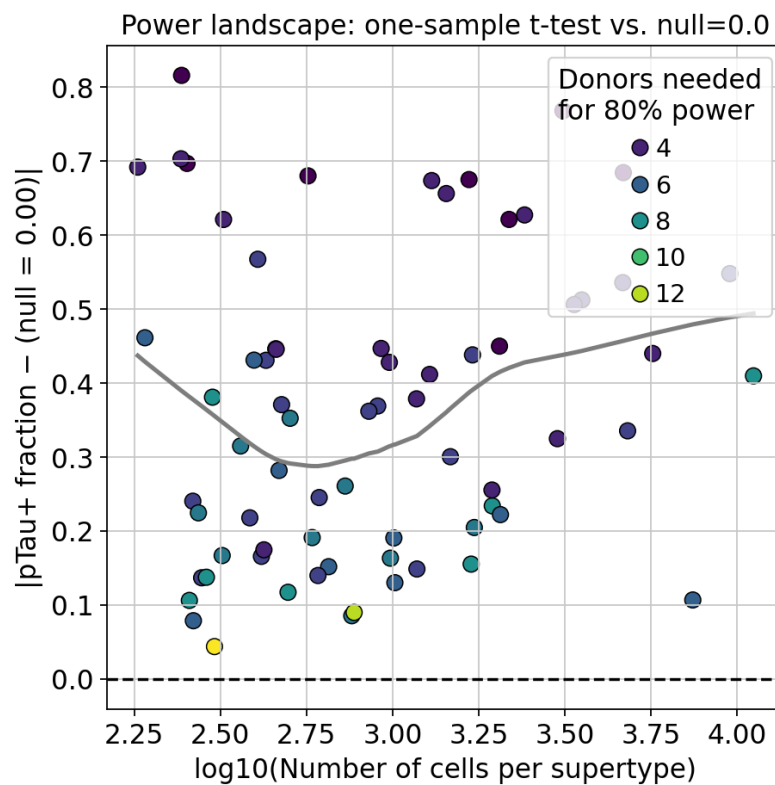


Figure 3.1. Power calculation for pTau-seq. Null value = 0, (i.e. at baseline healthy cells have zero pTau) suggest maximum sample size needed is N= 12 donors.

Single soma sample preparation

Single soma isolations were performed according to the methods published in Otero-Garcia et al., 2022. In brief, the protocol involves a gentle tissue dissociation, fluorescence activated cell sorting (FACS) to isolate pTau-positive and negative somas, and 10x Genomics

single cell processing. The dissociation was performed without detergents or enzymes, using a specialized tissue grinder (Potter Elvehjem, eight slow strokes), which reduces shear force due to its wider clearance. Next, sucrose-iodixanol gradient centrifugation was used to separate the debris from an enriched soma fraction. Samples were reviewed under a bright field microscope to confirm the presence of well-preserved somas, some with small processes still intact, and confirm little or no nuclei were present. Isolated somas were then immunohistologically stained for AT8 (pTau) and MAP2 (Microtubule Associated Protein 2, a marker of mature neurons) in a PBS-based buffer.

Below is the adapted protocol from Otero-Garcia et al., 2022.

Protocol

- *Aims for minimal mechanical damage (as little centrifugation and pipetting as possible)*
- *Material and tools needed: Forceps, spatula, blades, dounce tissue grinder (Potter-Elvehjem tissue grinder), cell strainers, petri dishes*
- *All solutions and materials should be RNase free and kept iced or at 4 °C at all times*
- *Glassware and metal tools are sealed with aluminum foil and baked at 220 °C for 6 h*

(In advance) Prepare Isolation Medium

1. *Prepare in a 50 ml Falcon tube and store at 4°C up to 6 months.*
 2. *7500 µl 1.5M Sucrose (final concentration 250mM)*
 3. *1125 µl 1M KCl (25mM)*
 4. *225 µl 1M MgCl₂ (5mM)*
 5. *450 µl 1M Tris pH8 (10mM)*
 6. *35.7 ml H₂O*
-
1. *Prepare Homogenization Buffer and cool it on ice.*
 - a. *Homogenization buffer (3mL per sample)*
 - i. *Prepare FRESH and keep ICED or at 4°C. Discard after use.*
 - ii. *2955 µl IM1*
 - iii. *3 µl DTT 1mM (final concentration 1µM)*

- iv. 30 μ l 50x Protease Inhibitor (0.5x)
 - v. 15 μ l RNaseIn 40U/ μ l (0.2U/ μ l)
 - vi. No Triton!
- b. Pre-cool the dounce tissue grinder on ice
- c. (Potter-Elvehjem tissue grinder; 8 mL tubes, 0.1–0.15 mm clearance space between pestle and tubes (Kimble, K886000-0022))
- d. Add 2.4 ml of Homogenization buffer to the dounce tissue grinder.
- 2. Collect the brain chunk (~200 mg) and transfer to a Petri dish on ice.
 - a. Cut out into small pieces using a chilled scalpel or blade.
- 3. Transfer all pieces into the dounce tissue grinder.
 - a. Slowly homogenize the tissue, on ice(8~15 strokes; slow).
 - b. Check on a hemacytometer while homogenizing and adjust the number of strokes (too few strokes – you will see clumps; too many strokes – you will get a higher proportion of naked nuclei).
- 4. Filter homogenate using a 100 μ m Corning cell strainer to remove clumps.
 - a. Take sample for hemocytometer.
- 5. Transfer the homogenate into two precooled 1.5 ml Eppendorf tubes.
 - a. Centrifuge at 400g for 5 min at 4°C.
- 6. Slowly aspirate the supernatant from the side of the tube.
 - a. Avoid disturbing the pellets, up to
 - b. 50 μ l of supernatant can be left.
 - c. Gently resuspend each pellet in 225 μ l (final volume) of cold Homogenization buffer and pool both tubes (final volume 450 μ l).
- 7. Iodixanol solutions
 - a. Iodixanol dilutions
 - b. Prepare in a 50 ml Falcon tubes and store at 4°C up to 6 months. Accuracy with Iodixanol and
 - c. sucrose concentrations is critical.
 - d. Iodixanol medium (IDM), 45 ml
 - e. 7500 μ l 1.5M Sucrose (final concentration 250mM)
 - f. 3750 μ l 1M KCl (83mM)

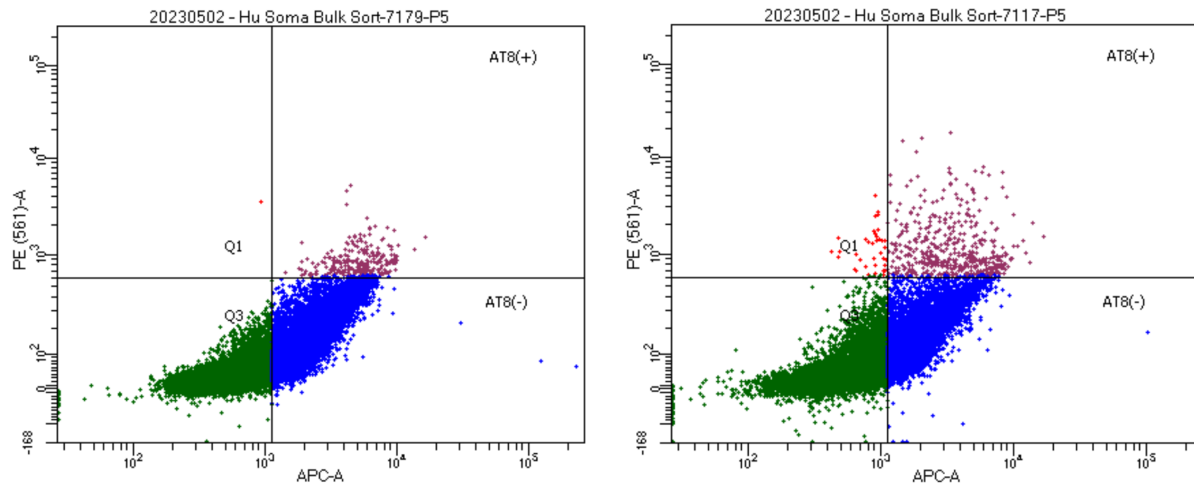
- g. 750 μ l 1M MgCl₂ (17mM)
 - h. 1500 μ l 1M Tris pH8 (33mM)
 - i. 31.5 ml H₂O
 - j. Iodixanol 42%(v/v), 20 ml
 - k. 14 ml Iodixanol 60%(v/v) + 6 ml IDM
 - l. Iodixanol 25%(v/v), 30 ml
 - m. 12.5 ml Iodixanol 60%(v/v) + 17.5 ml IDM
8. Add an equal volume of cold 42%v/v iodixanol. (Critical! Be very precise with volumes.
- a. 450 μ l suspension + 450 μ l 50% iodixanol; Final iodixanol concentration is 21%).
 - b. Gently pipette mix.
9. Add an equal volume of 25% iodixanol (900 μ l) into a 2 ml Eppendorf tube.
- a. Then, slowly layer off the 21% iodixanol suspension mix over the 25% iodixanol, without mixing them.
 - b. Centrifuge at 8,000g for 15 min at 4°C.
10. Prepare buffer for Immunostaining
- a. Buffer for Immunostaining, 10ml
 - b. 9400 μ l RNase-free PBS pH7.4
 - c. 500 μ l RNase-free 10% BSA (final concentration 0.5%)
 - d. 50 μ l 1M MgCl₂ (5mM)
 - e. 10 μ l DNase I (2000U/ml) (2U/ml)
11. After centrifugation, remove and discard the top myelin-rich debris layer (use a 1 ml pipette with the tip cut, or cotton swabs).
- a. Then, remove and discard the aqueous supernatant, without disrupting the somas pellet.
 - b. Avoid contaminating with the top layer.
 - c. Use a small amount of buffer for Antibody Incubation to resuspend the pellet and transfer the solution to a new tube.
12. Gently resuspend in 200 μ l of buffer for Antibody Incubation.
- a. Buffer for Antibody Incubation
 - i. Add to 1ml of Buffer for Immunostaining:

- ii. 5 μ l RNaseIN 40U/ μ l (0.2U/ μ l)
- b. Incubate for 15 min, at 4°C or iced, for blocking nonspecific staining.
- c. Take sample for hemocytometer.
- 13. Take a sample for unstained control.
- 14. Take sample for 2AB-only (nonspecific binding control).
- 15. Add primary antibodies (Ms-a-AT8 1:150 and Rb-a-MAP2 1:40)
 - a. Incubate on a shaker in cold room (4°C) for 40 min.
 (MAP2 antibody: Millipore #AB5622, p-Tau (AT8) antibody: Thermo Fisher #MN1020)
- 16. Prep Eppendorfs coated with BSA for collection
 - a. (to coat the Eppendorf tubes, fill them with 10% BSA solution in PBS for 5 min, rinse with PBS, and dry at 4°C overnight).
- 17. Washing. Add 500 μ l of buffer for Immunostaining and invert several times.
 - a. Centrifuge at 400g for 5 min at 4 °C. Carefully transfer all the supernatant (don't leave any supernatant!
 - b. It's ok to drag some pellet) to a new Eppendorf tube and resuspend the pellet in 400 μ l of buffer for Antibody Incubation.
 - c. Centrifuge the supernatant at 500g for 5 min at 4 °C to recover non-pelleted somas, resuspend them in 200 μ l of buffer for Antibody Incubation and pool for a final volume of 600 μ l.
- 18. Take samples for single color staining.
- 19. Add secondary antibodies (G-a-Ms Alexa350 1:500 and G-a-Rb Alexa647 1:500) and Sytox Green 1:400 (1:100 stock, final concentration 1:40.000)
 - a. Check the quality of the sample on hemocytometer. For 200 mg of tissue, we should have ~ 1 Million somas in ~0.6 mL.
- 20. Transfer to 7ml culture tubes.
 - a. Place on ice and bring them to the FACS facility for sorting.
- 21. Collect in BSA-coated Eppendorfs containing collection medium
 - a. Volume of collection medium should be ~1/5 of the expected final volume after collection.

22. Add BSA after collection for 1% BSA final concentration (tested for 10x Chromium v2 and v3). Different Single Cell assays tolerate different BSA concentrations, but lowering it may increase nuclei aggregation.

Fluorescent activated cell sorting (FACS)

FACS was used to isolate and collect single somas from each donor that were either pTau-positive or negative. All runs were performed on a FACSARIAII and processed with FACSDiva (version 8.0.2). Controls including secondary only, 647 isotype control, and PE isotype controls were included on every run. Samples were collected in a 1% BSA buffer. Two cellular populations were collected from each donor: AT8+/MAP2+ and AT8-/MAP2+ (Figure 3.1). Gates for MAP2 were relatively lenient as the intention was only to enrich neuronal types. The pTau-positive yield per sample ranged from (2,900 - 90,000 cells (cutoff)) (Table 3.2). Somas were collected in Eppendorf tubes that were precoated in BSA to avoid somas sticking to the tube walls. Samples with low initial yields were re-run with a larger tissue sample



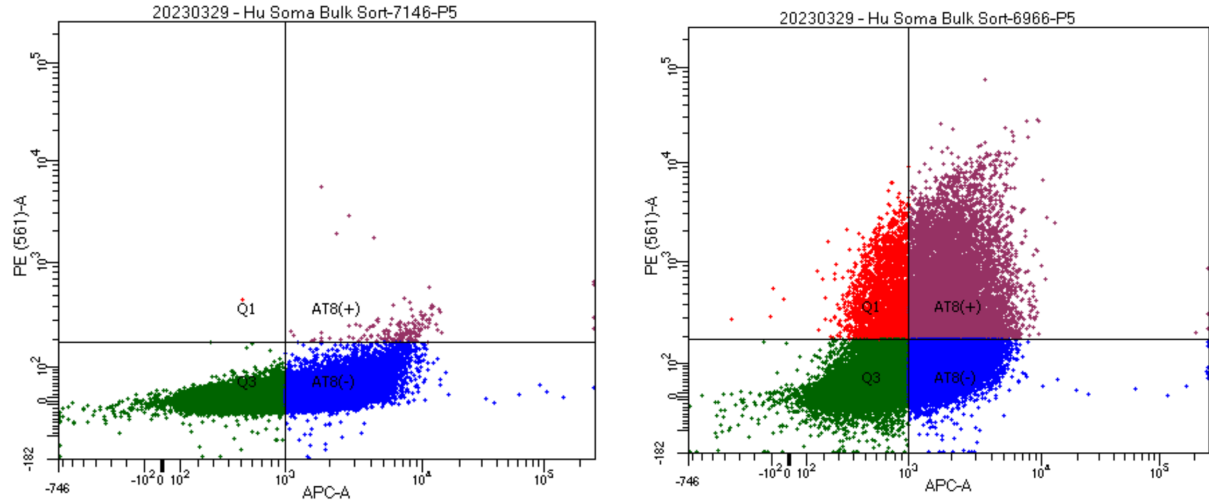


Figure 3.2. Representative FACS plots from two different batches. Top includes donor 7179 (left) and 7117 (right) from the same batch and bottom includes donor 7146 (left) and 6966 (right) from a different processing batch. Note that the two donors on the lefthand side are Braak stage IV and the two donors on the righthand side are Braak stage VI. The y-axis (PE) indicates pTau levels (AT8 antibody) and the x-axis (APC) indicates neuronal levels (MAP2 antibody).

Donor	Date	FACS (pTau-positive)	FACS (pTau-negative)
7117	3/27	3774 events from AT8(+)	90K events from AT8(-)
7179	3/27	3121 events from AT8(+)	90K events from AT8(-)
7370	3/27	8563 events from AT8(+)	90K events from AT8(-)
6966	3/29	90096 events from AT8(+)	90K events from AT8(-)
7067	3/29	9424 events from AT8(+)	90K events from AT8(-)
7101	3/29	4684 events from AT8(+)	90K events from AT8(-)
7146	3/29	2900 events from AT8(+)	90K events from AT8(-)
7041	4/4	23664 events from AT8(+)	90K events from AT8(-)
7045	4/4	90084 events from AT8(+)	90K events from AT8(-)
6999	4/4	20646 events from AT8(+)	90K events from AT8(-)
7043	4/11	22100 events from AT8(+)	90K events from AT8(-)
7048	4/11	16387 events from AT8(+)	90K events from AT8(-)

7117*	5/2	33689 events from AT8(+)	90K events from AT8(-)
7370*	5/2	25429 events from AT8(+)	90K events from AT8(-)
7179*	5/2	13681 events from AT8(+)	90K events from AT8(-)
7101*	5/3	4667 events from AT8(+)	90K events from AT8(-)
7146*	5/3	4922 events from AT8(+)	90K events from AT8(-)
7067*	5/3	50011 events from AT8(+)	90K events from AT8(-)

Table 3.2. FACS cell counts following pTau (AT8) sorting. Asterisk indicates repeated soma isolation.

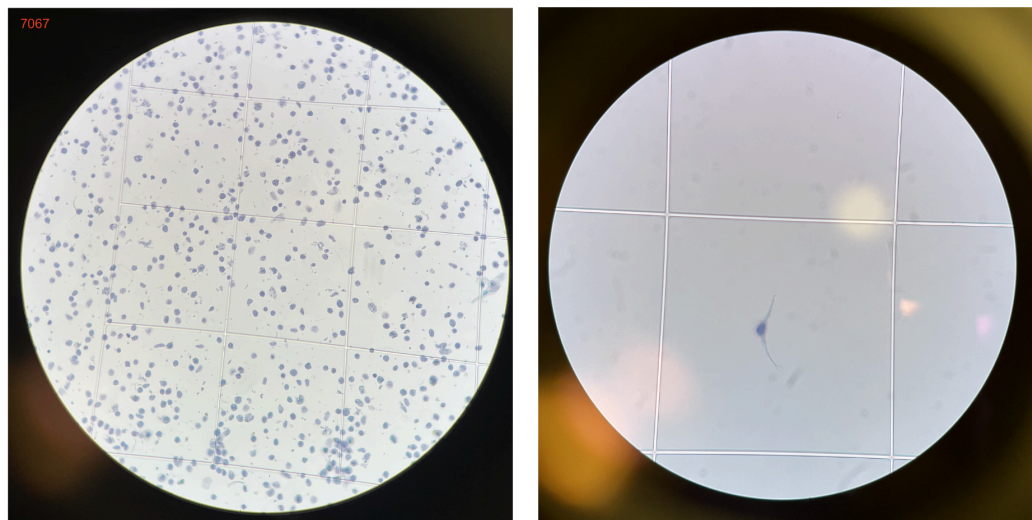


Figure 3.3. Representative bright field image pre and post FACS sorting. Donor 7067 after the soma isolation procedure only (left) and the soma isolation procedure followed by FACS and collection of the AT8+ fraction (right). Noting the presence of an intact appearing soma and retention of two processes.

10x Genomics sample processing

10x Genomics chip loading and postprocessing of the emulsions to the sequencing libraries were done with the Chromium Next GEM Single Cell 3' Gene Expression v.3.1, Nuclei concentration was calculated either manually using a disposable hemocytometer (DHC-NO1, INCYTO) or using the NC3000 NucleoCounter.

10x sequencing and preprocessing

For each sample, 10,000 somas were targeted for capture. This value was met for all pTau-negative samples and due to the expected variable amounts of pTau-positive somas per donor depending on their pathological burden, sequenced samples range from approximately 2,000 - 10,000 captured somas (Table 3.3). All 10x libraries were sequenced according to the manufacturer's specifications on a NovaSeq 6000 using either a NovaSeq X or S4 flow cell. Reads were demultiplexed to FASTQ files using BCL Convert (v.4.2.7) for libraries run on NovaSeq X flow cells and bcl2fastq (v.2-20-0) for libraries run on S4 flow cells. Reads from snRNA-seq libraries were mapped to the 10X Genomics official human reference (GRCh38-2020-A); UMIs per gene were counted using the Cell Ranger (v.6.1.1) pipeline with the --include-introns parameter included. Sequencing depth for all samples targeted 120k reads per cell. In line with the Otero-Garcia et al., pTau-positive cells had higher genes (median genes, pTau+: 5740 and pTau-: 4805) and UMIs (median UMIs - pTau+: 21102 and pTau-: 15189). With total genes being 19.5% higher in pTau-positive cells and UMIs being 38.9% higher in pTau-positive cells.

Donor	Date	10X Load (pTau-positive)	10X Load (pTau-negative)
7117 [not sequenced]	3/27	569 somas captured	10698 somas captured
7179 [not sequenced]	3/27	962 somas captured	10680 somas captured
7370 [not sequenced]	3/27	4159 somas captured	10700 somas captured
6966	3/29	10698 somas captured	10660 somas captured
7067 [not sequenced]	3/29	5704 somas captured	10724 somas captured
7101 [not sequenced]	3/29	1268 somas captured	10725 somas captured
7146 [not sequenced]	3/29	1073 somas captured	10680 somas captured
7041	4/4	7020 somas captured	10715 somas captured
7045	4/4	10689 somas captured	10725 somas captured
6999	4/4	10716 somas captured	10724 somas captured
7043	4/11	10691 somas captured	10760 somas captured
7048	4/11	10683 somas captured	10683 somas captured

7117*	5/2	10691 somas captured	8480 somas captured
7370*	5/2	10698 somas captured	10724 somas captured
7179*	5/2	10701 somas captured	10725 somas captured
7101*	5/3	2145 somas captured	10647 somas captured
7146*	5/3	2113 somas captured	10715 somas captured
7067*	5/3	10698 somas captured	10715 somas captured

Table 3.3. Capture rate of single somas following 10x processing.

Dataset integration and label transfer

To improve standardization and assess the results in a broader context, the pTau-seq dataset was integrated with the publicly accessible SEA-AD MTG snRNA-seq dataset of approximately 2 million cells using the Single Cell Variational Inference (scVI) package⁴. As described by Lopez et al., 2018, ScVI is based on a hierarchical Bayesian model with conditional distributions specified by deep neural networks, which can be trained very efficiently on very large datasets. The transcriptome of each cell is encoded through a nonlinear transformation into a low-dimensional latent vector of normal random variables. This latent representation is then decoded by another nonlinear transformation to generate a posterior estimate of the distributional parameters of each gene in each cell. The transformation assumes a zero-inflated negative binomial distribution, which accounts for the observed overdispersion and limited sensitivity. The model takes in raw gene counts (cell x gene matrix) as an input, where each cell is a vector of counts $x_i = [x_{i1}, x_{i2}, \dots, x_{ig}]$, where g = the number of genes. The observed expression x_{ng} of each gene g in each cell n as a sample drawn from a zero-inflated negative binomial (ZINB) distribution conditioned on the batch annotation sn of each cell (if available - scANVI), as well as two additional, unobserved random variables. The variational autoencoder (VAE), learns a distribution over the latent variables from which it can decode or reconstruct gene counts. This robust approach accounts for inherent noise in the data and batch effects, leading to very effective integration of samples and accurate label transfer.

As an extension of this approach, a custom adaptation of the Single Cell Annotation using Variational Inference (scANVI) referred to as ‘iterative_scANVI’ was used for label transfer to provide highly accurate and standardized cell-type levels down to supertype resolution as previously published. This adaptation is referred to as iterative because it iterates three times through the cell type label hierarchy. First, it applies cell type labels at the class-level, then proceeds to subclass labeling considering only subclasses that would be present in each class, followed by supertype labeling, considering only supertypes that would be present in that subclass. This approach was established as a part of SEA-AD and yields higher confidence cell type labels than when the entire dataset is reconsidered at each labeling level. Assigned cell type labels align as expected with common cell type marker genes (Figure 3.3).

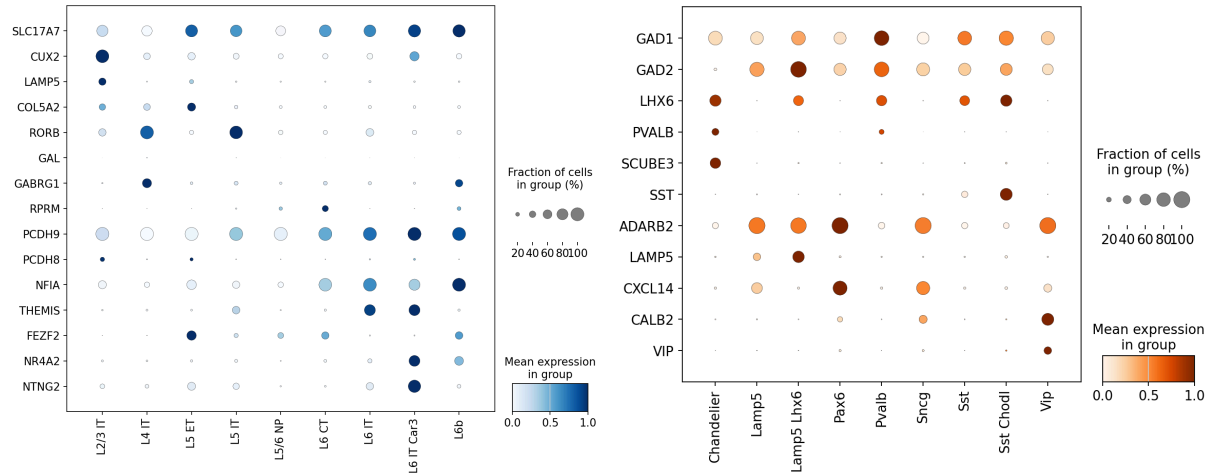


Figure 3.4. Subclass assignments align with standard cell-type marker genes. Excitatory cell types (left) and inhibitory cell types (right).

The custom implementation, iterative_scANVI, was executed as follows: the SEA-AD MTG dataset set as the reference dataset for cell type labels, the pTau-seq dataset as the query (i.e., no cell type labels at ingest), the batch effect as ‘Source’ (denotes SEA-AD versus pTau-seq, as this is the expected largest source of technical variation), the categorical covariate as ‘library_prep’ (this was selected instead of ‘Donor_ID’ because each donor has two samples, each from different library preparations for sequencing), and lastly dispersion set to ‘gene-batch’ (which dictates how the dispersion parameter (α) in the Negative Binomial (NB) model is estimated - defining ‘gene-batch’ sets the expectation that the variability (overdispersion) of gene

expression likely differs between donors even if the mean expression of a given gene is similar). The code to execute `iterative_scANVI` is shown below.

```
iterative_scANVI_kwargs = {
    "batch_key": "Source",
    "categorical_covariate_keys": ["library_prep"],
    "plot_latent_space": True,

    "scVI_model_args": {
        "n_layers": 2,
        "dispersion": "gene-batch"
    },
    "scANVI_model_args": {
        "n_layers": 2,
        "dispersion": "gene-batch"
    }
}

isc_mapping.iteratively_map(
    adata_ptau,
    adata_ref,
    output_dir=os.path.join(filepath, "models", region + "_" + dataset),
    labels_keys=["class", "subclass", "supertype"],
    **iterative_scANVI_kwargs
```

Cross-dataset integration varies in feasibility depending on how many core features of the dataset diverge. Examples of different types of integration include snRNA-seq data from different sequencing platforms (10x versus SMARTseq), different modalities (snRNA-seq vs snATAC-seq) or the same modality but different biological entities (somat (pTau-seq) versus nuclei (SEA-AD)) as in the present study. The current integration was successful given the parameters described above, with even mixing of donors and representation of all expected classes and subclasses with high mapping confidence (Figure 3.5). Due to the reduced overall and per donor dataset size relative to SEA-AD, not all 139 supertypes were represented across all donors in the dataset. Importantly, the twelve donors (nine post-QC) included in pTau-seq were also included in SEA-AD, thereby increasing the power of downstream analyses and enabling development of cross-dataset tools.

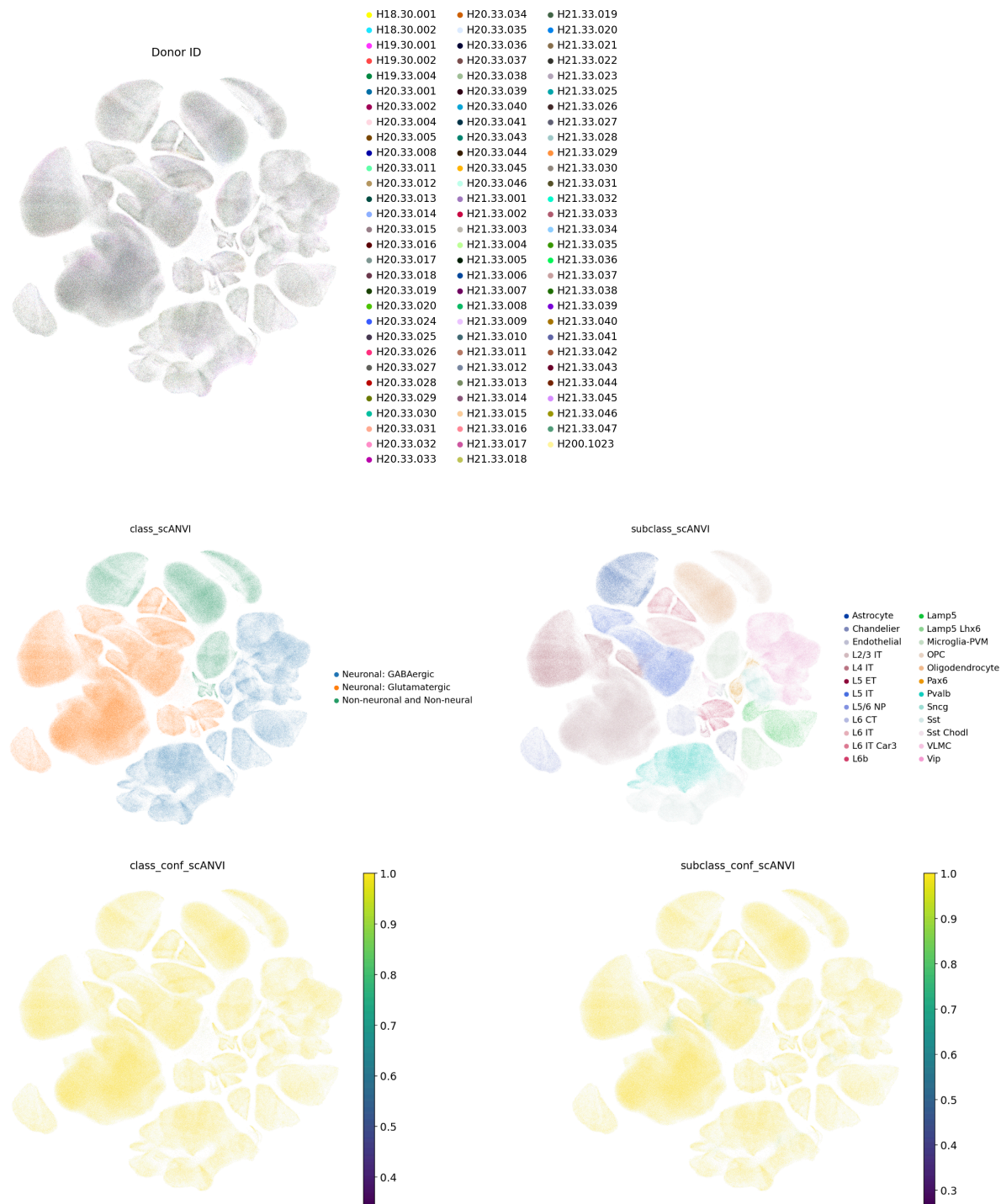


Figure 3.5. scVI integration of SEA-AD and pTau-seq datasets with high label transfer confidence. Uniform mixing of donors following integration with scVI. The integrated SEA-AD dataset includes non-AD reference donors, making the overall number of donors greater than 84. Model confidence scores (softmax probability) of class and subclass level label transfer.

