Title
Uncovering the molecular architecture of Alzheimer’s disease

Permalink
https://escholarship.org/uc/item/4435w86z

Author
Miyoshi, Emily

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,

IRVINE

Uncovering the molecular architecture of Alzheimer's disease

DISSERTATION

Submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Emily Miyoshi

Dissertation Committee:
Professor Vivek Swarup, Chair
Professor Kim Green
Professor Leslie Thompson

2022

Dedication

to

Popo (Mildred Ho)

4/4/1930 - 5/13/2014

Table of Contents

List of Figures

Acknowledgements

I first would like to thank my advisor Dr. Vivek Swarup. His mentorship and guidance were instrumental in the culmination of this dissertation, and I am grateful to have had an advisor with so much faith and confidence in me. In addition, I would like to thank the members of the lab, especially my work wife Dr. Neethu Michael and fellow graduate student Samuel Morabito (also soon-to-be Dr.!), who were with us at the beginning of the lab. It has been such a rewarding journey to see the lab grow over the years. I am grateful for my committee members Dr. Kim Green and Dr. Leslie Thompson for their guidance and insight over the years. Their lab members were also immensely helpful in bringing the lab to where it is now, in addition to being some great friends. I would like to thank Dr. Elizabeth Head and her lab, as well as the core, for their help; none of this would have been possible without their hard work. Additionally, without my undergraduate advisor Dr. Karen Gylys and mentor Dr. Tina Bilousova, I would not have even considered graduate school. They both inspired me with their passion for research and provided the foundation for my career. Furthermore, I would like to thank my family and friends for their unending support throughout these years—those who pushed me to go outside of my comfort zone and those who understood when I couldn't get out of lab. Of course, I also must thank my partner Mikhail Golub for being my personal chef and handyman and our dogs Kiwi and Nugget. Mike, I am so proud of you and your budding business.

This work was supported by UCI and UCI MIND startup funds, NIA 1RF1AG071683, Adelson Medical Research Foundation funds and an American Federation of Aging Research Young Investigator Award to V.S.; NIA 3U19AG068054-02S2 (ABC-DS); NIA U54 AG054349 (MODEL-AD); and NIA T32 AG00096-38 to E.M. The text of Chapter Two is a reprint of the material as it appears in Morabito, S. Miyoshi, E., Michael, N. *et al.* Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nature Genetics*. 2021. Springer Nature.

Vita

Emily Miyoshi

Education

University of California, Irvine
PhD, Biological Sciences Expected 2022
MSc, Biological Sciences 2021

University of California, Los Angeles
BS, Neuroscience 2017

Awards

- James L. McGaugh Award, UCI Neurobiology and Behavior, 2022.
- NRSA F31, NIA, 2022.
- First place, UCI GPS-STEM Business Pitch Competition, 2022.
- Carl W. Cotman Scholar's Award, UCI MIND, 2022.
- T32 (Aging and Alzheimer's Disease), NIA, 2021.
- William D. Redfield Graduate Fellowship Award, UCI, 2021.
- Best Poster Presentation, UCI Neurobiology and Behavior, 2020.
- Magna Cum Laude, UCLA, 2017.
- Departmental Highest Honors, UCLA, 2017.
- Undergraduate Research Scholarship Program Scholar, UCLA, 2016.
- Chapter Scholarship, UCLA Alpha Lambda Delta & Phi Eta Sigma, 2015.
- National Residence Hall Honorary Student of the Month, UCLA, 2014.
- Dean's List, UCLA, 2013-2017.

Publications

- Shi Z*, Das S*, Morabito S, **Miyoshi E**, Swarup V. (2022) Protocol for single-nucleus ATAC sequencing and bioinformatic analysis in frozen human brain tissue. *STAR Protocols.* 3(3): 101491.
- Shabestari SK, Morabito S, Danhash EP, McQuade A, Sanchez JR, **Miyoshi E**, Chadarevian JP, Claes C, Coburn MA, Hasselmann J, Silva J, Tran KN, Martini AC, Rothermich WC, Pascual J, Head E, Hume DA, Pridans C, Davtyan H, Swarup V, Blurton-Jones M. (2022) Absence of microglia promotes diverse pathologies and early lethality in Alzheimer's disease mice. *Cell Reports.* 39(11): 110961.
- **Miyoshi E***, Morabito S*, Swarup V. (2021) Systems biology approaches to unravel the molecular and genetic architecture of Alzheimer's disease and related tauopathies. *Neurobiology of Disease.* 160: 105530.
- **Miyoshi E***, Bilousova T*, Melnik M, Fakhrutdinov D, Poon WW, Vinters HV, Miller CA, Corrada M, Kawas C, Bohannan R, Caraway C, Elias C, Maina KN, Campagna JJ, John V, Gylys KH. (2021) Exosomal tau with seeding activity is released from Alzheimer's disease synapses, and seeding potential is associated with amyloid beta. *Laboratory Investigation.* 101(12): 1605-1617.
- Morabito S*, **Miyoshi E***, Michael N*, Shahin S, Martini AC, Head E, Silva J, Leavey K, Perez-Rosendahl M, Swarup V. (2021) Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nature Genetics.* 53(8): 1143-11555.

- **Miyoshi E**, Swarup V. (2021) Rogue networks gone awry in Alzheimer's disease. *Neural Regeneration Research.* 16(12):2415.
- Morabito S, **Miyoshi E**, Michael N, Swarup V. (2020) Integrative genomics approach identifies conserved transcriptomic networks in Alzheimer's disease. *Human Molecular Genetics.* 29(17): 2899-2919.
- Bilousova T, Melnik M, **Miyoshi E**, Gonzalez BL, Poon WW, Vinters HV, Miller CA, Corrada MM, Kawas C, Hatami A, Albay R 3rd, Glabe C, Gylys KH. (2019) Apolipoprotein E/Aβ Complex Accumulates in AD Cortical Synapses via apoE Receptors and Is Enhanced by APOE4. *American Journal of Pathology.* 189(8): 1621-1636.
- Bilousova T, Elias C, **Miyoshi E**, Alam MP, Zhu C, Campagna J, Vadivel K, Jagodzinska B, Gylys KH, John V. (2018) Suppression of tau propagation using an inhibitor that targets the DK- switch of nSMase2. *Biochemical and Biophysical Research Communications.* 499(4):751-757.

Talks

- "Spatially resolved transcriptomics in Alzheimer's disease," UCI Neurodegeneration Community Workshop, 2022.
- "Single-nucleus epigenomics and transcriptomics identifies oligodendrocyte-specific SREBF1 dsyregulation in Alzheimer's disease," REMIND Emerging Scientists Symposium, 2021.

Poster Presentations

- **Miyoshi E***, Morabito S*, Henningfield CM, Michael N, Shabestari SK, Das S, Shahin S, MODEL-AD, Green KN, Swarup V. Spatiotemporal transcriptomic characterization of an amyloid mouse model of Alzheimer's disease. Alzheimer's Association International Conference, San Diego, CA. 2022.
- **Miyoshi E*,** Morabito S*, Henningfield CM, Michael N, Shabestari SK, Das S, Wright S, Silva J, Leavy K, Shahin S, Perez-Rosendahl M, Head E, Green KN, Swarup V. Spatially resolved transcriptomics in Alzheimer's disease. REMIND Emerging Scientists Symposium, Irvine, CA. 2022.
- Morabito S*, **Miyoshi E***, Michael N*, Shahin S, Head E, Silva J, Leavy K, Perez-Rosendahl M, Swarup V. Single-nuclei epigenomics and transcriptomics to uncover the gene regulatory landscape in Alzheimer's disease. Neurobiology and Behavior Departmental Retreat, Virtual. 2020.
- **Miyoshi E*,** Bilousova T*, Alam, M, Campagna JJ, Elias CJ, Hatami A, Fakhrutdinov D, John V, Gylys KH. Extracts from AD synapses accelerate in vitro tau propagation. Society for Neuroscience annual meeting, Washington D.C. 2017.
- **Miyoshi E**, Bilousova T, Gylys KH. Characterizing the depolarization- dependent synaptic release of tau. Society for Neuroscience annual meeting, San Diego, CA. 2016.

Professional Development

- Industry Insights for STEM Scientists Certification, 2022.
- Activate to Captivate Public Speaking Certification, 2020.

Abstract of the Dissertation

Uncovering the molecular architecture of Alzheimer’s disease

by

Emily Miyoshi

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2022

Professor Vivek Swarup, Chair

Alzheimer’s disease (AD) is a devastating, progressive neurodegenerative disorder that results in dementia, with care for those with dementia estimated to cost the U.S. $321 billion in 2022. Although many years of research have uncovered notable findings about AD biology, it is clear that we still have a limited understanding of the disease as evidenced by the lack of effective therapeutics against AD. For example, genome-wide association studies (GWAS) have uncovered multiple genetic risk variants, revealing novel genes and pathways for study in AD. However, it remains a challenge to ascertain the functional significance of these risk variants. Multiple studies have attempted to clarify the role of AD risk variants with “bulk”-tissue RNA-sequencing (RNA-seq), but they have been hindered by the vast cellular heterogeneity of the brain. Single-cell (scRNA-seq) and single-nucleus RNA-seq (snRNA-seq) performed on the 5XFAD mouse model of AD, as well as human AD samples, identified disease-associated glial subpopulations, but it is unclear what regulates these subtypes. Additionally, studies have revealed region-specific glial subpopulations existing in the healthy brain, suggesting potential regional differences in the disease phenotype. This dissertation aimed to clarify the molecular

landscape of AD with spatial and cellular resolution with an emphasis on the analysis and integration of different data modalities and model systems.

We first sought to define the cell-type specific gene regulatory programs dysregulated in AD to identify potential regulators of disease-associated cell subpopulations and to further unravel AD genetic risk (**Chapter Two**). Recent advances in sequencing methods now allow interrogation of the transcriptome and epigenome at the single cell resolution, and AD epigenetic data has been limited. We generated paired single-nucleus transcriptomic and epigenomic data from postmortem human brain tissue of late-stage AD and cognitively healthy controls. In addition to being the first epigenetic dataset of human AD with single-cell resolution, we directly integrated the two different data modalities, allowing us to define disease-associated glial subpopulations at the transcriptome and epigenome. We identified cell-type specific, disease-associated candidate cis-regulatory elements (cCREs) and their candidate target genes. Although this is possible with single-cell epigenetic data alone, paired gene expression data provides additional functional evidence of cCREs. We also revealed cell-type specific transcription factors dysregulated with disease, like SREBF1 in oligodendrocytes, altogether identifying both cis- and trans-gene regulatory mechanisms that may regulate AD cell states. Furthermore, we characterized the cis-regulatory landscape at AD GWAS loci in specific cell-types, providing insight into the cell-types relevant to specific AD risk variants.

On the other hand, while recent transcriptomic studies have revealed both brain region- and cell-type-specific gene expression changes in AD, “bulk”-tissue and scRNA-seq do not retain spatial information for gene expression changes without careful microdissection and sequencing separate samples. The human brain’s spatially complexity at both macro- and microscopic levels underlies brain function and thus is a critical feature to consider in disease pathophysiology. We generated spatial transcriptomic data from postmortem human brain tissue from cognitively healthy controls, early-, and late-stage AD to investigate the spatial relationship of disease-associated transcriptomic changes (**Chapter Three**). Additionally, we performed a comparative analysis of AD in Down Syndrome (DS) and the general population by generating both spatial and single-nucleus transcriptomic data from AD in DS. To date, no published spatial or single-nucleus studies have explored the concordance between these two populations, although AD in DS may serve as an advantageous group for preclinical and clinical studies of AD. We identified regional and cell-type specific transcriptomic changes shared between both AD populations. Further, we present a time-course analysis of the spatial transcriptome of the amyloid mouse model 5XFAD to assess the AD transcriptome across multiple brain regions and identify evolutionary-conserved gene expression changes. In addition to surveying different model systems, we integrated imaging data with spatial transcriptomic data. Amyloid beta (Aβ) pathology is one of the classical hallmarks of AD and proposed as a critical driver of AD pathogenesis (amyloid cascade hypothesis). We identified transcripts spatially localized to Aβ pathology conserved between both human and mouse. We also integrated spatial and single-nucleus transcriptomic data to discover spatially defined cell signaling pathways dysregulated with disease and highlight regional heterogeneity of astrocytes. Altogether, this dissertation reveals regional and cellular molecular changes occurring in AD and contextualizes them in a systems-level framework to uncover pathways for further study in AD.

Chapter One: Multi-omics to disentangle the molecular architecture of Alzheimer's disease

Genetics of Alzheimer's disease

Alzheimer's disease (AD) is one of the most common neurodegenerative disorders, clinically manifesting as cognitive decline and memory loss. Pathologically it is defined by the accumulation of amyloid beta (Aβ) plaques and neurofibrillary tangles—abnormal protein aggregates consisting of Aβ and tau, respectively. The genetic components of AD can be divided into two forms: deterministic gene mutations and risk conferring genetic variants. The former includes rare, autosomal dominant mutations in *APP*, *PSEN1*, and *PSEN2*, constituting a small proportion of AD cases, where age of onset may be as early as 35 years old (early-onset familial AD, EOFAD), compared to the typical >65 age of onset (late-onset AD, LOAD) (*1*). *APP* encodes amyloid precursor protein, which is cleaved by β- and γ-secretase to produce Aβ, and both *PSEN1* and *PSEN2* encode subunits of γ-secretase (*1*). The introduction of these mutations results in extensive amyloid deposition and gliosis, as well as cognitive deficits, in mice (*2–4*), indicating a critical role of Aβ in AD pathophysiology.

However, most AD cases (LOAD) do not have a clear genetic etiology and are labeled as "sporadic," although LOAD heritability is reported to be between 58-79% (*5*). It is suggested that LOAD may be the result of a combination of genetic variants with smaller effect sizes, and therefore, there have been large efforts to determine the genetic landscape of LOAD. The *APOE* E4 allele was the first genetic variant identified to confer risk for LOAD and influences the age of onset (*6*, *7*). Many genome-wide association studies (GWAS), relying on microarray or whole genome or exome sequencing, have discovered additional genetic risk variants, with the most recent study reporting a total of 75 LOAD risk loci (*8–15*). Although risk variants may act on genes large distances away, GWAS-identified risk genes have been typically annotated by the nearest gene. More recent studies are attempting to determine candidate causal genes with gene prioritization analyses, which should help to finetune mechanistic follow-up studies. As a whole however, AD GWAS has highlighted several biological pathways of interest, such as APP metabolism (*ADAM10*, *APP*, *SORL1*), lipid homeostasis (*ABCA1*, *ABCA7*, *APOE*, *CLU*), the immune response (*CD33*, *CR1*, *HLA*, *MS4A*, *SPI1*, *TREM2*), and endocytosis (*BIN1*, *PICALM*). Protective genetic factors also reside within some of the same loci associated with risk (*16–19*), suggesting these are key AD genes.

Another particularly interesting genetic form of AD is trisomy of chromosome 21 (chr21), Down syndrome (AD in DS). Notably, chr21 includes *APP*, and most individuals with DS accumulate AD pathology by 40 years old and in a predictable manner (*20*, *21*). At ages >65 years old, individuals with DS have an 80% risk for dementia (*22*). Biomarker changes in AD in DS follow those in both LOAD and EOFAD (*23*), and *APOE* E4 may also be a genetic risk factor for AD in DS (*24*). Additionally, a study identified risk variants in *APP*, *CST3*, and *MARK4* for AD in DS (*25*)—all three genes have been associated with AD risk in the general population. This smaller and more genetically and phenotypically homogeneous population may be advantageous to study AD, considering the shared features between AD in DS and LOAD.

Global transcriptomic changes in Alzheimer's disease

It remains a challenge, however, to understand the molecular underpinnings of AD with genetic variation alone, as the downstream functional outcomes of these AD-associated risk variants still need to be determined. Researchers therefore turned to a more functional assessment, gene expression measurements from microarray and RNA-sequencing (RNA-seq), to dissect the relationship between genetic variation and disease phenotype, and multiple studies have now identified both upregulated and downregulated genes in AD at the tissue-level with differential expression analysis (*26–35*). Gene expression changes reflected the known pathological progression of AD in the brain (*32–34*). While a few studies identified expression quantitative trait loci (*36–38*), several studies also revealed potential posttranscriptional regulatory changes, such as microRNAs over- or underrepresented in AD (*31*, *39*, *40*), as well as differential splicing or editing events (*31*, *41*, *42*). Further, it was discovered that transposable elements, which also may influence regulation of gene expression, are differentially expressed in AD (*43*, *44*).