Subclass-level leiden-cluster based quality control

An extensive per-subclass QC process was performed on the integrated data, which involved subsetting into individual subclass-level objects (24 objects total) and re-running iterative_scANVI on the subclass-level objects. Then, leiden clustering (resolution=5) was performed to aid in downstream filtering. Three key QC features - ‘Doublet score’, ‘Fraction mitochondrial UMIs’ and ‘Genes detected’ - were calculated on a per-cell basis. Subclass-wide distributions were plotted using histograms to aid in empirically selecting thresholds for subclass-specific QC cutoff values (Table 3.4). Doublet score assesses the probability that two cells were captured and falsely designated as a single cell. Genes detected is the total number of unique genes in a given cell. Fraction mitochondrial reads is a measure of cell quality, where a high fraction of mitochondrial reads indicates a lower quality cell. All excitatory neurons were assessed with the same thresholds. Inhibitory neuron

Cell Type	Doublet Score	Genes detected	Fraction Mito
Excitatory (all)	0.25	2500	0.04
Inhibitory - PAX6	0.17	2000	0.04
Inhibitory - LAMP 5	0.10	2000	0.020
Inhibitory - LAMP5 LHX6	0.15	2000	0.020
Inhibitory -PVALB	0.19	2000	0.050
Inhibitory -SNCG	0.095	2000	0.030
Inhibitory -SST	0.14	2000	0.030
Inhibitory - SST CHODL	0.18	2000	0.019
Inhibitory - VIP	0.15	2000	0.020
Astrocyte	0.08	1500	0.020
Endothelial	0.04	1000	0.03

Micro-PVM	0.04	1000	0.020
Oligo	0.05	2000	0.010
OPC	0.06	2000	0.040
VLMC	0.05	2000	0.030

Table 3.4. Empirically defined subclass level QC thresholds. Thresholds were established as a part of the SEA-AD study based on QC metrics from multiple cortical regions.

Remapping of supertypes at the subclass level results in very clearly delineated supertype-level clusters. Across the different subclasses, lower quality cells were consistently observed along the periphery of each cluster. This can be seen in the case of the supertype mapping confidence where the center of each supertype-level cluster in UMAP space has uniformly high mapping confidence (yellow), whereas the periphery has lower mapping confidence, generally defined as 60% of lower (Figure 3.6).

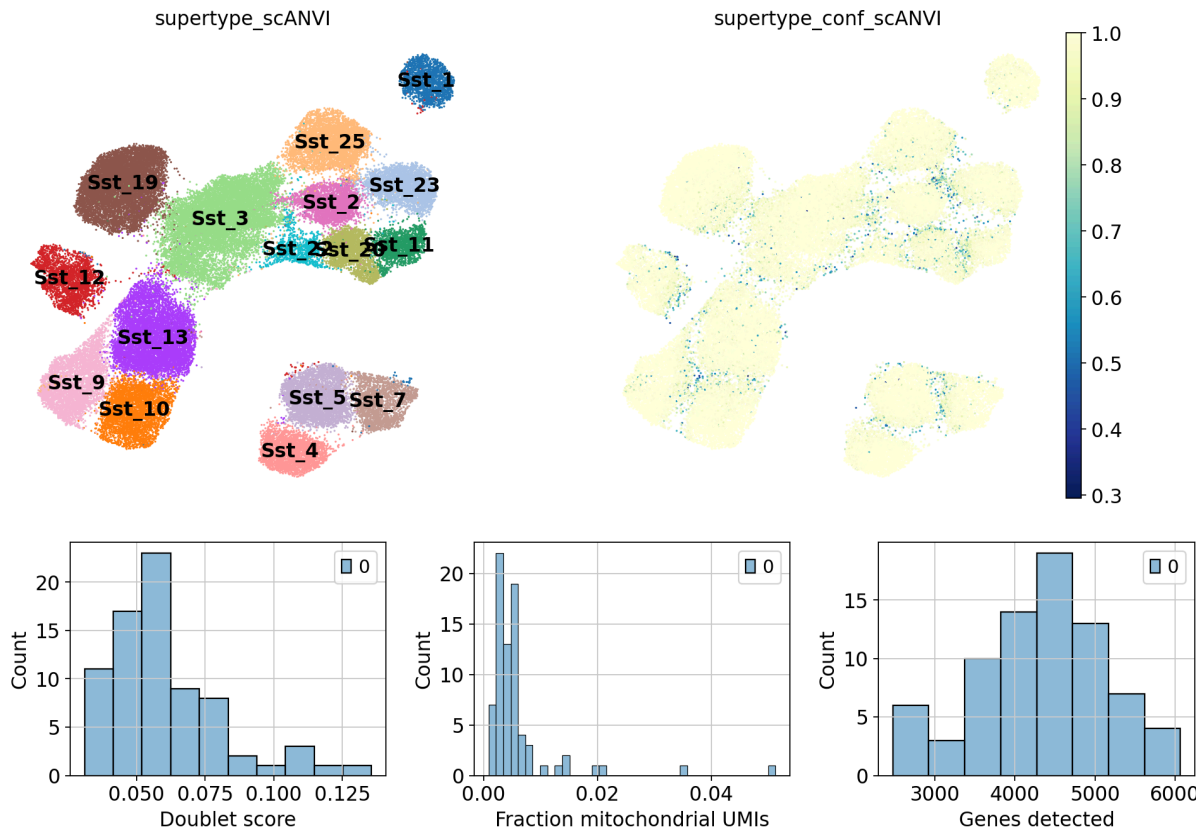


Figure 3.6. Representative subclass-level QC workflow. Recomputed UMAP on the subsetting Sst only object shows clear separation of each supertype and the confidence of supertype labeling is high with an even distribution of lower confidence calls along the periphery of clusters. Histograms of the three QC metrics to visualize subclass distribution patterns.

After each subclass is individually assessed and individual cell IDs are assigned a flag to be removed or retained in the dataset, the lists are merged with the complete dataset-level object and filtered. As is common in snRNA-seq datasets, an area of lower quality cells appeared to aggregate in the center of the UMAP space with tails connecting them to certain subclasses with a larger proportion of lower quality cells. Following the subclass-level QC, this area of spatially undefined cells with mixed cell type labels was largely removed (Figure 3.7).

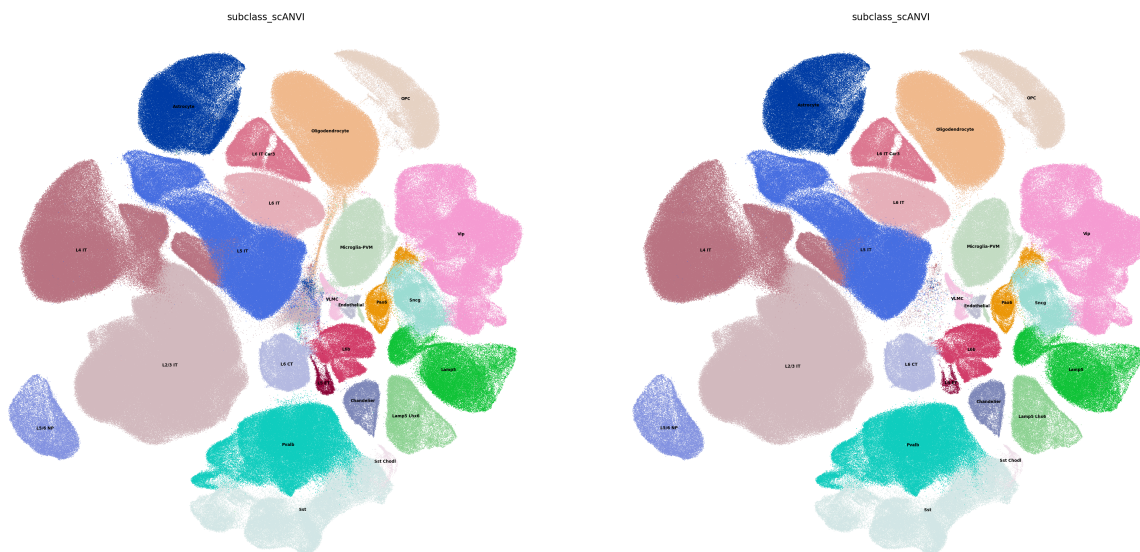


Figure 3.7. Subclass-level UMAP representation of integrated dataset pre-QC and post-QC. Pre-QC data (left) and post-QC data (right), note the absence of poorly defined cells in the center of the UMAP. The SEA-AD dataset underwent QC prior to integration using a less refined method. As such, the majority, but not all, of cells removed were from the pTau-seq dataset.

pTau-seq dataset sample size and representation of low abundance supertypes

To more specifically probe pTau-related changes, the larger integrated dataset was subset to contain only cells collected from the pTau-seq protocol. The pTau-seq dataset comprises 117,776 single somas with an approximately 40% pTau positivity rate. When assessing the nine

donors that passed QC at the donor and library preparation level, this is relatively uniform mixing with some gradation, perhaps driven by pathology-related differences in the cellular composition.

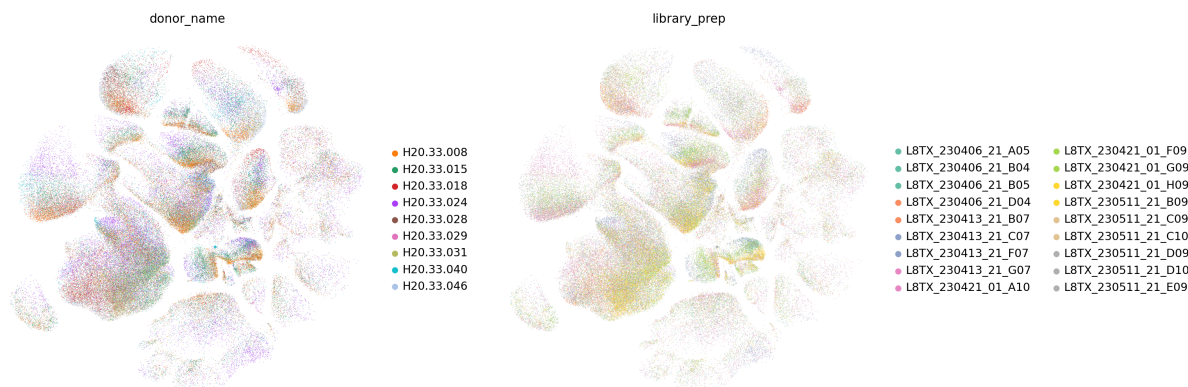


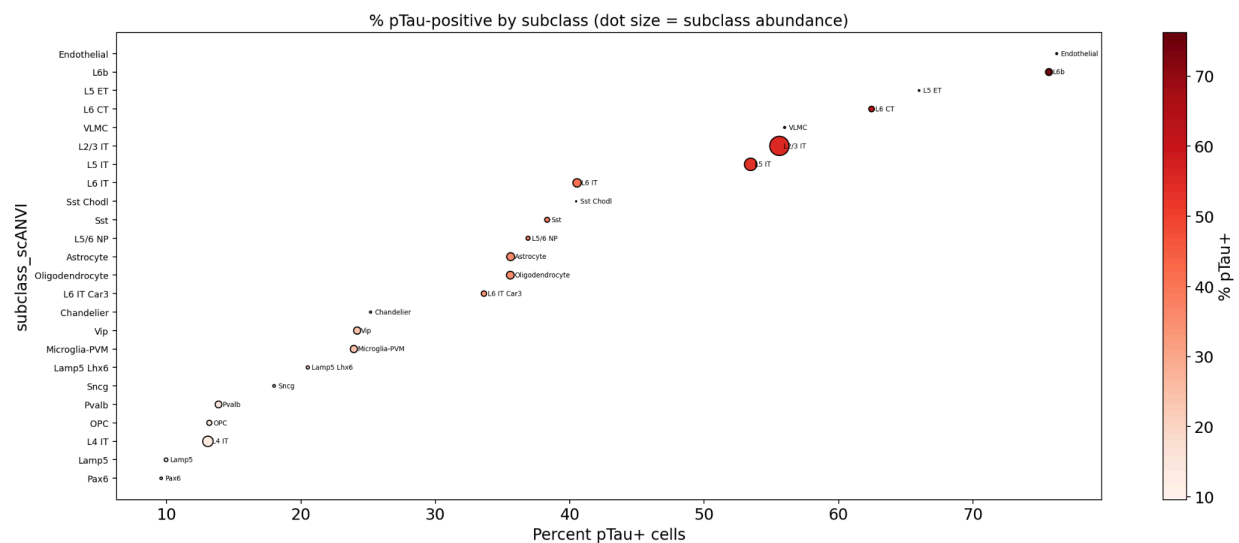
Figure 3.8. pTau-seq UMAP representation of donor and library preparation. Subset dataset including only cells from the pTau-seq method. The nine donors that passed initial quality control are shown on the left and library prep (two library preparations per donor given that each donor had a pTau-positive and pTau-negative fraction collected following FACS).

As expected, due to the small sample size, several of the less abundant supertypes are not well represented in the dataset. This is a limitation of the highly granular nature of the supertype-level labels; not all datasets (depending on sample size) will contain sufficient representation of all cell types depending on the number of cells sequenced per donor. Unfortunately, while power analysis performed before the onset of the pTau-seq study, the focus was on confidently mapping to the supertype-level, not whether each supertype would contain sufficient pTau-positive and pTau-negative fractions.

Post-hoc power analysis indicates that an increase in sample size at the donor level would improve pTau-based assessments, particularly relative abundance testing, at the supertype level. Several supertypes had fewer than 20 cells in either their pTau-positive or pTau-negative population (Endo_1, L6b_3, Lamp5_1, Lamp5_3, Lamp5_4, Lamp5_6, Micro-PVM_2_1-SEAAD, Monocyte, OPC_1, Oligo_3, Pax6_1, Pax6_2, Pax6_3, Pax6_4, Pericyte_2-SEAAD, Pvalb_6, Pvalb_7, Pvalb_10, Pvalb_12, Pvalb_13, SMC-SEAAD, Sncg_2, Sncg_3, Sncg_5, Sncg_8, Sst_1, Sst_5, Sst_7, Sst_11, Vip_5, Vip_14, Vip_21). A subset of these had fewer than 10 cells per group (Endo_1, L6b_3, Lamp5_1, Lamp5_4, Lamp5_6,

Micro-PVM_2_1-SEAAD, Monocyte, OPC_1, Oligo_3, Pax6_3, Pericyte_2-SEAAD, Pvalb_7, Pvalb_12, SMC-SEAAD, Sncg_2, Sncg_3, Sst_1, Sst_11). It is not unexpected to have low abundances of the non-neuronal cell types, because the FACS step involved neuronal enrichment (MAP2-positive population). Several relevant vulnerable supertypes from SEA-AD fall into the above category, most notably the Sst and Pvalb supertypes - results pertaining to cells in these low abundance groups are being interpreted with caution.

This limitation was particularly problematic for assessing pTau-driven supertype abundance changes, as even the t-test based calculations were not possible if this minimum criterion was not met (>3 donors with both the pTau-positive and negative fraction >0 for a given supertype). More involved approaches, such as scCODA require an even larger sample size to yield credible results, further motivating an increase in sample size for future study. Importantly, all types at the subclass-level were well represented in the dataset and sufficient for statistical analysis (Figure 3.9).



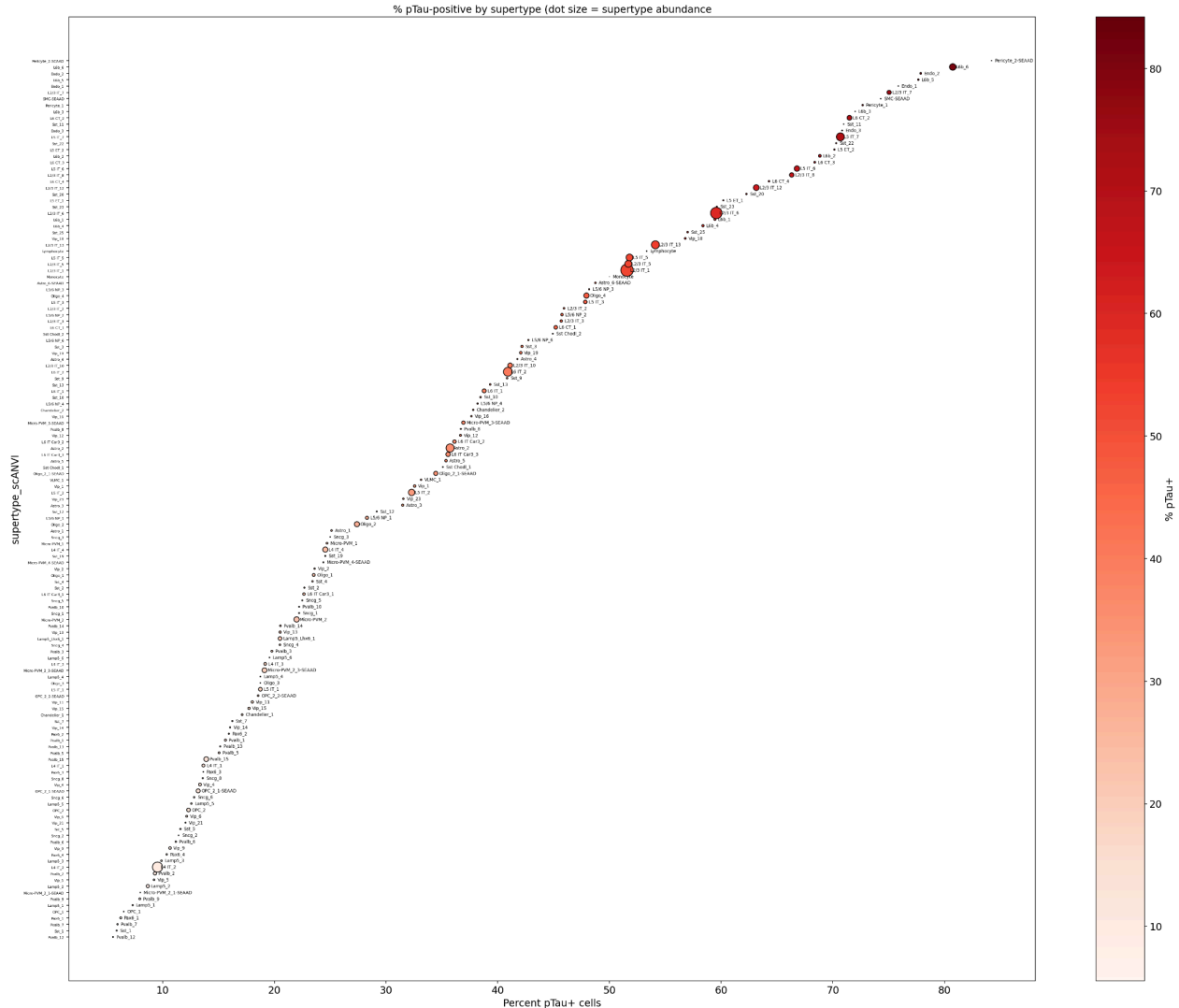


Figure 3.9. Relative cellular abundances and pTau percentage by supertype and subclass.

pTau-seq donor-level QC assessment

Across donors, all subclasses had roughly equal proportions. The donor-level subclass proportions in the pTau-negative data were very consistent across donors. However, the donor-level subclass portions were less consistent in the pTau-positive data. Donor H20.33.029 had a disproportionate increase of L6b cells only in their pTau-positive fraction which, due to the sample size, is difficult to interpret. Donor H20.33.008 had a disproportionate increase in Lamp5_Lhx6 cells and to a lesser degree, L6 IT Car3, only in their pTau-positive fraction. This highlights the increased variability in the pTau-positive population across all donors, while

noting several examples of more donor-specific effects. A larger dataset is needed to more conclusively delineate whether disease-driven processes rather than donor-specific variation is driving these proportional changes (Figure 3.10).

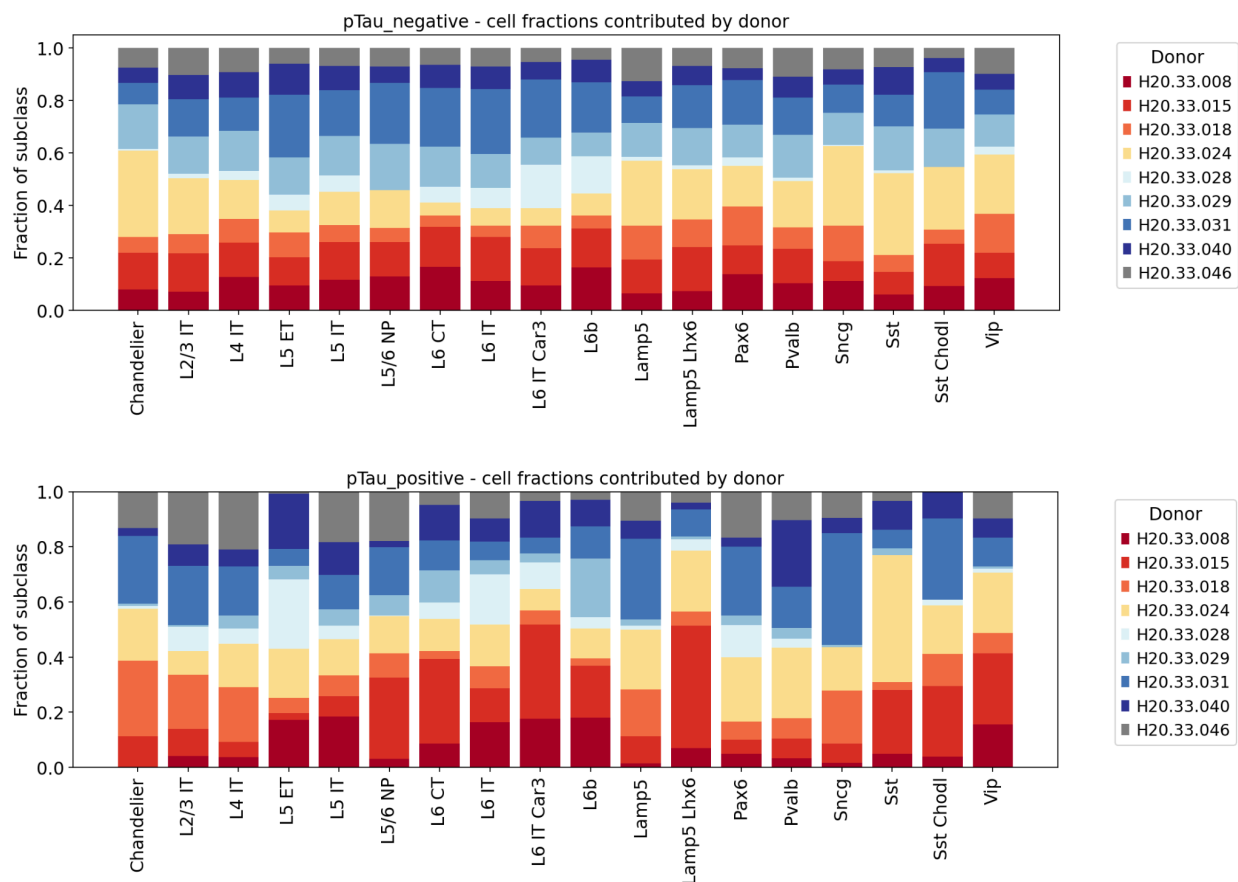
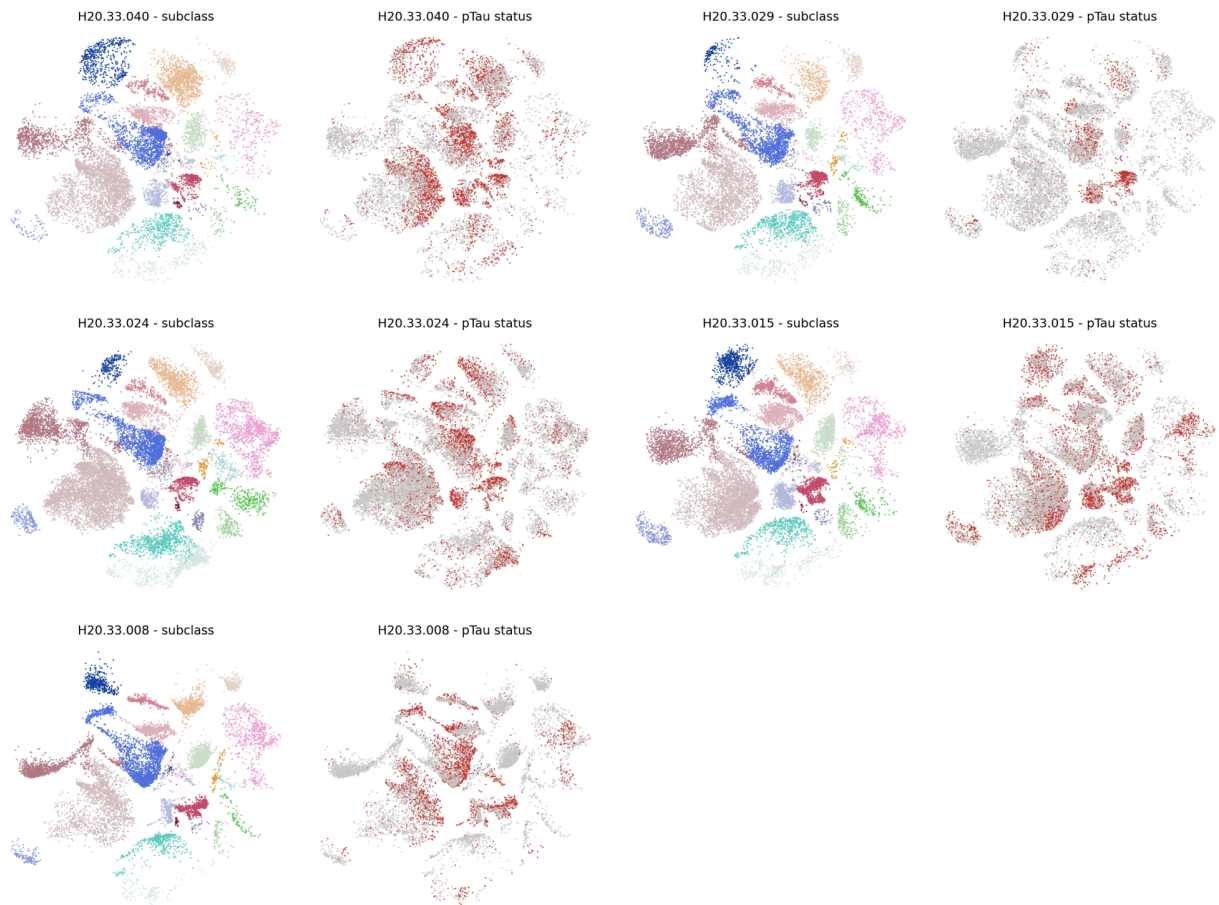


Figure 3.10. Relative contribution of each donor to subclass proportions.

Due to each donor's unique pathological state and cellular composition, the total number of cells per donor is highly variable. This is driven by expected differences in the number of pTau-positive cells captured across donors spanning the spectrum of moderate (Braak IV) to high pathology (Braak VI). This can be visualized by varying per-supertype representation filling out the clusters on the subclass-level UMAP and variable pTau-positive cells (red) superimposed on the pTau-status UMAP (Figure 3.11). As highlighted in the FACs portion of the methods section, the yield of pTau-positive cells ranged from 2,900 - 90,000 (cutoff). This inherent variability in pTau-positivity tracks closely with the donor Braak stage and therefore is not unexpected in this type of dataset. The total number of pTau-negative cells was capped at 90,000, so this value is

consistent across donors with the only minor differences in cell number emerging from minor technical variability inherent to the 10x processing workflow.



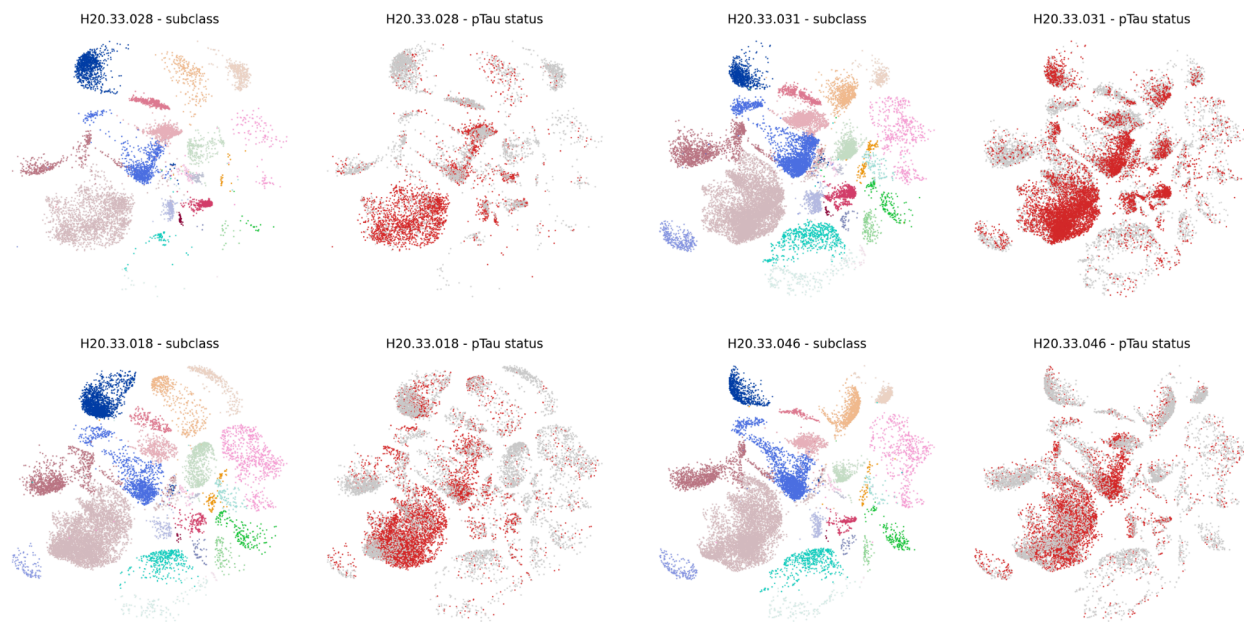


Figure 3.11. Donor-level pTau distributions in UMAP space. Pseudocolored by subclass (left) and pTau status (right).

The pTau-seq dataset is missing a subset of the highly granular supertypes defined by the SEA-AD cellular taxonomy. Given that the pTau-seq dataset is a subset of the same donors and was processed using anatomically adjacent samples of brain tissue, it is likely that this dropout is a consequence of dataset size, rather than sampling differences or true biological variability. This is further evidenced by observation that the missing supertypes are known low abundance types. As is apparent in the heatmap, certain supertypes consistently have dropout across most or all donors. There is one case of disproportionate donor-level dropout, specificity in the case of H20.33.028. This is a Braak stage V, high ADNC donor, though pathology is likely not the sole explanation for this dropout given that the SEA-AD taxonomy was built including high pathology (up to Braak VI) donors and other donors within the pTau-seq dataset with higher staging (Braak VI), exhibit substantially lower supertype-level dropout (Figure 3.12).

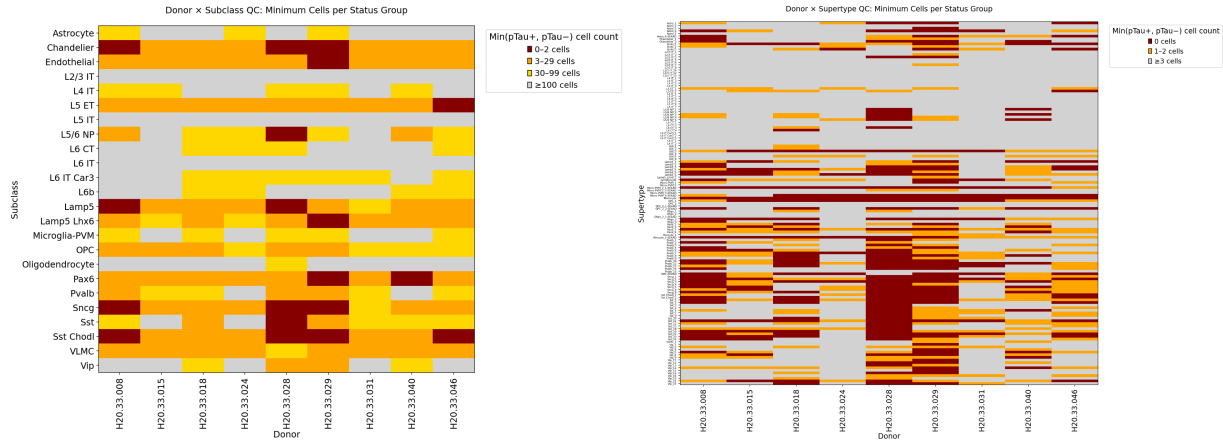


Figure 3.12. Heatmap highlighting low abundance subclasses and supertypes per donor.

Differential gene expression analysis

Nebula is an R-based package for differential expression analysis of single cell data⁵. It is particularly well suited for data with a nested structure, such as donors that each contain many individual cells, as is the case with single-cell RNA sequencing data. NEBULA uses negative binomial mixed models (NBMMs) to model gene expression counts. It does this in two key steps, (1) Model overdispersion at cell-level (to account for variability across cells) and subject-level (to account for between-donor variability), and (2) Efficient estimation by introducing an approximation method (NEBULA-LN and NEBULA-HL) that allows fast computation, even for large datasets, without losing much accuracy. A negative binomial model is used to account for overdispersion in gene counts. In Poisson models, the variance equals the mean — but this doesn't hold true for scRNA-seq data. In real scRNA-seq data, overdispersion is observed where the variance is greater than the mean. The mixed model aspect of the model is important to handle random effects such as those that arise from donors. The model selected for the soma seq data was with 'pTau_status' (pTau-positive or pTau-negative) as the only covariate (test) and 'donor_name' (individual donors) as the random effect. The code to execute nebula is shown below.

```
for supertype in adata.obs[cluster_key].unique():
    print(supertype)
    adata_sub = adata[adata.obs[cluster_key] == supertype]
    try:
```

```

print(f"\n--- {supertype} ---")
print("pTau_status distribution:\n",
      adata_sub.obs["pTau_status"].value_counts(dropna=False))
print("Donor count:", adata_sub.obs["donor_name"].nunique())
print("Cell count:", adata_sub.shape[0])

results = nebula_deg_analysis(adata_sub,
                              ['pTau_status'],
                              'donor_name',
                              counts_layer = "counts_layer"
                              )

print(results)
results.to_csv(f"{output_dir}/
               {supertype.replace('/', '_')}_supertype_pTau_nebula_June10.csv")
except Exception as e:
    print(f"{supertype} Failed")
    print(str(e))
    continue

```

Gene set enrichment analysis

GSEA is a method to determine whether a predefined set of genes (e.g., genes in a biological pathway or predefined gene ontology (GO) category) is enriched at the top or bottom of a ranked list of genes — for example, genes ranked by differential expression between two conditions (e.g., pTau⁺ and pTau⁻ cells). First, the list of differentially expressed genes from differential expression analysis is ranked based on a defined metric (e.g, Log fold change - ‘logFC_pTau_statuspTau_positive’). Next, each gene within a GO category (e.g., GO Biological Process 2025) is assessed to calculate an enrichment score (ES) by walking down the ranked list of genes and increasing a running-sum statistic when you see a gene from the gene set or decreasing it when you see a gene not in the set. Lastly, the data is permuted and an ES score is calculated to generate a null distribution then compares that against the ES true scores from the dataset to obtain p-values. Finally, multiple testing correction, such as false discovery rate (FDR) is applied.

Results

Integration and mapping to a standardized AD taxonomy highlights shared features

To improve standardization and assess the results in a broader context, the pTau-seq dataset was integrated with the larger SEA-AD MTG snRNA-seq dataset. Our custom implementation of the Single Cell Annotation using Variational Inference (scANVI) referred to as iterative_scANVI was used for label transfer to provide highly accurate and standardized cell-type levels down to supertype resolution, as previously published. The integrated dataset includes 1,361,219 million cells (post-QC), with representation of expected classes, subclasses, and supertypes with high mapping confidence (Figure 3.13). Of note, the nine donors included in pTau-seq are a subset of the 84 donors included in the SEA-AD study, which can be leveraged to increase the power of downstream analyses and enable the development of cross-dataset tools.

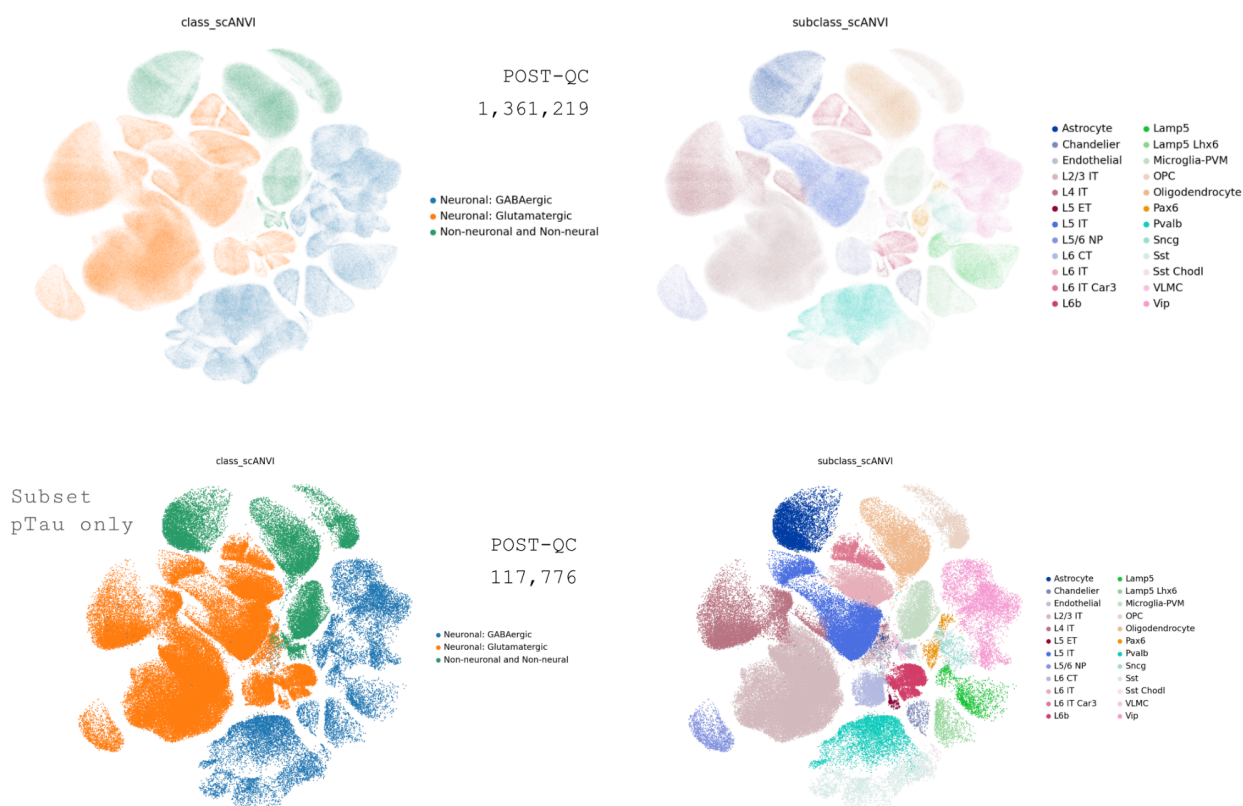


Figure 3.13. Integrated SEA-AD and pTau-seq dataset pseudocolored by class and subclass. The top panel represents the integrated dataset of approximately 1.3M cells at the class level (left) and subclass level (right). The bottom panel shows the same UMAP space of the pTau-seq dataset only with class level labels (left) and subclass level labels (right).

Cell type composition and pTau enrichment in the human MTG

To dive deeper into the composition of cells harboring pTau pathology, downstream analyses were performed on the pTau-seq dataset of 117,776 cells, where the pathological (pTau) state of each single cell is known. Due to the sampling strategy and variable pTau levels across the donors (Braak IV - Braak VI), the dataset contains fewer pTau-positive cells than pTau-negative cells (Figure 3.14). Because neurons (MAP2+) and pTau-positive (AT8+) cells were enriched during FACS, the dataset proportions are also skewed towards neurons (82.8% - 97,501 / 117,776) and particularly, excitatory neurons (69.1% - 81,344 / 117,776 (69.1%)), which are more abundant in the cortex compared to inhibitory populations and known to preferentially accumulate pTau.

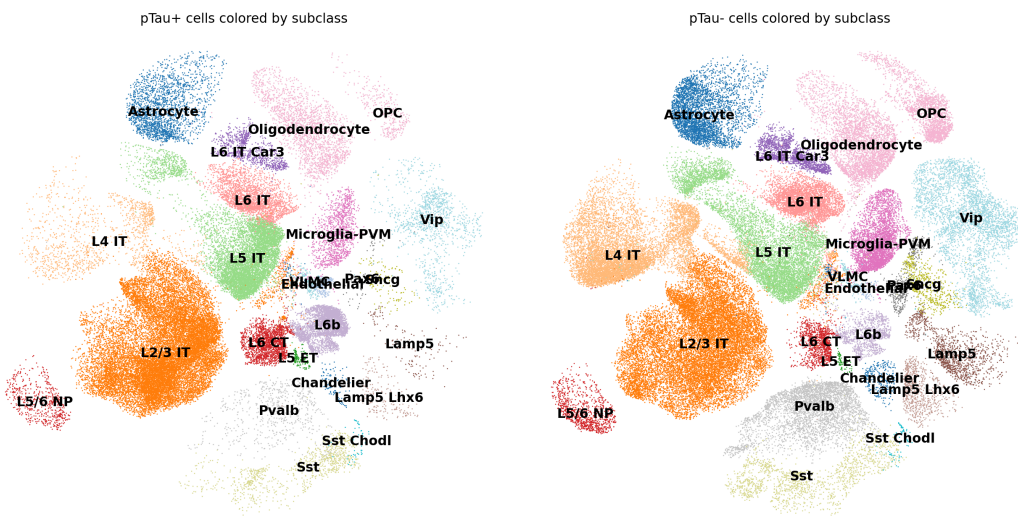


Figure 3.14. pTau-seq dataset split by pTau status pseudocolored by subclass.

By applying standardized and highly granular cell type labels we were able to disentangle precisely which supertypes were preferentially vulnerable to pTau accumulation. Across the dataset, 41.89% (49,342 / 117,776) cells were pTau-positive. Most notably, the L2/3 IT subclass is highly enriched in pTau. In terms of total cell numbers, L2/3 IT has the most

pTau-positive cells followed by L4 IT and L5 IT. Importantly, when considering relative proportions, the enrichment for L2/3 IT types becomes more pronounced, especially in the pTau-positive population. Interestingly, the L4 IT enrichment drops off, most notably in the pTau-positive population. This is in line with expected increased pathological pTau burden in L2/3 and L5, with sparing of L4. Across the class of excitatory neurons, the range of pTau-positivity within a subclass was 13.1% – 75.7%, while the range in the inhibitory neuron class was 9.6% – 40.5%. As expected, cortical L2/3 had the highest percentage of pTau-positive cells (~55%) in the dataset at the subclass level. In line with the known pTau cortical banding patterns, the second highest percentage of pTau-positive cells (~53%) at the subclass level was L5. There is clear sparing of L4, with only ~13% pTau-positive cells (Figure 3.15).

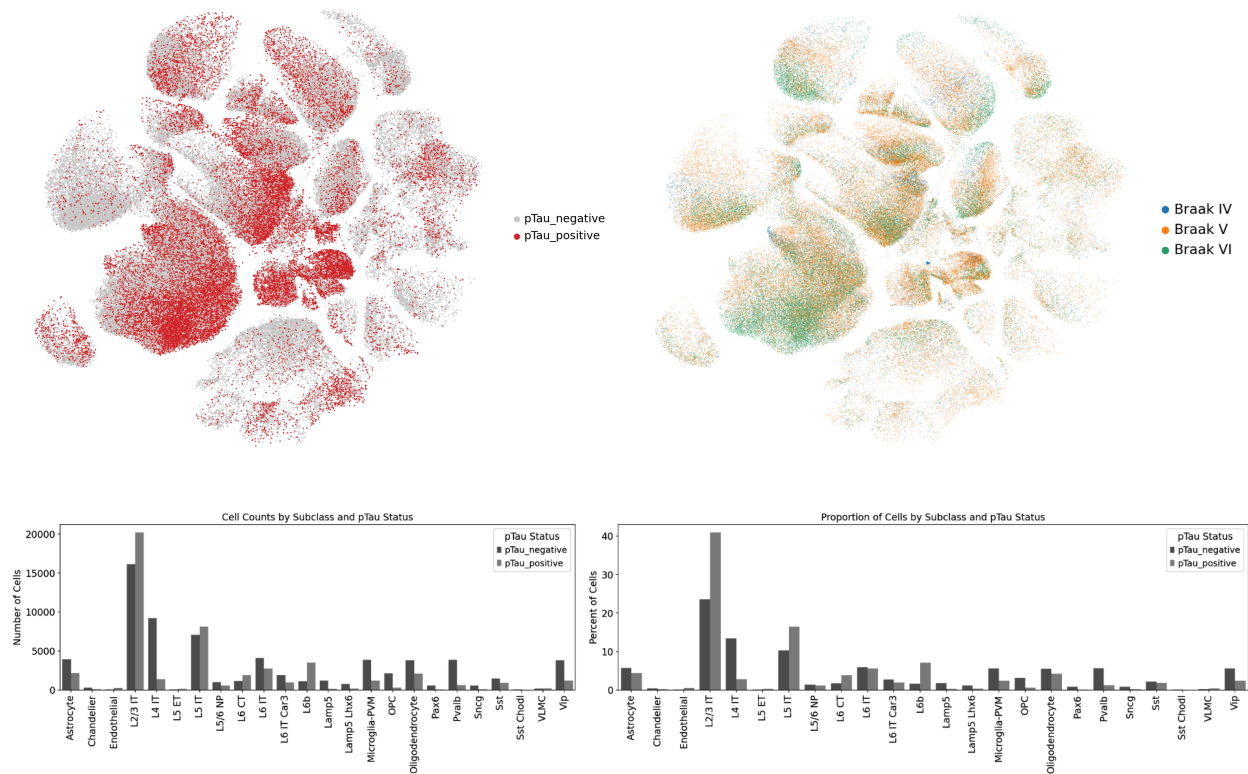


Figure 3.15. pTau-positive fractions track with Braak Stage. Top panel UMAP representations of pTau-seq dataset showing pTau-positive cells (red) and pTau-negative cells (grey) and Braak Stages (Braak IV – blue, Braak V – orange, and Braak VI – green). Bottom panel barplots of number (left) and percent (right) of pTau-positive and pTau-negative cells by subclass.

Methodological comparison to the Otero-Garcia publication

There are several methodological differences between the MTG pTau-seq data discussed in this chapter and the data originally published in Otero-Garcia et al., 2022. First, the current dataset profiles an earlier affected brain region in AD progression; the accumulation of pTau in the MTG marks the progression to Braak Stage IV. The Otero-Garcia dataset profiles the PFC which is one of the latest affected regions in AD and marks the progression to Braak Stage VI. Next, the current dataset includes tissue from donors spanning moderate to high spectrum of AD neuropathologic change (Braak IV - Braak VI) whereas the Otero-Garcia only included tissue from the highest pathology donors (Braak VI). The dissociated tissue samples were also microdissected differently, resulting in potentially higher cell yields in the current dataset. Finally, the dissociations in the current study were performed more gently, with fewer (eight) grinder strokes versus fifteen in the Otero-Garcia method. This approach resulted in very consistent yields of pure somas with no visibly detectable nuclei.

Unsurprisingly, the differences in experimental design result in a different number of total somas classified as pTau-positive in each dataset. The percentage of pTau-positive cells across the current pTau-seq dataset is 41.89% (49,342 / 117,776 total), whereas the Otero-Garcia describes the percentage of NFTs as $6.3\% \pm 1.15\%$. In line with this, the percent of excitatory neurons that are pTau-positive in the current study ranges from 13 - 75%, whereas in Otero-Garcia it was 1 - 11.7%. Across the interneuron subclasses, the pTau-positive percentages for a given subclass ranged from 9.6 - 40.5% in the current study and 0.5 - 6.8% in Otero-Garcia. In line with published quality control metrics in Otero-Garcia, the median UMIs per soma were higher in pTau-positive somas relative to pTau-negative somas (19.5% in pTau-seq and 18% in Otero-Garcia).

In terms of pTau-positive subclass (or equivalent) level proportions across the dataset, Otero-Garcia observed a similar enrichment in L2/3 IT cells in both overall number and percent pTau positivity. This was followed by an enrichment in L5 IT, also observed in the current dataset. These layer-specific increases have long been known in the field and represent the characteristic AD pTau banding pattern discussed in Chapter 2. Across inhibitory interneuron

subclasses, the overall percentages for pTau positivity are lower, but present with enrichment in Sst and Pvalb subclasses. Distinguishing between the fraction of pTau-positive subclass relative to the entire pTau-positive population (yellow bars) and the fraction of pTau-positive cells within a subclass (blue bars), highlights both the cell types which are abundant and enriched for pTau and those that are lowly abundant and enriched for pTau (Figure 3.16).

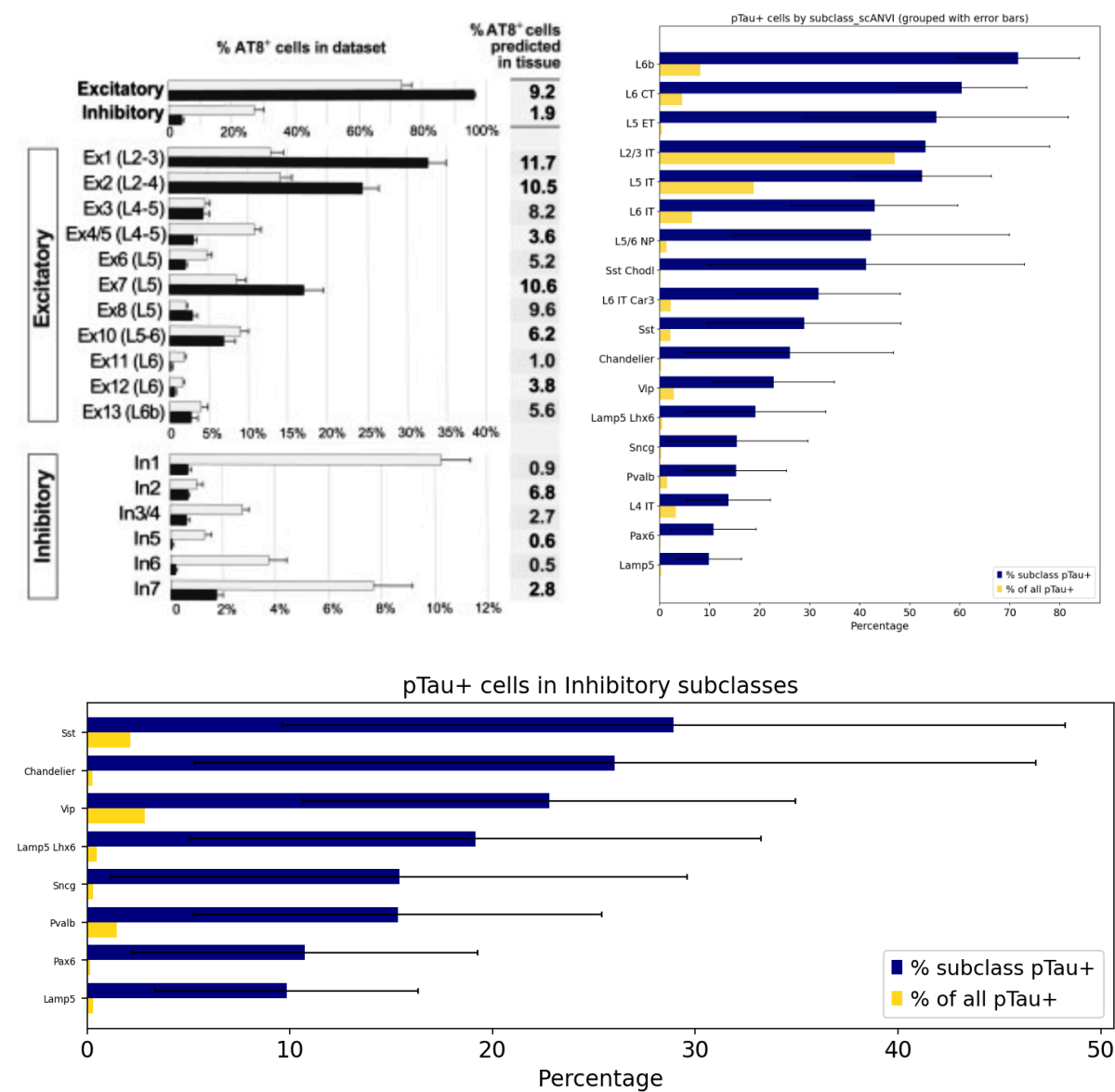


Figure 3.16. Comparison of pTau positivity between Otero-Garcia et al., 2022 and pTau-seq. Grayscale figure on upper left extracted from Otero-Garcia et al., 2022.

As was observed in Otero-Garcia et al., a gradient-like expression pattern of several putative vulnerability-related genes that have previously been associated with AD and

senescence (SERPINE2/LAMP5 and COL5A2), were observed in the L2/3 IT cluster (CUX2+ Ex1 and Ex2 clusters in Otero-Garcia). With the increased cellular resolution of the current dataset, it was possible to discern that the areas of the subclass-level UMAP that have increased expression of these genes represent several supertypes found to be selectively vulnerable to cell death in the SEA-AD dataset. This area of the UMAP is also enriched for pTau-positive cells, suggesting a relationship between the expression of these genes, vulnerability to accumulation of pTau, and associated vulnerability to cell death (Figure 3.17).

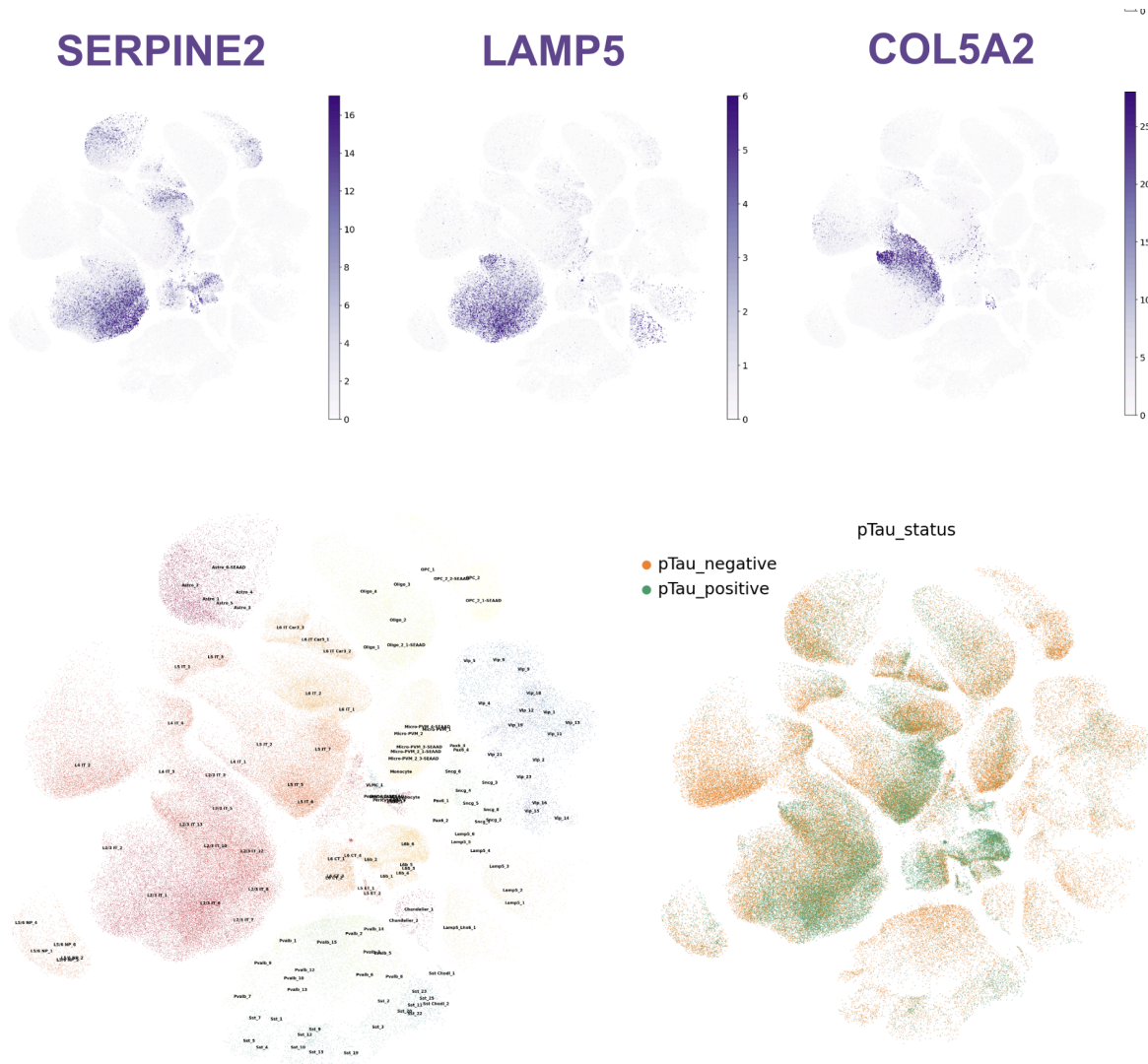


Figure 3.17. Known AD and senescence-related genes associated with pTau accumulation.

Selective vulnerability to pTau accumulation is widespread but unevenly distributed across excitatory cortical supertypes

The pTau-seq dataset adds an important layer of resolution to the SEA-AD snRNA-seq results - single-cell pTau status - enabling a direct comparison between cell types vulnerable to pTau accumulation and those vulnerable to cell loss in the SEA-AD dataset. While it might be expected that pTau accumulation would closely track with cell loss, the relationship appears to be more nuanced. While many supertypes with high pTau fractions generally exhibit reduced relative abundances in SEA-AD, several supertypes stood out as clear exceptions to this pattern, suggesting that pTau accumulation may not always be pathogenic. Conversely, some supertypes that were vulnerable to relative abundance changes in SEA-AD, were not vulnerable to pTau accumulation in pTau-seq, suggesting that other pathological factors are likely contributing to vulnerability in these cell types.

Two upper-layer excitatory neuron supertypes, L2/3 IT_1 and L2/3 IT_6, are the most abundant pTau-bearing cells in the dataset. Notably, these L2/3 IT types are also selectively vulnerable to cell death based on the results of the SEA-AD study. Two deeper-layer excitatory neuron supertypes, L5 IT_7 and L6b_6 were the second most abundant pTau-bearing cells in the dataset. While excitatory neurons within these layers are associated with the accumulation of pTau based on histological studies and may be selectively vulnerable to death, this was not reflected as a relative abundance change in the SEA-AD study. These observations are consistent with neuropathological understanding that pTau preferentially accumulates in large pyramidal neurons within specific cortical layers (e.g., L2/3 and L5/6), but adds finer grained (supertype-level) cellular resolution.

Extending upon this, given that traditional neuropathological assessments have historically lacked the cellular resolution to distinguish beyond classifying cells as neuronal (typically based on NeuN-immunoreactivity), this opens up the possibility that rarer populations of neuronal cells, such as inhibitory interneurons, may be overlooked due to their laminar location and sparsity, leading to the incorrect assumption that all affected neurons in high pathology areas are excitatory neurons. It's possible that a small, but functionally relevant,

fraction of cortical interneurons are both preferentially accumulating pTau and are selectively vulnerable to cell loss in AD, but are being overlooked or misclassified as excitatory neurons.

Inhibitory neuron supertypes exhibit unexpectedly high pTau fractions

Some of the most vulnerable supertypes early in disease, as defined by the SEA-AD study, are those with the highest fraction of pTau-positive cells in the pTau-seq dataset. Of particular interest were four vulnerable Somatostatin (Sst) supertypes (Sst_20, Sst_22, Sst_23, and Sst_25) which exhibited an unexpectedly high fraction of cells that were pTau-positive. These supertypes are not highly abundant in the cortex and as an extension, have relatively few overall counts of pTau-positive cells relative to other neuronal cells in the dataset. This sparsity may contribute to this cell type being easily overlooked in the context of AD pathophysiology. This is especially relevant to analyzing discerning cell type abundance changes in AD, as lower abundance populations require substantially larger datasets to be sufficiently powered to discern real changes using statistical tests. Further, in the context of selective vulnerability to pTau pathology, their sparsity may also play a role in them being overlooked using histological approaches as the lack of cellular granularity and inherently smaller sample (Figure 3.18).

The observation that a subset of Sst supertypes harbor high fractions of pTau-positive cells was unexpected given that Sst cells were historically not considered vulnerable in the context of AD. However, over the past year significant progress unfolded in this area, as several AD omics studies simultaneously published that select Sst types appear to be preferentially vulnerable to cell death early in AD. However, the mechanisms underlying this vulnerability to cell death remain completely unknown. pTau-accumulation is one possible explanation, though published evidence that Sst cells preferentially accumulate pTau in AD, especially in neocortical areas, is extremely limited. Intriguingly, while not emphasized, the Otero-Garcia study noted two inhibitory clusters with increased susceptibility to pTau pathology, with one type also associated with vulnerability to cell death, independently implicating inhibitory neurons in the pathological progression of AD. The pTau-seq dataset adds validation and crucial cell-type level granularity to this novel AD-relevant finding.

current study does not attempt to distinguish NFT-bearing cells, and as such it is not surprising that even higher percentages of cells were found to be pTau-positive in the pTau-seq dataset, given it captures a wider spectrum of pTau states ranging from initial phosphorylation, to pretangles, to mature NFTs. Taken together, in order to assess whether individual cell types had significantly more pTau-positive cells than expected in a non-diseased state, the supertype with the lowest relative fraction of pTau-positive cells across the dataset (L4 IT_2) was selected as the baseline for supertype-level statistical tests.

For an initial assessment of which cell types preferentially accumulate pTau, a paired t-test with FDR correction ($p < 0.05$) was performed at the supertype-level. These results exhibited close alignment with vulnerable supertypes in SEA-AD, but highlighted several interesting examples of divergence between vulnerability to pTau accumulation and vulnerability to cell death (Figure 3.19).

Direct alignment was observed in a subset of supertypes that were vulnerable both to cell loss in SEA-AD and pTau accumulation. This list consists of primarily known vulnerable cell types in AD (Vip_15, Vip_2, Vip_11, Sst_22, Sst_25, L2/3 IT_1, L2/3 IT_6, L2/3 IT_7, L2/3 IT_13, L2/3 IT_5, L2/3 IT_10, L2/3 IT_8, L2/3 IT_12, L6 IT_2, L5 IT_3, L2/3 IT_3, L5 IT_5, L5 IT_7, L6 CT_2). With the exception of Sst, which has only recently been implicated in AD, these types represent well-established cell types in the AD literature associated with selective vulnerability to AD-associated pTau pathology and cell loss - large excitatory pyramidal neurons in L2 and L3, and L5, with sparing of L1 and L4. Deeper L6 neurons exhibit some variability in their vulnerability in the AD literature; the improved resolution of this dataset suggests that this may be explained by contrasting dynamics at the supertype-level.