In addition to differential analysis, many transcriptomic studies applied network analysis to help decipher the biological significance of differentially expressed genes (DEGs) by clarifying the relationships between these genes and other co-regulated genes. Across these studies, key biological pathways, such as the immune response, synaptic function, myelination, and metabolism, were dysregulated in AD (*30*, *33*, *39*, *45–49*). We performed a meta-analysis of three independent, large-scale RNA-seq cohorts and found many of these biological pathways are robustly preserved across brain regions and independent datasets (*45*). Interestingly, the expression of synaptic genes appears to increase in early stages of disease but decreases at the later stages (*26*, *29*, *33*, *48*). Myelination-related genes also decrease in AD (*28*, *47*), and a recent study discovered transcriptomic subtypes of AD, where *PLP1*, encoding a major myelin protein, is a key driver of a major subtype (B) (*34*). A study also revealed *TYROBP*, which interacts with *TREM2*, is a regulator of immune-related processes in AD (*30*), while we identified the transcription factor SPI1 in our meta-analysis (*45*). These tissue-level transcriptomic analyses have not only reaffirmed postulated biological processes involved in AD, but also highlighted candidate genes for further investigation.

Cellular diversity of the brain

The adult human brain is composed of multiple cell populations and even cell subpopulations; the existence of subpopulation-level diversity was revealed in the drawings of Santiago Ramón y Cajal, illustrating morphologically distinct neurons and astrocytes. From sequencing a single mouse blastomere (*50*) to thousands of cells per sample (*51*, *52*), single-cell RNA-seq (scRNA-seq) has provided us the opportunity to unbiasedly discover molecularly distinct cell subpopulations in the brain, and the development of profiling of single-nucleus suspensions (snRNA-seq) (*53*, *54*) allowed us to utilize archived postmortem human brain tissue samples, in addition to reducing technical artifacts in mouse studies. Several studies have greatly propelled our understanding of the brain's cellular heterogeneity (*51*, *54–62*). For example, the Allen Institute identified over 40 inhibitory neuron clusters within the human middle temporal gyrus (*56*), and more recently more than 200 excitatory neuron clusters across multiple brain regions in the adult mouse (*58*). Through their careful microdissection of the cortical layers, we now have the transcriptomic signatures of layer-specific and multiple layer-spanning neuronal subpopulations, and interestingly, many cluster-specific markers are noncoding RNAs, reinforcing that gene regulation is largely cell-type specific.

Furthermore, several studies revealed transcriptionally distinct astrocytes and microglia that are region-specific; this includes those specific to grey or white matter, in addition to individual cortical layers (*63–67*). Oligodendrocytes are also surprisingly heterogeneous. scRNA-seq performed on mouse brain tissue, as well as snRNA-seq on human brain tissue, discovered several transcriptionally distinct mature oligodendrocyte subpopulations, majorly separating by myelinating or non-myelinating state (*60*, *68–70*). A few studies also characterized the epigenome of single cells in the mouse and human brain and likewise identified epigenetically distinct cell subpopulations (*71–73*), and more recently vascular subpopulations in the brain were transcriptionally defined (*74*). It remains unclear, however, what is the functional significance of these heterogeneous subsets of cells, as well as how do different cell subpopulations respond to disease.

Single-cell studies of Alzheimer's disease

While tissue-level transcriptomic studies helped to uncover biological pathways and genes dysregulated in AD, they are limited by their inability to resolve cell-specific gene expression changes, which may greatly influence our understanding of a gene's role in disease. Results from "bulk" RNA-seq ultimately are skewed based on a tissue's cellular composition, and this was confirmed in the first single-nucleus transcriptomic study of the human AD brain (*75*) with 48 prefrontal cortical samples from the Religious Order Study and the Rush Memory and Aging Project (ROSMAP, 24 controls and 24 AD based on pathological staging). They notably found that most AD DEGs are cell-type specific and that bulk-tissue RNA-seq can fail to capture AD gene expression changes unique to smaller cell populations.

Moreover, single-cell profiling provides the opportunity to examine cell subpopulation-specific changes in disease. Another snRNA-seq study identified *RORB*$^+$ excitatory neuronal subpopulations in the entorhinal cortex that are selectively vulnerable in AD, as well as similar populations in the superior frontal gyrus (*76*). They interestingly did not observe this selective vulnerability among inhibitory neurons. In contrast, many studies have found expansion of select cell subpopulations with AD (*75*, *77–84*). scRNA-seq studies of the amyloid mouse model 5XFAD discovered disease-associated microglia (DAMs), astrocytes (DAAs), and oligodendrocytes (DOLs) (*77*, *79*, *82*). These subpopulations notably have overlapping gene markers, like *Serpina3n*, *B2m*, and *Ctsb*, and are localized to amyloid plaques, suggesting similarities in the response to amyloid pathology across glia. In addition, Keren-Shaul *et al.* proposed that microglia transition to DAM in two stages, where one stage is *Trem2* dependent (*77*), and this was validated in a separate snRNA-seq study of 5XFAD and 5XFAD-*Trem2* deficient mice (*78*). It is unknown what regulates the other DAM stage, as well as DAAs and DOLs. However, it is also uncertain if these subpopulations are relevant in clinical AD. For example, DAMs have been only partially recapitulated in human AD studies (*78*, *85*). The analysis of both human and mouse datasets, rather than the analysis of one alone, may help to deepen our understanding of disease by allowing us to identify species-conserved, disease-associated changes that we can manipulate and study *in vivo*.

Further, there is presently a paucity of epigenetic data for AD, especially at the single-cell resolution. A few studies identified AD-associated, tissue-level changes in DNA methylation and histone modifications (*86*–*89*); however, the cell-type specificity of gene regulation underlines the need for single-cell profiling of epigenetic features. This is exemplified in a study that examined chromatin accessibility changes in sorted cells (*90*) and functionally validated a microglia-specific enhancer for *BIN1*. Recent advances in single-cell technologies, such as single-cell assay for transposase-accessible chromatin using sequencing (ATAC-sequencing, open chromatin) (*91*, *92*), 10x Genomics Multiome (paired open chromatin and gene expression), and Paired-Tag (paired histone modifications and gene expression) (*93*), allow us to explore the single-cell epigenome in a high throughput manner, and leveraging these will help to provide a comprehensive knowledge of the molecular changes occurring in AD and subsequent delineation of the underlying biological processes.

Spatially resolved molecular profiling to study Alzheimer’s disease

Unsurprisingly sc- and snRNA-seq have their own limitations; tissue homogenization into a single-cell suspension results in a loss of associated spatial information. Therefore, researchers have developed several approaches to investigate gene expression while retaining spatial context, such as multiplexed *in situ* hybridization (*94*), *in situ* sequencing (*95*), and spatial transcriptomics

(*96*, *97*). In comparison to methods requiring pre-determined gene targets (probes), spatial transcriptomics functions similarly to RNA-seq by utilizing reverse transcription to allow unbiased profiling of the transcriptome. Arrays of “spots” with primers including a spot-specific barcode and poly(dT) primer result in transcripts uniquely barcoded by their spatial location within the array. Maynard *et al*. performed spatial transcriptomics on human prefrontal cortical samples, identifying layer-specific gene signatures without the need for microdissection (*98*). A few studies also used spatial transcriptomics to study AD and amyotrophic lateral sclerosis to examine region-specific gene expression changes in disease (*99–101*).

However, the resolution of spatial transcriptomics is limited to the size of the spot—currently 55μm in the commercially distributed form, 10x Genomics Visium—and thus each spot is likely a mixture of 2-3 cells, depending on the tissue. This subsequently prompted the development of methods to integrate single-cell and spatial datasets, like SPOTlight (*102*) and BayesSpace (*103*). Two studies from the Liddelow lab analyzed both data modalities to further interrogate astrocyte subpopulations and their spatial distributions within the brain (*104*, *105*), and Kamath and Abdulraouf *et al*. identified *AGTR1*+ neurons within the ventral substantia nigra pars compacta are selectively vulnerable in Parkinson’s disease (*106*). Altogether these studies, as well as those of other tissue types (*107–109*), highlight the utility of integrating single-cell and spatial data to understand specific cell subpopulations in disease.

In addition, spatial data should not be constrained to the transcriptome. Recent adaptations allow us to profile the epigenome in a spatial manner (*110–112*), and in proteinopathies, like AD, it is important to examine molecular changes in relation to pathological protein aggregates. Previous studies have relied on global pathological scoring, but resolving the localization of pathological features and associated transcripts within the local environment will provide higher confidence for pathology-associated genes. Spatial transcriptomics was originally limited to hematoxylin and eosin staining, so Chen *et al*. immunostained adjacent sections for Aβ to investigate Aβ-associated gene expression changes and identified "plaque-induced genes" (*99*). Immunostaining, however, can now be performed on the same tissue sections, to more concretely associate the spatial localization of pathology to gene transcripts. Moreover, several promising "hi-plex" multi-omic approaches, such as the commercially available Nanostring CosMx and Akoya Phenocycler, allow single-cell profiling of hundreds to thousands of proteins and transcripts within the same tissue section. Overall, new spatial methods provide incredibly exciting opportunities to finely map disease-associated molecular changes within the microenvironment and uncover spatially restricted disease signatures that may be linked to brain circuitry.

Chapter Two: Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease

Introduction

The human brain is composed of multiple heterogeneous subsets of cells; both neuronal and nonneuronal cells work in concert to perform simple and higher-order tasks. Recent studies have provided more precise molecular characterization and identification of neuronal and nonneuronal cell populations in the cognitively normal brain (*54*, *56*, *57*, *113*). However, our understanding of heterogeneous cell populations within the diseased brain is still largely limited, hindering our understanding of the biological processes underlying disease. Neurodegenerative disorders, like Alzheimer's disease (AD), are marked with massive neuronal loss, accompanied by gliosis, and the role of specific neuronal and glial cell populations in AD pathophysiology remains unclear. Several single-cell and single-nucleus RNA-sequencing (snRNA-seq) studies have been performed on both mouse and human tissue to study AD, revealing cell-type specific transcriptional changes (*75*, *76*, *78*, *114*, *115*), but the regulators of these disease-associated cell subtypes have yet to be defined.

Moreover, a slew of genetic studies have been performed on AD, identifying multiple associated genetic risk variants (*8*, *9*, *11*, *15*, *116*–*118*). Genome-wide association studies (GWAS) of complex diseases such as AD show that a substantial proportion of genetic risk from common variants partitions to distal regulatory elements, which are often cell-type specific regions in disease-relevant tissues. While much work has gone into intersecting GWAS signals with functional genomics assays, including bulk-tissue RNA-seq and assay for transposase accessible-chromatin with high-throughput sequencing (ATAC-seq) (*90*), the resolution of such studies is noticeably limited by cell-type heterogeneity. A prerequisite for linking GWAS hits to cell-types is a map that links distal regulatory elements with their target genes.

ATAC-seq profiles the open chromatin regions within a tissue and has recently been adapted for single cell resolution (*91*). To date, single-cell chromatin accessibility techniques, such as single-nucleus ATAC-seq (snATAC-seq) have been seldom used in primary samples of diseased tissues, with only two published studies of single-cell chromatin accessibility in the cognitively normal human brain (*71*, *72*). Therefore, we performed snATAC-seq and snRNA-seq in the same AD postmortem human brain tissue samples to define AD-associated gene regulatory programs at the epigenomic and transcriptomic level, providing a powerful lens into the cellular heterogeneity of the brain and allowing us to unravel novel biological pathways underlying neurodegeneration in specific cell populations.

Here, we present a multi-omic analysis of 191,890 nuclei from postmortem human brain tissue of AD and cognitively healthy controls at the single-nucleus resolution, in which we directly integrated snRNA-seq and snATAC-seq datasets, thus providing a more complete understanding of the molecular changes in AD. We identified cell-type specific candidate cis-regulatory elements (cCREs) based on chromatin accessibility and found disease-associated cell subpopulation-specific transcriptomic changes. We identified transcription factors (TFs) that may be regulating AD gene expression changes. Further, we applied pseudotime trajectory analysis on our integrated dataset to extensively characterize disease-associated glial cell states at the epigenomic and transcriptomic level, expanding on previous work exploring gene expression in diverse glial subtypes. We integrated fine-mapped GWAS signals at selected AD risk loci with our snATAC-seq data to link AD risk signals to the specific cell-types in which they are accessible and defined the cis-regulatory chromatin accessibility networks at these loci. Moreover, since network analysis has been effective at clarifying disease transcriptomic signatures in tissue-level RNA-seq data, we designed a co-expression network analysis pipeline, integrating single-cell and bulk-tissue RNA-seq datasets, that robustly identified AD-associated co-expression networks within each cell-type. Altogether, we have clarified the gene regulatory landscape of AD, highlighting the role of glia in AD pathophysiology and identify several genes, namely SREBF1 in oligodendrocytes, for further study in the context of AD. Finally, we provide an online interface for exploration of these datasets (https://swaruplab.bio.uci.edu/singlenucleiAD).

Materials and Methods

Human Samples

Human prefrontal cortex brain samples were obtained from UCI MIND's Alzheimer's Disease Research Center (ADRC) tissue repository and under UCI's Institutional Review Board (IRB). Postmortem human brain tissue from the Religious Orders Study and Memory and Aging Project (ROSMAP) study was obtained under the IRB of Rush University Medical Center. Informed consent was obtained for all human participants. Samples were dissected, homogenized on a dry ice pre-chilled isolating platform and aliquoted for snRNA- and snATAC-seq. For details on human samples used in this study (AD n = 6 males and 6 females, controls n = 5 males and 3 females, all 74-90+ years old), please see Supplementary Tables 1-2[i]. ROSMAP RNA-seq data and details can be found on synapse.org website using corresponding synapse (syn) ID syn3219045.

Bulk RNA-seq

Total RNA was isolated from human prefrontal cortex using Mini Nucleospin RNA kit (Cat #740955.250, Takarabio). RNA integrity was assessed using 2100 Bioanalyzer (Agilent). Total RNA was quantified using Qubit RNA HS assay kit (Cat# Q32852, Invitrogen). ~500ng total RNA was used to prepare the cDNA library using SMARTer Stranded Total RNA Sample Prep kit-HI Mammalian (Cat#634874, Takarabio). cDNA library concentration was calculated using Qubit dsDNA HS assay kit (Cat#Q32851, Invitrogen). Library quality was assessed using either High sensitivity DNA assay kit (Cat# 5067-4626) on 2100 Bioanalyzer or D5000 HS kit (Cat#5067-

5589, 5067-5588) on 4200 Tapestation (Agilent). Libraries were multiplexed with 96 and 95 samples in 2 lanes on an Illumina Novaseq S4 for 100-bp paired-end reads.

Single-nucleus RNA-seq

Single nucleus suspensions were isolated from ~ 50mg frozen human prefrontal cortex. Samples were homogenized in Nuclei EZ Lysis buffer (Cat#NUC101-1KT, Sigma-Aldrich) and incubated for 5 min. Samples were passed through a 70µm filter and incubated in additional lysis buffer for 5 min and centrifuged at 500 g for 5 min at 4°C before two washes in Nuclei Wash and Resuspension buffer (1xPBS, 1% BSA, 0.2U/µl RNase inhibitor). Nuclei were FACS sorted with DAPI (NucBlue Fixed Cell ReadyProbe Reagent, Cat#R37606, Thermo) before running on the 10x Chromium™ Single Cell 3' v3 platform. cDNA library quantification and quality were assessed as in bulk RNA-seq. Libraries were sequenced using Illumina Novaseq 6000 S4 platform at the New York Genome Centre, using 100bp paired-end sequencing.

Single-nucleus ATAC-seq

Single nucleus suspensions were isolated from ~ 50mg frozen human prefrontal cortex according to the 10x Genomics Nuclei Isolation from Mouse Brain Tissue protocol (CG000212, Rev A) with an additional sucrose purification step. Before resuspending our nuclei in Diluted Nuclei Buffer, we removed cellular debris by preparing a sucrose gradient (Nuclei PURE Prep Nuclei Isolation Kit, Cat#NUC201-1KT, Sigma). Nuclei were spun at 13,000 g for 45 minutes at 4°C and then washed once and filtered before running on the 10x Chromium™ Single Cell ATAC platform. Library quantification and quality check were performed according to manufacturer’s recommendations. Libraries were sequenced using Illumina Novaseq 6000 S4 platform at the UCI Genomics core facility, using 100bp paired-end sequencing.