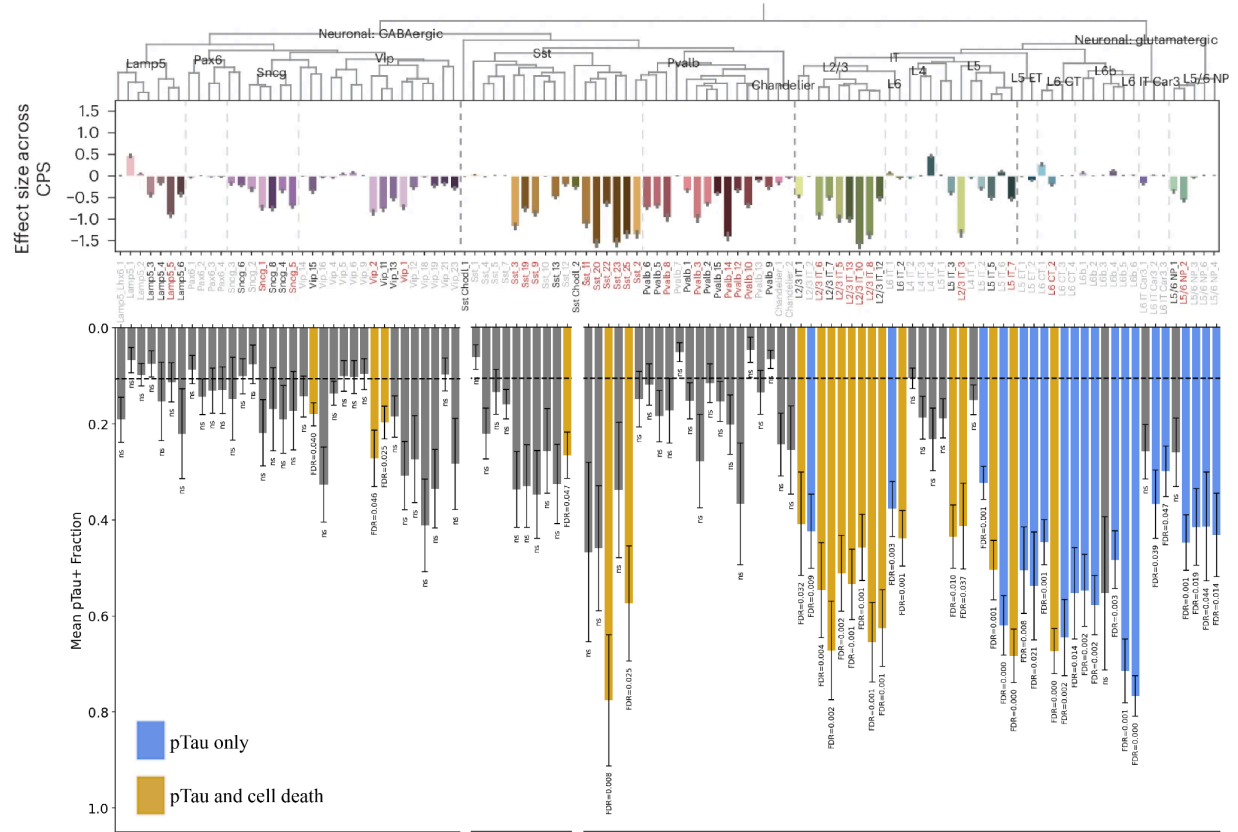


Figure 3.19. Vulnerable supertypes in SEA-AD track with significant pTau enrichment in pTau-seq. Top portion adapted from SEA-AD study indicating effect sizes of relative abundance changes at the supertype-level across CPS. Red text indicates a significantly changed supertype in both the MTG and BA9 (frontal cortex), while black indicates significance in MTG only. Bottom barplot indicates the FDR-corrected paired t-test results assessing pTau accumulation at the supertype-level. Colored bars highlight significant FDR-corrected significant p-values, significant pTau accumulation only (blue) or significant pTau accumulation and cell death (yellow). Non-significant results are indicated by ns. The dashed line indicates the biological baseline of a non-vulnerable supertype (L4 IT_2).

Several supertypes that were vulnerable to cell loss in SEA-AD were not vulnerable to pTau accumulation. This list consists primarily of Sst and Pvalb interneurons (Lamp5_3, Lamp5_4, Lamp5_5, Lamp5_6, Sncg_6, Sncg_1, Sncg_8, Sncg_4, Sncg_5, Vip_13, Vip_1, Sst_3, Sst_19, Sst_9, Sst_13, Sst_11, Sst_20, Sst_23, Sst_2, Pvalb_6, Pvalb_5, Pvalb_8, Pvalb_1, Pvalb_3, Pvalb_2, Pvalb_15, Pvalb_14, Pvalb_12, Pvalb_10, and Pvalb_9, L5/6 NP_1). **Conversely, several supertypes that were not vulnerable in SEA-AD, were found to be vulnerable to pTau accumulation** This list consists almost exclusively of deeper-layer excitatory neuron types (Sst_12, L2/3 IT_2, L6 IT_1, L5 IT_2, L5 IT_6, L5 ET_1, L5 ET_2, L6 CT_1, L6 CT_3, L6 CT_4, L6b_1, L6b_2, L6b_4, L6b_5, L6b_6, L6 IT Car3_2, L6 IT Car3_3,

L5/6 NP_2, L5/6 NP_3, L5/6 NP_6, L5/6 NP_4), somewhat unsurprisingly given their general association with pTau accumulation based on traditional histological approaches. This finding makes the apparent resilience of L6b to cell death an intriguing area for further investigation.

The remaining supertypes were not vulnerable to either cell loss in SEA-AD or pTau accumulation (L4 IT_1, L4 IT_2, L4 IT_3, L4 IT_4, L5 IT_1, L6b_3, L6 IT Car3_1, Lamp5_1, Lamp5_2, Lamp5_Lhx6_1, Pax6_1, Pax6_2, Pax6_3, Pax6_4, Pvalb_7, Pvalb_13, Sncg_2, Sncg_3, Sst_1, Sst_4, Sst_5, Sst_7, Sst_10, Chandelier_1, Chandelier_2, Vip_12, Vip_14, Vip_16, Vip_18, Vip_19, Vip_21, Vip_23, Vip_4, Vip_5, Vip_6, Vip_9). The unaffected populations include several well-established spared cell types in AD, particularly L4 excitatory neurons, as well as inhibitory interneuron types that have not been previously associated with AD pathophysiology.

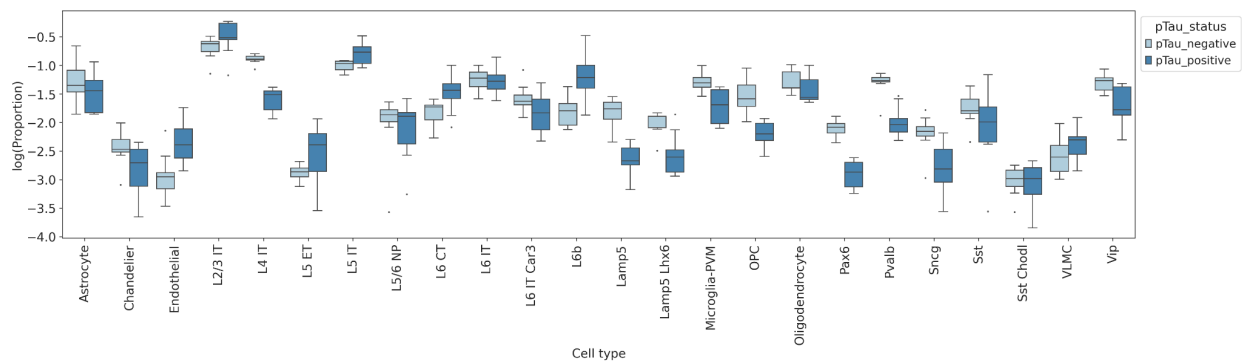
T-tests are not a preferred statistical approach for snRNA-seq data given that these data do not comply with the necessary assumptions (e.g., normal distribution, equal variance, and independent samples). An alternative, more nuanced statistical approach, tailored to the complexities of snRNA-seq data, was used to assess whether select supertypes had significantly increased pTau-positive populations. Single cell compositional data analysis (scCODA) was the approach for relative abundance testing used in SEA-AD, but was modified to accommodate the expectations of the pTau-seq dataset by setting a single reference (e.g., L4, rather than average across all types as was performed in SEA-AD).

scCODA is a Bayesian statistical framework designed to detect differential changes in cell type composition across conditions in scRNA-seq data⁶. An important distinguishing feature relative to simpler statistical tests is that the approach models the inherent compositional nature of single cell data (e.g., cell type proportions sum to one per donor). It uses a multinomial logistic-normal model to account for the relative proportions and employs a spike-and-slab prior to induce sparsity (as a way to identify which cell types exhibit real changes). Of note, approaches such as scCODA typically require an even larger sample size per group, a standard not met by the pTau-seq dataset. As such, the results from scCODA in this context are

interpreted as trends. Despite this caveat, the scCODA results of supertypes with increased pTau fractions (positive effect sizes) are mostly overlapping with the FDR-corrected t-test results.

At the subclass level, four broad types exhibit increased pTau-positive fractions, L2/3 IT, L5 IT, L6 CT, and L6b. This aligns with the t-test results, demonstrating the robustness of the observed pattern of increased pTau accumulation in cortical layers traditionally associated with AD-associated pTau pathology (L2/3, L5, and L6). While L5 IT types did have an inclusion probability above 0.8 (threshold), the effect size was smaller and the fraction of credible values did not meet threshold.

Interestingly, two subclasses exhibited decreased pTau-positive populations (Lamp5 and Pvalb) relative to the L4 IT subclass baseline (Figure 3.20). Several Pvalb supertypes were implicated in the SEA-AD study as having decreased relative abundance and are now thought to be vulnerable in AD - the finding that this type appears to be relatively spared from pTau accumulation makes for an interesting point of comparison, especially as it relates to Sst interneurons which appear to be vulnerable to both cell loss and pTau accumulation. The relationship between vulnerability to cell death and pTau accumulation appear to be more complex for the Lamp5 population, and required supertype-level resolution to decipher meaningful dynamics, which will be discussed in the next section.



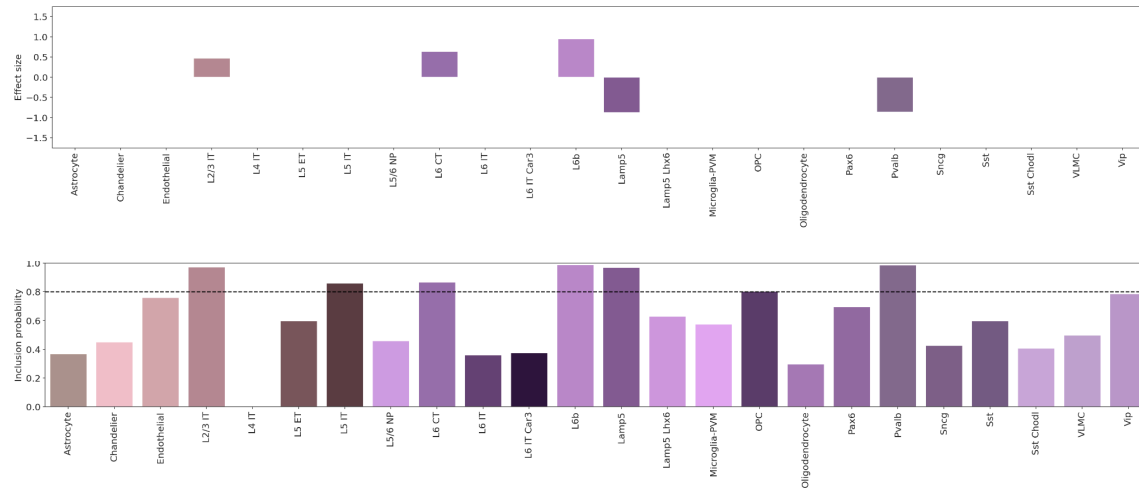


Figure 3.20. Relative abundance changes in pTau-seq at the subclass-level using scCODA. subclass level assessment of differential abundance.

The increased resolution afforded by running scCODA at the supertype-level highlighted important subtle changes that would be otherwise overlooked by subclass-level comparisons. First, within the L2/3 IT subclass, three supertypes had substantially increased pTau-positive fractions with an inclusion probability above threshold - L2/3 IT_7, L2/3 IT_8, and L2/3 IT_12. Interestingly, one supertype (L2/3 IT_1) had a reduced pTau-positive fraction, though this supertype exhibited high donor-to-donor variability, and would require broader donor representation to assess confidently. Three L4 IT types had reduced pTau-positive fractions, L4 IT_1, L4 IT_2, and L4 IT_3, supporting the relative sparing of excitatory neurons to pTau pathology in this cortical layer and supporting the approach of using L4 as a representative a biological baseline of non-pathological pTau accumulation levels (Figure 3.21).

Interestingly, based on the scCODA results within L5, an interesting pattern of directionality emerged. Two supertypes had increased pTau-positive fractions (L5 IT_1 and L5 IT_2) while two had decreased pTau-positive fractions (L5 IT_6 and L5 IT_7), highlighting a potential reason why this cell type did not reach threshold for either direction of change at the aggregate subclass level. Based on the t-test results, L5 IT_2, L5 IT_6, and L5 IT_7 all had increased reduced pTau-positive fractions with L5 IT_1 as non-significant.

All L6b supertypes had increased pTau fractions in scCODA, but only L6b_6 had an inclusion probability of greater than 0.8. However, this L6b_6 increase appears to be driven primarily by one donor, as discussed in the methods. Several Pvalb and Vip supertypes had decreased pTau-fractions relative to baseline, and though the Vip trends are more subtle.

The Lamp5 inhibitory interneuron population had more mixed results in the SEA-AD study related to cellular vulnerability with several subtypes exhibiting vulnerability to cell death (Lamp5_3, Lamp5_5, Lamp5_4, and Lamp5_6) and three supertypes not vulnerable to cell death (Lamp5_Lhx6_1, Lamp5_1, Lamp5_2) - all of which were not vulnerable to pTau accumulation (and as noted, had significantly fewer pTau positive cells) in the pTau-seq dataset.

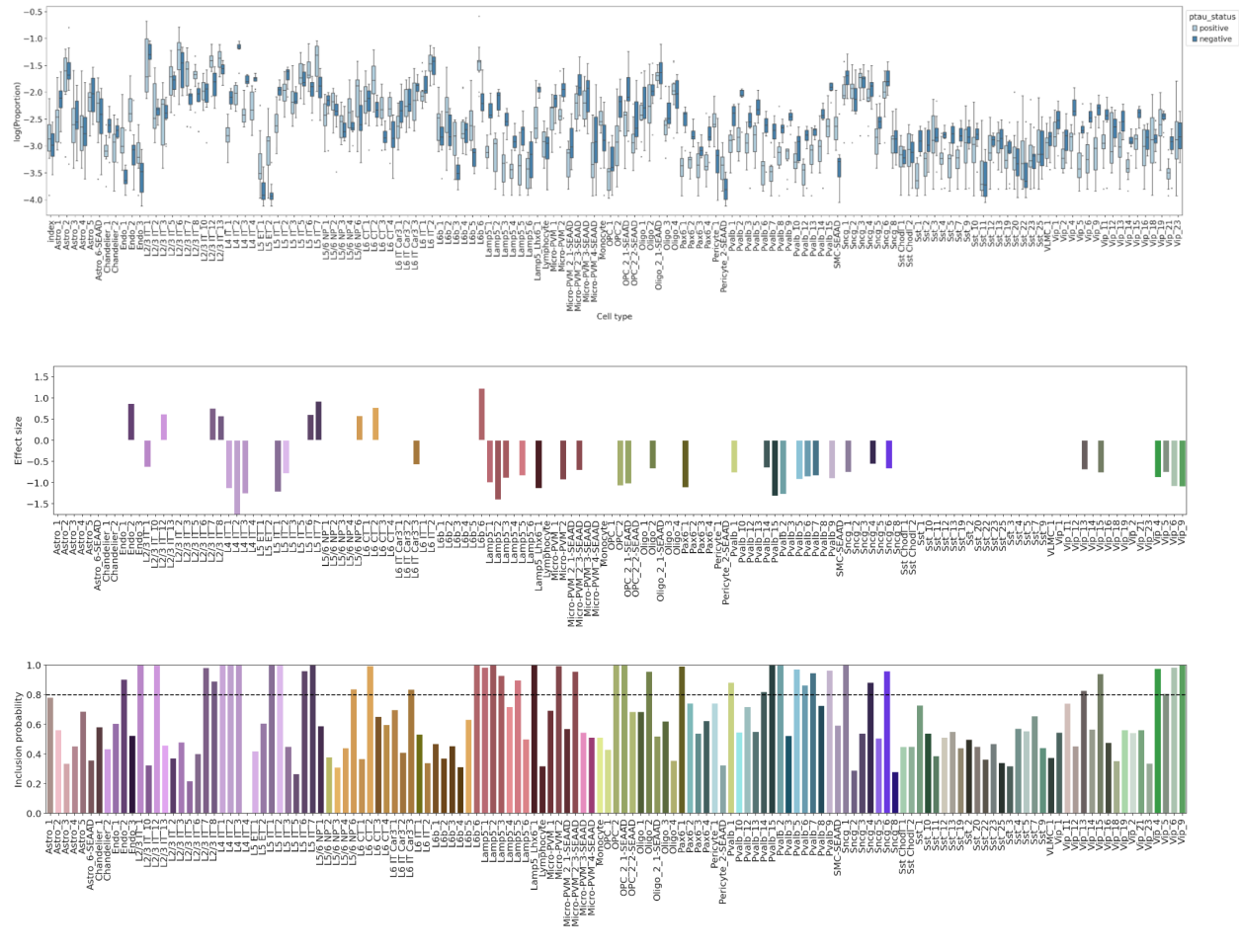


Figure 3.21. Relative abundance changes in pTau-seq at the supertype-level using scCODA. Supertype level assessment of differential abundance. The top panel shows the proportion of pTau-positive and pTau-negative cells within each supertype. The middle panel shows the effect size of the relative abundance change. The bottom panel highlights which supertypes cross the threshold of 0.8.

Shared transcriptomic signatures link pTau accumulation to cell loss

Differential expression analysis was performed using a negative binomial mixed model to assess the pTau-driven changes in gene expression within a subclass or supertype. The supertypes with the highest number of differentially expressed genes (both upregulated and downregulated) not only had high overall levels of pTau-positive cells, but also represented those with highest fraction of pTau-positive cells in the pTau-seq dataset - specifically, L2/3 IT_1, L3/3 IT_6, L2/3 IT_13, L2/3 IT_5, and L5 IT_7 (Figure 3.22). Notably, these supertypes show complete overlap with those that exhibited selective vulnerability (decreased relative abundance with increasing CPS) in the SEA-AD dataset. Together, the transcriptomic differences between pTau-positive and pTau-negative cells within a supertype suggest that pTau accumulation may underlie the changes in transcriptomic state that precede cell loss in AD.

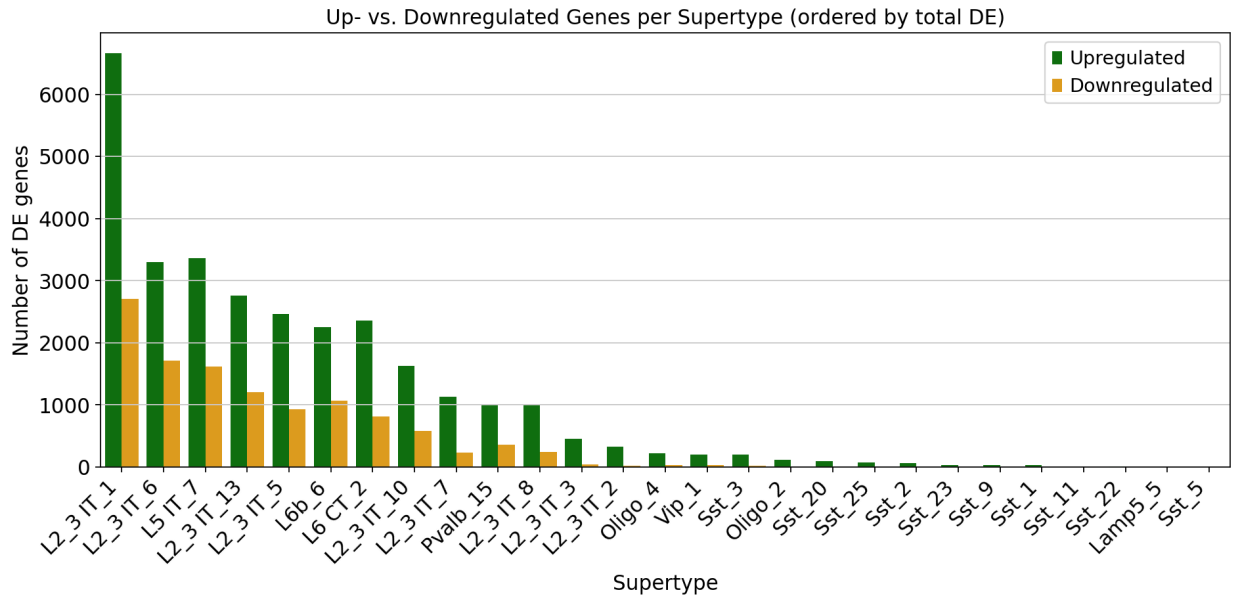


Figure 3.22. Supertypes with top upregulated and downregulated genes across pTau status.

Consistent cross dataset signatures of tau-related dysfunction

To further investigate the supertypes that are preferentially vulnerable to both pTau accumulation and cell death in AD, the top differentially expressed upregulated and downregulated genes, were ranked according to significance and fold-change. Both broad pTau-associated transcriptomic signatures as well as more specific pTau-associated transcriptomic states were observed. While there were AD-relevant genes present in the downregulated gene sets, there were no broader patterns detectable. However, a consistent set of top upregulated genes were found across all vulnerable supertypes and several neuronal and non-neuronal subclasses and exhibited partial overlap with the published Otero-Garcia pTau-associated gene signature (Figure 3.23). The four excitatory L2/3 IT types with the highest overall differentially expressed genes (L2/3 IT_1, L2/3 IT_5, L2/3 IT_6, and L2/3 IT_13) also had the highest fold-change values in the dataset for each of the most significant differentially expressed genes.

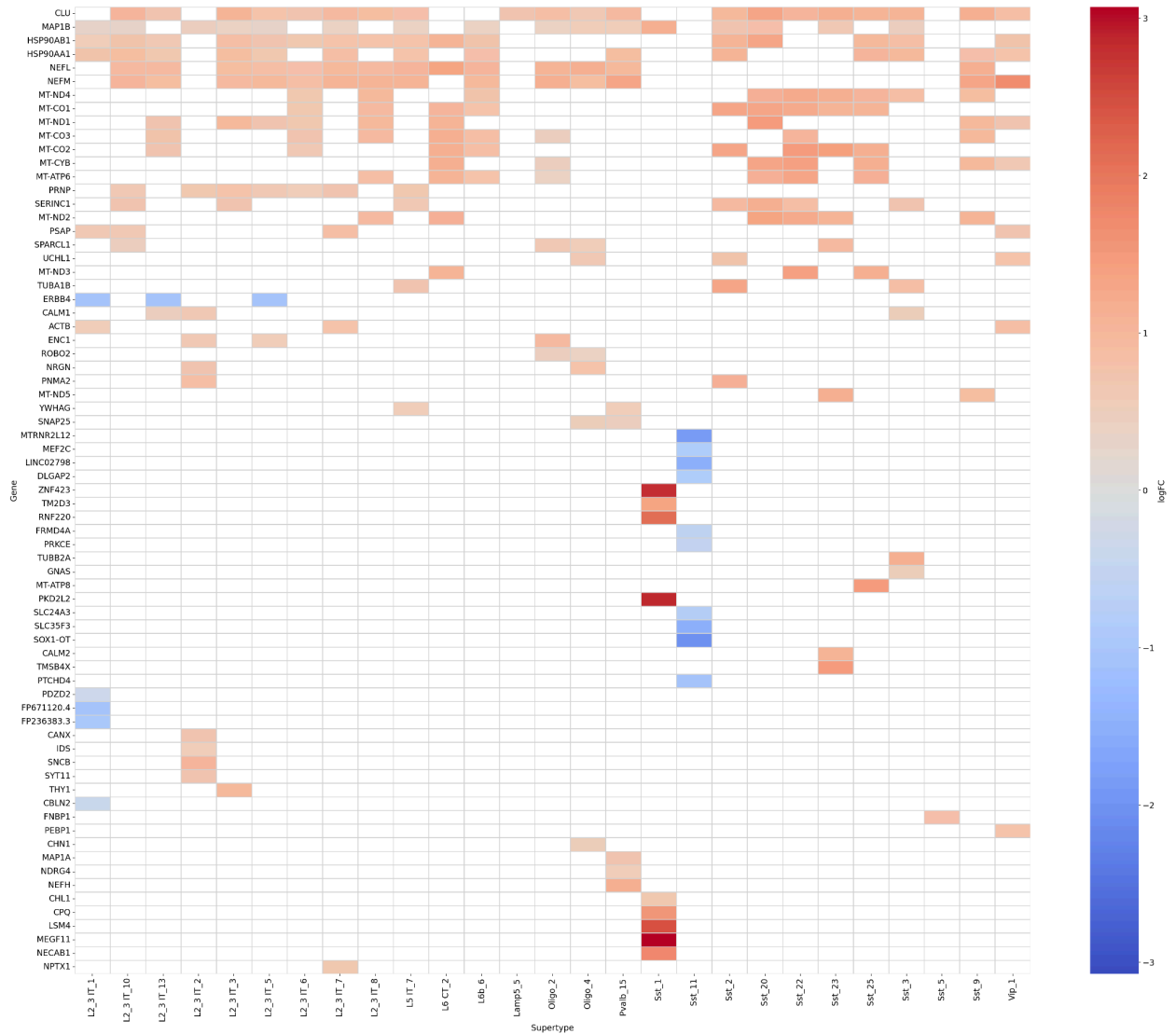


Figure 3.23. Heatmap of top differentially expressed genes across selected vulnerable supertypes.

When assessing the most significant ($FDR < 0.05$) and highest fold change ($LogFC > 0.2$) genes across the selected vulnerable supertypes, several broad patterns emerged. Most notably, CLU remains among the most significantly upregulated genes across all types, including glia. This gene was not considered significant in the Otero-Garcia study due to its relatively high expression across cell types, however while this is true, when using the nebula-based approach, CLU is significantly upregulated only in pTau-positive cells, suggesting its relevance as a non-cell-type-specific marker of pTau accumulation. This is unsurprising given that CLU, which stands for Clusterin, encodes a secreted glycoprotein that has a well-established link to AD; in fact, according to GWAS, CLU is one of the strongest risk loci for late onset AD⁷.

Another broad pTau-associated gene found to be significant across most vulnerable cell types was HSP90AA1/HSP90AB1. Importantly, this gene was also noted as one of the commonly upregulated genes in pTau-positive cells in the Otero-Garica dataset. This gene encodes the protein HSP90, a molecular chaperone protein that assists in numerous functions including folding, stabilization, and degradation of proteins. HSP90AA1 is highly expressed in neurons where it has a more specific role in maintaining cytoskeletal components. More directly, it's been shown to chaperone tau and is thought to contribute to tau aggregation under conditions of cell stress. HSP90 inhibitors have been shown to promote the degradation of hyperphosphorylated tau via the ubiquitin-proteasome system, suggesting a direct role in cellular pathophysiology and a potential route for pTau-directed therapeutics⁸.

Lastly, MAP1B, which encodes the protein microtubule associated protein 1B, plays an important role in regulating the cytoskeleton and more specifically, microtubule dynamics⁹. MAP1B was found to have a broad, non-cell-type specific pTau signature, as it was upregulated in the pTau-positive fraction of the majority of cell types in the dataset. Otero-Garcia et al. noted increased MAP1A expression in pTau-positive excitatory neurons in their dataset, which is a closely related gene in the same family with similar functions.

SERINC1 (Serine Incorporator 1) encodes a transmembrane protein which incorporates serines into membrane lipids. This gene is ubiquitously expressed with low regional specificity; however, in the pTau-seq dataset this gene is most upregulated in a subset of vulnerable L2/3 IT (L2/3 IT_3 and L2/3 IT_10), L5 IT (L5 IT_7) and Sst (Sst_2, Sst_3, Sst_20, and Sst_22). The link between the functional role of this protein and AD are limited; however transcriptomic studies, including SEA-AD and AMP-AD, have found this gene to be differentially expressed in different AD-related contexts¹⁰. In SEA-AD, emerged as DE gene when comparing affected and unaffected (in terms of relative abundance changes along CPS) Sst supertypes, suggesting that pTau accumulation may be driving the ultimate cellular vulnerability initially observed in SEA-AD.

Two genes were specifically upregulated in the excitatory types (L2/3 IT) but not inhibitory (Sst) - NEFL and NEFM. These genes were among the top upregulated genes in pTau-positive cells in three of the vulnerable L2/2 IT supertypes, L2/3 IT_5, L2/3 IT_6 and L2/3 IT_13 (but not L2/3 IT_1). These genes encode two variations of neurofilament proteins (light chain and heavy chain), which are structural proteins that play an integral role in maintaining neuronal, and particularly axonal structure¹¹. Both NEFL and NEFM transcripts are present in these supertypes and among the most significant high fold change genes. Both transcripts are widely expressed in neurons, particularly cortical excitatory neurons. Importantly, NEFL protein levels in CSF and plasma are an established biomarker for neurodegeneration. Upregulation of NEFL was also observed in the corresponding pTau-positive cell types by Otero-Garcia et al. demonstrating the robustness of this association across institutes, methodological differences, and brain regions. Upregulation of these genes is notably absent from affected Sst types, highlighting an interesting divergence.

Across L2/3 IT types, the most consistently downregulated gene was ERBB4. This gene is a member of the ErbB family of receptor tyrosine kinases, which are activated by ligands such as neuregulins (NRGs) that trigger intracellular signaling cascades. This gene is highly expressed in the brain and has been implicated in several functions, including synaptic maintenance through the binding of NRG-1¹². Along this vein, dysregulation of this receptor-ligand interaction is thought to contribute to synaptic loss and cognitive decline¹³. Further, some data suggest that ERBB4 may modulate neurotoxicity through JNK/tau pathway activation¹⁴.

The most consistently upregulated gene that was specific only to L2/3 IT types was PRNP. Prion Protein is a highly conserved cell surface glycoprotein involved in regulating neuronal excitability and synaptic function along with neuroprotective and anti-apoptotic roles¹⁵. Critically, misfolded versions of this protein form self-propagating beta sheets that are extremely neurotoxic that emerge in cases inherited autosomal dominant diseases (e.g., Familial Creutzfeldt–Jakob disease and fatal familial insomnia) and acquired infection (e.g., mad cow disease). As it relates to AD, several studies have suggested that prion protein is directly involved in key AD pathophysiological processes, such as binding to BACE1 causing dysregulation to amyloid preprocessing towards increased A β oligomer formation¹⁶. The top ten

differentially expressed genes ($FDR < 0.05$ and $\text{LogFC} > 0.2$) were visualized as volcano plots for a subset of key excitatory supertypes, with the top ten most significant genes denoted (Figure 3.24).

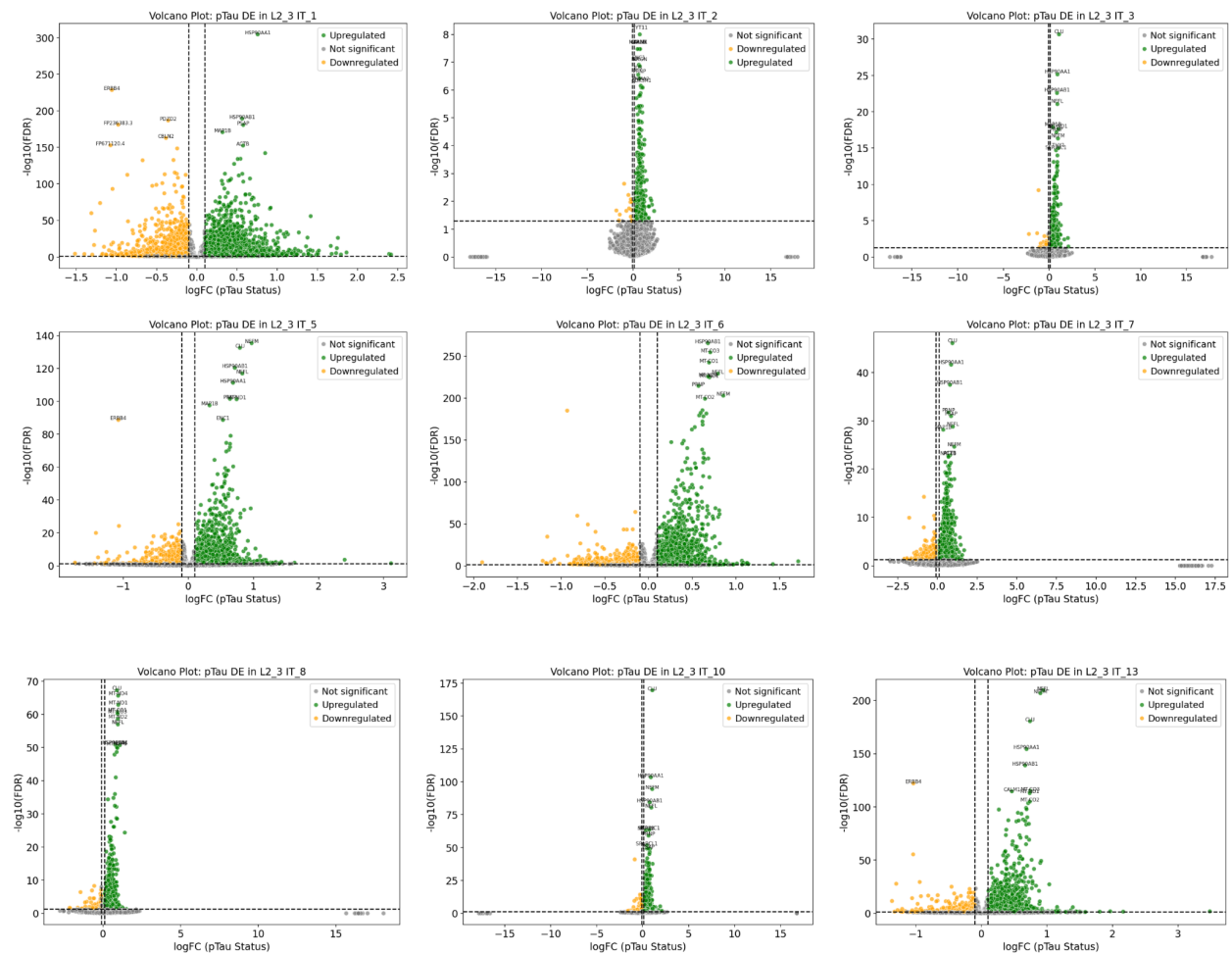


Figure 3.24. Top differentially expressed genes in vulnerable excitatory populations. Selected excitatory supertypes selectively vulnerable pTau accumulation. Volcano plots visualizing top 10 differentially expressed genes from nebula differential expression analysis.

Despite fewer overall numbers of inhibitory neurons in the dataset due to their low abundance in the cortex, several supertypes identified as vulnerable in SEA-AD were found to have several interesting and unique significantly upregulated and downregulated genes relative to pTau status. First, two Sst supertypes had unique patterns of top DE genes that were not present in other supertypes - Sst_1 and Sst_11. While the AD-relevance of many of these genes

were intriguing, these supertypes were particularly low in abundance, which must be taken into account when assessing the results.

Sst_1 exhibited a unique pattern of transcriptomic changes, which is particularly interesting given that this was the only Sst supertype with substantially fewer pTau-positive cells than pTau-negative cells (though not statistically significant). It is also one of only a few Sst supertypes that was found not to be vulnerable in SEA-AD. This population had nine highly significant and high fold-change upregulated genes that were not represented in other supertypes (NECAB1, MEGF11, LSM4, CPQ, CHL1, PDK2L2, RNF220, TM2D3, SNF423).

The protein NECAB1 (N-terminal EF-hand calcium binding protein 1) plays a central role in regulating intracellular calcium signaling¹⁷. MEGF11 (Multiple EGF-like-domains 11) is a transmembrane protein involved in cell-to-cell interactions and possibly implicated in synaptic pruning¹⁸. CHL1 (Close homolog of L1) encodes a protein involved in axonal function and are cleaved by BACE1¹⁹. ZNF423 (Zinc Finger Protein 423) is a protein involved in DNA damage response and implicated in ciliopathies²⁰. Together this suggests that Sst_1 is undergoing a distinct stress response that implicates changes to synaptic function; however this state may be induced by factors other than intracellular pTau accumulation.

Sst_11 is one of the only selected vulnerable supertypes with a significant subset of downregulated genes. Several genes have more direct relevance to AD. First, PKCE (Protein Kinase C Epsilon Type) is a serine/threonine kinase involved in intracellular signaling and has been implicated in regulating tau pathways²¹. FRMD4A (FERM Domain Containing 4A) is a scaffold protein that has been associated with late-onset AD risk and modulation cellular release of tau, making this association with pTau-positive cells particularly interesting^{22,23}. DLGAP2 (Discs Large Associated Protein 2) is a scaffold protein at postsynaptic densities and is thought to be important for glutamatergic synaptic function and is linked to synaptic degeneration in neurodegenerative disease²⁴. MTRNR2L12 (MT-RNR2-like 12 / Humanin-like peptide 12) is thought to be neuroprotective, as it plays a role modulating oxidative stress and apoptosis; interestingly, it has also been proposed as an early biomarker of AD^{25,26}. Together, this suggests

that Sst_11 is undergoing pTau-related transcriptomic changes that involve synaptic alterations, tau-specific regulation, and neuroprotective pathways.

The remaining Sst supertypes, visualized in the volcano plots below, exhibit an extremely consistent pattern of gene upregulation in response to pTau, with few exceptions (Figure 3.25). The gene encoding TUBB2A is significantly upregulated in Sst_3, which is notable given that the beta-tubulin protein is a core component of microtubules and a tau binding partner. Mutations in TUBB2A are associated with neurodevelopmental malformations and developmental disorders, and recently, an in vitro primary neuron model found that primary neurons with pTau inclusions were enriched for this gene²⁷. CALM2 is also noted as an exception, significantly upregulated in Sst_23 as the protein Calmodulin 2 has a well-established role as an intracellular calcium sensor and second messenger regulating several downstream pathways²⁸. With relevance to AD, Calmodulin can localize in pTau tangles and activates kinases and phosphatases that directly phosphorylate and dephosphorylate pTau²⁹.

Consistent changes include the non-cell-type specific pTau-associated genes mentioned above (CLU, MAP1B, and HSP90AA1/HSP90AB1) and well as several groups of mitochondrial (MT) genes. While MT genes have historically been thought of as an indicator of low quality snRNA-seq preparations, the updated interpretation is more nuanced, and this is especially true given the pTau-seq dataset retains somas, which theoretically would contain more MT-genes. Broadly, mitochondrial dysfunction has been implicated as a contributor and potential biomarker of AD^{30, 31, 32, 33}. The pTau-associated upregulated MT genes include MT-ND1 / MT-ND2 / MT-ND4, MT-CO1 / MT-CO2 / MT-CO3, MT-CYB, MT-ATP6. The MT-NDs encode mitochondrially encoded NADH dehydrogenases that serve as subunits (Complex I) of oxidative phosphorylation complexes in the mitochondrial electron transport chain. MT-ND1 has specifically been implicated in AD^{34, 35}.

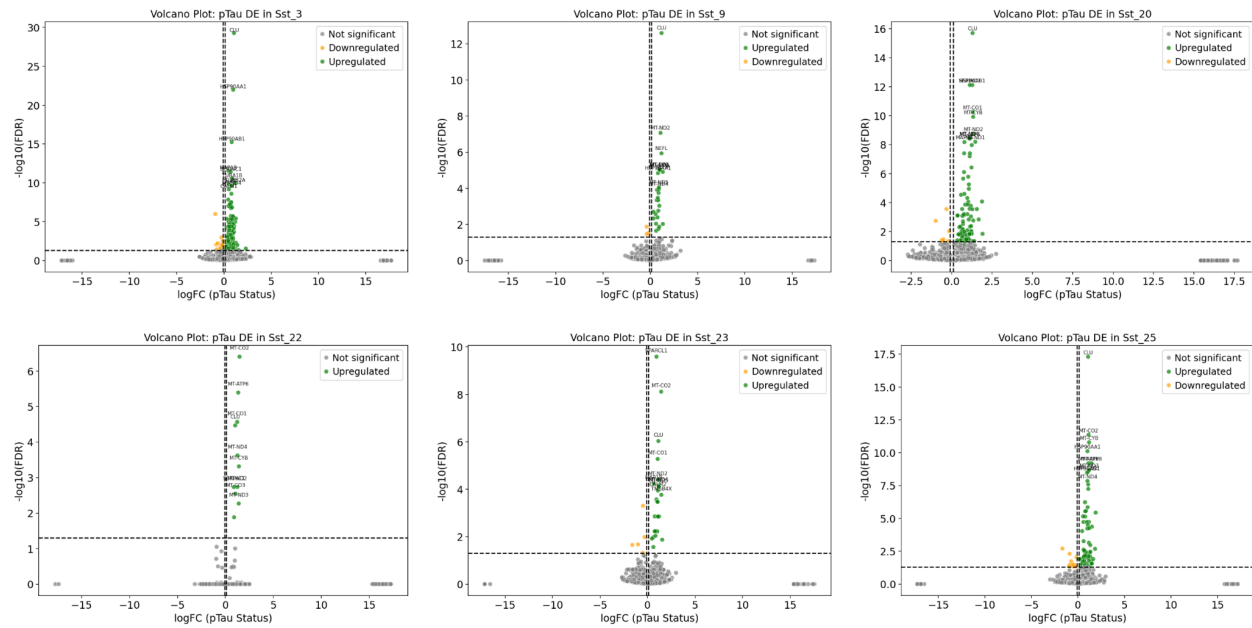


Figure 3.25. Top differentially expressed genes in vulnerable inhibitory populations. Selected inhibitory supertypes. Volcano plots visualizing top 10 differentially expressed genes from nebula differential expression analysis.

pTau-seq enabled the identification of several disease-relevant cellular populations with unexpectedly high overall fraction of pTau-positive cells. These populations, such as somatostatin-positive interneurons, have not previously been appreciated as cortical pathology-bearing vulnerable cell types in AD. Our results suggest that temporal and cell-type specific dynamics of intracellular pTau pathology is more nuanced than previously appreciated, which must be considered when studying cellular vulnerability and developing pTau-based therapeutics for AD.

<i>Supertypes NOT vulnerable in SEA-AD</i>	<i>Supertypes vulnerable in SEA-AD</i>
<i>Not vulnerable to pTau or cell death</i> <i>L4 IT_1</i> <i>L4 IT_2</i> <i>L4 IT_3</i> <i>L4 IT_4</i> <i>L5 IT_1</i> <i>L6b_3</i> <i>L6 IT Car3_1</i> <i>Lamp5_1</i> <i>Lamp5_2</i> <i>Lamp5_Lhx6_1</i> <i>Pax6_1</i> <i>Pax6_2</i>	<i>Not vulnerable to pTau but vulnerable to cell death</i> <i>Lamp5_3</i> <i>Lamp5_4</i> <i>Lamp5_5</i> <i>Lamp5_6</i> <i>Sncg_6</i> <i>Sncg_1</i> <i>Sncg_8</i> <i>Sncg_4</i> <i>Sncg_5</i> <i>Vip_13</i> <i>Vip_1</i> <i>Sst_3</i>

Pax6_3 Pax6_4 Pvalb_7 Pvalb_13 Sncg_2 Sncg_3 Sst_1 Sst_4 Sst_5 Sst_7 Sst_10 Chandelier_1 Chandelier_2 Vip_12 Vip_14 Vip_16 Vip_18 Vip_19 Vip_21 Vip_23 Vip_4 Vip_5 Vip_6 Vip_9	Sst_19 Sst_9 Sst_13 Sst_11 Sst_20 Sst_23 Sst_2 Pvalb_6 Pvalb_5 Pvalb_8 Pvalb_1 Pvalb_3 Pvalb_2 Pvalb_15 Pvalb_14 Pvalb_12 Pvalb_10 Pvalb_9 L5/6 NP_1
Vulnerable to pTau but not cell death Sst_12 L2/3 IT_2 L6 IT_1 L5 IT_2 L5 IT_6 L5 ET_1 L5 ET_2 L6 CT_1 L6 CT_3 L6 CT_4 L6b_1 L6b_2 L6b_4 L6b_5 L6b_6 L6 IT Car3_2 L6 IT Car3_3 L5/6 NP_2 L5/6 NP_3 L5/6 NP_6 L5/6 NP_4	Vulnerable to pTau and cell death Vip_15 Vip_2 Vip_11 Sst_22 Sst_25 L2/3 IT_1 L2/3 IT_6 L2/3 IT_7 L2/3 IT_13 L2/3 IT_5 L2/3 IT_10 L2/3 IT_8 L2/3 IT_12 L6 IT_2 L5 IT_3 L2/3 IT_3 L5 IT_5 L5 IT_7 L6 CT_2

Table 3.5. Breakdown of cell-type specific vulnerability to pTau accumulation and cell death (relative abundance changes observed in SEA-AD). Not vulnerable: no relative abundance changes in SEA-AD. Vulnerable SEA-AD: relative abundance changes in MTG only. Vulnerable SEA-AD BOTH: relative abundance changes in MTG and PFC. Ns: not significantly different with FDR-corrected t-test L4 IT_4 null baseline.

Discussion

Identification of cells types vulnerable to pTau pathology

To disentangle the relationship between pTau pathology and cellular vulnerability, pTau-seq was performed on a subset of extensively characterized donors from the SEA-AD study. Assessing protein status and the complete transcriptome of individual cells has provided unprecedented insight into which highly resolved, transcriptomically-defined cell types preferentially accumulate pTau. Further, comparing and contrasting vulnerability to cell death versus vulnerability to pTau accumulation has revealed several broad patterns of affected cell types and associated gene expression patterns as well as intriguing differences.

Excluding cell types with low representation in the dataset (<20 cells in either pTau status group), the most abundant pTau bearing cell types in the dataset are L2/3 IT (~41%), L5 ET (~16%), L6b (~7%), and L5 IT (~6%). The Sst subclass only represents approximately 2% of the pTau positive cells, however the fraction of pTau positive cells relative to the Sst population is markedly higher (~38%). The high pTau fraction is driven by particular Sst supertypes, which exhibit strikingly high pTau-positive populations relative to the pTau-negative population.

As an initial approach to test for significant increases or decreases in pTau accumulation within supertypes, FDR-corrected t-tests were performed on each supertype. This yielded 40 statistically significant supertypes, all with increased pTau-fractions relative to baseline (Sst_12, L2/3 IT_2, L6 IT_1, L5 IT_2, L5 IT_6, L5 ET_1, L5 ET_2, L6 CT_1, L6 CT_3, L6 CT_4, L6b_1, L6b_2, L6b_4, L6b_5, L6b_6, L6 IT Car3_2, L6 IT Car3_3, L5/6 NP_2, L5/6 NP_3, L5/6 NP_6, L5/6 NP_4, Vip_15, Vip_2, Vip_11, Sst_22, Sst_25, L2/3 IT_1, L2/3 IT_6, L2/3 IT_7, L2/3 IT_13, L2/3 IT_5, L2/3 IT_10, L2/3 IT_8, L2/3 IT_12, L6 IT_2, L5 IT_3, L2/3 IT_3, L5 IT_5, L5 IT_7, L6 CT_2). This list captures the cell types typically associated with cellular vulnerability and pTau pathology in AD, but adds a critical layer to resolution as not all supertypes within a subclass conform to the same direction of change or transcriptomic alterations.

Determining a biological baseline for the expected amount of pTau positivity a healthy or otherwise non-diseased supertype should have was unexpectedly difficult. Given the high sensitivity of FACS (relative to prior methods of assessing pTau positive cells such as immunohistochemistry) and the novelty of combining pathological protein assessment with

snRNA-seq data, there is no precedent for the expected levels. Prior to a FACS-based dataset with relatively high basal levels of pTau positivity across all cell types, the biological expectation may be that hyperphosphorylation is pathological and therefore the biological baseline would be set to zero. Ultimately, the well-established sparing of L4 cells and the observation that this was the only supertype with consistently lower pTau-positive fractions led to the decision to set this as the expected baseline for statistical tests.

Approaches such as Cellular Indexing of Transcriptomes and Epitopes by sequencing (CITE-seq), which allows for simultaneous measurement of transcriptomes and protein expression by using oligo-conjugated antibodies that are targeted to the surface protein of interest (e.g., AT8 for pTau as in pTau-seq), directly address this challenge among others. The antibody-conjugated oligos are cleaved, barcoded, and sequenced just as individual RNA transcripts are, allowing for a quantitative readout of single cell pTau binding in the context of the whole transcriptome. This added layer of quantitative information has the potential to be extremely valuable in unpacking how pTau levels may differentially impact transcriptomic state.

Further, eliminating the FACS step also eliminates a step that involves somewhat arbitrary user-defined cutoffs (FACS gating). The potential of assessing precise changes in transcriptomic state of highly granular cell types in the context of varying levels of pathological protein expression is the key to disentangling the complex relationship between pathology and vulnerability in the brain. The risk of implementing this adapted approach (relative to the standard snRNA-seq workflow) is quite low, given that the RNA measurements retain the same quality and robustness. This approach can and should be applied to any brain disease involving intracellular protein accumulation.

Assessing the overlap between vulnerability to cell death (SEA-AD) and pTau accumulation

Of the 40 superotypes vulnerable to pTau accumulation, 19 were also vulnerable to relative abundance changes in the SEA-AD study (Vip_15, Vip_2, Vip_11, Sst_22, Sst_25, L2/3 IT_1, L2/3 IT_6, L2/3 IT_7, L2/3 IT_13, L2/3 IT_5, L2/3 IT_10, L2/3 IT_8, L2/3 IT_12, L6 IT_2, L5 IT_3, L2/3 IT_3, L5 IT_5, L5 IT_7, L6 CT_2).

As was noted in the description of the complete list of supertypes with significantly increased pTau fractions, these supertypes exhibiting both forms of vulnerability, correspond even more closely with well-established understanding that AD pathophysiology preferentially impacts large excitatory pyramidal neurons in L2 and L3, and L5, with sparing of L1 and L4. Deeper L6 neurons exhibit some variability in their vulnerability in the AD literature and the improved resolution of this dataset suggests that this may be explained by contrasting dynamics at the supertype-level. The finding that Sst neurons are vulnerable in AD and particularly relevant to early pathophysiological processes is very new and the current work adds an essential missing piece to understanding early vulnerability - increased susceptibility to pTau pathology.

Several interesting points of divergence were observed between supertypes vulnerable to cell death in SEA-AD and vulnerable to pTau accumulation in pTau-seq. The first was the inhibitory neuron supertype Vip_18, which was not vulnerable to relative abundance changes in SEA-AD, but has significant pTau accumulation in pTau-seq. Despite being another low abundance cell type in the brain, at the subclass level more cells were captured for Vip than Sst or Pvalb, such that at the supertype-level all types except for Vip_21 had sufficient representation in the dataset. Vasoactive intestinal peptide (Vip) cells are GABAergic inhibitory neurons that localize in superficial layers of the cortex and functionally inhibit other inhibitory populations, such as Sst, leading to disinhibition of excitatory pyramidal neurons. Their role in AD is not well understood, though recent snRNA-seq studies, including SEA-AD, have implicated certain subtypes as being vulnerable to cell loss (Vip_1, Vip_2, Vip_11, Vip_13, and Vip_15).

The second interesting divergence is the pronounced vulnerability of L6b cells to pTau accumulation, despite not being vulnerable to cellular abundance changes in SEA-AD. Cortical layer 6 (L6) is a relatively unexplored layer of the cortex with a broad cell type composition¹⁸. The deepest portion of this layer, L6b, resides along the white matter boundary and contains primarily excitatory neurons, as well as a subset of residual subplate neurons retained from early neocortical development¹⁹.

The transcriptomically-defined L6b subclass describes a population of excitatory cells, primarily pyramidal neurons, whose function and projection patterns remain largely unknown²⁰.

Given their anatomical location near the white matter boundary, it is possible that pTau-seq sampled more generously along this boundary, thereby capturing more L6b cells. Sampling variability may also explain why one donor had disproportionate increases in this subclass only. The enrichment for pTau-positive cells in the pTau-seq method and the finding that this population has a particularly high fraction of pTau-positive cells, may also explain the increased sampling of L6b cells in pTau-seq relative to other studies such as SEA-AD. The robustness of the pTau accumulation across L6b supertypes is in alignment with the findings of Otero-Garcia et al., who found one of their broader excitatory clusters associated with L6b as having increased pTau proportions. This intriguing observation suggests that this relatively understudied population of excitatory neurons preferentially accumulates pTau but appears to be relatively resistant to cell death in the context of AD.

Affected cell types exhibit both shared and cell-type-specific transcriptomic changes

Across vulnerable supertypes, both broad and cell-type specific pTau-associated transcriptomic signatures were observed. Interestingly, however, only upregulated genes exhibited broader signatures, while downregulated genes appeared to be more supertype-specific. The top upregulated genes across subclasses were CLU, HSP90AA1 / HSP90AB1, NEFL / NEFM, MAP1B, and SERINC1. Several mitochondrial genes show broad upregulation (e.g., MT-NDs, MT-COs, MT-CYB, and NT-ATP). While MT genes have historically been viewed as an indicator of low-quality snRNA-seq preparations, the updated interpretation is more nuanced. Broadly, mitochondrial dysfunction has been implicated as a contributor and potential biomarker of AD²¹. An updated perspective suggests that MT transcripts may elicit a protective signaling response to enhance nuclear DNA repair in states of stress and disease²². The pTau-seq dataset is particularly well suited to capture biologically relevant MT-gene changes because unlike snRNA-seq, it retains somas, which theoretically would contain more MT-genes.

Differential expression across subclasses indicates a shared patterns of dysregulation to protein homeostasis (e.g., HSP90AA1 upregulation suggests disrupted tau proteostasis), cytoskeletal instability (e.g., MAP1B, NEFL/NEFM, and TUBB2A), and mitochondrial

dysfunction (MT-gene upregulation suggestive of energy stress and reactive oxygen species pathways). Specific to excitatory neurons were genes implicating synaptic and membrane signaling (e.g., ERBB4 downregulation and PRNP upregulation, which could suggest dysfunctional neuregulin signaling and increased vulnerability to pathological aggregates). Gene expression in inhibitory neurons suggests intracellular stress potentially driven by calcium dysregulation (e.g., CALM2, NECAB1). GSEA with GO gene sets (including GO_Biological_Process_2025, GO_Cellular_Component_2025, GO_Molecular_Function_2025, KEGG_2021_Human, MSigDB_Hallmark_2020, Reactome_2022, SynGO_2024) confirmed these broad and cell-type specific associations.

Addressing weaknesses and opportunity for improvement with pTau status tool development

There are several key points of improvement for the pTau-seq dataset that highlight exciting future directions for this work. One limitation of the pTau-seq dataset is the small sample size. While the sample size of pTau-seq alone is sufficient to map to highly refined supertypes, it is not sufficient to test for intra-supertypes pTau-associated changes - a critical assessment to understand which cells preferentially accumulate pTau. An important design consideration of the pTau-seq dataset was that the donors and brain regions have complete overlap with the SEA-AD study, such that it could be possible to develop a pTau status tool that could infer predicted pTau status in standard snRNAs-seq datasets, such as SEA-AD. This impactful tool could not only increase the scientific utility of many existing AD scRNA-seq datasets, but it would also readily address the sample size limitation in the current dataset.

The second limitation is the FACS-based binary thresholding of pTau. This can be readily addressed by adapting pTau-seq protocol to be compatible with cellular indexing of transcriptomes and epitopes sequencing (CITE-seq), which is a sequencing-based method that simultaneously quantifies cell surface proteins and transcriptomic data within a single cell readout. This is an exciting future direction that may be further adapted for multiplexed protein readouts, enabling quantitative assessment of multiple intracellular pathological peptides and transcriptomic data from individual cells. This addition would not only increase the sample size (if performed on the same donors), enabling more rigorous statistical analysis, but it would

provide quantitative single-cell measure of pTau levels, providing a level of resolution that is not available from pTau-seq alone. The addition of CITE-seq data would not only serve to validate the initial FACS-based sorting strategy in pTau-seq, but importantly, it would allow for a continuous metric of pTau pathology to be applied across highly resolved cell types.

Chapter 4. DISENTANGLING THE RELATIONSHIP BETWEEN VULNERABLE CELL TYPES AND PROTEIN PATHOLOGY IN AD USING SPATIAL TRANSCRIPTOMICS

Introduction

Spatial transcriptomics has been one of the fastest evolving areas in biological research over the past decade, particularly in the fields of cancer and neuroscience. Around 2015, the groundwork for the technology was being established with methods such as in situ RNA sequencing¹. These early assays demonstrated that intact tissue sections were amenable to genome-wide profiling. Shortly thereafter in 2016, Stahl et al., published a newer approach in Nature, in which they coined the term ‘spatial transcriptomics’². These early advances created excitement in the field and were highly specific, but were low throughput and varied in their sensitivity and resolution.

The commercial race to develop the best spatial transcriptomics technology kicked off around 2018 when VizGen released MERSCOPE and 2019 when 10x Genomics acquired ST Technology and launched Visium. In 2019, NanoString entered the scene by introducing their first spatial instrument, GeoMx, which was also the first to provide an FFPE-compatible workflow. This compatibility is game-changing when it comes to assessing postmortem human brain tissue, which still is typically preserved in FFPE blocks. More recently, oligo-tagging approaches combined with cyclic high resolution imaging has enabled both Nanosting (CosMx) and 10x Genomics (Xenium) to push this technology further. Both platforms will soon offer whole tissue, whole transcriptome profiling capabilities at sub-cellular resolution. The multi-omic capabilities of these platforms are also continually improving, such that high plex protein profiling can now be run on the same tissue section as RNA assay, though this is an area of active development and improvement is still needed.

The fast-paced technological advancement enabling the rapid growth of spatial transcriptomic-based technologies has led to numerous breakthroughs in understanding biology and disease. As it relates to topics in this dissertation, this technology has proven to be a key

player in the modern human brain atlasing effort. The early Allen Institute brain atlases were based on in situ hybridization technology to probe and visualize individual RNAs localization throughout a tissue sample. Understanding which brain regions and substructures certain RNA's are expressed was an important advancement at the time, though the goal of assaying the RNA composition of single cells was still far from reality. The rapid development and widespread adoption of droplet based scRNA-sequencing technologies has made this next level of classification possible. Across all fields of biology, this has since shifted the definition of a cell type to a primarily RNA-based classification system.

Modern spatial transcriptomic approaches provide the promise of bridging the spatial atlases that were previously limited to only a handful of RNA probes, with the current state-of-the-art single cell whole transcriptome profiling the field has become accustomed to. Now, in a single piece of tissue, every single cell can be identified transcriptomically, in situ - in its physical location in the brain. This not only provides the spatial context of cellular compositions, but also how the tissue microenvironment fits into this complex picture. This context is especially true for cells like interneurons, which based on their transcription alone, cannot be assigned a physical location in the brain. As described in Chapter 3, it's recently been discovered that only certain subsets of interneuron supertypes are selectively vulnerable in disease. In our SEA-AD study, we were able to determine that these vulnerable supertypes are consistently localized to the layers of the cortex where vulnerable excitatory neurons are also found, which was made possible by probing this cell type using spatial transcriptomic approaches with a custom panel of RNAs specifically designed to identify these highly granular transcriptomic supertypes. With this added information one can generate much more targeted hypotheses as to how very specific cell types may be impacting one another in health and disease.

Early iterations of the technology offered lower plex (typically <100 probes) and were unable to combine RNA and protein in the same slide. As described above, the combination of protein (to establish pathological state of a cell) and RNA (for both cell type labeling and differential gene expression analysis) is critical to understanding how protein-based pathological processes may contribute to cellular changes and selective vulnerability. The pTau-seq assay

described in Chapter 3, contributes substantially to this understanding as the first to probe the MTG in AD and across varying pathological stages (Braak IV - VI). This approach, however, lacks any spatial context. Spatial context is the third, essential piece of the puzzle to disentangling how pathology contributes to cellular vulnerability in AD.

In order to probe cell-type specific vulnerability and gene-expression changes while retaining relevant spatial and protein pathology (pTau and A β) context, two custom antibodies were developed for the NanoString CosMx platform (AT8 - pTau, and 6e10 - A β). These antibodies were selected for their robustness in postmortem human brain tissue and for consistency across modalities in the larger project. As described in detail in the methods, pTau and A β -specific antibodies were added to the standard NanoString protein panel for morphological assessment for cell segmentation in place of a secondary nuclear marker (histone) and membrane marker (PanCK). While there was some risk of losing a membrane marker and redundant nuclear marker, the NanoString R&D team was confident that the remaining markers (DAPI, rRNA, and GFAP) would be sufficient for their cellpose-based cell segmentation pipeline. Following successful segmentation, our modified CosMx protocol yielded an incredibly high-quality proof-of-concept dataset to assess cell-type specific vulnerability with protein and spatial context.

Methods

Brain processing

The tissue was processed exactly as described in Chapter 2, given the donors were consistent across the larger project. As it pertains to this chapter, the brain is hemisected by performing a midsagittal cut through the cerebrum, cerebellum, and brainstem. One hemisphere is embedded in dental alginate and transferred to a slicing sled for uniform 4mm coronal sectioning anterior to posterior (first slice is the frontal pole). Each slab is photographed. Alternating slabs are either fixed in 10% neutral buffered formalin (NBF) or frozen in a dry ice isopentane slurry. The formalin fixed slabs are then subsampled into blocks to be paraffin embedded (FFPE) for 5 μ m sectioning and histological staining. For NanoString specific application, the sections are cut fresh, less than three days before the intended run date on the

NanoString CosMx instrument. 5um MTG sections were isolated from an FFPE block with STG-MTG and mounted on superfrost slides. Tissue for three donors was included on each slide.