RNAscope (fluorescent in situ hybridization)

Fresh frozen human postmortem tissue was sectioned at 20µm on a cryostat at −20°C. Slides were stored airtight at −80°C until use. Immediately after removing from −80°C, slides were dried for 20 minutes at room temperature and then fixed in 4% paraformaldehyde/PBS for 15 minutes at 4°C. Slides were then washed in RNase-free PBS for 5 minutes at room temperature 3 times. For single labeling experiments, slides were incubated in PBS with an LED light for 96 hours at 4°C to quench autofluorescence (*119*), and for dual labeling, autofluorescence was quenched with TrueBlack (Biotium) for 30 seconds before coverslipping. Slides were processed following the RNAscope Multiplex Fluorescent Reagent Kit v2 Assay (ACD) instructions for fresh frozen tissue, except protease IV incubation was 15 minutes. Probes used were NEAT1 (Cat#411531), PLP1 (Cat#499271), CNP (Cat#509131-C2), SREBF1 (Cat#469871), ACSL4 (Cat#408301), MOG (Cat#543181-C2), and AQP4 (Cat#482441-C2). Fluorophores used were TSA Plus Cy5 (1:200, Perkin Elmer) and Opal 570 (1:200, Perkin Elmer) to avoid autofluorescence. Images were taken on ZEISS Axio Scan.Z1 at 20x magnification. Four regions per section were analyzed using QuPath. We used a trained object classifier to identify MOG+ or AQP4+ nuclei, except for ASCL4/MOG dual staining, which required manual assignment of MOG+ nuclei due to high background. Subcellular detection was used to count RNA punctae. We used linear mixed effects model to account for random effects (age, sex) and fixed effects (multiple regions from the same individual).

Immunofluorescence

Fixed and cryoprotected human postmortem tissue was sectioned at 40µm using a cryotome (Leica). For 6E10, Iba-1, MAP2, and GFAP, brain sections were treated with 90% formic acid for

4 min. For PDGFRA and Olig2, sections in sodium citrate buffer were heated at 80°C in a bead bath for 30 min. Sections were then washed before blocking (PBS with 5% goat or donkey normal serum, respective to the antibodies, and 0.2% TritonX-100) for 1 hour at room temperature. Primary antibodies were incubated at 4°C overnight (6E10-1:1000, Cat#803001, Biolegend; Iba-1-1:1000, Cat#019-19741, Wako; MAP2-1:500, Cat#ab32454, Abcam; GFAP-1:500, Cat#G3893, Sigma; PDGFRA-1:50, Cat#AF-307, R&D Systems; Olig2-1:200, Cat#ab109186, Abcam). Secondary antibodies (Goat anti-mouse 555, Cat#A-21422; Goat anti-rabbit 488, Cat#A11034; Goat anti-rabbit 488, Cat#A11034; Goat anti-mouse 555, Cat#A-21422; Donkey anti-goat 488, Cat#A-11055; Donkey anti rabbit 555, Cat#A31572; all from ThermoFisher) were diluted 1:200 and incubated for 1 hour. Slides were treated with 0.3% Sudan Black in 70% EtOH for 4 min to reduce autofluorescence and imaged on a confocal microscope (Leica). Images from 3 randomly selected areas were used for volume analysis of amyloid plaques using IMARIS. We used linear mixed effects model as previously stated.

Annotation of major cell-types

Major cell-type annotations were assigned to UMAP partitions and initial clusters in snATAC-seq and snRNA-seq datasets respectively through manual inspection of canonical marker gene signals. ‘Pseudo-bulk’ chromatin accessibility coverage profiles of gene body and upstream promoter regions were visualized using the Signac (*120*) (v0.2.0) function CoveragePlot, while gene expression signals were visualized using Seurat (*121*, *122*) (v3.1.2) snRNA-seq cell-type assignments were further validated by integration with the Mathys *et al.* (*75*) dataset.

Integrated analysis of snRNA-seq and snATAC-seq data

A unified dataset of both chromatin accessibility and gene expression was constructed using Seurat's integration framework. Canonical Correlation Analysis (CCA) was used to generate a shared dimensionality reduction of the 'query' snATAC-seq gene activity and the 'reference' snRNA-seq gene expression. MNNs were then identified in this shared space, effectively identifying pairs of corresponding cells that anchor the two datasets together. To confirm major cell-type annotations in snATAC-seq cell populations, we used Seurat's label transfer algorithm, which leverages these anchors to predict cell-types in snATAC-seq data, with cell-type annotations in snRNA-seq cells as the reference and LSI reduced chromatin accessibility as the weights. We achieved a max prediction score >= 0.5 in 94% of cells, demonstrating high correspondence between the two data modalities. Next, we used these shared anchors to impute gene expression signals in snATAC-seq data. Following imputation, we merged gene expression in cells from the snRNA-seq dataset with snATAC-seq cells whose max prediction score >= 0.5. The merged dataset was then centered, dimensionally reduced with PCA using 30 dimensions, batch corrected with MNN (monocle3, v0.2.0) and embedded with UMAP. Clusters and UMAP partitions were identified using Leiden clustering (monocle3). We visualized correspondence of major cell-types from their dataset of origin to their joint UMAP partitions using ggalluvial (v 0.11.1).

Cell-type specific dimensionality reduction and cluster analysis

Cell-type specific analyses were performed for snATAC-seq and snRNA-seq by subsetting each major cell-type from the fully processed dataset followed by re-embedding with UMAP. Subpopulations of each cell-type used for all downstream analysis were then identified using Leiden clustering (monocle3). Clusters smaller than 100 cells were removed as outliers. We then used the addReproduciblePeakSet function from the R package ArchR (v1.0.0) (*123*) with default parameters to call accessible chromatin peaks using MACS2 (v2.2.7.1) in each cell-type subcluster. For snRNA-seq and snATAC-seq clusters, we performed a bootstrapped cluster composition analysis to robustly assess the composition of each cluster with respect to AD diagnosis. Over 25 iterations, 20% of cells were sampled from the whole dataset, and the proportion of cells from AD or control samples were computed for each cluster. A two-sided Wilcoxon rank sum test was used to compare the proportion of AD and control samples in each cluster using the wilcox.test R (v3.6.1) function with default parameters and Benjamini-Hochberg multiple testing correction.

Annotation of cell subpopulations

snRNA-seq subpopulations for astrocytes, microglia, neurons, and oligodendrocyte progenitors were annotated in a similar way to the major cell-types, using canonical marker gene signals as well as differentially expressed genes. snATAC-seq subpopulations for astrocytes, microglia, neurons, and oligodendrocyte progenitors were annotated using Seurat label transfer prediction scores with the snRNA-seq clusters as a reference annotation. We annotated the snRNA-seq oligodendrocytes by hierarchically clustering oligodendrocyte and oligodendrocyte progenitor clusters based on the gene expression matrix of the top 25 DEGs (by average log fold-change) from each oligodendrocyte subpopulation, grouping oligodendrocytes into major lineage classes such as progenitor, intermediate, and mature. We used the same approach to annotate the snATAC-seq oligodendrocytes, hierarchically clustering the gene activity matrix using the same DEGs. The R package ComplexHeatmap (v 2.7.6.1010) (*124*) was used for hierarchical clustering and visualization of these gene expression and gene activity matrices.

Single-nucleus Transcription Factor (TF) binding motif analysis

Single-nucleus TF motif enrichment was computed for a set of 452 TFs from the JASPAR 2018 database (*125*) using the Signac wrapper for chromVAR (v 1.12.0) (*126*). The motif accessibility matrix was first computed, describing the number of peaks that contain each TF motif for all cells. chromVAR then uses this motif accessibility matrix to compute deviation Z-scores for each motif by comparing the number of peaks containing the motif to the expected number of fragments in a background set that accounts for confounding technical factors such as GC content bias, PCR amplification, and variable Tn5 tagmentation. To further analyze specific TFs of interest, we used the getFootprints function in ArchR to perform TF footprinting analysis in pseudo-bulk aggregates of single nuclei in the same cluster or cell-type, splitting nuclei from control or late-stage AD samples where appropriate.

Chromatin Cis Co-Accessibility Network (CCAN) analysis

The correlation structure of chromatin accessibility data was analyzed using the R package cicero (v1.3.4.7) (*127*). Briefly, cicero quantifies ‘co-accessibility’ between pairs of genomic regions in a population of cells by correlating accessibility signals aggregated from several cells at a time, penalizing for distance between regions using a graphical LASSO with a maximum interaction constraint of 500 kb. Importantly, prior to correlation and regularization, a bootstrap approach was used to generate metacells by aggregating 50 cells at a time using k-nearest neighbors, circumventing the sparsity of single-cell chromatin data. Finally, networks of cis co-accessible regions (CCANs) were identified through community detection. We applied this procedure in each major cell-type as well as splitting each cell-type into control and AD cells for CCAN analysis.

Analysis of gene-linked candidate cis-regulatory elements (gl-cCREs)

We sought to further contextualize co-accessible chromatin regions by linking them to likely target genes using an accessibility-expression correlation strategy stratified by major cell-type and disease status of each sample. First, we identified pairs of co-accessible peaks where one of the peaks overlaps a gene's promoter, which serves as a candidate target gene for that particular cCRE. We then computed the Pearson correlation between the expression of the candidate target gene from snRNA-seq with the chromatin accessibility of the linked cCRE from snATAC-seq, where expression and accessibility values have been averaged for all cells within a given cell population. This correlation analysis was performed iteratively across all promoter-cCRE co-accessible links identified separately in each major cell with regard to AD diagnosis status. Retaining links with Pearson correlation coefficient in the 95th percentile and p-value <= 0.01, we defined gene-linked candidate cis-regulatory elements (gl-cCREs) as genomic regions with a significant correlation to at least one target gene, and we defined cCRE-linked genes as genes with a significant correlation to at least one cCRE. We used non-negative matrix factorization (NMF v 0.23.0) as implemented in the R NMF package (*128*) using k=30 matrix factors on the gl-cCRE accessibility matrix averaged by each snATAC-seq cluster split by cells from control and AD samples, yielding 30 gl-cCRE modules. The NMF basis matrix (W) was used to assign each gl-cCRE to its top associated module, and the NMF coefficient matrix (H) was used to determine which cell cluster that each module was most associated to. To identify biological process associated with these gl-cCRE modules, we used the enrichR (v 3.0) (*129*, *130*) package to query enriched GO terms for the set of target genes in each gl-cCRE module in the GO Biological Processes 2018, GO Cellular Component 2018, and GO Molecular Function 2018 databases.

TF regulatory network construction

Using snATAC-seq and snRNA-seq data in one cell-type, we identified candidate TF regulatory target genes and used this information to construct cell-type specific TF regulatory networks. We used the same set of TF binding motifs as in our single-cell TF motif enrichment analysis (JASPAR 2018 motifs). For a given TF, we defined candidate target genes as those with an accessible promoter containing the TF binding motif, or an accessible gl-cCRE linked to the target gene's promoter, allowing us to distinguish between TFs that regulate genes through promoter or enhancer binding events. We used this information to construct a directed TF regulatory network using the R package igraph (v 1.2.6), where each vertex represents a TF or target gene, and each edge represents a promoter or linked-cCRE binding event, overlaying additional information onto the network such as DEG or AD GWAS gene status.

Estimating GWAS enrichment using cluster specific accessible chromatin regions

To estimate heritability of a variety of complex traits, we used LDSC (v 1.0.1) (*131*). GWAS summary statistics were input to LDSC, which then computes enrichment of heritability for an annotated set of SNPs conditioned on a baseline model in order to account for genomic features that influence heritability, and jointly modeling multiple annotations together. Sets of cluster specific peaks were constructed by extending peaks up and down stream by 2000 bp, identifying peaks that are accessible in 1% of all cells within each cluster, and removing all peaks that are accessible in more than one other cell-type. Cluster specific peaks were formatted for LDSC using the make_annotation.py script, and LD scores were computed for each set using the ldsc.py script. Publicly available GWAS summary statistics were collected for AD (*9*, *15*), Schizophrenia (*132*), Frontotemporal Dementia (FTD) (*133*), Progressive Supranuclear Palsy (PSP) (*134*), Multiple Sclerosis (MS) (*135*), Inflammatory Bowel Disease (IBD) (*136*), height (*137*), and cholesterol (*138*). Next, summary statistics were converted to hg38 coordinates using the UCSC liftover tool (v377) and formatted for LDSC using the munge_sumstats python script. We followed the recommended guidelines for cell-type specific partitioned heritability analysis, using HapMap3 SNPs and their provided hg38 baseline model (v2.2). The ldsc.py script was then used to compute cluster specific enrichments of GWAS heritability, with Benjamini-Hochberg multiple testing correction applied to enrichment p-values.

Single-nucleus Consensus Weighted Gene Co-expression Network Analysis (scWGCNA)

We developed a novel co-expression network analysis approach to single-cell data by integrating snRNA-seq and bulk-tissue RNA-seq datasets and called this approach Single-nucleus Consensus Weighted Gene Co-expression Network Analysis (scWGCNA). scWGCNA is based on a co-expression network analysis approach called Weighted Gene Co-expression Network Analysis (WGCNA), implemented using the WGCNA R package (v1.69) (*139*, *140*). For scWGCNA, we used multiple transcriptomic datasets comprising of our snRNA-seq data, Mathys *et al*. snRNA-seq data, bulk-tissue RNA-seq data from our UCI cohort and bulk tissue RNA-seq data from ROSMAP cohort (*46*). First, we integrated our snRNA-seq and Mathys *et al*. snRNA-seq datasets using iNMF approach, and then constructed metacells in a fashion similar to our CCAN analysis of chromatin accessibility data, in which we apply a bootstrapped aggregation process to single-nucleus transcriptomes. During metacell computation, we only pool cells within the same cell-type, and within the same AD diagnosis stage, in order to retain these metadata for scWGCNA. We then employed a signed consensus WGCNA approach (*45*) within a given cell-type, by calculating component-wise values for topological overlap for each dataset. First, bi-weighted mid-correlations were calculated for all pairs of genes, and then a signed similarity matrix was created. In the signed network, the similarity between genes reflects the sign of the correlation of their expression profiles. The signed similarity matrix was then raised to power β, varies between cell-types, to emphasize strong correlations and reduce the emphasis of weak correlations on an exponential scale. The resulting adjacency matrix was then transformed into a topological overlap matrix. Modules were defined using specific module cutting parameters which included minimum module size of 100 genes, deepSplit = 4 and threshold of correlation = 0.2. Modules with correlation greater than 0.8 were merged together. We used first principal component of the module, called the module eigengene, to correlate with diagnosis and other variables. Hub genes were defined using intra-modular connectivity (kME) parameter of the WGCNA package. Gene-set enrichment analysis was done using EnrichR.

Analysis of regulatory targets of SREBP

We downloaded a dataset of ENCODE ChIP-seq validated TF target genes from EnrichR, containing regulatory targets of SREBP. Fisher's enrichment tests were performed with the R function fisher.test to test whether oligodendrocyte modules were significantly over-represented with SREBP target genes, inferring which modules are regulated by SREBP. Module eigengenes were computed for the set of SREBP target genes, and the RNA expression as well as protein expression data from Inweb (*141*) and Biogrid (*142*) was also to analyze SREBP targets throughout the course of AD progression. A protein-protein interaction (PPI) network of SREBP target genes was constructed using SREBF1 ChIP-seq data from ENCODE and visualized using the STRING database (*143*), restricting the edges to known protein-protein interactions. In addition to bulk RNA-seq, we used a proteomics dataset from our group's previous study (*144*) of 685 samples representing AD, asymptomatic AD, and controls from the human PFC to interrogate the levels of SREBF1 target genes and target proteins in AD.

Statistics

All statistical methods and tests used in the manuscript were described in the figure legends, Methods, Supplementary Note[i], or main text as appropriate.