NanoString CosMx slide preparation

Day 0: Reagent Prep

- *Section and bake tissue overnight*
- *Prepare shelf stable reagents (leave at room temperature)*
 - *1x PBS (50mL 10x PBS + 450mL DEPC H2O)*
 - *2x SSC (50mL 20x SSC + 450mL DEPC H2O)*
 - *4x SSC (20mL 20x SSC + 80mL DEPC H2O)*
 - *NBF (1.181g Tris, 1.125g Glycine, 150mL DPEC H2O)*
 - *NSH Acetate (confirm 20mg aliquots ready at -20C)*

Day 1: Slide Prep

Two solutions to prep day of:

- *1X Target Retrieval Solution: 5mL 10X Target Retrieval Solution + 45mL DPEC water*
- *Digestion Buffer: serial dilution (20mg/mL stock → 3ug/mL), First: 2uL stock Proteinase K + 198uL 1X PBS, Second: from tube you just created take 30uL and combine with 1970uL 1X PBS to make final working solution*
- *Bring NHS-Acetate to room temp (from -20C) 1-2 hours before using*

Deparaffinize FFPE sections (20 minutes)

- *Prep materials*
 - *Staining jars (in clear bins in CosMx Instrument room - round tops RNA only)*
 - *Xylene*
 - *100% EtOH*
 - *Hybridization tray*
 - *Steamer*
 - *1X Target Retrieval Solution (combine in coplin jar)*

- 5mL 10x Target Retrieval solution
- 45mL DEPC water
- DEPC water
- Line hybridization tray with Kimwipes and wet them with DEPC water
- Preheat chamber to 40C (hybridization oven - red in CosMx instrument room)
- Set baking oven to 60C (should be already on from O/N bake, Amanda's room)
- Prepare jars with Xylene and Ethanol to cover tissue (SAFETY)
- Prepare steamer
 - Fill to reservoir up to line with water (regular water OK in basin, NOT jars)
 - Place two staining jars inside steamer (one with DEPC water and the other with 1x Target retrieval solution - DEPC water for thermometer checks only)
 - Loosely gap with aluminum foil
 - Preheat steamer to 100C
 - MAY TAKE UP TO 1HR TO HEAT LIQUID IN JARS
- Preheat target retrieval solution
- Remove slides from oven
- Perform washes
 - Xylene for 5 min (IN HOOD)
 - Move to new xylene jar for 5 min (wait for proper disposal)
 - Transfer to 100% ethanol for 2 minutes (down sink OK)
 - Move to new 100% ethanol jar for 2 minutes
- Dry slides in rack in 60C for 5 min
- Leave at RT until target solution has been preheated to 100C
 - Stick thermometer through aluminum foil and ensure 100C
- REMOVE NHS-Acetate from -20C freezer and leave at RT for 1-2hrs

Target Retrieval (50 minutes)

- When 1X Target Retrieval reaches 100C, remove and QUICKLY perform next steps
 - Remove foil cap
 - Transfer slides to jar (with 1X target retrieval)
 - Replace foil

- Put back in steamer
- Reinsert thermometer until returns to 99C
- Start timer and run steamer for 15 minutes (FFPE)
- Carefully remove steamer lid and staining jar
- Immediately transfer all slides to jar with fresh DEPC water (@RT)
- Move slides up and down for 15 seconds
- Transfer slides to fresh 100% ethanol for 3 minutes
- PREP BENCH SPACE WITH ETHANOL AND RNASE AWAY, LAY OUT KIMWIPE
- Remove slides from ethanol and dry at RT for 30 min - 1 hour
- DURING BREAK PREP INCUBATION FRAME AND DIGESTION BUFFER (p32/33)
 - Note: Proteinase K at -20C

Tissue Permeabilization (40 minutes)

- Prep materials
 - Staining jars
 - Hybridization oven
 - Hybridization tray (PREHEAT)
 - Incubation frames
 - Digestion buffer (ON ICE)
 - DEPC water
 - PCR tubes
- Prepare incubation frame by tearing along perforations
- Dry slide with Kimwipe
- !!! Apply incubation frame by removing the THIN sheet, place frame around tissue pressing lightly along border; use razor blade to trim around the edges (do NOT remove thick top plastic)
- Remove digestion buffer from ice and warm by hand for 3 minutes (RT)
- Retrieve preheated hybridization tray from oven
- Place slides into insert on tray and using P200 add 400uL digestion buffer to each slide
- Insert the hybridization tray into oven and incubate at 40C for 30 minutes

- *DURING INCUBATION REMOVE FIDUCIALS and 2x SSC-T from 4C and let reach RT (PROTECTED FROM LIGHT) FOR 10 minutes*
- *After incubation tap off excess buffer one side at a time and transfer to staining jar with DEPC water*
- *Move slides up and down 3-5 times to wash, transfer to new DEPC jar and repeat 3-5 times (slides can be left in water but try to limit)*

Fiducial Preparation and Application (20 minutes)

- *Prep materials*
 - *Ultrasonic bath*
 - *Vortex*
 - *2X SSC-T (should now be at RT, in foil)*
 - *Fiducials (should now be at RT, in foil)*
 - *Staining jars*
 - *Staining tray (clean and dry)*
 - *1X PBS*
- *Prepare fiducials*
 - *Vortex 1 minute*
 - *Sonicate 2 minutes*
 - *Vortex 1 minute*
 - *Sonicate 2 minutes*
 - *Vortex 1 minute*
- *Dilute fiducial stock (0.1%) to working concentration (0.001%) in 2X SSC-T*
- *Dilute 1:100 by adding 12uL of fiducial stock in 1188uL 2X SSC-T (more volume than needed to avoid clumping)*
- *Remove slide from DEPC water and tap tissue on Kimwipe to remove water, lay horizontal in staining tray*
- *Immediately before adding fiducials vortex for additional 1 minute*
- *Vortex fiducials for 30 seconds between slides to keep in suspension*

- *Apply 250uL of final fiducial solution ensuring covers entire incubation frame (MUST BE ON GLASS AREA TOO)*
- *Incubate covered in staining tray for 5 minutes at room temperature*
- *ALL DOWNSTREAM STEPS ARE LIGHT SENSITIVE*
- *DURING INCUBATION PREPARE STAINING JARS FOR NEXT STEP*
- *After incubation, tap on Kimwipe to remove water and transfer to jar containing 1X PBS*
- *Wash slides in 1x PBS for 1 minute*
- *PROCEED TO NEXT STEPS IMMEDIATELY*

Post-Fixation (20 minutes)

- *Prep materials*
 - *Staining jar*
 - *10% NBF*
 - *NBF stop buffer*
 - *1X PBS*
- *Transfer slides to 10% NBF and incubate for 1 minute at room temperature*
- *Immediately transfer to NBF stop buffer for 5 minutes*
- *Repeat with new staining jar containing NBF stop buffer*
- *Transfer slides to 1X PBS wash for 5 minutes (sitting OK while NHS-Acetate prepared)*
- *REMOVE FROM -20C AND THAW ON ICE*
 - *RNase Inhibitory*
 - *CosMx RNA probe mix*
 - *(any custom add ons)*
- *Remove Buffer R from 4C and allow to reach room temperature*

NHS-Acetate Application (20 minutes)

- *Prep materials*
 - *Staining jar*
 - *sulfo-NHS-Acetate powder (pre aliquoted 15mg at -20C)*
 - *NHS-Acetate Buffer (NanoString, 4C)*
 - *2X SSC*

- *Prepare 100mM NHS-Acetate mixture*
 - *Make sure powder has reached room temperature*
 - *15mg aliquots for 2 slide runs*
 - *Multiple weight of powder by 38.5*
 - $15\text{mg} * 38.5 = 577.5 \text{ uL}$
 - *Slowly pipette up and down to mix (bubbles OK but avoid)*
 - *Perform one slide at a time*
 - *Remove slide from 1X PBS tapping on Kimwipe to dry and transfer to clean staining tray*
 - *Apply 200uL of NHS-Acetate mixture onto tissue inside incubation frame*
 - *Repeat with the remaining slides*
 - *Incubate in covered staining tray for 15 minutes at room temperature*
 - *Following incubation, tap off liquid and wash slides in 2X SSC for >5 minutes*
 - *Repeat 2X SSC wash for 2 more washes (new container, keep in second wash while hybridization mix is prepared)*
 - *OK STOPPING POINT 1HR IN SSC OR 6HRS AT 4C*

In Situ Hybridization (overnight, actual prep ~20 Minutes)

- *Prep materials*
 - *Hybridization oven and tray*
 - *Intubation frame covers*
 - *Thermal cycler*
 - *Ice bucket with ice*
 - *Buffer R (@RT, in hood)*
 - *CosMx RNA Probe mix (-20C)*
 - *RNase Inhibitor (-20C)*
 - *DEPC water*
- *MUST BE WITHIN 16-18HRS OF DAY 2*
- *WARM BUFFER R TO RT BEFORE OPENING*
- *DO NOT VORTEX PROBES*
- *SET HYBRIDIZATION OVEN TO 37C*

- *Preheat PCR to 95C*
- *Remove incubation frame cover (top drawer in CosMx prep room) and clean with ethanol, dry with kimwipe, lay clean cover on clean kimwipe*
- *Flick to mix probes and centrifuge before using (DO NOT VORTEX)*
- *Denature CosMx RNA probe mix (probe mix and rRNA) by moving total volumes needed from stock to 0.2mL PCR tubes (KEEP PROBES AND ADD ONS SEPARATE)*
 - *32uL probe mix into new tube*
 - *12.8 uL rRNA into new tube*
- *Heat at 95C for 2 minutes on thermal cycler with heated lid and immediately transfer to ice for 1 minute*
- *Flick to mix and contribute tubes*
- *Make hybridization solution (for a 2 slide run final volume 320uL, table 11, in hood)*
 - *32uL probe mix*
 - *12.8uL rRNA segmentation marker*
 - *3.2uL RNase inhibitor*
 - *16uL DEPC H2O*
 - *256uL Buffer R*
- *Clean area with RNase away and allow to dry (OR WIPE WITH WATER FIRST)*
- *Arrange Kimwipes on bottom of chamber and add DEPC water*
- *Perform the following one slide at a time*
 - *Remove slides from 2X SSC, tap gently to remove excess liquid*
 - *Remove frame THICK backing to exposure top adhesive layer of frame (razor)*
 - *Lay flat on clean surface and add 150uL of hybridization solution to the tissue inside the frame area*
 - *Carefully apply frame cover (set edge down on frame then gradually bend to lay over rest of area, do NOT press on center with tissue)*
 - *Place horizontally inside hybridization tray*
 - *Repeat for second slide*
 - *Ensure good seal between frame and slide and cover*
 - *Close hybridization chamber, put in oven and incubate at 37C overnight (max 18 hrs)*

Before day 2

- *Prep formamide aliquots (40mL) and store at 4C*

Day 2: Slide Prep Cont.

Stringent Washes (90 minutes)

Steps 1 - 7 performed by Angela (reproductive hazard)

Start with sides in SSC wash (1hr window to sit in SSC)

Nuclear and Cell Segmentation (2 hours)

- *Prep materials*
 - *Four staining jars of 1X PBS*
 - *Blocking buffer (4C)*
 - *Nuclear stain stock (-80C), on ice*
 - *Segmentation marker kit (-80C), on ice*
- *Prepare the Nuclear Stain buffer (total volume 220uL per slide)*
 - *Vortex then centrifuge thawed nuclear stain for at least one minute*
 - *Be VERY CAREFUL not to touch pipette tip to bottom or will free the pellet*
(do not want this)
 - *Combine 11uL nuclear stock, 429uL blocking puffer (total volume 440uL)*
- *Add a new incubation frame to each slide (dry area, remove THIN side only)*
- *Apply 200uL of nuclear Stain Buffer to tissue ensuring covers entire tissue*
- *Incubate for 15 minutes at room temperature in dark area (staining tray)*
- *DURING INCUBATION PREPARE STAINING SOLUTION*

*SLIDE 1**SLIDE 2*

Histone (blue)	8uL	pTau 150 (blue)	2uL
GFAP (red)	8uL	GFAP (red)	8uL
6e10 (green)	2uL	6e10 (green)	2uL
Staining Buffer	182uL	Staining Buffer	188uL
DAPI (UV)	NA	DAPI (UV)	NA
rRNA (yellow)	NA	rRNA (yellow)	NA

- *Stock custom antibody (200ug/mL), want 200uL $\rightarrow$ (200ug/mL*0.002mL) /*
- *Note: DAPI (UV) and rRNA (yellow)*
- *Remove slides, tap off buffer and transfer to 1X PBS to wash for 5 minutes*
- *DURING WASH ADD DEPC WATER TO STAINING TRAY*
- *Remove from PBS and one slide at a time add 200uL of staining mix to completely cover tissue (use coverslip if desired to avoid evaporation)*
- *Incubate for 1 hour at room temperature*
- *Following incubation, wash 3 times in 1X PBS for 5 minutes*

See video for flow cell assembly! Remember slide is light sensitive

For instrument prep do NOT forget RNase in SMI imaging tray (puncture spot) and Catalase/Pyranose Oxidase need to be reconstituted and added to Buffer 4

NanoString CosMx instrument protocol and FOV selection

The instrument control center (Okta) is used for the following steps. Instrument runs were performed in batches of two slides. First, a new flow cell is selected and configured by setting pre-bleaching (Configuration B - Human Neural) and cell segmentation (Configuration B -

Human RNA Neuroscience Cell Segmentation) profiles. For channel configuration comparable with the segmentation kit (1.0 - Human RNA Neuroscience Cell Segmentation) , exposures were set accordingly: DAPI - 24 milliseconds(ms), rRNA - 100ms, AT8 - 33ms, GFAP - 50ms, and 6e10 - 25ms. Tissue thickness was set to 5um for all slides. After setting up the flow cell parameters, the slides were scanned and loaded into the instrument. All slides were processed with the NanoString 6k RNA gene panel (2025). The instrument then undergoes a deck validation (check reagent volumes and prime flow cells) and instrument preparation before generating a preview scan. The preview scan is a lower resolution fluorescent image of the protein-based markers that is used to select fields of view (FOVs) to profile. FOVs were manually selected for each donor with effort made to capture the complete cortical depth in boxes with an area of approximately 15x20um² resulting in a target of 60 FOVs per donor (200 FOVs per slide) not to exceed the recommendation of 400 FOVs maximum per instrument run.

Custom cell segmentation

This project was performed in close collaboration with NanoString (Seattle, WA) given a pre-existing collaborative relationship with the UW Neuropathology group. Given that the project involved the custom addition of two protein markers (pTau and A β), cell segmentation could not be performed with the standard settings on the instruments built in software (AtoMx). NanoString R&D already had a working pipeline for custom off-instrument segmentation, so the cell segmentation was performed at NanoString. The same underlying pipeline used in AtoMx (cellpose-based) was used; however, the channels associated with the custom antibodies (pTau and A β) were removed from being considered in the algorithm. This is critical because reading pTau and A β as generic cell morphology markers would drastically sway the algorithm given the neuritic localization of pTau and extracellular localization of A β . The resulting re-segmentation had clear and consistent overlap with neuronal marker (DAPI) and cytoplasmic marker (rRNA), with very few false positives or negatives by qualitative assessment (Figure 4.1).

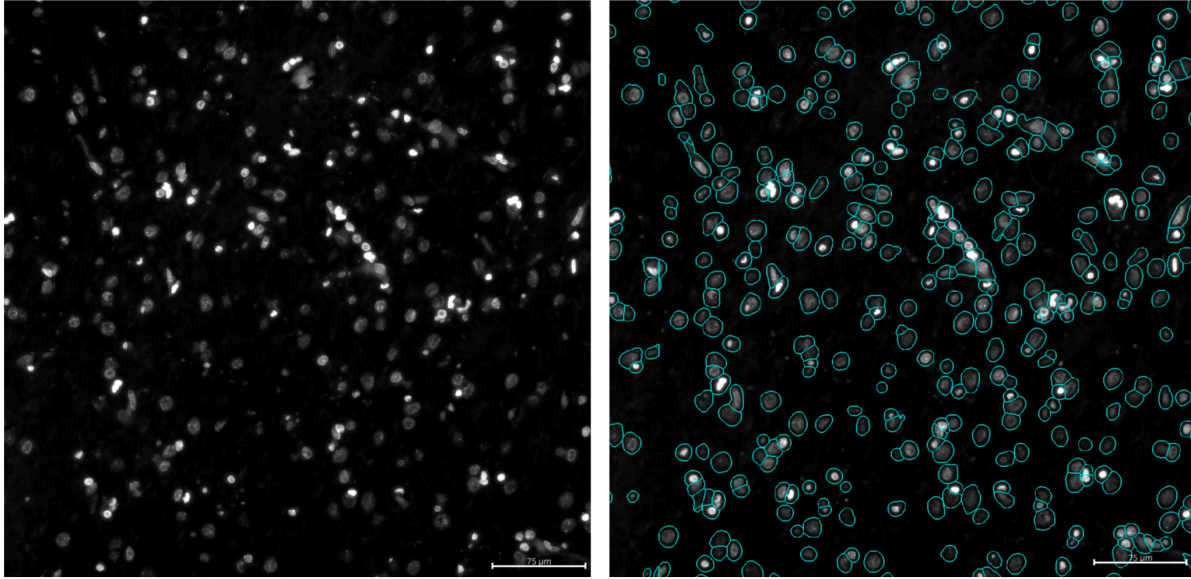


Figure 4.1. Cell segmentation boundaries following custom re-segmentation. DAPI stain in grey and updated cell segmentation ROIs shown in cyan.

Data ingest, quality control, and filtering

To maintain consistency with the other modalities (in particular pTau-seq) analysis, the raw .csv files were used as input and processed with a python-based workflow. Five input files (transcripts.csv, expression.csv, transcript.csv, cellmetadata.csv, and FOV.csv) were read into python as pandas dataframes before being merged and converted into anndata objects using cell_id as the common key. Spatial coordinates were converted from pixels to microns. All FOVs were assigned to the associated donor.

Across the slide and donors, the number of genes per cell (n_genes) correlated well with the total number of UMI counts (total_counts) per cell. When visualized in a scatterplot, a smooth upward curve and consistency across potential batch effects (e.g., donor ID). There is no vertical or horizontal spreading along the x or y-axis, confirming few cells with low overall genes and reads. In most instances, more than 2000 unique genes are detected per cell, which, taken together, is indicative of an exceptionally high quality spatial transcriptomics dataset. Across the dataset, the minimum number of per cell UMI counts was 50 (set cutoff), the median was 560, and the max was 19,792. As it relates to traditional scRNA-seq, there were 6115 cells with > 2,000 counts, which is approximately 12% of the dataset (Figure 4.2).

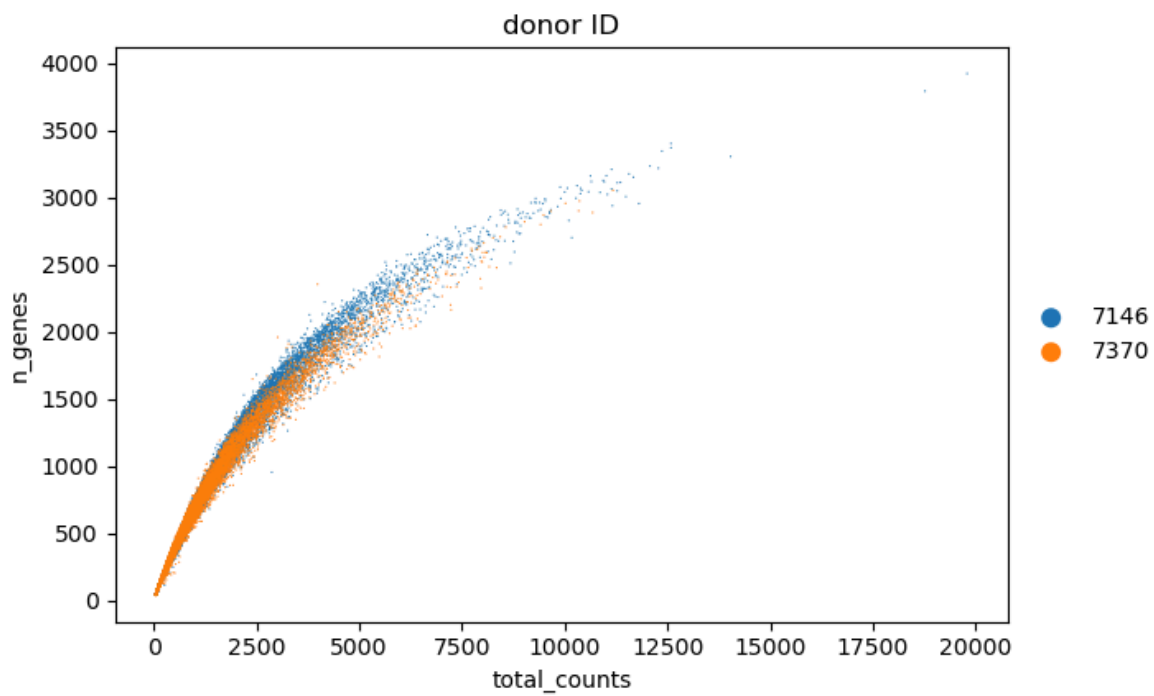
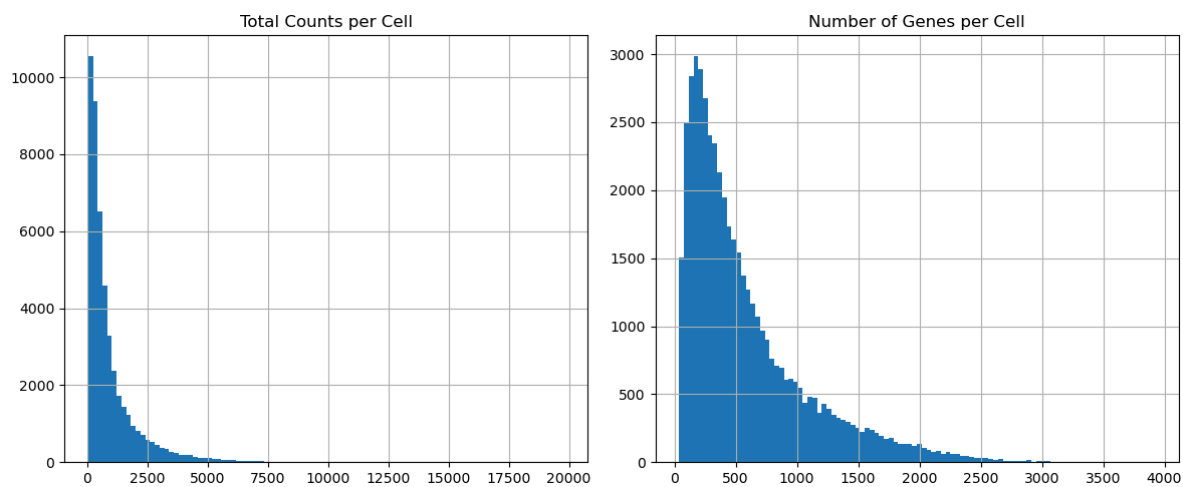
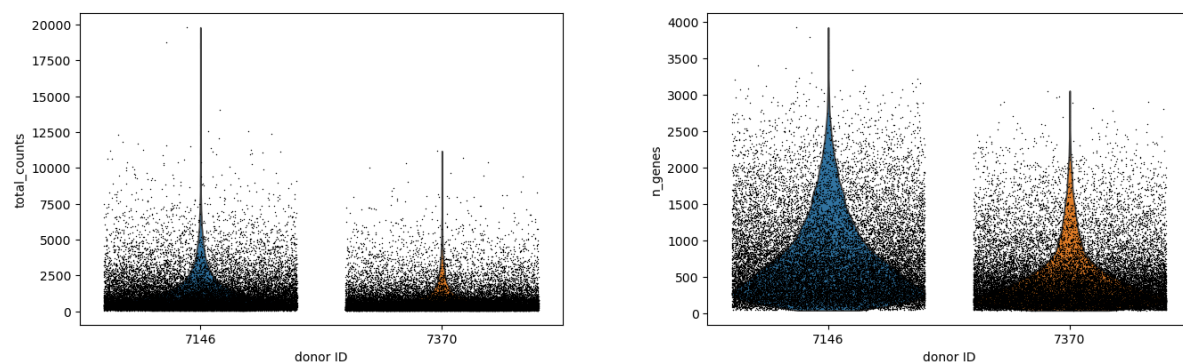


Figure 4.2. Consistent high-quality metrics across donors in proof-of-concept dataset. The top panel shows a violin plot representation of the number of genes per cell and UMIs per cell, which can also be visualized by the histogram in the middle panel. The lower plot illustrates the relationship between the number of genes and total UMI counts per cell.

Negative probes are important negative controls in the CosMx assay. They are synthetic sequences that are not known to bind to any known biological RNA sequences. They are a means of estimating background noise and assessing the limit of detection of the assay. The negative probe values should be low across the dataset and spatially uniform (not clustered in a specific area of the tissue). According to NanoString's guidelines, the negative probe count should be below five, ideally between 1-2 per object area. Spread is expected up to approximately 20 counts, though this occurs very rarely in the dataset (Figure 4.3).

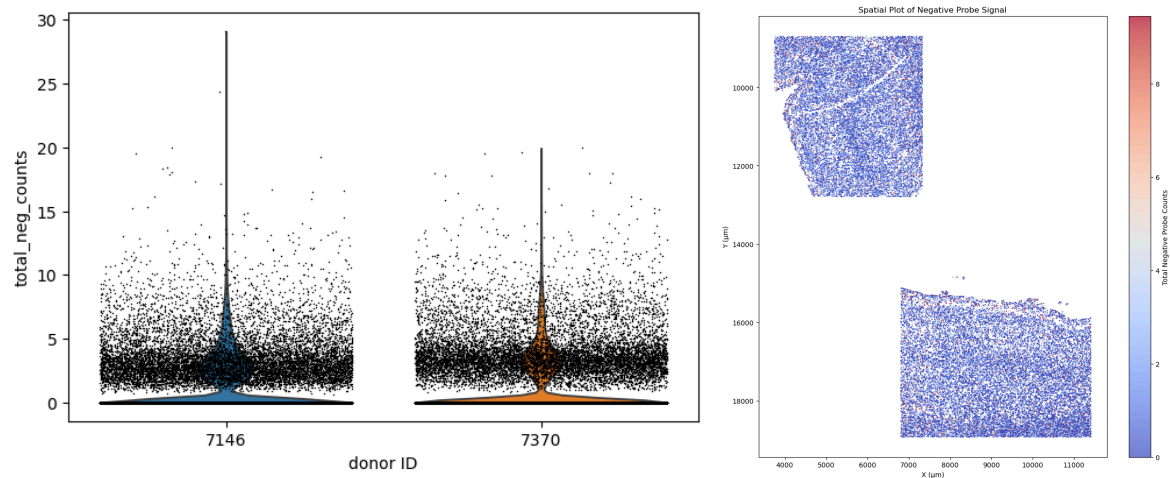


Figure 4.3. Negative probe detection indicates low background across tissue.

Standard QC metrics and filtering cutoffs were applied to the data. A lower limit of 50 UMI counts per cell was used to remove low quality cells. To remove genes that would be unreliable in downstream analysis, any gene that was present in fewer than three total cells in the dataset were removed. In both donors, the spatial plotting of the total count data shows that lower quality cells tend to be located along the periphery of the tissue (in this case, along cortical layer 1), which is not unexpected given prior meninges removal and the increased wear on that portion of the tissue during tissue processing. There does not appear to be any other spatial correlations with lower quality cells (Figure 4.4).

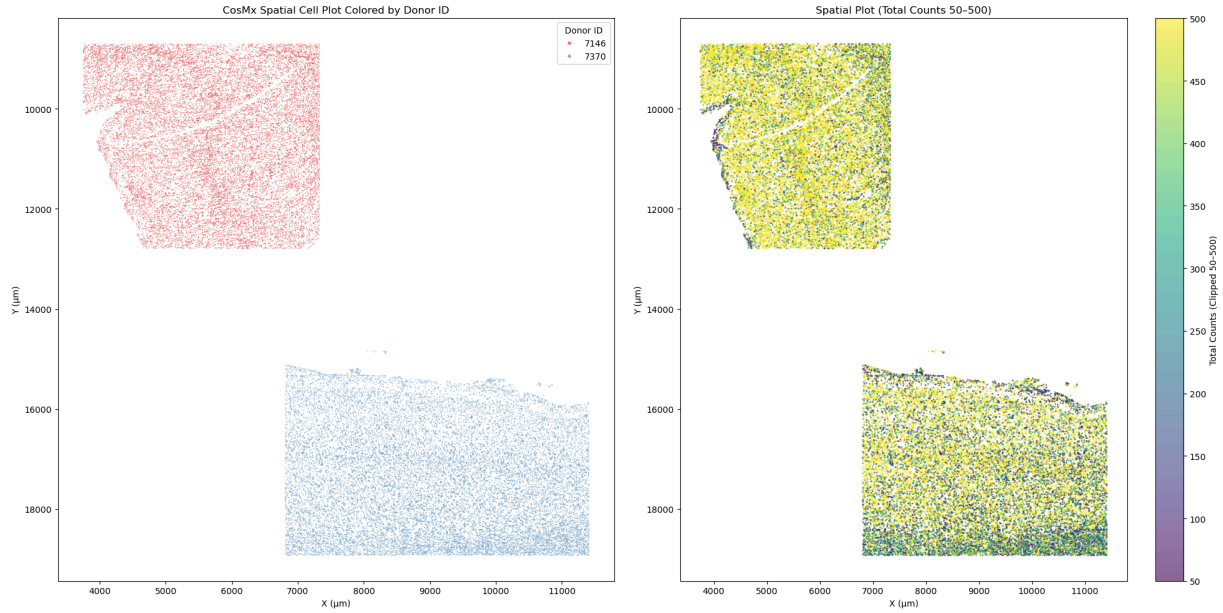


Figure 4.4. Spatial assessment of low-quality cells.

Cell-type mapping with MapMyCells

Following QC, the dataset of two donors (UWA 7370 and 7146) contained 25,918 and 22,580 cells, respectively. To generate cell-type labels that were consistent with the pTau-seq dataset, but without resource intensive dataset integration described in Chapter 3, were mapped to the same SEA-AD reference taxonomy, but using the publicly available Allen Institute tool MapMyCells. The deep learning (scVI-based) algorithm was selected for alignment across modalities. MapMyCells takes in a raw cell x gene matrix (.csv) and outputs a table with the unique cell ID for each cell, predicted cell type label, and corresponding mapping score. High mapping confidence (softmax probability from scANVI) was maintained across both donors down to the supertype level³. Of the approximately 50,000 total cells, only 5067 supertypes had a softmax probability below 60% (note that their subclass softmax probabilities above this threshold). Supertypes with a lower mapping confidence (<60% softmax probability) are biased towards lower abundance supertypes, which is not unexpected (Figure 4.5).