Results

Multi-omic analysis of the human prefrontal cortex

We performed both snATAC-seq (10x Genomics; n=12 late-stage AD; n=8 control) and snRNA-seq (10x Genomics v3; n=11 late-stage AD; n=7 control) on nuclei isolated from the prefrontal cortex (PFC) using postmortem human tissue from late-stage AD and age-matched cognitively healthy controls (74-90+ years old, Fig. 2.1a). We defined late-stage AD and controls based on both Braak and plaque staging (Supplementary Tables 1-2[i]). We specifically aimed to generate both transcriptomic and epigenetic data from the same tissue sample (aliquots of samples from the same dissection, see Methods) to minimize differences in cell-type composition between the two methods, thus allowing for meaningful downstream integrated analysis. After quality control filtering, we retained a total of 130,418 nuclei for snATAC-seq and 61,472 nuclei for snRNA-seq (Methods, Supplementary Fig. 1-2, Supplementary Table 3, Supplementary Note[i]). To ensure the rigor of our study, we applied batch correction methods to the data from both assays, since library preparation limitations required multiple batches. For snATAC-seq, we used mutual nearest neighbors (MNN) (*145*) to correct the Latent Semantic Indexing (LSI) reduced chromatin accessibility matrix, and for snRNA-seq we used integrative Non-negative Matrix Factorization (iNMF) (*146*) to reduce dimensionality while simultaneously eliminating batch effects (Methods, Extended Data Fig. 2.1, Supplementary Note[i]). We applied Uniform Manifold Approximation and Projection (UMAP) (*147*) dimensionality reduction and Leiden clustering (*148*) to the batch-corrected epigenomic and transcriptomic datasets, identifying distinct cell-type clusters in snATAC-seq (35) and snRNA-seq (34, Fig. 2.1b-c). With snATAC-seq, we profiled all major cell-types of the brain—excitatory neurons (24,076 nuclei, EX.a-e), inhibitory neurons (9,644 nuclei, INH.a-d), astrocytes (15,399 nuclei, ASC.a-f), microglia (12,232 nuclei, MG.a-e), oligodendrocytes (62,253 nuclei, ODC.a-m), and oligodendrocyte progenitor cells (4,869 nuclei, OPC.a)—annotated based on chromatin accessibility at the promoter regions of known marker genes (Fig. 2.1d, Extended Data Fig. 2.2). We used chromVAR (*126*) to compute TF motif variability in single nuclei by estimating the enrichment of TF binding motifs in accessible chromatin regions (Methods) and examined the enrichment of TF motifs by cell-type in respect to diagnosis, identifying several TF motifs with increased enrichment with disease in astrocytes, excitatory neurons, and microglia (Supplementary Fig. 3, Supplementary Data 1[i]). Moreover, we performed TF footprinting analysis to further clarify cell-type-specific TF regulation, highlighting the SOX9 TF footprint in oligodendrocytes. Interestingly, we noticed TF motif enrichment of oligodendrocyte-related TFs in excitatory neurons. Likewise, we detected similar cell-types using snRNA-seq— excitatory neurons (6,369 nuclei, EX1-5), inhibitory neurons (5,962 nuclei, INH1-4), astrocytes (4,756 nuclei, ASC1-4), microglia (4,126 nuclei, MG1-3), oligodendrocytes (37,052 nuclei, ODC1-13), and oligodendrocyte progenitor cells (2,740 nuclei, OPC1-2)—classified by the gene expression of cell-type markers (Fig. 2.1e). In both assays, oligodendrocytes were the most commonly profiled cell-type (Supplementary Fig. 3[i]). Additionally, while many differentially expressed genes (DEGs) in each major cell-type agreed with previous literature, we also found cluster-specific genes previously established as neuronal or glial subtype markers, such as *LINC00507* for L2-3 excitatory neurons (EX1) (*56*), *SV2C* for L3 interneurons (INH4) (*54*), and *CX3CR1* for homeostatic microglia (MG2) (*66*) (Fig. 2.1h, Supplementary Fig. 3-4, Supplementary Data 1[i]).

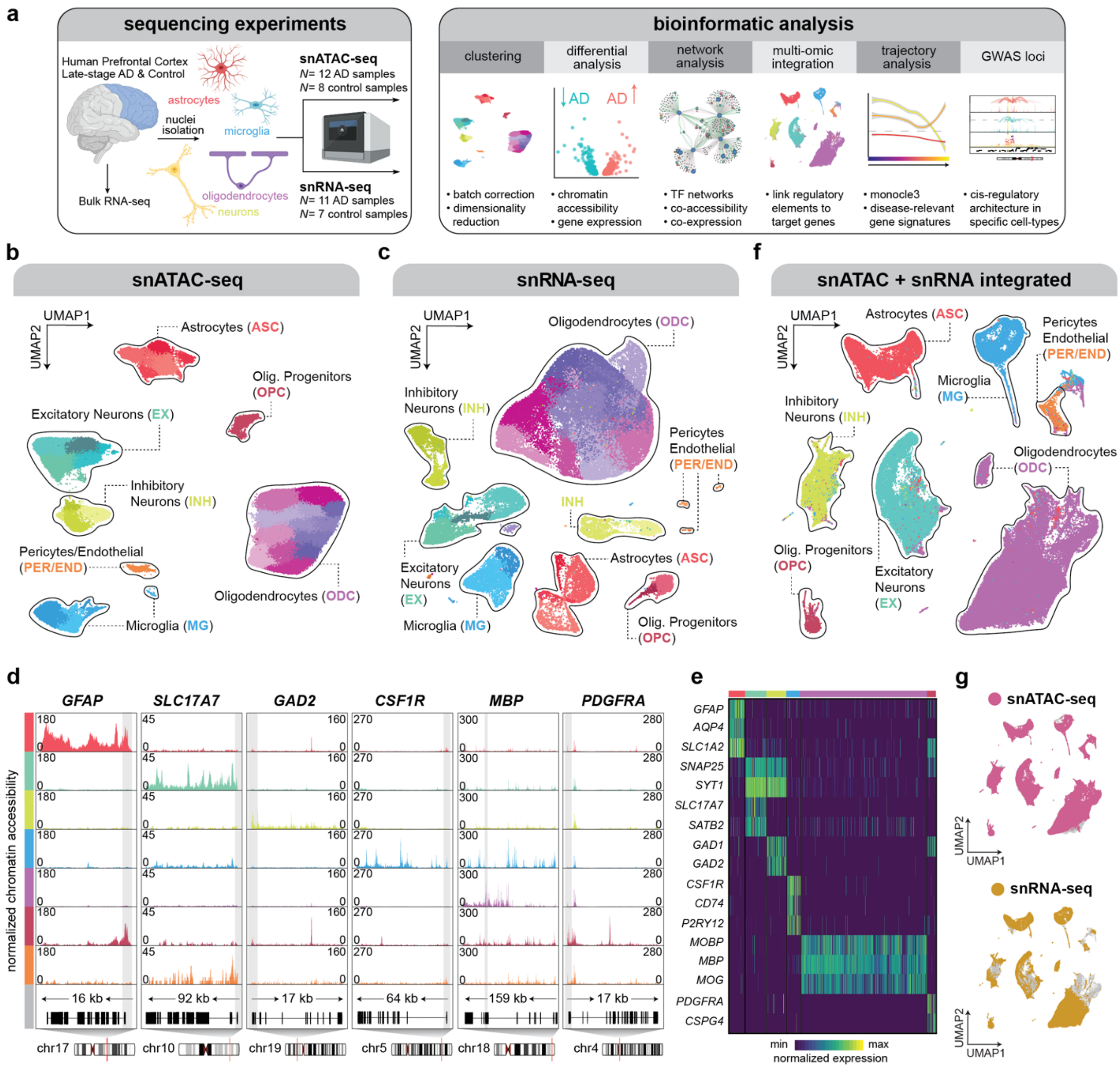
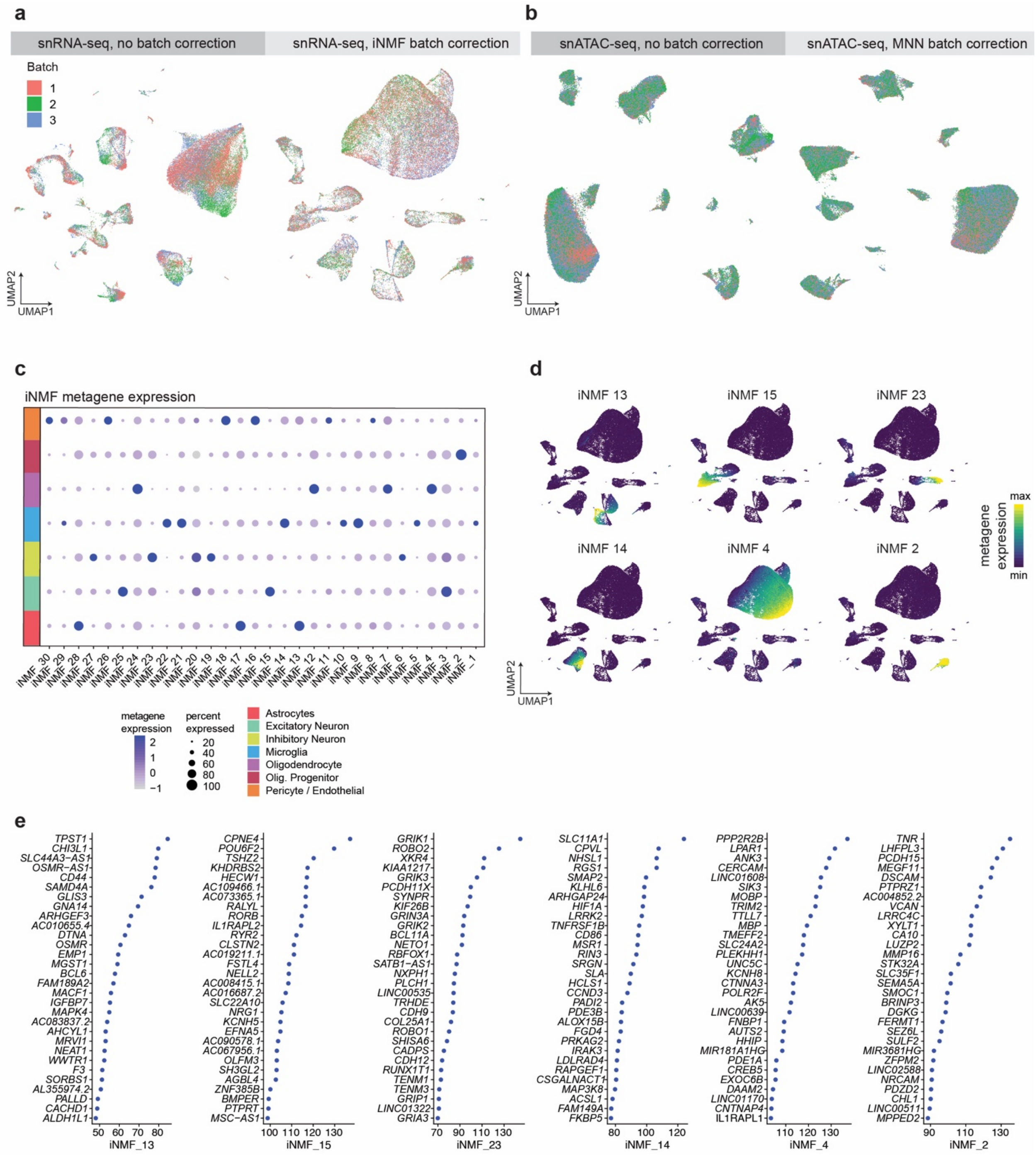


Figure 2.1. Single-nucleus ATAC-seq and single-nucleus RNA-seq to study cellular diversity in the diseased brain

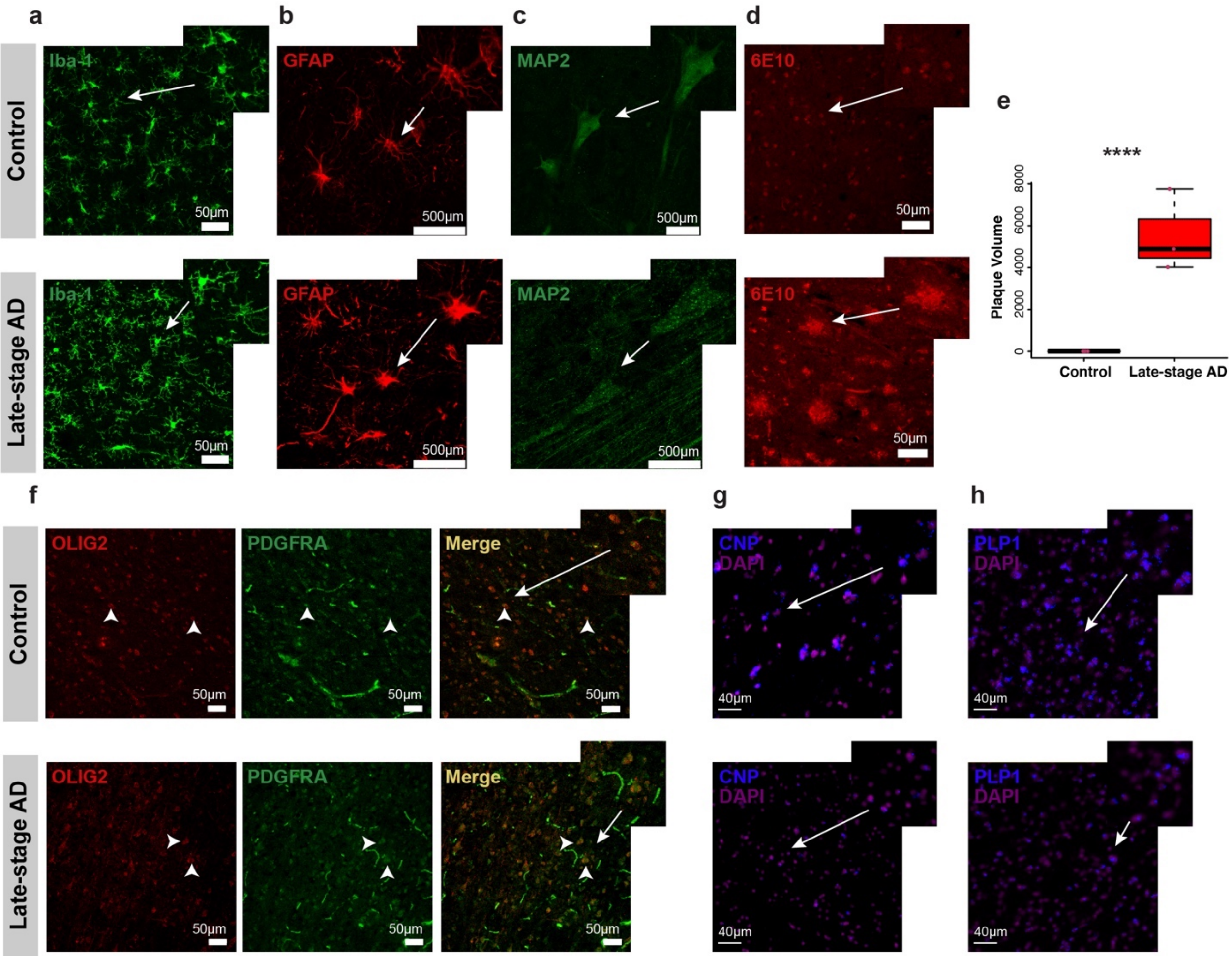
a, Schematic representation of the samples used in this study, sequencing experiments, and downstream bioinformatic analyses, created with BioRender.com. **b, c,** UMAP visualizations where dots correspond to individual nuclei for 130,418 nuclei profiled with snATAC-seq (**b)** and 61,472 nuclei profiled with snRNA-seq **(c)**, colored by Leiden cluster assignment and cell-type (ASC = astrocytes, EX = excitatory neurons, INH = inhibitory neurons, MG = microglia, ODC = oligodendrocytes, OPC = oligodendrocyte progenitor cells, PER/END = pericytes/endothelial cells). **d,** Pseudo-bulk chromatin accessibility profiles for each cell-type at canonical cell-type marker genes. For each gene, 1kb upstream and downstream are shown. Promoter/TSS highlighted in grey with gene model and chromosome position shown below. Chromosome coordinates are the following: *GFAP* chr17:44904008-44919937; *SLC17A7* chr19:49428401-49445360; *GAD2* chr10:26213307-26305558; *CSF1R* chr5:150052291-150116372; *MBP* chr18:76977827-77136683; *PDGFRA* chr4:54226097-54299247. **e,** Row-normalized single-nucleus gene expression heatmap of cell-type marker genes. **f,** UMAP plot of 186,167 nuclei from a jointly learned subspace of snATAC-seq and snRNA-seq, colored by cell-type assignment. **g,** Integrated UMAP as in **f**, colored by originating dataset. Smaller gray dots represent nuclei from the other data modality.

Since the epigenomic landscape is deeply intertwined with downstream gene expression signatures, we integrated our snATAC-seq and snRNA-seq datasets using Seurat's integration platform (*121*, *122*) (Methods, Fig. 2.1f, Extended Data Fig. 2.3, Supplementary Fig. 3[i]). Cell-types that were independently classified using chromatin data or transcriptome data overwhelmingly grouped together in the integrated UMAP space (Fig. 2.1g, Supplementary Fig. 3[i]). Using the same biological samples in snATAC-seq and snRNA-seq resulted in a high degree of overlap between nuclei from these two data modalities in the jointly constructed space. Additionally, we confirmed cell-type identities by gene activity and gene expression in a panel of canonical cell-type marker genes (Supplementary Fig. 3[i]) and used Seurat's label transfer algorithm to verify cell-type annotations in the snATAC-seq dataset using the snRNA-seq dataset as a reference (Supplementary Fig. 5[i]).



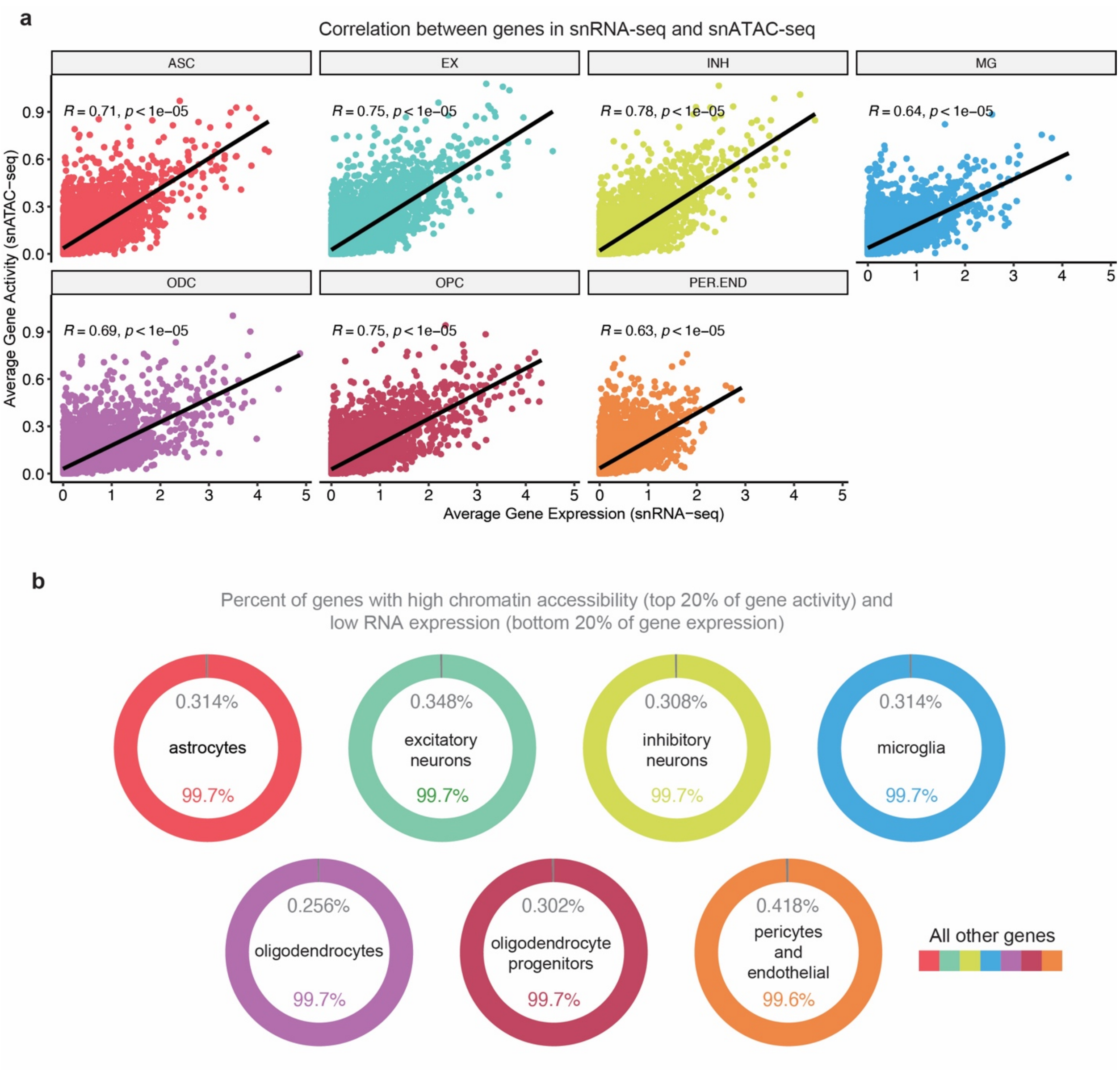
Extended Data Figure 2.1. Batch correction of snATAC-seq, snRNA-seq, and merged datasets.

a, snRNA-seq UMAPs before (left) and after iNMF batch correction (right), colored by sequencing batch. **b**, snATAC-seq UMAPs before (left) and after MNN batch correction (right), colored by sequencing batch. **c,** Dot plot of iNMF metagene expression in each snRNA-seq cluster. **d**, snRNA-seq UMAPs colored by metagene expression of selected iNMF metagenes. **e**, Dot plots showing the iNMF loading for the top 30 genes for the same metagenes in **d.**



Extended Data Figure 2.2. Cell-type immunostaining and in situ hybridization.

a-d, Representative immunofluorescence images from postmortem human brain tissue from control and late-stage AD cases for Iba-1 **(a)**, GFAP **(b)** MAP2 **(c)**, and 6E10 **(d)**. **e,** Quantification of 6E10-positive amyloid plaques. n = 3 cognitively healthy controls, 3 late-stage AD. Data is presented as the average of three different sections per sample. Linear mixed-effects model **** $p < 0.0001$. Box boundaries and line correspond to the interquartile range (IQR) and median respectively. Whiskers extend to the lowest or highest data points that are no further than 1.5 times the IQR from the box boundaries. **f,** Representative immunofluorescence images from postmortem human brain tissue from control and late-stage AD cases for OLIG2 with PDGFRA co-labeling. **g, h,** Representative RNAscope images from postmortem human brain tissue from control and late-stage AD cases for CNP **(g)** and PLP1 **(h)** with DAPI counterstain.



Extended Data Figure 2.3. Comparison of gene expression and gene activity.

a, Scatter plot comparing average gene activity from snATAC-seq and average gene expression from snRNA-seq by each major cell-type, with accompanying Pearson correlation statistics and linear regression lines. **b**, Donut chart showing the percent of genes with high chromatin accessibility and low gene expression in grey for each major cell-type. High chromatin accessibility was defined as genes in the top 20% of gene activity, while low gene expression was defined as genes in the bottom 20% of gene expression. Percent of all other genes are colored by the cell-type.

Multi-omic characterization of AD cellular heterogeneity

In both snATAC-seq and snRNA-seq, we discovered multiple neuronal and glial subpopulations, and we annotated the subpopulations from snRNA-seq based on previously identified marker genes (*54*, *56*) (Fig. 2.2, Supplementary Fig. 6-7, Supplementary Note[i]). For our snATAC-seq clusters, we used Seurat label transfer to calculate cluster prediction scores allowing for supervised annotation of our cell clusters, in which we mapped EX.a to EX1 and ASC.b to ASC2, for example (Supplementary Fig. 6-7[i]). We examined the composition of each cluster in the context of disease and found several that are significantly over- or under-represented in late-stage AD compared to control, in both data modalities (Fig. 2.2d-g, Methods). ASC3 (*GFAP*$^{\text{high}}$/*CHI3L*$^{+}$) significantly increased in proportion with disease (bootstrapped cluster proportion analysis using a two-sided Wilcoxon rank sum test, FDR = 8.63 x 10^{-5}), whereas ASC4 (*GFAP*$^{\text{low}}$/*WIF*1^{+}/*ADAMTS17*$^{+}$) significantly decreased (FDR = 4.68 x 10^{-7}), consistent with a recent snRNA-seq study of the 5XFAD mouse model of AD (*79*). We also found that the proportion of MG.a. and MG.b was increased in late-stage AD (FDR = 9.82 x 10–7, 8.88 x 10^{-10}), both of which mapped to the activated snRNA-seq cluster MG1 (*SPP1*$^{\text{high}}$/*CD163*$^{+}$), which was also increased with disease (FDR = 6.32 x 10^{-7}). Additionally, we found that immune oligodendrocyte cluster ODC13 was significantly increased in late-stage AD (FDR = 1.62 x 10^{-4}).

Further, we identified both differentially accessible chromatin regions (DARs) and differentially expressed genes (DEGs) in late-stage AD for each cell cluster and found high cluster specificity for GO term enrichment of distal and proximal DARs, as well as DEGs (Methods, Supplementary Fig. 7-9, Supplementary Data 1-6, Supplementary Note[i]). For example, we identified *NEAT1* as upregulated in astrocytes and oligodendrocytes, in agreement with previous findings in the entorhinal cortex (*114*), and we confirmed AD upregulation of *NEAT1* with in situ hybridization (Extended Data Fig. 2.4). Altogether, we found cluster-specific epigenetic and transcriptomic changes in late-stage AD, which may underlie the dysregulation of distinct biological pathways in different cell subpopulations in neurodegeneration.

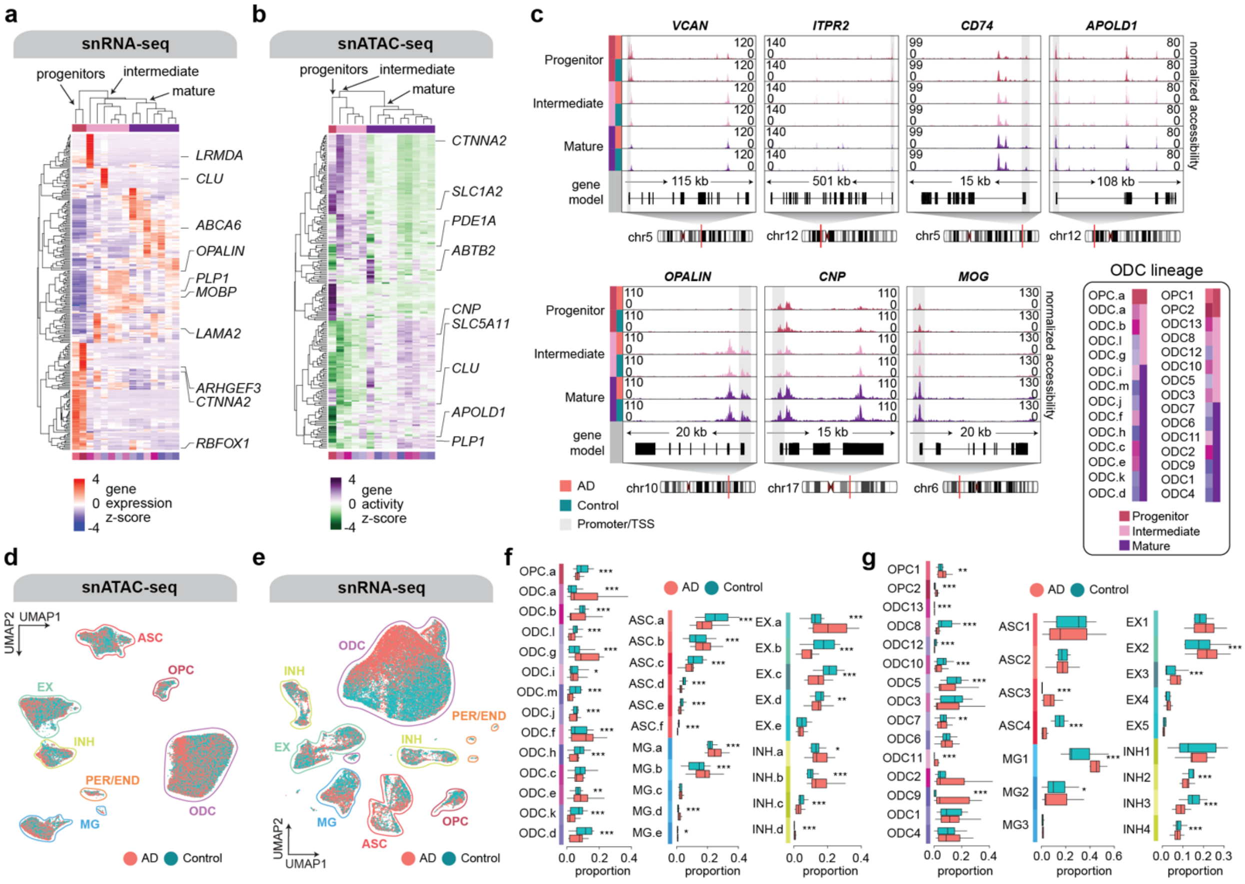
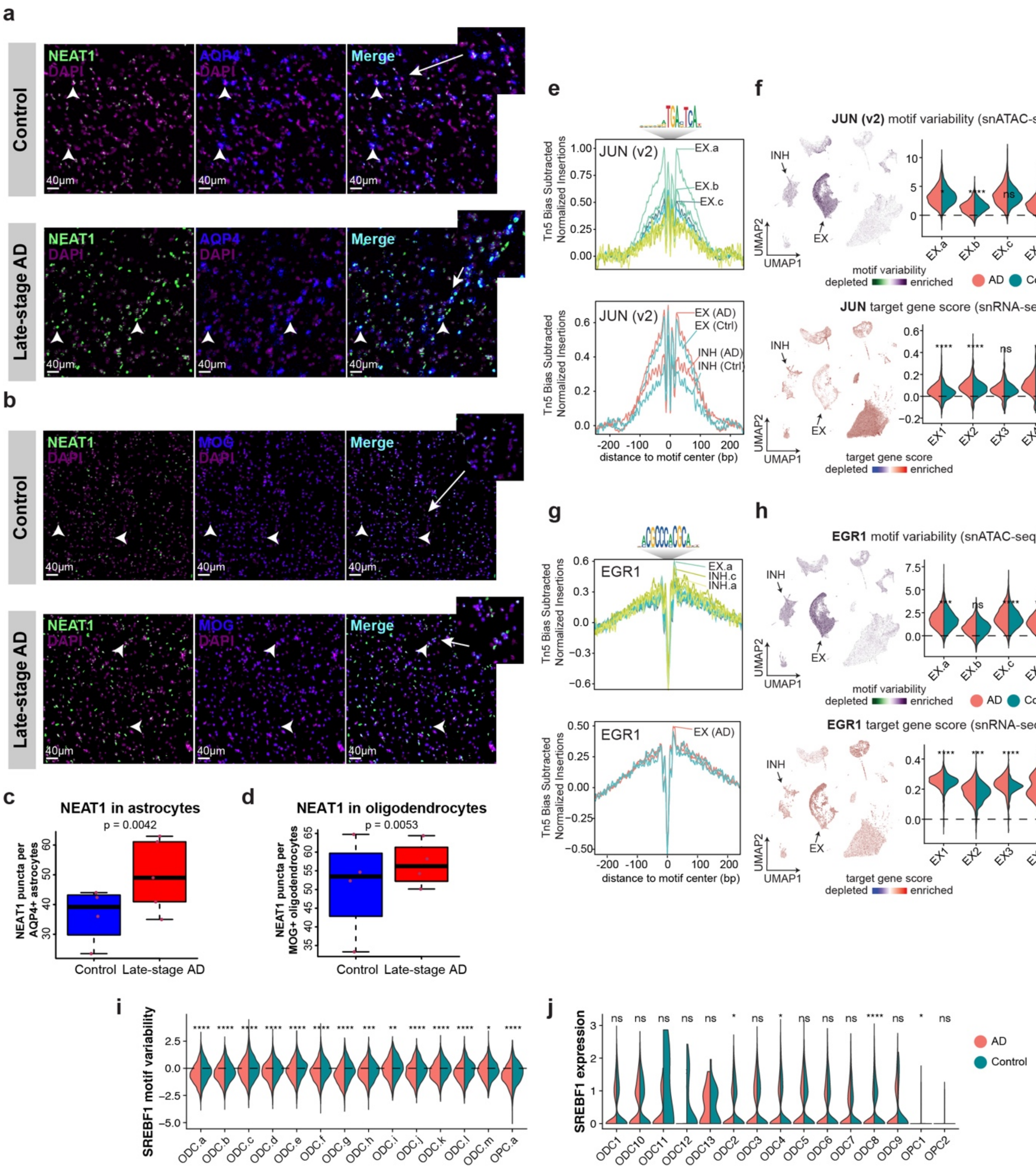


Figure 2.2. Epigenetically and transcriptionally distinct cell subpopulations in human AD prefrontal cortex

a,b, Hierarchically clustered heatmaps of row-normalized gene expression in snRNA-seq OPC and oligodendrocyte clusters **(a)** and gene activity in snATAC-seq OPC and oligodendrocyte clusters **(b)** for the top 25 upregulated DEGs (sorted by average log fold change) identified in each oligodendrocyte subpopulation. **c,** Pseudo-bulk chromatin accessibility coverage profiles for OPC (progenitor), intermediate oligodendrocyte and mature oligodendrocyte snATAC-seq clusters, assignments as in **b.** Promoter/TSS highlighted in grey with gene model and chromosome position shown below. Chromosome coordinates are the following: *VCAN* chr5: 83468465-83583303; *ITPR2* chr12: 26335515-26836198; *CD74* chr5: 150400637-150415929; *APOLD1* chr12: 12722917-12830975; *OPALIN* chr10: 96342216-96362365; *CNP* chr17: 41963741-41978731; *MOG* chr6: 29653981-29673372. **d,e,** snATAC-seq **(d)** and snRNA-seq **(e)** UMAPs as in Fig. 2.1, where nuclei are colored by AD diagnosis. Clusters annotated by cell-type. **f,g,** Box and whisker plots showing the proportion of nuclei mapping to each cluster for each sample, split by control and late-stage AD samples for snATAC-seq **(i)** and snRNA-seq **(j)** clusters, with measures of significance from bootstrapped cluster composition analysis (Wilcoxon test, see Methods, *** FDR <= 0.001, ** FDR <= 0.01, * 0.01 < FDR <= 0.05) and n as in Supplementary Tables 7-9[i]. For box and whisker plots, box boundaries and line correspond to the interquartile range (IQR) and median respectively. Whiskers extend to the lowest or highest data points that are no further than 1.5 times the IQR from the box boundaries.