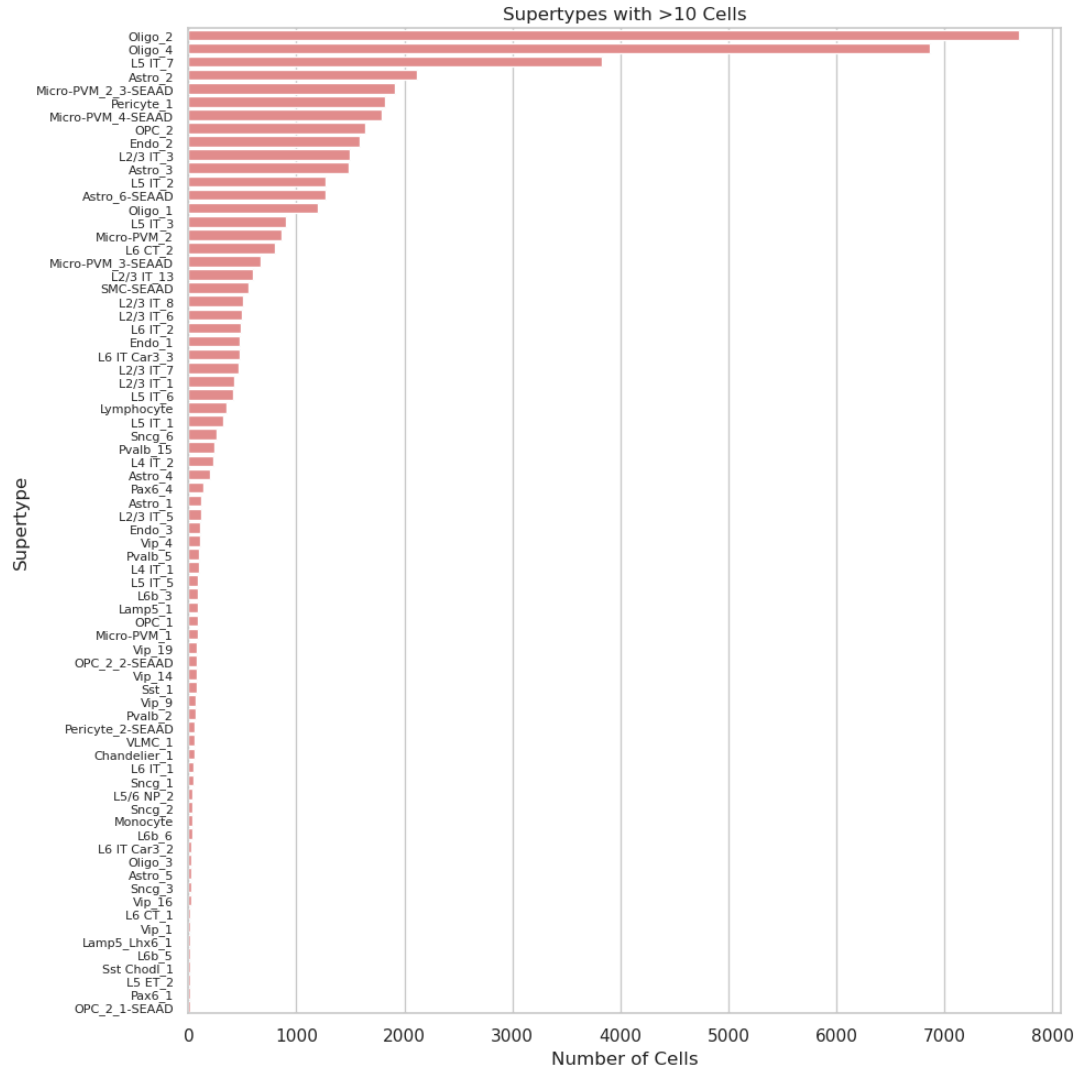


Figure 4.6 Ranked list of low supertypes (<10 cells).

Results

Cell type composition and spatial distribution

MapMyCells deep learning (scVI/scANVI-based) module was used to map single cells from the CosMx data cells to the supertype level. Subclass-level mapping level was robust across the dataset and showed clearly delineated laminar distribution when mapped to spatial coordinates (Figure 4.7). Localizing the transcriptomically-defined cortical layer subclasses to

their spatial coordinates is a strong validation that the cell type mapping is highly accurate. All the expected subclasses were identified at sufficient abundance levels (>10 cells) with the exception of Chandelier, Sst Chodl, L5 ET, L5/6 NP, Lamp5 Lhx6, which represent known sparse cell types. With the exception of Sst, these subclasses are generally not associated with AD-related cellular vulnerability.

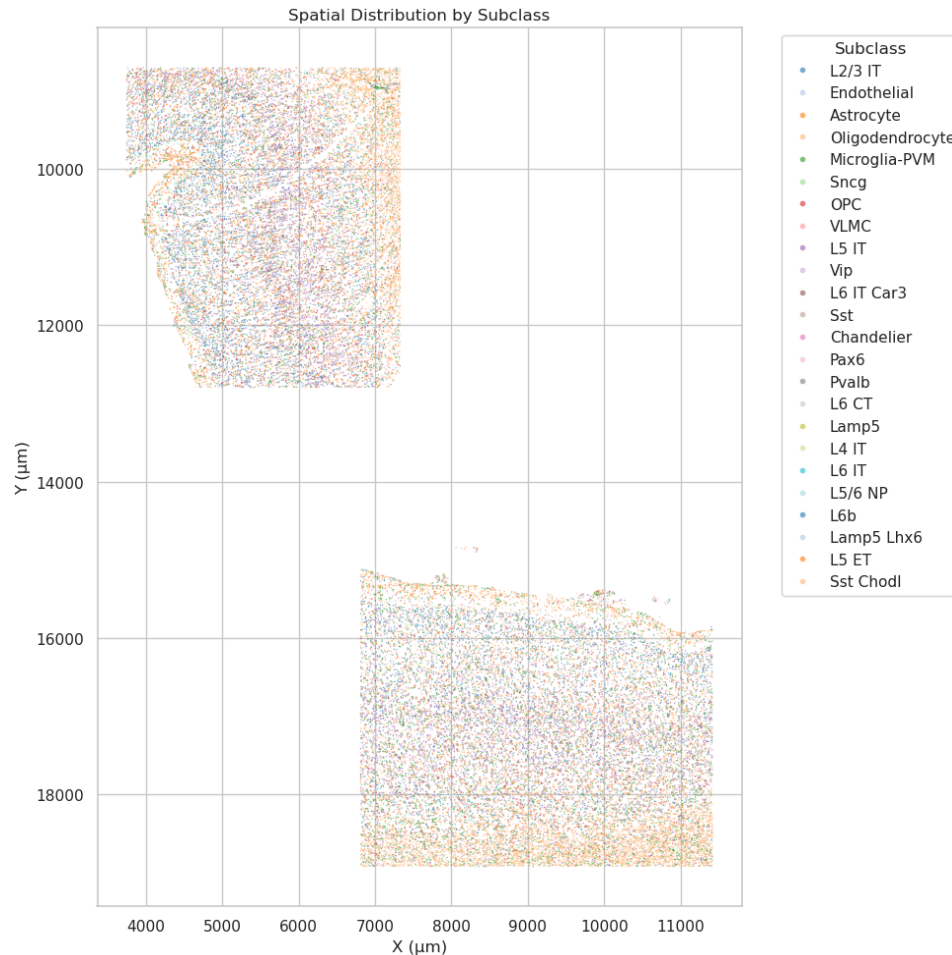


Figure 4.7. Spatial distribution of subclass-level laminar structure.

Mapping highlighted that the CosMx data contains substantially higher oligodendrocyte populations than in pTau-seq. This may be due to several contributing factors. Unlike the pTau-seq dataset, this dataset intentionally captures the adjacent white matter just beneath the cortical column, which is primarily composed of oligodendrocytes. Additionally, the pTau-seq dataset was intentionally enriched for neuronal cells (MAP2+ fractions were collected), so non-neuronal cells are intentionally depleted. This likely explains the increased relative amount

of microglia and astrocytes in the dataset also. The marked increase in oligodendrocytes, however, may involve a more complex explanation. In human brain tissue, disproportionate mapping to oligodendrocytes can also be an artifact of lower quality cells and incorrect cell type assignments. This dataset was mapped using the tool MapMyCells, which does not employ the complete algorithm that is run when iterative_scANVI is performed manually, as it was with the pTau-seq dataset. As discussed in Chapter 3, mapping across more divergent modalities (e.g., dissociated cells for snRNA-seq versus in-tact tissue for spatial transcriptomics) may result in decreased performance overall. Due to the proof-of-concept nature of this data (one slide, two donors), the sample size is not sufficient to disentangle the underlying factor driving the shift in cell type abundances towards oligodendrocytes (Figure 4.8).

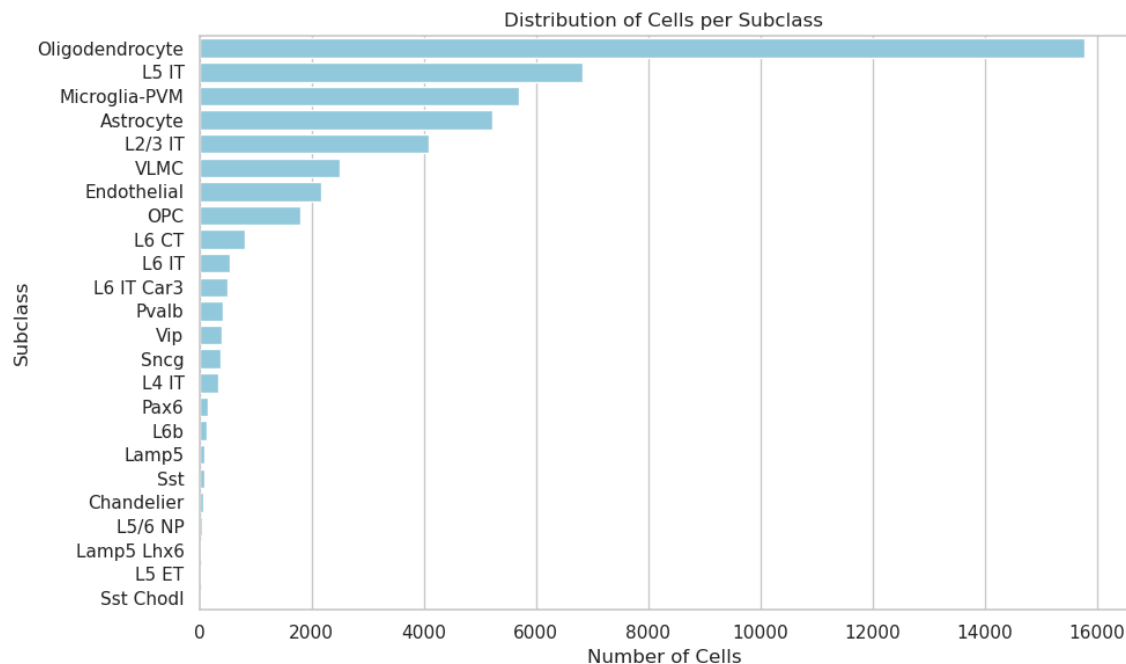


Figure 4.8. Ranked list of subclass proportions.

Interneuron types, Pvalb and Sst are sparse, but represented in the dataset at the subclass level and are spatially distributed throughout the cortical column, as expected, and track closely with high GAD2 expressing cells. However, in this pilot dataset consisting of only two donors, each with approximately 3mm x 6mm rectangular tissue area profiled, the number of cells underpowered to detect lower abundance supertypes. The vulnerable inhibitory interneuron supertypes are relatively low abundance types, so detecting them was expected to be a challenge

based on the outputs of similar sized tissue samples profiled via MERSCOPE in SEA-AD. Across both donors on the slide, a total of 82 Sst cells were present. However, these 82 were not equally represented across Sst supertypes, such that only Sst_1, Sst_2, and Sst_25 were present (Figure 4.9). The Pvalb population was slightly more abundant with a total of 414 cells with ten flagged as pTau-positive with the most stringent threshold (top 10%); this fraction is biologically plausible given the relative sparing of Pvalb neurons to pTau accumulation.

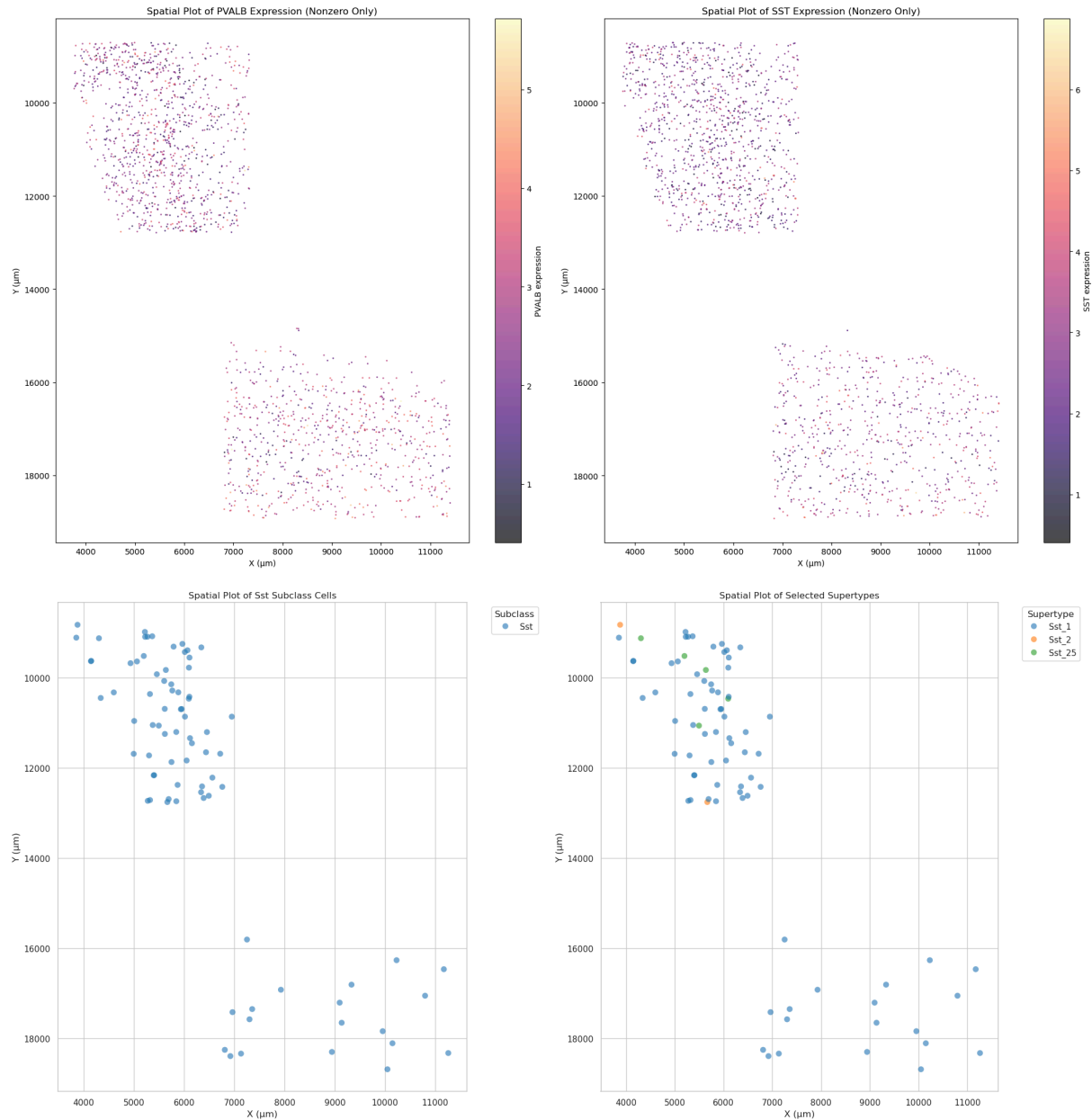


Figure 4.9. Spatial distribution of inhibitory interneuron subclass-level marker genes (Pvalb and Sst) and subclass and supertype. The Sst subclass consists of only 81 cells total across two donors.

Visualization of selected neuronal and non-neuronal subclasses also highlight the tissue boundary with high glial (typically astrocytic) involvement and the white matter (oligodendrocytes). Isolating cortical layer specific subclasses and visualizing along spatial coordinates, shows an even clear laminar distribution of transcriptomically-defined cortical layer types with clear boundaries (Figure 4.10).



Figure 4.10. Spatial distribution of selected neuronal and non-neuronal subclasses.

Latent representation of spatial transcriptomics data

When represented in low dimensional (UMAP) space, marked separation of class and subclass is apparent, with decreasing separation in UMAP space with increasing cell type granularity. Lower confidence cell type cells (softmax probability) are also associated with increasing cell type granularity with values close to 1 (complete confidence) at the class level and many more values around the cutoff range of 0.6 to 0.7 at the supertype level (Figure 4.11).

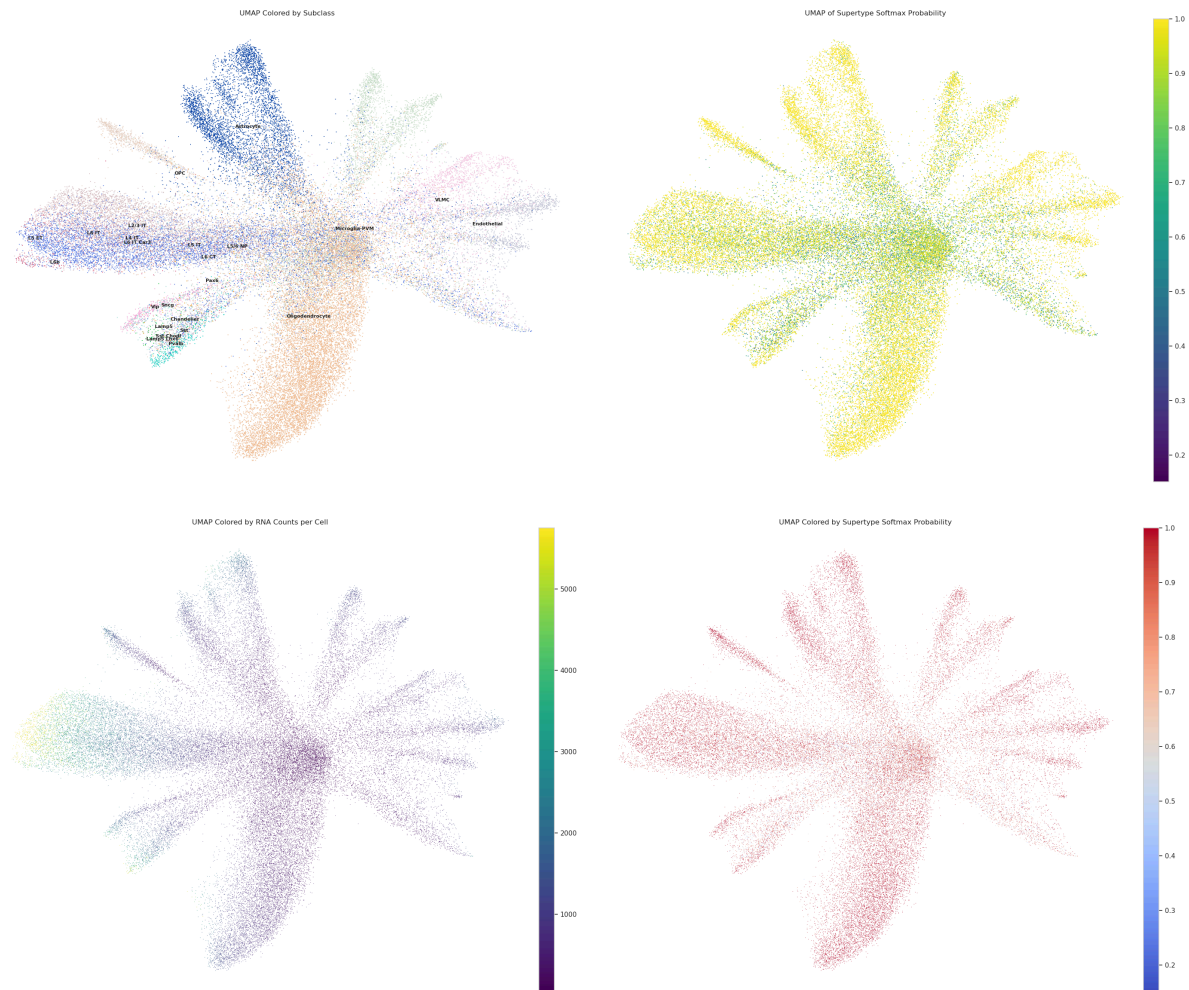


Figure 4.11. UMAP representation of subclasses with mapping confidence and quality metrics.

Spatial distribution of pTau-positive cells and A β plaques in the context of subclasses

The ultimate goal of the combined custom protein and spatial transcriptomics assay was to disentangle how selective vulnerability of transcriptomically and spatially defined cell types is related to pathological state. In this custom CosMx panel, antibodies against both pTau and A β were added to be able to assess intracellular pTau aggregates and extracellular A β plaques in the context of cellular vulnerability. On the CosMx platform, protein data is generated from proprietary fluorescently labeled antibodies. This fluorescence is read out in the first stage of the CosMx workflow, upstream of the cyclic imaging to quantify RNAs. Each cell is associated with a mean pTau intensity (across all bins in a cell boundary).

Only the pTau data can be assessed with the direct readouts because the CosMx pipeline only retains measurements from within (inside) cell boundaries. Quantifications for A β levels are only captured when they overlap with a cell boundary, which would not accurately capture the amount and distribution of A β plaques, extracellular and generally not overlapping with living cells. As such, custom image analysis pipelines will need to be developed to extract A β plaque quantifications.

Given the relatively high mean intensity for most subclasses, an initial stringent cutoff (top 10% mean fluorescence intensity values(cells)) was used to define pTau positive cells. This cutoff yielded 43,296 pTau-negative cells and 4,807 pTau-positive cells, yielding an overall pTau-positivity rate of 11%. Of the 81 detected Sst cells, only one was called pTau positive according to the more stringent criteria. Despite the flexibility of thresholding in the CosMx data (quantitative and mutable) relative to pTau-seq (binary and static), selecting the relevant threshold value for deeming a cell pTau-positive still remains a challenge. The addition of more specific pTau antibodies that label pre-tangles and mature tangles can help disentangle this thresholding problem.

A powerful advantage of this approach is the ability to utilize the 6,000 profiled genes to transcriptomically identify cell classes, subclasses, and in many cases supertypes, while incorporating pTau levels, all while retaining spatial context. The value of this is highlighted in the spatial plot, which shows each transcriptomically-defined subclass in their respective laminar location in the MTG, with the pTau status indicated by a red circle (Figure 4.12). Understanding spatial proximity of affected cells highlights potential local circuit interactions and microenvironmental contributions. Qualitatively, the L2/3 enrichment for pTau bearing cells can be seen. More donors are needed to begin quantifying relevant spatial relationships.

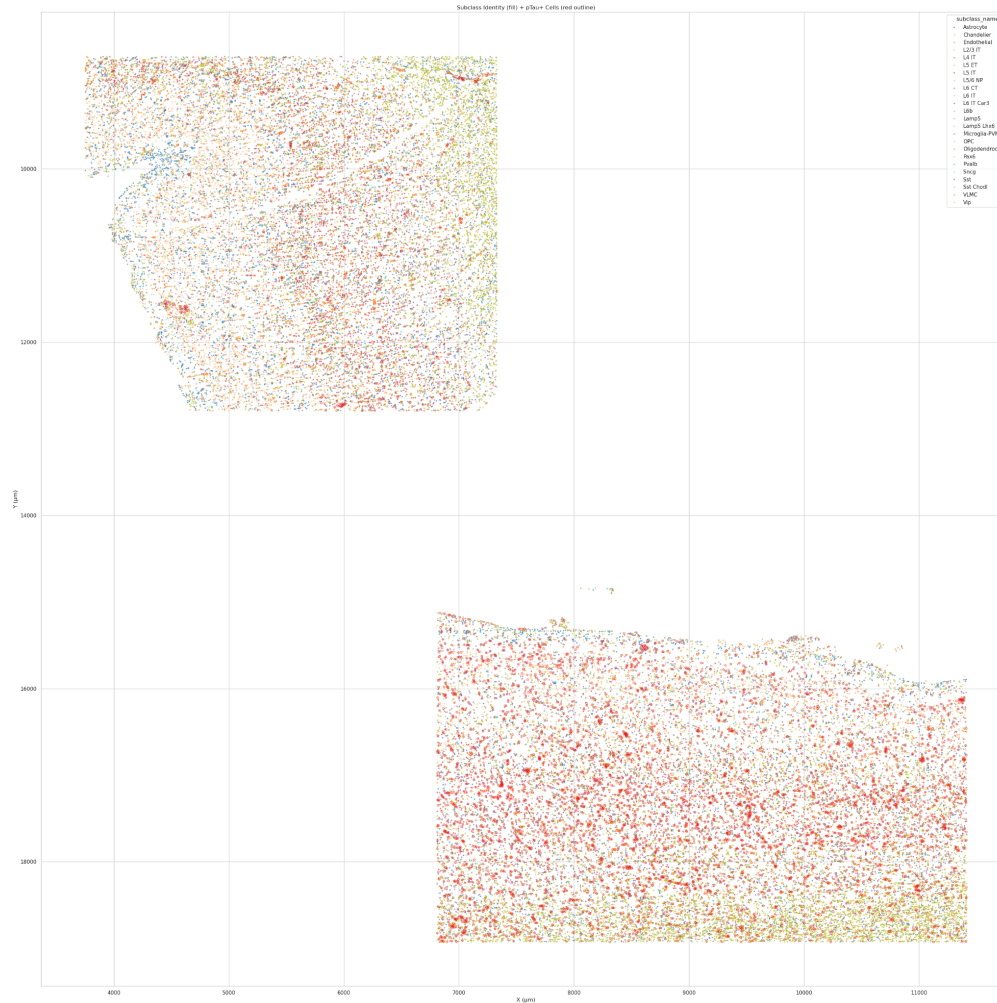


Figure 4.12. Spatial plot of subclasses highlighted by their pTau status. pTau-positive cells are indicated by a red circle around the cell boundary.

Differential expression analysis was performed on the pTau-seq dataset at the subclass and supertype-level. While the spatial transcriptomics pilot data is not sufficiently powered to perform differential expression, the top genes can be visualized in relation to the vulnerable cell types in spatial context, which may provide valuable insight into whether spatial location uniquely contributes to transcriptomic dysregulation in pTau positive cells (Figure 4.13).

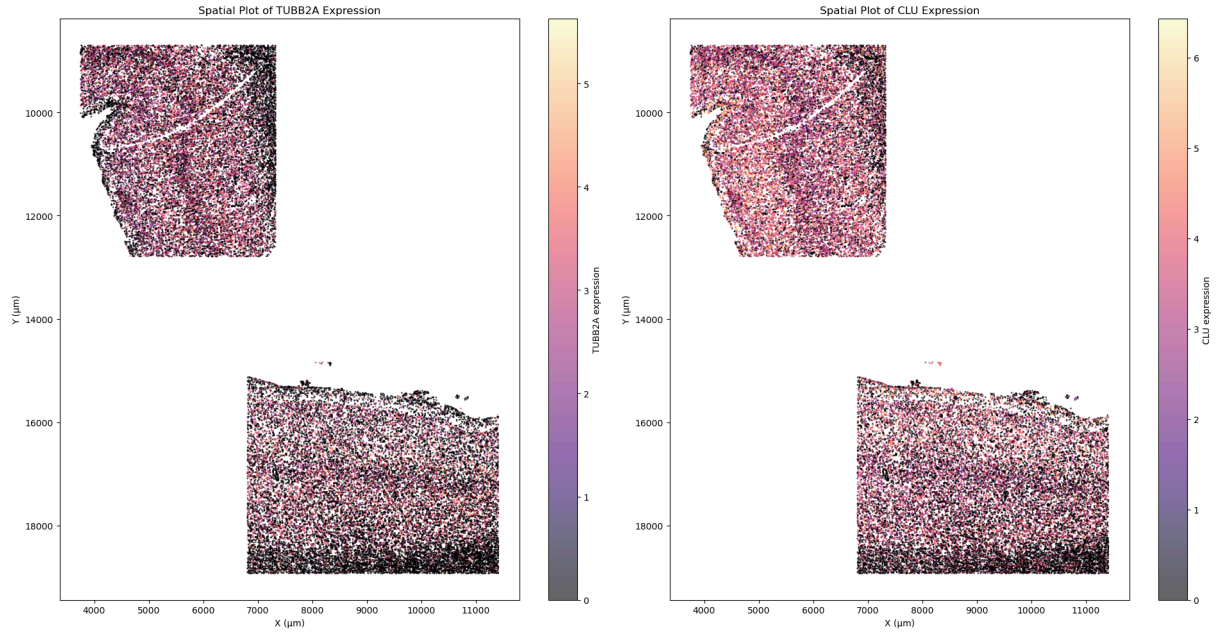


Figure 4.13. AD-related and DE genes visualized in spatial context.

The pronounced vulnerability of neurons in superficial cortical layers, such as Sst inhibitory interneurons and L2/3 intratelencephalic projecting excitatory neurons cells, suggest the microcircuitry in these layers may be dysregulated early in disease. The SEA-AD study leveraged the quantitative neuropathology-based continuous pseudoprogession score to order cellular abundance changes along a pseudotemporal scale, which indicated that loss of inhibitory interneurons, particularly Sst supertypes occurs at the earliest stage of disease progression, well before the observed vulnerability of L2/3 excitatory neurons. This suggests that early loss of Sst results in loss inhibition and dysregulation of L2/3 IT excitatory cells. This could result in hyperexcitability, which paired with pTau vulnerability, pushes these cells toward an ultimate fate of cell death⁴. Spatial context is needed to disentangle these potentially synergistic contributions to cellular vulnerability in the context of AD.

Discussion

In the pilot dataset described in this chapter, tissue from two donors on the same slide were profiled for protein expression (GFAP, AT8, and 6e10) and a panel of 6,000 genes. The profiled tissue area was 15-20mm², which captures approximately 25,000 cells per donor. Cells were mapped to the SEA-AD cellular taxonomy down to the supertype-level, though substantial

dropout was observed beyond the subclass level. Importantly, all cells in this dataset could be assigned a pTau pathology status based on the quantitative readout of (mean) fluorescent intensity from each cell. This direct quantitative readout has the potential to order cells along a continuum based on their pTau intensity, though establishing this requires a larger dataset. Further, it allows for precise and modifiable thresholding, a valuable feature to aid in establishing a relevant baseline for pTau levels across cell types.

Spatial multimonic approaches, such as custom pathological peptide and 6K gene panel developed for NanoString CosMx, used in this pilot data exploration provide exciting proof-of-concept results that highlight additional insight that can be gained from spatial context. First, these data are particularly well suited to establish ground truth cell-type distributions and abundances. While many transcriptomically-defined cell types can be associated with a specific anatomical or laminar location within the brain, several types with broader brain-wide distributions, such as inhibitory interneurons, require spatial resolution to determine location. This added value was highlighted in the SEA-AD study which confirmed that specifically the vulnerable Sst supertypes are localized to cortical L2/3, directly implicating them in relevant microcircuitry and association with adjacent vulnerable excitatory neuron types.

Given the subcellular resolution of modern spatial omic assays, these approaches can be used beyond simple anatomical localization of cell types. These data can be leveraged to ascertain precise location of potentially interacting cell types and combined with changes in gene expression (e.g., pathways involving synaptic dysregulation), can provide insight into potential cell-to-cell interactions and microcircuit dysregulation. While confirming direct functional connectivity requires higher resolution approaches, such as electron microscopy or electrophysiology with morphological reconstructions, the spatial approach brings us one step closer to this understanding compared to snRNA-seq assays which lose this relevant context.

Given this is a proof-of-concept pilot study, no quantitative assessments can be made and therefore no conclusions can be drawn from this dataset alone. However, this dataset exemplifies how exceptionally high quality spatial omic data can lead to highly accurate cell type mapping, down to the supertype level - provided a sufficiently large section of tissue is profiled. This

observation can be applied practically to future datasets, as rather than trying to combine three donors tissues per slide and setting smaller profile areas, combining only two donors per slide and doubling the cortical tissue area profiled is essential to increasing the captured numbers of less abundant cell types and beginning to assess variations in pathological state. It also highlights the value of obtaining continuous quantitative reads of relevant protein expression levels for dynamic thresholding and establishing a continuous metric of pTau pathology.

Chapter 5. CLOSING REMARKS

The work discussed in this dissertation confirmed the core hypothesis that a subset of selectively vulnerable populations of cells preferentially accumulate pTau and can be identified transcriptomically based on changes to relative abundance and cell state associated with disease progression. Leveraging quantitative neuropathology data to continuously order donors according to their aggregate pathological state revealed precisely which transcriptomically-defined cell types are lost across distinct stages of AD pathological progression. These disease-associated relative abundance changes, particularly in low abundance cell types, are notoriously difficult to resolve due to sample size requirements. Leveraging a continuous pathology score, rather than discrete stages, to test for subtle changes, bolstered statistical power aiding in the identification of a core set of highly granular cell types that are preferentially vulnerable to AD-associated cell loss. Further, the pseudotemporal ordering of these cellular changes provide unprecedented insight into the early versus late patterns of cell loss in AD - resulting in the novel finding that Sst interneurons are preferentially lost in early stages of AD pathophysiology, before the onset of widespread pathological burden in the brain. The work of this study and others, are beginning to set the groundwork for a new era of AD research that approaches the disease from a cellular perspective.