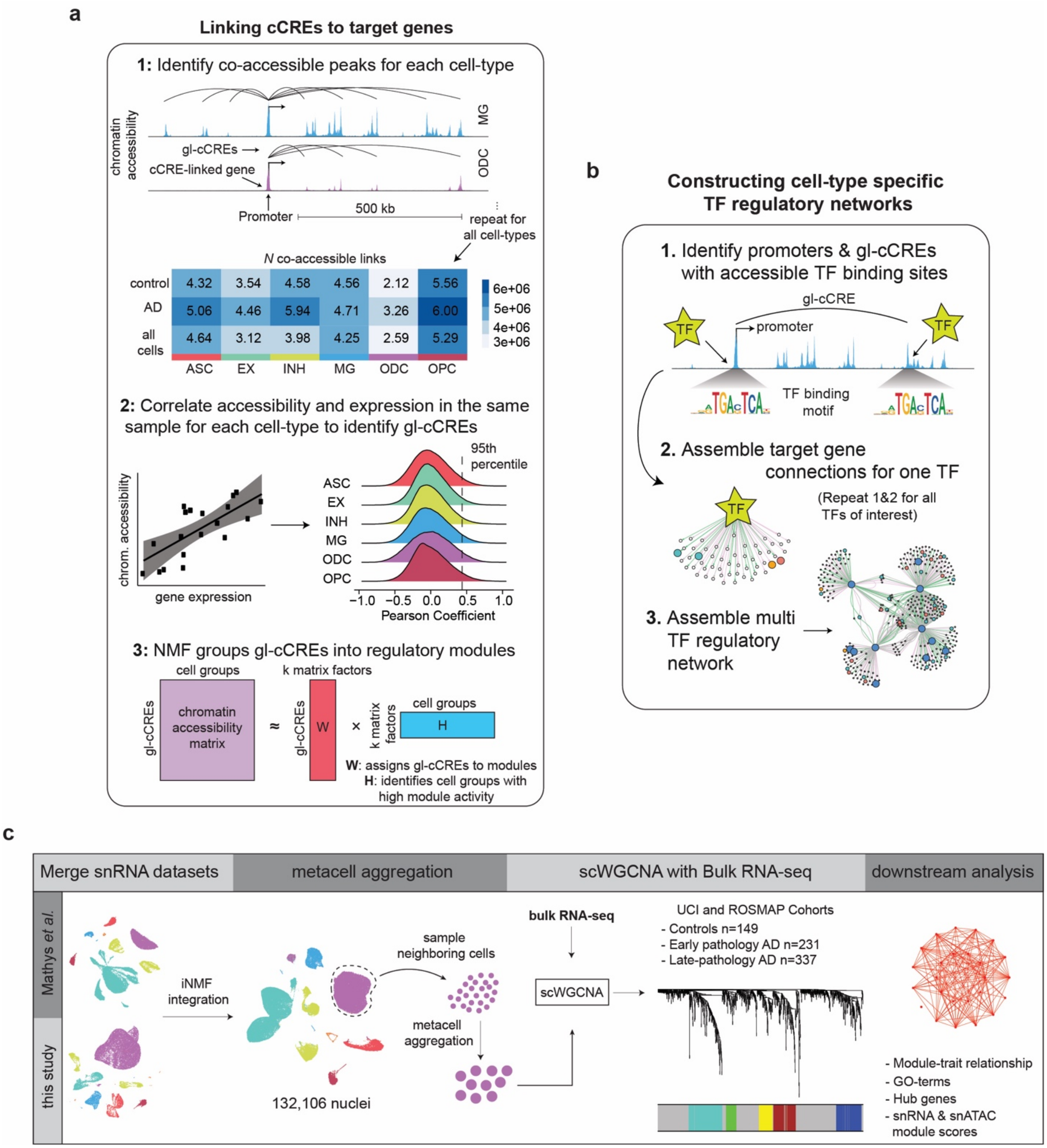


Extended Data Figure 2.4. NEAT1 validation and neuronal TFs.

a, b, Representative RNAscope images from postmortem human brain tissue for *NEAT1* and *AQP4* staining **(a)** and *NEAT1* and *MOG* staining **(b)** with DAPI nuclear counterstain. **c,** Boxplots showing quantification of *NEAT1* puncta per *AQP4+*astrocyte as in **a**. n = 4 cognitively healthy controls, 5 late-stage AD. **d,** Boxplots showing quantification of *NEAT1* puncta per *MOG+* oligodendrocyte as in **b**. n = 4 cognitively healthy controls, 4 late-stage AD. Data is represented as the mean of four equally sized regions per sample. Linear mixed-effects model **e,** Tn5 bias subtracted TF footprinting for JUN by snATAC-seq neuron cluster (top) and by AD diagnosis (bottom), with TF binding motif logo above and Tn5 bias insertions below. **f,** Left: Co-embedding UMAP colored by JUN motif variability (top) and JUN target gene score (bottom). Right: Violin plots of JUN motif variability (top) and JUN target gene score (bottom) in excitatory neuron clusters, split by diagnosis. Wilcoxon test (ns: $p > 0.05$, *: $p <= 0.05$, **: $p <= 0.01$, ***: $p <= 0.001$, ****: $p <= 0.0001$). **g,** Tn5 bias subtracted TF footprinting for EGR1 by snATAC-seq neuron cluster (top) and by AD diagnosis (bottom), as in **e**. **h,** Left: Co-embedding UMAP colored by EGR1 motif variability (top) and EGR1 target gene score (bottom). Right: Violin plots of EGR1 motif variability (top) and *EGR1* target gene score (bottom) in excitatory neuron clusters, split by diagnosis, as in **f. i,**Violin plots of SREBF1 motif variability in oligodendrocyte snATAC-seq clusters, as in **f. j**, Violin plots of SREBF1 gene expression in oligodendrocyte snRNA-seq clusters, as in **i**. For boxplots, box boundaries and line correspond to the interquartile range (IQR) and median respectively. Whiskers extend to the lowest or highest data points that are no further than 1.5 times the IQR from the box boundaries.

Cell-type-specific cis-gene regulation in late-stage AD

Based on our experimental design utilizing both snATAC-seq and snRNA-seq in the same samples, we reasoned that we could identify the target genes of cCREs in specific cell populations (Extended Data Fig. 2.5a, Methods). To this end, we sought to elucidate the cis-regulatory architecture of the PFC in late-stage AD by constructing cis co-accessibility networks (CCANs) (*127*) separately for late-stage AD and control in each cell-type (Methods). To identify target genes of cCREs, we focused on the subset of co-accessible peaks where one of the peaks lies in a promoter element, yielding a set of cCREs and candidate target genes. For this set of co-accessible links, we correlated the expression of the candidate target gene to the chromatin accessibility of the cCRE, strengthening the evidence of a potential regulatory relationship beyond co-accessibility alone. Finally, we used NMF to analyze and cluster these gene-linked cCREs (gl-cCREs) based on their chromatin accessibility in each cell cluster. In sum, this process results in a set of candidate enhancer elements (gl-cCREs) grouped into functional modules, as well as a set of cCRE-linked genes, for each major cell-type in late-stage AD and control.



	ASC	EX	INH	MG	ODC	OPC
control	4.32	3.54	4.58	4.56	2.12	5.56
AD	5.06	4.46	5.94	4.71	3.26	6.00
all cells	4.64	3.12	3.98	4.25	2.59	5.29

Extended Data Figure 2.5. Schematics of analyses.

a, Schematic diagram linking cCREs to target genes and downstream analysis. First, we identify co-accessible chromatin peaks in each cell-type for control and late-stage AD. Second, we identify pairs of co-accessible peaks where one peak overlaps a gene promoter and correlate the expression of that gene with the chromatin accessibility of the other peak. Third, NMF is used to group gl-cCREs into functional modules. **b,** Schematic of construction of TF regulatory networks for each cell-type. **c,** Schematic representation of scWGCNA analysis, including iNMF integration with the Mathys *et al.* 2019 dataset, *metacell* aggregation, construction of co-expression networks, and downstream analysis of gene modules.

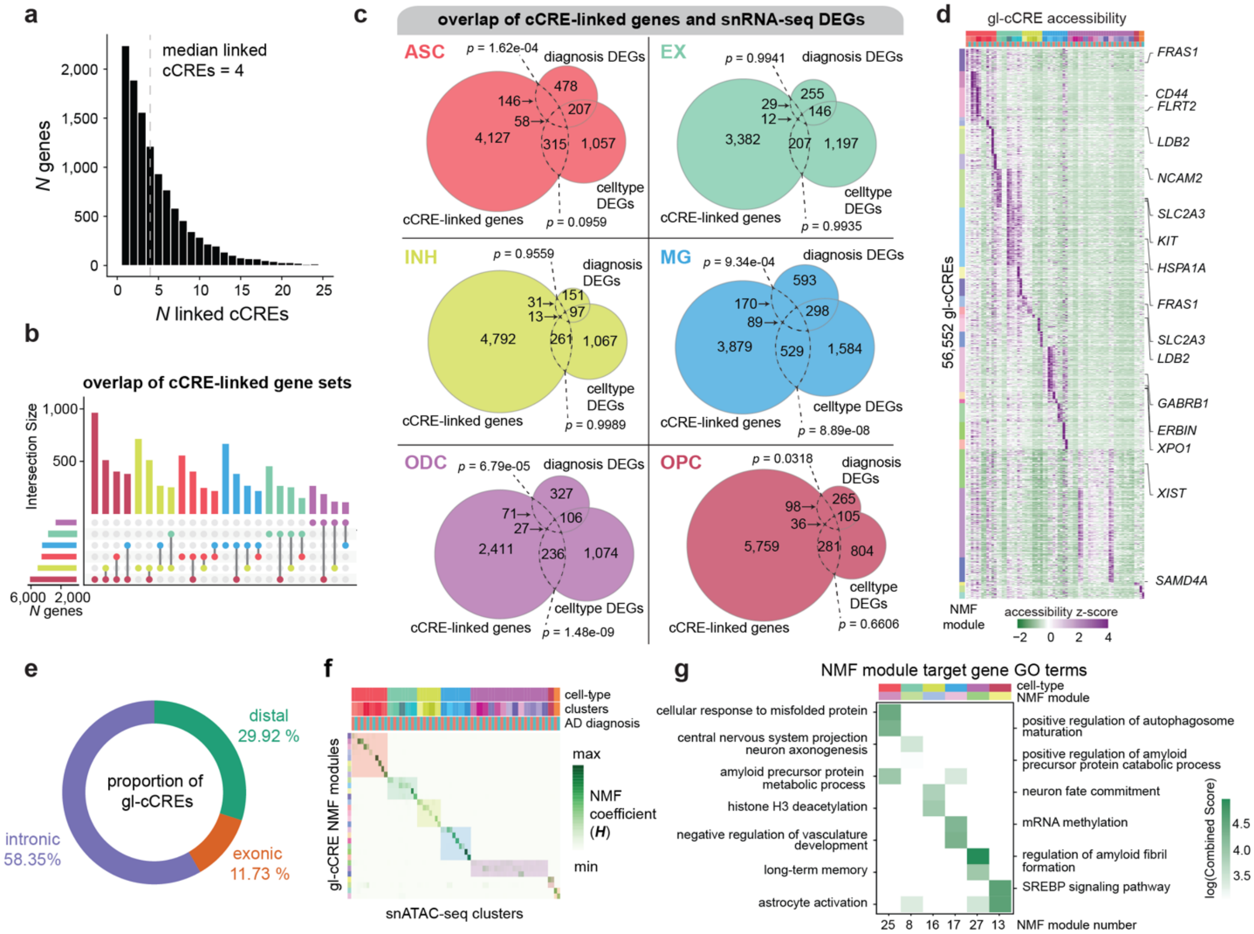


Figure 2.3. Linking cis-regulatory elements to downstream target genes in specific cell-types

a, Histogram showing the number of genes that have 1 through 25 linked cCREs. **b**, Upset plot showing the size of overlaps between the sets of cCRE-linked genes identified in each cell-type. The barplot on the left shows the set size of cCRE-linked genes for each cell-type, and the barplot on the top shows the number of overlapping genes between two sets, or the number of unique genes in one set. **c,** Venn diagrams for each major cell-type showing the overlaps between the set of cCRE-linked genes and genes upregulated in that cell-type (celltype DEGs) and genes upregulated in AD within this cell-type (diagnosis DEGs). A one-sided Fisher's exact test was used for gene set overlap significance (*** $p <= 0.001$, ** $p <= 0.01$, * $p < 0.05$). **d**, Heatmap showing row-normalized pseudo-bulk chromatin accessibility in each snATAC-seq cluster split by nuclei from control and late-stage AD samples. Rows (cCREs) are organized based on NMF module assignment. Annotations correspond to genes from DGE analysis that are upregulated in AD in at least one cell-type. **e**, Donut chart showing the percentage of gl-cCREs that map to intronic, exonic, or distal regions. **f,** Heatmap showing NMF coefficients in each snATAC-seq cluster split by nuclei from control and late-stage AD samples. **g**, Heatmap showing log transformed enrichR combined scores for GO terms for gene sets of selected NMF modules.

In total, using this approach we identified 56,552 gl-cCREs and 11,440 cCRE-linked genes, with a median of 4 cCREs linked to each of these genes (Fig. 2.3a, Supplementary Tables 4-5[i]). By examining the overlap between sets of cCRE-linked genes identified in each cell-type, we observed a substantial number of genes with linked cCREs that are shared across multiple cell-types, in addition to those that are cell-type specific (Fig. 2.3b). For several cell-types, we found a significant overlap between the set of cCRE-linked genes and cell-type marker DEGs, as well as genes that are upregulated in AD within that cell-type, highlighting a critical role of cCREs in disease-related gene expression changes (Fig. 2.3c). We also investigated the chromatin accessibility in each snATAC-seq cluster for these gl-cCREs and noted a high degree of cell-type and cluster specificity (Fig. 2.3d). The majority of the gl-cCREs mapped to intronic regions (58.35%) (Fig. 2.3e). Moreover, by inspecting the NMF coefficient matrix (H), we were able to identify which cluster or cell-type each NMF module corresponds to, and we annotated several modules that are specific to control or late-stage AD nuclei within a given cluster (Fig. 2.3f-g, Supplementary Note[i]). Additionally, we found that some of the cCRE target genes that are common to more than one cell-type are regulated by different cCREs in each cell-type.

Cell-type-specific transcription factors in late-stage AD

To complement our analysis of cis-regulatory elements, we sought to identify cell-type specific trans-regulatory elements in late-stage AD. TFs tightly control cell fate in neurodevelopment and have been implicated in neurodegenerative processes. We examined the regulatory role of microglial TF SPI1 (also known as PU.1) and nuclear respiratory factor 1 (NRF1) in oligodendrocytes (Figure 4a-f, Supplementary Fig. 10, Supplementary Note[i]). SPI1 motif variability in our snATAC-seq microglia clusters was significantly increased in only upregulated clusters MG.a and MG.b, but SPI1's targets were significantly downregulated in only MG1 (Fig. 2.4a-b, Supplementary Fig. 10[i]). We also identified NRF1 is dysregulated in select oligodendrocyte clusters (Fig. 2.4d-f, Supplementary Fig. 10[i]). These results indicate that SPI1 acts as a transcriptional repressor in late-stage AD, providing insight into how SPI1 contributes to AD pathophysiology. Additionally, NRF1 has previously been associated with mitochondrial function, and impaired mitochondrial function (*149*), mediated by NRF1 dysregulation, may contribute to neuronal dysfunction in late-stage AD through the disruption of myelination. TF analyses in neuronal populations and Fos related antigen 2 (FOSL2) in astrocytes are shown in Extended Data Fig. 2.4 and Supplementary Fig. 10[i].

To gain further insight into TF-mediated gene regulation in late-stage AD, we constructed cell-type specific TF regulatory networks. For a given TF, we identified candidate target genes as those whose promoters or linked cCREs are accessible and contain the TF's binding motif in the cell-type of interest, and we repeated this for several select TFs, generating microglia-specific and oligodendrocyte-specific TF regulatory networks (Fig. 2.4g-h, Extended Data Fig. 2.5b, Supplementary Note[i]). Within these networks we identified multiple AD DEGs, in addition to genes located at known AD GWAS loci, regulated by SPI1 in microglia and NRF1 in oligodendrocytes.

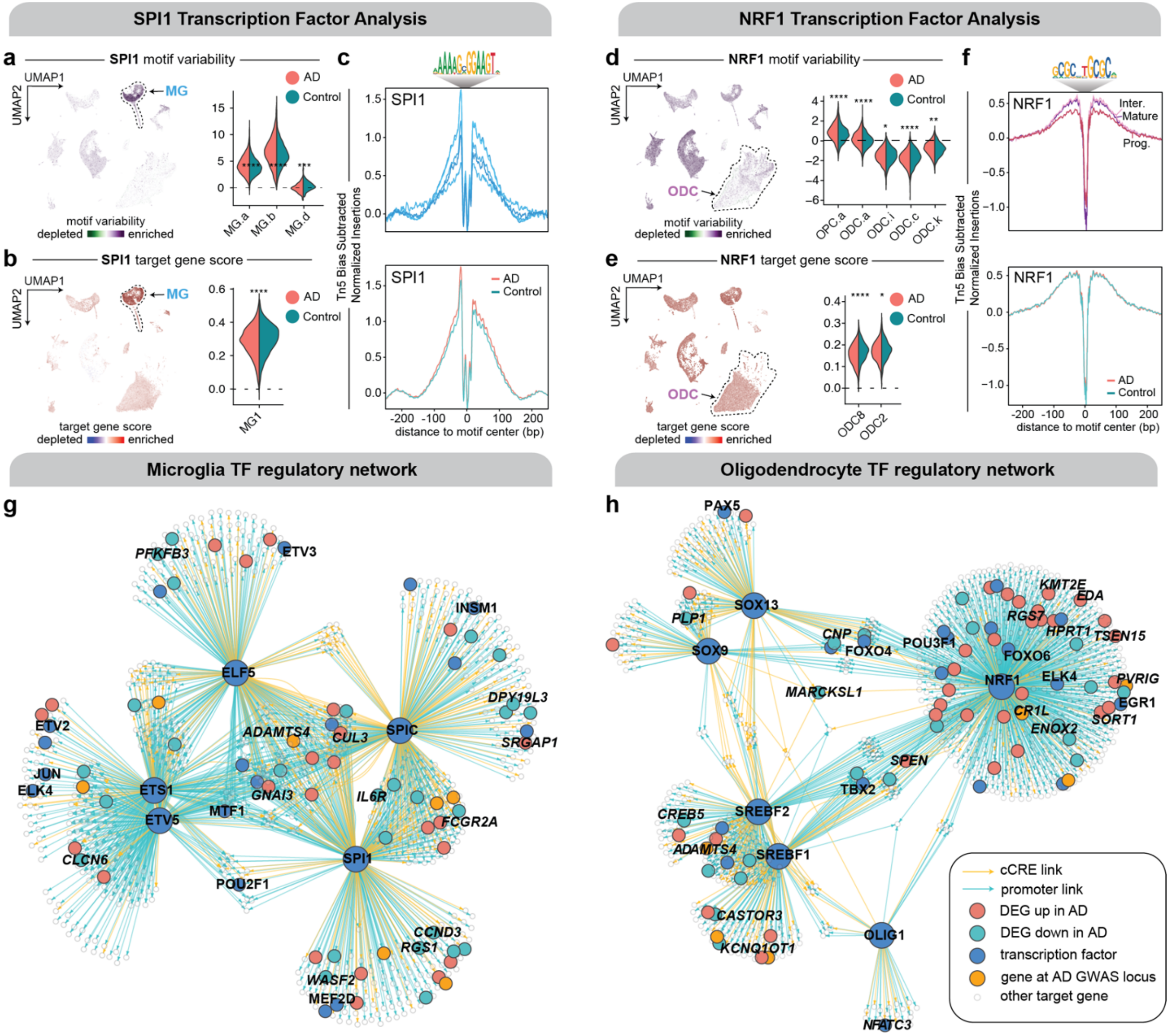


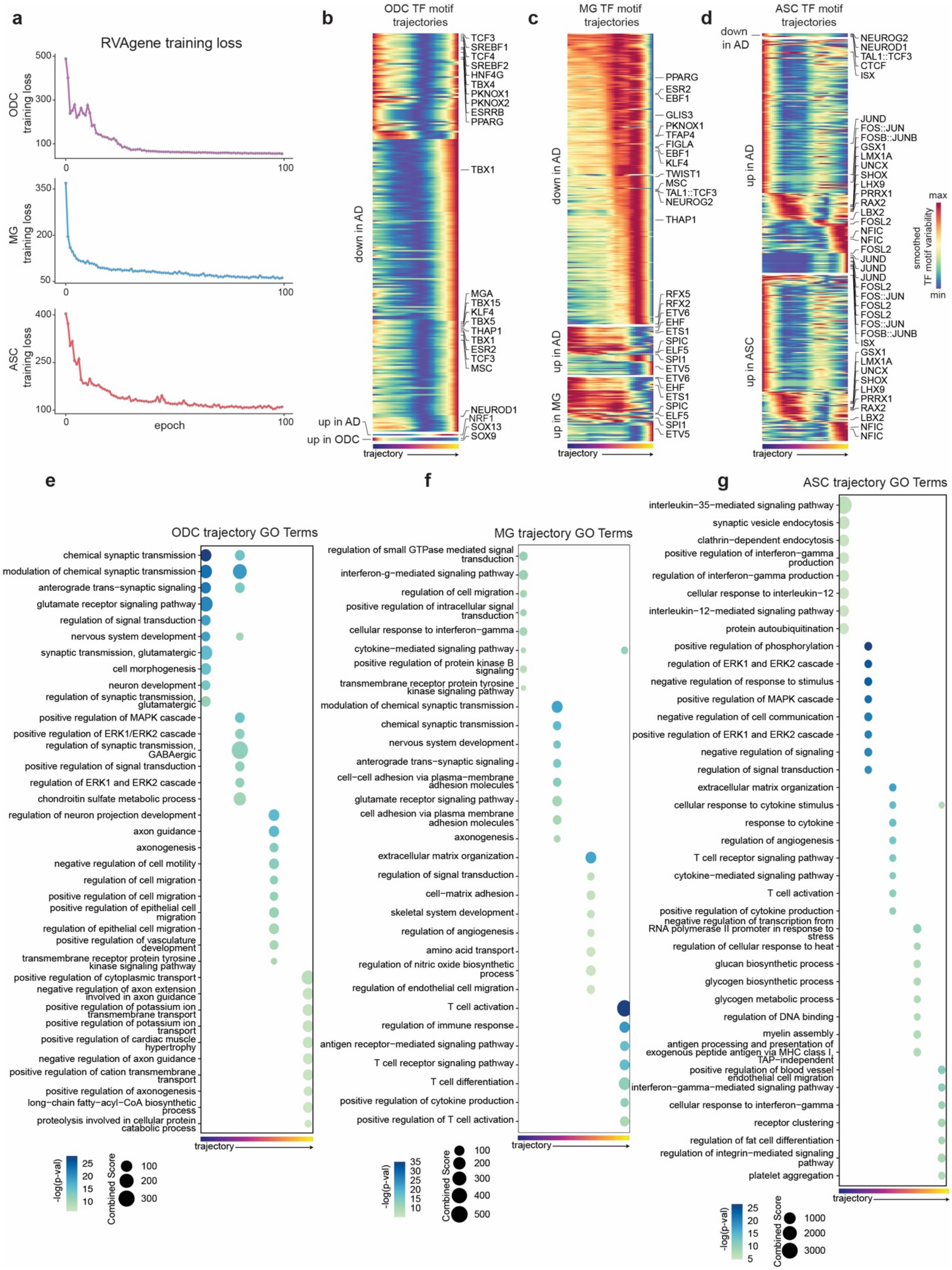
Figure 2.4. Cell subpopulation-specific transcription factor regulation in late-stage AD

a, Left: snATAC-seq and snRNA-seq integrated UMAP colored by SPI1 motif variability with microglia circled. Right: Violin plots of SPI1 motif variability in significant snATAC-seq microglia clusters, split by diagnosis. **b,** Left: Integrated UMAP colored by SPI1 target gene score with microglia circled. Right: Violin plots of SPI1 target gene score in significant snRNA-seq microglia clusters, split by diagnosis as in **a**. **c,** Tn5 bias subtracted TF footprinting for SPI1 by snATAC-seq microglia cluster (top) and by AD diagnosis (bottom). TF binding motif shown as motif logo above. **d,** Left: Integrated UMAP colored by NRF1 motif variability with oligodendrocytes circled. Right: Violin plots of NRF1 motif variability in significant snATAC-seq oligodendrocyte clusters, split by diagnosis as in **a**. **e,** Left: Integrated UMAP colored by NRF1 target gene score with oligodendrocyte circled. Right: Violin plots of NRF1 target gene score in significant snRNA-seq oligodendrocyte clusters, split by diagnosis as in **a**. **f,** Tn5 bias subtracted TF footprinting for

NRF1 by snATAC-seq oligodendrocyte cluster (top) and by AD diagnosis (bottom) as in **b**. **g, h,** TF regulatory networks showing the predicted candidate target genes for the following TFs: ELF5, ETS1, ETV5, SPIC, and SPI1 in microglia **(g)**; SOX9, SOX13, SREBF1, SREBF2, OLIG1, and NRF1 in oligodendrocytes **(h)**. For violin plots, two-sided Wilcoxon test was used to compare control versus AD, ns: $p > 0.05$, *: $p <= 0.05$, **: $p <=0.01$, ***: $p <= 0.001$, ****: $p <= 0.0001$.

Integrated trajectory analysis of disease-associated glia

To further uncover molecular mechanisms driving glial heterogeneity in AD, we performed pseudotime trajectory analysis using monocle3 (*150–152*) on the integrated snATAC-seq and snRNA-seq data in oligodendrocytes, microglia, and astrocytes (Supplementary Note[i]). Multi-omic trajectory analysis allows us to investigate the dynamics of gene expression, chromatin accessibility, and TF motif variability throughout a continuum of cell-state transitions. We modeled gene expression and chromatin accessibility dynamics using a recurrent variational autoencoder (RVAE) (*153*). Briefly, RVAE is an encoder-decoder neural network framework that uses long short-term memory (LSTM) units to effectively model temporal biological data, yielding a two-dimensional latent representation of the input features as well as a de-noised reconstructed version of the original input (Supplementary Note[i]). For each cell-type, we identified genes that are differentially expressed along the trajectory (t-DEGs, Supplementary Data 7[i]) and used these genes as features to train the RVAE until the loss function converged (Supplementary Note[i], Extended Data Fig. 2.6).


a
RVAgene training loss
ODC training loss
500
300
100
0
100
MG training loss
350
200
50
0
100
ASC training loss
400
250
100
0
epoch
100
b
ODC TF motif trajectories
down in AD
up in AD
up in ODC
trajectory
TCF3
SREBF1
TCF4
SREBF2
HNF4G
TBX4
PKNOX1
PKNOX2
ESRRB
PPARG
TBX1
MGA
TBX15
KLF4
TBX5
THAP1
TBX1
ESR2
TCF3
MSC
NEUROD1
NRF1
SOX13
SOX9
c
MG TF motif trajectories
down in AD
up in AD
up in MG
trajectory
PPARG
ESR2
EBF1
GLIS3
PKNOX1
TFAP4
FIGLA
EBF1
KLF4
TWIST1
MSC
TAL1::TCF3
NEUROG2
THAP1
RFX5
RFX2
ETV6
EHF
ETS1
SPIC
ELF5
SPI1
ETV5
ETV6
EHF
ETS1
SPIC
ELF5
SPI1
ETV5
d
ASC TF motif trajectories
down in AD
up in AD
up in ASC
trajectory
NEUROG2
NEUROD1
TAL1::TCF3
CTCF
ISX
JUND
FOS::JUN
FOSB::JUNB
GSX1
LMX1A
UNCX
SHOX
LHX9
PRRX1
RAX2
LBX2
FOSL2
NFIC
NFIC
FOSL2
JUND
JUND
JUND
FOSL2
FOS::JUN
FOSL2
FOSL2
FOS::JUN
FOSB::JUNB
ISX
GSX1
LMX1A
UNCX
SHOX
LHX9
PRRX1
RAX2
LBX2
NFIC
NFIC
max
smoothed TF motif variability
min
e
ODC trajectory GO Terms
chemical synaptic transmission
modulation of chemical synaptic transmission
anterograde trans−synaptic signaling
glutamate receptor signaling pathway
regulation of signal transduction
nervous system development
synaptic transmission, glutamatergic
cell morphogenesis
neuron development
regulation of synaptic transmission, glutamatergic
positive regulation of MAPK cascade
positive regulation of ERK1/ERK2 cascade
regulation of synaptic transmission, GABAergic
positive regulation of signal transduction
regulation of ERK1 and ERK2 cascade
chondroitin sulfate metabolic process
regulation of neuron projection development
axon guidance
axonogenesis
negative regulation of cell motility
regulation of cell migration
positive regulation of cell migration
positive regulation of epithelial cell migration
regulation of epithelial cell migration
positive regulation of vasculature development
transmembrane receptor protein tyrosine kinase signaling pathway
positive regulation of cytoplasmic transport
negative regulation of axon extension involved in axon guidance
positive regulation of potassium ion transmembrane transport
positive regulation of potassium ion transport
positive regulation of cardiac muscle hypertrophy
negative regulation of axon guidance
positive regulation of cation transmembrane transport
positive regulation of axonogenesis
long−chain fatty−acyl−CoA biosynthetic process
proteolysis involved in cellular protein catabolic process
trajectory
-log(p-val)
25
20
15
10
Combined Score
100
200
300
f
MG trajectory GO Terms
regulation of small GTPase mediated signal transduction
interferon-g-mediated signaling pathway
regulation of cell migration
positive regulation of intracellular signal transduction
cellular response to interferon−gamma
cytokine−mediated signaling pathway
positive regulation of protein kinase B signaling
transmembrane receptor protein tyrosine kinase signaling pathway
modulation of chemical synaptic transmission
chemical synaptic transmission
nervous system development
anterograde trans−synaptic signaling
cell−cell adhesion via plasma−membrane adhesion molecules
glutamate receptor signaling pathway
cell adhesion via plasma membrane adhesion molecules
axonogenesis
extracellular matrix organization
regulation of signal transduction
cell−matrix adhesion
skeletal system development
regulation of angiogenesis
amino acid transport
regulation of nitric oxide biosynthetic process
regulation of endothelial cell migration
T cell activation
regulation of immune response
antigen receptor−mediated signaling pathway
T cell receptor signaling pathway
T cell differentiation
positive regulation of cytokine production
positive regulation of T cell activation
trajectory
-log(p-val)
35
30
25
20
15
10
Combined Score
100
200
300
400
500
g
ASC trajectory GO Terms
interleukin−35−mediated signaling pathway
synaptic vesicle endocytosis
clathrin−dependent endocytosis
positive regulation of interferon−gamma production
regulation of interferon−gamma production
cellular response to interleukin−12
interleukin−12−mediated signaling pathway
protein autoubiquitination
positive regulation of phosphorylation
regulation of ERK1 and ERK2 cascade
negative regulation of response to stimulus
positive regulation of MAPK cascade
negative regulation of cell communication
positive regulation of ERK1 and ERK2 cascade
negative regulation of signaling
regulation of signal transduction
extracellular matrix organization
cellular response to cytokine stimulus
response to cytokine
regulation of angiogenesis
T cell receptor signaling pathway
cytokine−mediated signaling pathway
T cell activation
positive regulation of cytokine production
negative regulation of transcription from RNA polymerase II promoter in response to stress
regulation of cellular response to heat
glucan biosynthetic process
glycogen biosynthetic process
glycogen metabolic process
regulation of DNA binding
myelin assembly
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP−independent
positive regulation of blood vessel endothelial cell migration
interferon−gamma−mediated signaling pathway
cellular response to interferon−gamma
receptor clustering
regulation of fat cell differentiation
regulation of integrin−mediated signaling pathway
platelet aggregation
trajectory
-log(p-val)
25
20
15
10
5
Combined Score
1000
2000
3000

Extended Data Figure 2.6. Pseudotime trajectory analysis to identify dysregulated TFs and gene expression in glia.
a, Line plot showing the RVAgene training loss at each epoch for oligodendrocyte (ODC), microglia (MG), and astrocyte (ASC) RVAE models. **b-d,** Heatmaps showing TF motif variability smoothed using loess regression and scaled to minimum and maximum values for TFs up- and down-regulated in AD as well as cell-type marker TFs along the oligodendrocyte trajectory **(b)**, microglia trajectory **(c)**, and astrocyte trajectory **(d)**. TFs are ordered by trajectory rank (point in trajectory where of 75% maximum value is reached). **e-g,** Dot plot showing the enrichR combined score for the top enriched GO terms in oligodendrocyte **(e)**, microglia **(f)**, and astrocyte **(g)** t-DEGs.