This is not to say that protein pathology, a hallmark of AD, should be overlooked. In fact, this is a critical feature missing from the snRNA-seq data assessed in SEA-AD. In standard snRNA-seq datasets, the potential role of pathological protein aggregation in conferring cellular vulnerability cannot be assessed directly, only indirectly through approaches such as pathway analysis. To address this, pTau-seq was performed on a subset of the donors in the SEA-AD study. This approach involves gentle isolation of cell somas, rather than nuclei, to retain intracellular pTau aggregates that can be fluorescently labeled and isolated with FACS. This yields a binary pTau status, a critical added layer of information, for every cell in the RNA sequencing dataset that maintains the same quality and depth as a standard snRNA-seq dataset. This strategy enables the direct assessment of which highly granular cell types preferentially accumulate pTau and how these patterns align or diverge from the cell types the exhibited relative abundance changes in SEA-AD.

The results highlight that the role of pTau in cell-type specific vulnerability in the AD brain is more nuanced than previously appreciated. While there is tight overlap between the observed

vulnerability to pTau accumulation and vulnerability to cell loss in well-established vulnerable cells in AD, such as L2/3 IT and L5 IT excitatory neurons, there are several intriguing examples of divergence. Namely, the striking vulnerability of deep L6b neurons to pTau accumulation that appears to impact all supertypes within the subclass, despite SEA-AD results suggesting that this cell type not affected in terms of relative abundance changes with increasing CPS. Additionally, pTau-seq identified L5 IT and L5 ET types as vulnerable to pTau accumulation, despite the observation from SEA-AD that they do not exhibit relative abundance changes with increasing CPS. Conversely, there are several examples of supertypes vulnerable to cell loss that do not exhibit significant pTau accumulation, such as select Lamp5, Vip, Sncg, and Pvalb supertypes, however they do trend towards increased pTau fractions, suggesting the changes are more subtle or these types remain underpowered due to their low abundance.

Spatial multiomic approaches have the potential to further aid in disentangling pathology and cellular vulnerability by adding a key missing piece of information, the cells spatial context. Cells in the brain are highly organized into distinct anatomical and topographical patterns. Understanding the local environment of vulnerable cells in the brain will provide important additional insight into factors contributing to cellular vulnerability. The pilot NanoString CosMx dataset discussed here provides an exciting example of how these data can be used to strengthen biological insight. For example, Sst interneurons are dispersed throughout the context, but spatial omics can be used to highlight that the vulnerable Sst supertypes have a very relevant physical association - they are preferentially localized cortical layers L2/3, where pathological burden is generally higher and where the selectively vulnerable L2/3 IT neurons reside. This added context illuminates the possibility that local circuit dysregulation in L2/3, implicating Sst and L2/3 IT cells, may be underlying early pathophysiological changes in AD.

Given that assessing cellular protein pathology in the context of transcriptomic data is a new and relatively uncharted area, it is important to highlight the need for additional studies aimed at establishing what the basal levels of pTau are expected to be in these highly resolved cell types, in the absence of pathology. Traditional approaches, such as immunohistochemistry, suggested that pathological accumulation falls into one of three discrete groups (e.g., no immunoreactivity, pre-tangle, or mature NFT); However, the more sensitive approaches of FACS (and likely oligo-tagged antibody approaches such as CITE-seq), clearly demonstrate that the levels of pTau are

continuous and adhere to different dynamics in different subclasses of cells. Concretely establishing these cell-type specific baselines by including more control tissue and increasing the sample size (both donor number and cell number) is critical to move these proof-of-concept studies to robust datasets that can be used to leverage novel biological insights on the pathophysiology of AD.

5.1. Future directions for quantitative neuropathology and multiomics in AD

To begin disentangling pathology and cellular vulnerability, pathological proteins and RNA must be assessed in the same cell, ideally while retaining spatial context. The results of this work showcase the exciting opportunity of these new technologies, if applied to a broader donor base with deeper sampling of cells per donor. These approaches must be applied at scale in order to discern the more subtle, yet highly biologically relevant, changes in low abundance cell types. Oligo-based methods of single-cell protein quantification, such as CITE-seq, enable higher throughput, higher plex proteomics, and provide quantitative pTau readouts, making this the preferred approach for future work in this area.

First, efforts must be made to establish the expected baseline levels of pTau are expected to be across brain cell types, as this sets an important biological baseline and reference for statistical tests. The quantitative neuropathology data in SEA-AD, highlighted the substantial variation in pTau levels within a single discrete Braak stage, serving as an early indicator that assessment of pTau levels at higher resolution is much more nuanced. More sensitive readouts for pTau, such as FACS or oligo-tagging, further support the updated understanding that most cells express moderate basal levels of pTau, potentially even in a non-pathological context. Given that this unexpected observation has an outsized impact on how the results of such studies can be interpreted, it is essential to intentionally and systematically address this unknown by increasing sample size and extending these approaches to assess pTau-levels in non-AD brains, rather than just those with moderate to high pTau pathology based on Braak Staging.

Given the growing efforts to develop pTau-based therapies for AD, it is critical not only to consider the structure, accumulation, and brain-wide distribution pTau itself but critically, which cell types are preferentially affected by pTau-driven pathological processes. This work

serves as one of the first datasets in the field linking protein pathology to transcriptomically-defined cell types, and the first ever in the human MTG. The nuanced relationship between the vulnerability to pTau and vulnerability to cell death highlight that this variable cell type selectivity is not something to be overlooked.

Major advancements in the development of cell-type specific enhancer-AAVs which can selectively target and deliver molecular cargo to any cell type in the brain, hold great promise in this area. The catalog of validated enhancer-AAVs is ever expanding, with the goal to ultimately provide coverage for all brain cell types. Unlike targeted therapies for A β , which aggregates extracellularly, tau protein and its aberrant hyperphosphorylation and aggregation occurs exclusively intracellularly. In its non-pathologic form, tau performs several known roles within the cell, traditionally being thought of as a microtubule stabilizing protein and more recently being appreciated as an RNA binding protein. Given the involvement of tau in such central roles to cellular function, developing a cell-type specific pTau-targeting therapy is essential to avoid toxicity and unwanted side effects. With the massive efforts towards developing a brain-wide catalog of cell type specific enhancer-AAVs, targeting of pTau in a cell-specific manner may soon become a reality. However, before systemic development of a cell-type specific pTau-targeting therapy can begin, a comprehensive map of which cell types bear pathological pTau burden must be generated. This work serves as a first step in this effort that will ultimately require more sensitive measurements for protein accumulation and vastly larger sample sizes.

While there is still much work to be done, the topics discussed in this dissertation serve as an exciting window into what lies ahead with regard to understanding cell-type specific selective vulnerability in AD. Taken together, this work highlights the nuanced role that pTau accumulation plays in selective vulnerability; imparting variable outcomes down to the supertype level. While a core set of well-established cell types are vulnerable both to pTau accumulation and cell loss, some cell types appear to be vulnerable to cell loss without apparent pTau accumulation, while others accumulate pTau without evidence of cell loss. As efforts are made to map every cell type in the human brain in health and disease, it would be prudent to consider employing an approach such as CITE-seq, as this would add valuable quantitative single-cell

protein information while retaining exceptional quality transcriptomic data in building these foundational atlases.

References

CHAPTER 1

1. Carlos, Arenn Faye, Tino Emanuele Poloni, Valentina Medici, Maia Chikhladze, Antonio Guaita, and Mauro Ceroni. "From brain collections to modern brain banks: A historical perspective." *Alzheimer's & Dementia: Translational Research & Clinical Interventions* 5 (2019): 52-60.
2. Tourtellotte, W. W., I. P. Rosario, A. Conrad, and K. Syndulko. "Human neuro-specimen banking 1961-1992. The National Neurological Research Specimen Bank (a donor program of pre-and post-mortem tissues and cerebrospinal fluid/blood; and a collection of cryopreserved human neurological specimens for neuroscientists)." *Journal of neural transmission. Supplementum* 39 (1993): 5-15.
3. Danner, Benjamin, Angelique D. Gonzalez, William Cole Corbett, Mohammad Alhneif, Shahroo Etemadmoghadam, Julie Parker-Garza, and Margaret E. Flanagan. "Brain banking in the United States and Europe: Importance, challenges, and future trends." *Journal of Neuropathology & Experimental Neurology* 83, no. 4 (2024): 219-229.
4. Vonsattel, J. P., H. Aizawa, P. Ge, M. DiFiglia, A. C. McKee, M. MacDonald, J. F. Gusella, G. B. Landwehrmeyer, E. D. Bird, and E. P. Richardson Jr. 1995. "An Improved Approach to Prepare Human Brains for Research." *Journal of Neuropathology & Experimental Neurology* 54 (1): 42–56
5. Bussolati, Gianni, Laura Annaratone, Enzo Medico, Giuseppe D'Armento, and Anna Sapino. "Formalin fixation at low temperature better preserves nucleic acid integrity." *PLoS One* 6, no. 6 (2011): e21043.
6. Bussolati, Gianni, Laura Annaratone, Enzo Medico, Giuseppe D'Armento, and Anna Sapino. "Formalin fixation at low temperature better preserves nucleic acid integrity." *PLoS One* 6, no. 6 (2011): e21043.
7. Allen Institute for Brain Science. 2021. "Human Tissue Slicing and Dissections for Nuclear Isolations V.3." protocols.io. <https://doi.org/10.17504/protocols.io.bq6ymzfw>.
8. Latimer, Caitlin S., Erica J. Melief, Jeanelle Ariza-Torres, Kim Howard, Amanda R. Keen, Lisa M. Keene, Aimee M. Schantz et al. "Protocol for the systematic fixation, circuit-based sampling, and qualitative and quantitative neuropathological analysis of human brain tissue." In *Alzheimer's Disease: Methods and Protocols*, pp. 3-30. New York, NY: Springer US, 2022.

9. NIH NeuroBioBank. 2021. "Best Practices for Human Brain Collection, Processing, and Distribution." National Institutes of Health. <https://neurobiobank.nih.gov/about-best-practices>.
10. Freund, Michelle, Anna Taylor, Cathy Ng, and A. Roger Little. "The NIH NeuroBioBank: creating opportunities for human brain research." In *Handbook of clinical neurology*, vol. 150, pp. 41-48. Elsevier, 2018.
11. Gabitto, Mariano I., Kyle J. Travaglini, Victoria M. Rachleff, Eitan S. Kaplan, Brian Long, Jeanelle Ariza, Yi Ding et al. "Integrated multimodal cell atlas of Alzheimer's disease." *Nature Neuroscience* 27, no. 12 (2024): 2366-2383.
12. McKenzie, Andrew T., Emma L. Thorn, Oge Nnadi, Borys Wróbel, Emil Kendziorra, Kurt Farrell, and John F. Crary. "Cryopreservation of brain cell structure: a review." *Free Neuropathology* 5 (2024): 35.
13. Rosene, Douglas L., and Kenneth J. Rhodes. "Cryoprotection and freezing methods to control ice crystal artifact in frozen sections of fixed and unfixed brain tissue." In *Methods in Neurosciences*, vol. 3, pp. 360-385. Academic Press, 1990.
14. Kumarasami, Ramdayalan, Richa Verma, Karthika Pandurangan, Jivitha Jyothi Ramesh, Sathish Pandidurai, Stephen Savoia, Jaikishan Jayakumar et al. "A technology platform for standardized cryoprotection and freezing of large-volume brain tissues for high-resolution histology." *Frontiers in Neuroanatomy* 17 (2023): 1292655.
15. Bacon, Eliza R., Kena Ihle, Peter P. Lee, and James R. Waisman. "Building a rapid autopsy program—a step-by-step logistics guide." *Translational Medicine Communications* 5, no. 1 (2020): 23.
16. Hodge, Rebecca D., Trygve E. Bakken, Jeremy A. Miller, Kimberly A. Smith, Eliza R. Barkan, Lucas T. Graybuck, Jennie L. Close et al. "Conserved cell types with divergent features in human versus mouse cortex." *Nature* 573, no. 7772 (2019): 61-68.
17. Siletti, Kimberly, Rebecca Hodge, Alejandro Mossi Albiach, Ka Wai Lee, Song-Lin Ding, Lijuan Hu, Peter Lönnerberg et al. "Transcriptomic diversity of cell types across the adult human brain." *Science* 382, no. 6667 (2023): eadd7046.
18. Mathys, Hansruedi, Carles A. Boix, Leyla Anne Akay, Ziting Xia, Jose Davila-Velderrain, Ayesha P. Ng, Xueqiao Jiang et al. "Single-cell multiregion dissection of Alzheimer's disease." *Nature* 632, no. 8026 (2024): 858-868.
19. Chen, Xinyue, Yin Huang, Liangfeng Huang, Ziliang Huang, Zhao-Zhe Hao, Lahong Xu, Nana Xu et al. "A brain cell atlas integrating single-cell transcriptomes across human brain regions." *Nature medicine* 30, no. 9 (2024): 2679-2691.

20. Miller, Jeremy A., Kyle J. Travaglini, Tain Luquez, Rachel E. Hostetler, Aaron Oster, Scott Daniel, Bosiljka Tasic, and Vilas Menon. "Annotation Comparison Explorer (ACE): connecting brain cell types across studies of health and Alzheimer's Disease." *bioRxiv* (2025).

CHAPTER 2

1. World Health Organization. "Global status report on the public health response to dementia." (2021).
2. Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica*, 82, 239–259.
3. Braak, H., Alafuzoff, I., Arzberger, T., Kretschmar, H., & Del Tredici, K. (2006). Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. *Acta Neuropathologica*, 112, 389–404.
4. Braak, H., Thal, D. R., Ghebremedhin, E., & Del Tredici, K. (2011). Stages of the pathologic process in Alzheimer disease: Age categories from 1 to 100 years. *Journal of Neuropathology & Experimental Neurology*, 70, 960–969.
5. Hyman, Bradley T., Creighton H. Phelps, Thomas G. Beach, Eileen H. Bigio, Nigel J. Cairns, Maria C. Carrillo, Dennis W. Dickson et al. "National Institute on Aging–Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease." *Alzheimer's & dementia* 8, no. 1 (2012): 1-13.
6. Montine, Thomas J., Creighton H. Phelps, Thomas G. Beach, Eileen H. Bigio, Nigel J. Cairns, Dennis W. Dickson, Charles Duyckaerts et al. "National Institute on Aging–Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach." *Acta neuropathologica* 123, no. 1 (2012): 1-11.
7. van der Kant, Rik, Lawrence SB Goldstein, and Rik Ossenkoppele. "Amyloid- β -independent regulators of tau pathology in Alzheimer disease." *Nature Reviews Neuroscience* 21, no. 1 (2020): 21-35.
8. Van Dyck, Christopher H., Chad J. Swanson, Paul Aisen, Randall J. Bateman, Christopher Chen, Michelle Gee, Michio Kanekiyo et al. "Lecanemab in early Alzheimer's disease." *New England Journal of Medicine* 388, no. 1 (2023): 9-21.
9. Lopes, M. Beatriz S. "Diagnostic histochemistry in neuropathology." In *Seminars in Diagnostic Pathology*, vol. 36, no. 1, pp. 71-82. WB Saunders, 2019.

10. Jack, Clifford R. "Advances in Alzheimer's disease research over the past two decades." *The Lancet Neurology* 21, no. 10 (2022): 866-869.
11. Jack, Clifford R., and David M. Holtzman. "Biomarker modeling of Alzheimer's disease." *Neuron* 80, no. 6 (2013): 1347-1358.
12. Latimer, Caitlin S., and Nicole F. Liachko. "Tau and TDP-43 synergy: a novel therapeutic target for sporadic late-onset Alzheimer's disease." *Geroscience* 43, no. 4 (2021): 1627-1634.

CHAPTER 3

1. Kampmann, Martin. "Molecular and cellular mechanisms of selective vulnerability in neurodegenerative diseases." *Nature Reviews Neuroscience* 25, no. 5 (2024): 351-371.
2. Mattsson, Niklas, Jonathan M. Schott, John Hardy, Martin R. Turner, and Henrik Zetterberg. "Selective vulnerability in neurodegeneration: insights from clinical variants of Alzheimer's disease." *Journal of Neurology, Neurosurgery & Psychiatry* 87, no. 9 (2016): 1000-1004.
3. Otero-Garcia, Marcos, Sameehan U. Mahajani, Debia Wakhloo, Weijing Tang, Yue-Qiang Xue, Samuel Morabito, Jie Pan et al. "Molecular signatures underlying neurofibrillary tangle susceptibility in Alzheimer's disease." *Neuron* 110, no. 18 (2022): 2929-2948.
4. Gayoso, Adam, Romain Lopez, Galen Xing, Pierre Boyeau, Valeh Valiollah Pour Amiri, Justin Hong, Katherine Wu et al. "A Python library for probabilistic analysis of single-cell omics data." *Nature biotechnology* 40, no. 2 (2022): 163-166.
5. He, Liang, Jose Davila-Velderrain, Tomokazu S. Sumida, David A. Hafler, Manolis Kellis, and Alexander M. Kulminski. "NEBULA is a fast negative binomial mixed model for differential or co-expression analysis of large-scale multi-subject single-cell data." *Communications biology* 4, no. 1 (2021): 629.
6. Buettner, Maren, Johannes Ostner, Christian L. Mueller, Fabian J. Theis, and Benjamin Schubert. "scCODA is a Bayesian model for compositional single-cell data analysis." *Nature communications* 12, no. 1 (2021): 6876.
7. Lambert, Jean-Charles, Simon Heath, Gael Even, Dominique Campion, Kristel Slegers, Mikko Hiltunen, Onofre Combarros et al. "Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease." *Nature genetics* 41, no. 10 (2009): 1094-1099.

8. Lackie, Rachel E., Andrzej Maciejewski, Valeriy G. Ostapchenko, Jose Marques-Lopes, Wing-Yiu Choy, Martin L. Duennwald, Vania F. Prado, and Marco AM Prado. "The Hsp70/Hsp90 chaperone machinery in neurodegenerative diseases." *Frontiers in neuroscience* 11 (2017): 254.
9. Yokota, T., M. Mishra, H. Akatsu, Y. Tani, T. Miyauchi, T. Yamamoto, K. Kosaka, Y. Nagai, T. Sawada, and K. Heese. "Brain site-specific gene expression analysis in Alzheimer's disease patients." *European journal of clinical investigation* 36, no. 11 (2006): 820-830.
10. Li, Qingqin S., and Louis De Muynck. "Differentially expressed genes in Alzheimer's disease highlighting the roles of microglia genes including OLR1 and astrocyte gene CDK2AP1." *Brain, behavior, & immunity-health* 13 (2021): 100227.
11. Zhou, Maotian, Rafi U. Haque, Eric B. Dammer, Duc M. Duong, Lingyan Ping, Erik CB Johnson, James J. Lah, Allan I. Levey, and Nicholas T. Seyfried. "Targeted mass spectrometry to quantify brain-derived cerebrospinal fluid biomarkers in Alzheimer's disease." *Clinical proteomics* 17, no. 1 (2020): 19.
12. Mei, Lin, and Klaus-Armin Nave. "Neuregulin-ERBB signaling in the nervous system and neuropsychiatric diseases." *Neuron* 83, no. 1 (2014): 27-49.
13. Woo, Ran-Sook, Ji-Hye Lee, Ha-Nul Yu, Dae-Yong Song, and Tai-Kyoung Baik. "Expression of ErbB4 in the apoptotic neurons of Alzheimer's disease brain." *Anatomy & Cell Biology* 43, no. 4 (2010): 332.
14. Zhang, Heng, Ling Zhang, Dongming Zhou, Hongfei Li, and Yang Xu. "ErbB4 mediates amyloid β -induced neurotoxicity through JNK/tau pathway activation: Implications for Alzheimer's disease." *Journal of Comparative Neurology* 529, no. 15 (2021): 3497-3512.
15. Wulf, Marie-Angela, Assunta Senatore, and Adriano Aguzzi. "The biological function of the cellular prion protein: an update." *BMC biology* 15, no. 1 (2017): 34.
16. Kellett, Katherine AB, and Nigel M. Hooper. "Prion protein and Alzheimer disease." *Prion* 3, no. 4 (2009): 190-194.
17. Sugita, S., A. Ho, and T. C. Südhof. "NECABs: a family of neuronal Ca^{2+} -binding proteins with an unusual domain structure and a restricted expression pattern." *Neuroscience* 112, no. 1 (2002): 51-63.
18. Kay, Jeremy N., Monica W. Chu, and Joshua R. Sanes. "MEGF10 and MEGF11 mediate homotypic interactions required for mosaic spacing of retinal neurons." *Nature* 483, no. 7390 (2012): 465-469.

19. Zhou, Lujia, Soraia Barao, Mathias Laga, Katrijn Bockstael, Marianne Borgers, Harry Gijssen, Wim Annaert et al. "The neural cell adhesion molecules L1 and CHL1 are cleaved by BACE1 protease in vivo." *Journal of Biological Chemistry* 287, no. 31 (2012): 25927-25940.
20. Chaki, Moumita, Rannar Airik, Amiya K. Ghosh, Rachel H. Giles, Rui Chen, Gisela G. Slaats, Hui Wang et al. "Exome capture reveals ZNF423 and CEP164 mutations, linking renal ciliopathies to DNA damage response signaling." *Cell* 150, no. 3 (2012): 533-548.
21. Sen, Abhik, Thomas J. Nelson, Daniel L. Alkon, and Jarin Hongpaisan. "Loss in PKC epsilon causes downregulation of MnSOD and BDNF expression in neurons of Alzheimer's disease hippocampus." *Journal of Alzheimer's Disease* 63, no. 3 (2018): 1173-1189.
22. Yan, Xu, Niko-Petteri Nykänen, Cecilia A. Brunello, Annakaisa Haapasalo, Mikko Hiltunen, Riikka-Liisa Uronen, and Henri J. Huttunen. "FRMD4A–cytohesin signaling modulates the cellular release of tau." *Journal of cell science* 129, no. 10 (2016): 2003-2015.
23. Lambert, Jean-Charles, Benjamin Grenier-Boley, Denise Harold, Diana Zelenika, Vincent Chouraki, Yoichiro Kamatani, Kristel Slegers et al. "Genome-wide haplotype association study identifies the FRMD4A gene as a risk locus for Alzheimer's disease." *Molecular psychiatry* 18, no. 4 (2013): 461-470.
24. Seto, Mabel, Rebecca L. Weiner, Logan Dumitrescu, and Timothy J. Hohman. "Protective genes and pathways in Alzheimer's disease: moving towards precision interventions." *Molecular neurodegeneration* 16, no. 1 (2021): 29.
25. Bik-Multanowski, Mirosław, Jacek J. Pietrzyk, and Alina Midro. "MTRNR2L12: A candidate blood marker of early Alzheimer's disease-like dementia in adults with down syndrome." *Journal of Alzheimer's Disease* 46, no. 1 (2015): 145-150.
26. Hashimoto, Yuichi, Takako Niikura, Hirohisa Tajima, Takashi Yasukawa, Haruka Sudo, Yuko Ito, Yoshiko Kita et al. "A rescue factor abolishing neuronal cell death by a wide spectrum of familial Alzheimer's disease genes and A β ." *Proceedings of the National Academy of Sciences* 98, no. 11 (2001): 6336-6341.
27. Ficulie, Elena, Sarubini Kananathan, David Airey, Severine I. Gharbi, Neil Humphries-Kirilov, James Scherschel, Charlotte Dunbar et al. "A human tau seeded neuronal cell model recapitulates molecular responses associated with Alzheimer's disease." *Scientific Reports* 12, no. 1 (2022): 2673.
28. Ghosh, Anshua, and Karl Peter Giese. "Calcium/calmodulin-dependent kinase II and Alzheimer's disease." *Molecular brain* 8, no. 1 (2015): 78.

29. Wei, Yu-Ping, Jin-Wang Ye, Xiong Wang, Li-Ping Zhu, Qing-Hua Hu, Qun Wang, Dan Ke, Qing Tian, and Jian-Zhi Wang. "Tau-induced Ca²⁺/calmodulin-dependent protein kinase-IV activation aggravates nuclear tau hyperphosphorylation." *Neuroscience Bulletin* 34, no. 2 (2018): 261-269.
30. Bell, Simon M., Katy Barnes, Matteo De Marco, Pamela J. Shaw, Laura Ferraiuolo, Daniel J. Blackburn, Annalena Venneri, and Heather Mortiboys. "Mitochondrial dysfunction in Alzheimer's disease: a biomarker of the future?." *Biomedicines* 9, no. 1 (2021): 63.
31. Reddy, P. Hemachandra. "Abnormal tau, mitochondrial dysfunction, impaired axonal transport of mitochondria, and synaptic deprivation in Alzheimer's disease." *Brain research* 1415 (2011): 136-148.
32. Kim, Kyoung Mi, Qiong Meng, Olivia Perez de Acha, Maja Mustapic, Aiwu Cheng, Erden Eren, Gautam Kundu et al. "Mitochondrial RNA in Alzheimer's disease circulating extracellular vesicles." *Frontiers in cell and developmental biology* 8 (2020): 581882.
33. Mastroeni, Diego, Omar M. Khmour, Elaine Delvaux, Jennifer Nolz, Gary Olsen, Nicole Berchtold, Carl Cotman, Sidney M. Hecht, and Paul D. Coleman. "Nuclear but not mitochondrial-encoded oxidative phosphorylation genes are altered in aging, mild cognitive impairment, and Alzheimer's disease." *Alzheimer's & Dementia* 13, no. 5 (2017): 510-519.
34. Terada, Tatsuhiko, Joseph Therriault, Min Su Peter Kang, Melissa Savard, Tharick Ali Pascoal, Firoza Lussier, Cecile Tissot et al. "Mitochondrial complex I abnormalities is associated with tau and clinical symptoms in mild Alzheimer's disease." *Molecular neurodegeneration* 16, no. 1 (2021): 28.
35. Thubron, Elisabeth B., Hannah S. Rosa, Angela Hodges, Sobha Sivaprasad, Paul T. Francis, Ilse S. Pienaar, and Afshan N. Malik. "Regional mitochondrial DNA and cell-type changes in post-mortem brains of non-diabetic Alzheimer's disease are not present in diabetic Alzheimer's disease." *Scientific reports* 9, no. 1 (2019): 11386.

CHAPTER 4

1. Lee, Je Hyuk, Evan R. Daugharthy, Jonathan Scheiman, Reza Kalhor, Thomas C. Ferrante, Richard Terry, Brian M. Turczyk et al. "Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues." *Nature protocols* 10, no. 3 (2015): 442-458.

2. Ståhl, Patrik L., Fredrik Salmén, Sanja Vickovic, Anna Lundmark, José Fernández Navarro, Jens Magnusson, Stefania Giacomello et al. "Visualization and analysis of gene expression in tissue sections by spatial transcriptomics." *Science* 353, no. 6294 (2016): 78-82.
3. Allen Institute for Brain Science. 2025. MapMyCells (RRID: SCR_024672). Accessed August 1, 2025. <https://knowledge.brain-map.org/mapmycells/process/>.
4. Song, You-Hyang, Jiwon Yoon, and Seung-Hee Lee. "The role of neuropeptide somatostatin in the brain and its application in treating neurological disorders." *Experimental & Molecular Medicine* 53, no. 3 (2021): 328-338.

ADDITIONAL RESOURCES

1. Yiannopoulou, K. G., Anastasiou, A. I., Zachariou, V., & Pelidou, S.-H. (2019). Reasons for failed trials of disease-modifying treatments for Alzheimer disease and their contribution in recent research. *Biomedicines*, 7.
2. Rubin, R. (2021). Recently approved Alzheimer drug raises questions that might never be answered. *JAMA*, 326, 469–472.
3. He, X., Memczak, S., Qu, J., Belmonte, J. C. I., & Liu, G.-H. (2020). Single-cell omics in ageing: A young and growing field. *Nature Metabolism*, 2, 293–302.
4. Gómez-Isla, T., et al. (1996). Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. *Journal of Neuroscience*, 16, 4491–4500.
5. Mayeux, R., & Stern, Y. (2012). Epidemiology of Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine*, 2.
6. Nichols, E., et al. (2019). Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology*, 18, 88–106.
7. Hou, Y., et al. (2019). Ageing as a risk factor for neurodegenerative disease. *Nature Reviews Neurology*, 15, 565–581.
8. Global Dementia Cases Forecasted to Triple by 2050. https://www.alz.org/aaic/releases_2021/global-prevalence.asp.

9. Drummond, E., & Wisniewski, T. (2017). Alzheimer's disease: Experimental models and reality. *Acta Neuropathologica*, 133, 155–175.
10. Kametani, F., & Hasegawa, M. (2018). Reconsideration of amyloid hypothesis and tau hypothesis in Alzheimer's disease. *Frontiers in Neuroscience*, 12, 25.
11. Darmanis, S., et al. (2015). A survey of human brain transcriptome diversity at the single cell level. *Proceedings of the National Academy of Sciences USA*, 112, 7285–7290.
12. Stranahan, A. M., & Mattson, M. P. (2010). Selective vulnerability of neurons in layer II of the entorhinal cortex during aging and Alzheimer's disease. *Neural Plasticity*, 108190.
13. Leng, K., et al. (2021). Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. *Nature Neuroscience*, 24, 276–287.
14. Mathys, H., et al. (2019). Single-cell transcriptomic analysis of Alzheimer's disease. *Nature*, 570, 332–337.
15. Hawrylycz, M. J., et al. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, 489, 391–399.
16. Hawrylycz, M., et al. (2015). Canonical genetic signatures of the adult human brain. *Nature Neuroscience*, 18, 1832–1844.
17. Lake, B. B., et al. (2016). Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science*, 352, 1586–1590.
18. Lake, B. B., et al. (2018). Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nature Biotechnology*, 36, 70–80.
19. Garcia, A. D., & Buffalo, E. A. (2020). Anatomy and function of the primate entorhinal cortex. *Annual Review of Vision Science*, 6, 411–432.
20. Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, 2, 913–919.
21. Davey, J., et al. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *NeuroImage*, 137, 165–177.

22. Gabrieli, J. D., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences USA*, 95, 906–913.
23. Friedman, N. P., & Robbins, T. W. (2022). The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology*, 47, 72–89.
24. Convit, A., et al. (2000). Atrophy of the medial occipitotemporal, inferior, and middle temporal gyri in non-demented elderly predict decline to Alzheimer's disease. *Neurobiology of Aging*, 21, 19–26.
25. Davila-Velderrain, Jose, Hansruedi Mathys, Shahin Mohammadi, Brad Ruzicka, Xueqiao Jiang, Ayesha Ng, David A. Bennett, Li-Huei Tsai, and Manolis Kellis. "Single-cell anatomical analysis of human hippocampus and entorhinal cortex uncovers early-stage molecular pathology in Alzheimer's disease." *bioRxiv* (2021): 2021-07.
26. Olah, M., et al. (2020). Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. *Nature Communications*, 11, 6129.
27. Lein, E., Borm, L. E., & Linnarsson, S. (2017). The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. *Science*, 358, 64–69.
28. Campbell, M. J., & Morrison, J. H. (1989). Monoclonal antibody to neurofilament protein (SMI-32) labels a subpopulation of pyramidal neurons in the human and monkey neocortex. *Journal of Comparative Neurology*, 282, 191–205.
29. Oberstein, T. J., et al. (2015). Astrocytes and microglia but not neurons preferentially generate N-terminally truncated A β peptides. *Neurobiology of Disease*, 73, 24–35.
30. Langseth, C. M., et al. (2021). Comprehensive in situ mapping of human cortical transcriptomic cell types. *Communications Biology*, 4, 1–7.