Oligodendrocyte trajectory reveals SREBF1 dysregulation

We constructed an integrated oligodendrocyte trajectory using 58,221 nuclei from snATAC-seq and 36,773 nuclei from snRNA-seq (Fig. 2.5a), noting that the proportion of nuclei from late-stage AD samples appears to increase along the trajectory (Fig. 2.5b, Pearson correlation R = 0.32, p-value = 0.022). To clarify the functional state of oligodendrocytes associated with late-stage AD, we examined the gene expression signatures (*68*, *70*) of newly formed oligodendrocytes (NF-ODC), myelin-forming oligodendrocytes (MF-ODC), and mature oligodendrocytes (mature ODC) (Fig. 2.5c, see Supplementary Note[i] for gene signature lists). Interestingly, we found that the mature oligodendrocyte gene expression signature increased at the end of the trajectory, whereas the myelin-forming oligodendrocyte gene signature decreased. In addition, the newly formed oligodendrocyte gene signature decreased throughout the trajectory, altogether suggesting that the oligodendrocyte pseudotime trajectory appears to recapitulate oligodendrocyte maturation. Chromatin accessibility of 9,231 oligodendrocyte gl-cCREs and gene expression of 1,563 oligodendrocyte t-DEGs reconstructed with RVAE showcases the vast amount of chromatin remodeling and transcriptional reprogramming that may be underlying oligodendrocyte maturation (Fig. 2.5d).

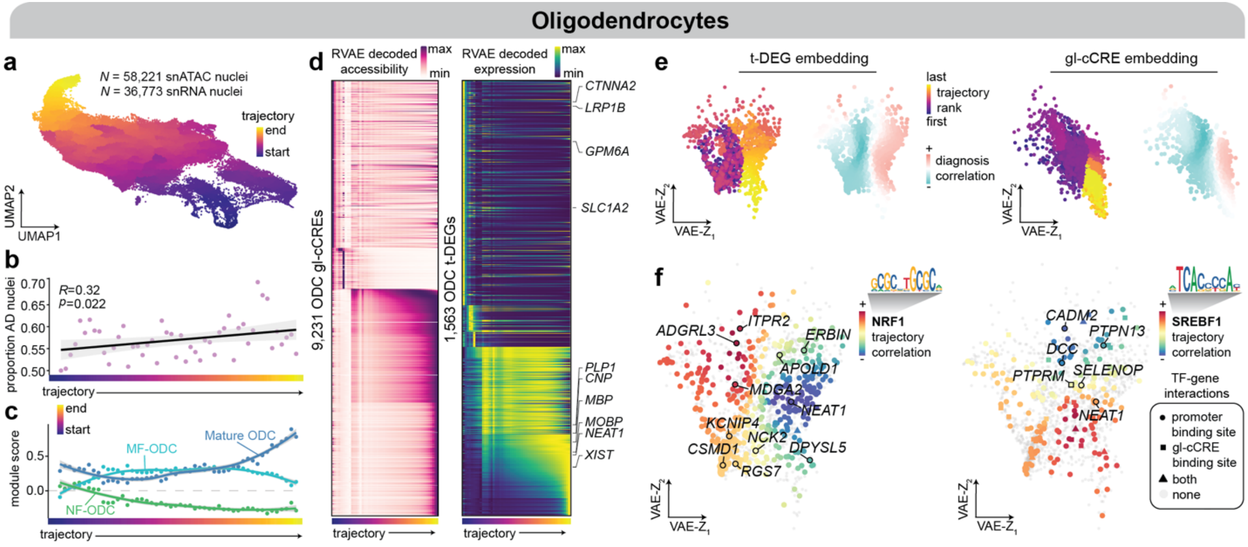


Figure 2.5. Multi-omic oligodendrocyte trajectory analysis

a, UMAP dimensionality reduction of oligodendrocytes from the integrated snATAC-seq (n=58,221 nuclei) and snRNA-seq (n=36,773 nuclei) analysis. Each cell is colored by its pseudotime trajectory assignment. **b,** Scatter plot showing the proportion of oligodendrocyte nuclei from AD samples at 50 evenly sized bins across the trajectory. The black line shows a linear regression, and the gray outline represents the 95% confidence interval. Pearson correlation coefficient and p-value from two-sided test are shown. **c,** Scatter plot of module scores for newly formed oligodendrocyte (NF-ODC), myelin forming oligodendrocyte (MF-ODC) and mature oligodendrocyte gene signatures (*68*, *70*) (see Supplementary Note[i] for full gene lists) averaged for nuclei in each of the 50 trajectory bins. Solid colored lines represent loess regressions for each signature, and the gray outlines represent 95% confidence intervals. **d,** Left: heatmap of chromatin accessibility at 9,231 oligodendrocyte gl-cCREs reconstructed using RVAE. Right: heatmap of gene expression for 1,563 oligodendrocyte trajectory DEGs (t-DEGs) reconstructed using RVAE. Annotated genes are DEGs in oligodendrocytes, in respect to other cell-types, or AD upregulated genes in oligodendrocytes. **e,** 2D latent space learned by RVAE modeling of oligodendrocyte t-DEGs (left) and gl-cCREs (right), where each dot represents one gene. Left: genes colored by trajectory rank, the point in the trajectory where the gene reaches 75% of max expression. Right: genes colored by correlation of RVAE reconstructed expression with AD diagnosis proportion as in **b**. **f,** Oligodendrocyte t-DEG latent space colored by correlation of reconstructed gene expression to NRF1 (left) and SREBF1 (right) motif variability. The shape of each point represents the regulatory relationship between the TF and each gene, while genes without regulatory evidence are shown as small gray dots. Annotated genes are AD upregulated genes in oligodendrocytes (AD DEGs). TF binding motifs are shown as motif logos.

Additionally, the latent feature space (Z) learned by the RVAE provides further biological insight into the pseudotime trajectory and gene regulation in disease (Fig. 2.5e). Here, each dot represents a single feature (gene or chromatin region), and they are organized in 2D space based on their pseudotemporal dynamics learned by the RVAE. We ranked each feature based on the point in the trajectory that it reaches 75% of its maximum value, which we termed as the feature's "trajectory rank". We then correlated the reconstructed feature trajectories, as in Fig. 2.5d, to the proportion of late-stage AD nuclei, as in Fig. 2.5b, to see which features consistently change with AD. For both genes (t-DEGs) and chromatin regions (gl-cCREs), the latent space clearly groups features together that are positively or negatively correlated with the proportion of late-stage AD nuclei and groups features together with similar trajectory ranks, demonstrating the power of this RVAE model for the analysis and interpretation of multi-omic pseudotemporal dynamics.

We showcase two key TFs in oligodendrocytes: NRF1 and sterol regulatory element binding transcription factor 1 (SREBF1). SREBF1 is critical in regulating the expression of genes involved in cholesterol and fatty acid homeostasis (*154*), and it is proposed that Aβ inhibits SREBF1 activation (*155*). We found that NRF1 motif variability is upregulated in oligodendrocytes in late-stage AD (Bonferroni adjusted p-value = 5.13×10^{-20}, Fig. 2.4g), and SREBF1 motif variability is downregulated with disease in oligodendrocytes (Bonferroni adjusted p-value = 2.67×10^{-191}, Extended Data Fig. 2.4). We correlated TF motif variability trajectories (Extended Data Fig. 2.6) with the reconstructed t-DEG expression trajectories and visualized the correlation between the TF and each gene within the 2D latent space, identifying candidate target genes activated or repressed by TF binding events (positive or negative trajectory correlation, respectively) (Fig. 2.5f, Supplementary Note[i]). We found that NRF1 is negatively correlated with target genes at the end of the trajectory, while SREBF1 is positively correlated with target genes at both the beginning and the end of the trajectory, indicating that SREBF1 acts as a transcriptional activator throughout the trajectory.

Microglia trajectory to define disease-associated microglia

Using the same analytical approach as our oligodendrocyte trajectory analysis, we constructed an integrated microglia trajectory using 10,768 nuclei from snATAC-seq and 4,119 nuclei from snRNA-seq (Fig. 2.6a). The proportion of nuclei from late-stage AD samples significantly increased throughout the microglia trajectory (Fig. 2.6b, Pearson correlation R = 0.53, p-value = 6.9 x 10–5). We next sought to investigate gene signatures of disease-associated microglia (DAMs), which were introduced in Keren-Shaul *et al.*'s single-cell transcriptomic study (*77*) of 5XFAD mice and are highly debated in the field of AD genomics. DAMs are described as AD associated phagocytic microglia that are sequentially activated in *TREM2*-independent and -dependent stages (stage 1 and stage 2, respectively). We found that the integrated microglia trajectory follows a decrease in the homeostatic signature, an increase in the stage 1 DAM signature, and a distinct global depletion of the stage 2 *TREM2*-dependent DAM signature (Fig. 2.6c, see Supplementary Note[i] for gene signature lists), suggesting that this microglia trajectory describes the transcriptional and epigenetic changes during the transition from a homeostatic to disease-associated cell-state.

To further dissect the microglia trajectory, we modeled the chromatin accessibility and gene expression dynamics of 9,163 microglia gl-cCREs and 2,138 microglia t-DEGs, respectively, using RVAE (Fig. 2.6d-e). We highlight two ETS family TFs, SPI1 and ETS variant 5 (ETV5), both of which showing upregulated motif variability in late-stage AD (Bonferroni adjusted p-values 1.19 x 10^{-20}, 6.68 x 10^{-19} respectively), and their candidate target genes along the trajectory (Fig. 2.6f, Supplementary Note[i]). We observed that the SPI1 motif trajectory is negatively correlated with genes at the end of the trajectory, supporting our previous findings that SPI1 acts as a repressor in late-stage AD.

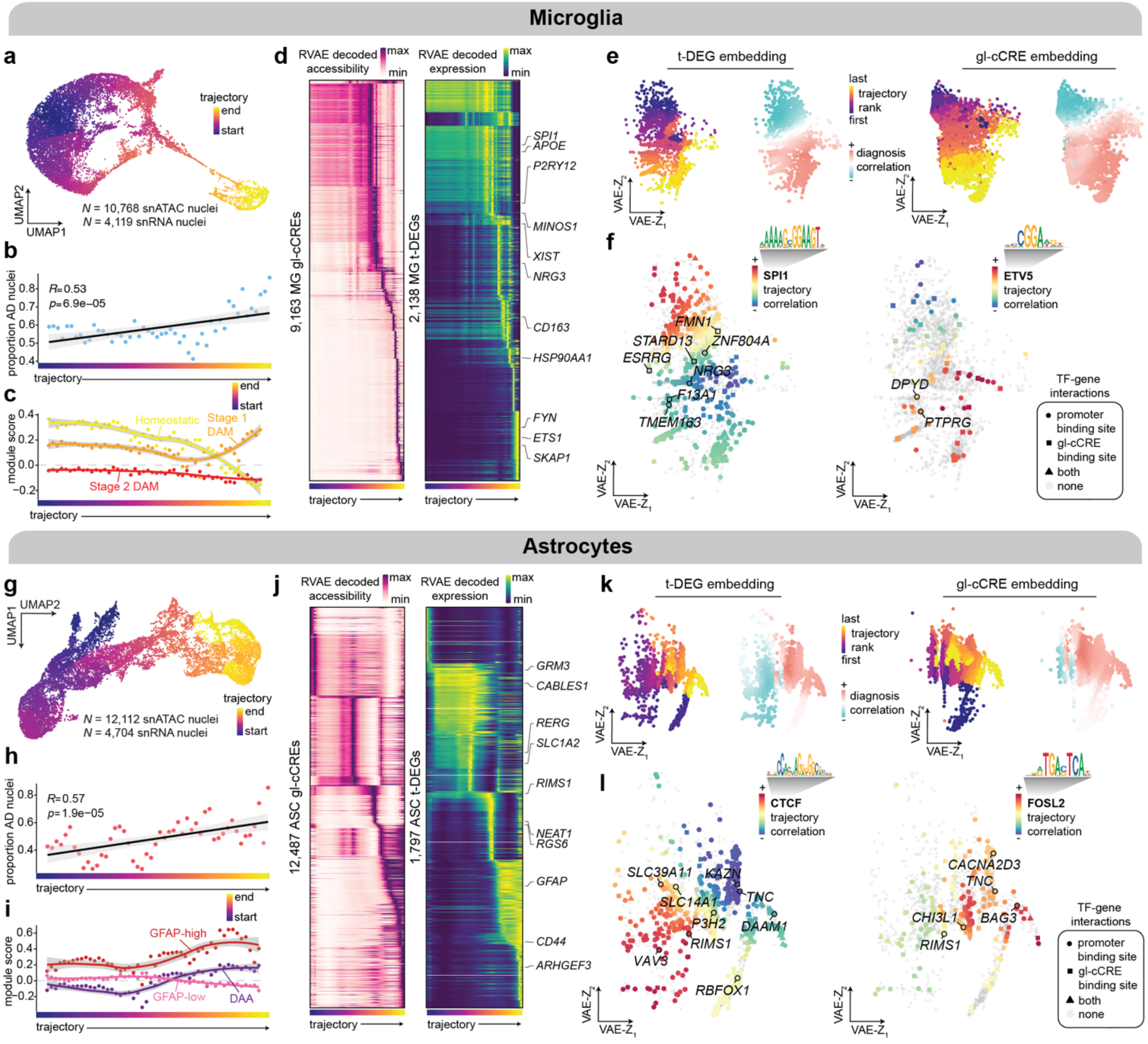


Figure 2.6. Multi-omic microglia and astrocyte trajectory analyses

a, UMAP dimensionality reduction of microglia from the integrated snATAC-seq (n=10,768 nuclei) and snRNA-seq (n=4,119 nuclei) analysis. **b,** Scatter plot of the proportion of AD microglia nuclei as in Fig. 2.5b. **c,** Scatter plot of module scores as in Fig. 2.5c for gene signatures from Keren-Shaul *et al.* (*77*): homeostatic microglia, Stage 1 disease-associated microglia (DAM), and Stage 2 DAM (see Supplementary Note[i] for full gene lists). **d,** Heatmaps of RVAE reconstructed chromatin accessibility and gene expression as in Fig. 2.5d, for 9,163 microglia gl-cCREs (left) and 2,138 microglia t-DEGs (right). **e,** 2D latent space learned by RVAE modeling of microglia t-DEGs (left) and gl-cCREs (right), as in Fig. 2.5e. **f,** Microglia t-DEG latent space colored by correlation of gene expression to SPI1 (left) and ETV5 (right) motif variability, as in Fig. 2.5f. **g,**UMAP dimensionality reduction of astrocytes from the integrated snATAC-seq (n=12,112 nuclei) and snRNA-seq (n=4,704 nuclei) analysis. **h,** Scatter plot of the proportion of AD astrocyte nuclei as in **b**. **i,** Scatter plot of module scores as in **c** for gene signatures from Habib et al. 2020 (*79*):GFAP-low, GFAP-high, and Disease Associated Astrocytes (DAA, see Supplementary Note[i] for full gene lists). **j,** Heatmaps of RVAE reconstructed chromatin accessibility and gene expression as in **d** for 12,487 astrocyte gl-cCREs (left) and 1,797 astrocyte t-DEGs (right). **k,** 2D latent space learned by RVAE modeling of astrocyte t-DEGs (left) and gl-cCREs (right), as in **e**. **l,** Astrocyte t-DEG latent space colored by correlation of gene expression to CTCF (left) and ETV5 (right) motif variability, as in **f**.

Disease-associated astrocytes in human AD

We also constructed an integrated astrocyte trajectory using 12,112 nuclei from snATAC-seq and 4,704 nuclei from snRNA-seq (Fig. 2.6g), and we again found that the proportion of late-stage AD nuclei significantly increases throughout the trajectory (Fig. 2.6h, Pearson correlation R = 0.57, p-value = 1.9 x 10−5). In a similar fashion to our analysis of the DAM signature in the microglia trajectory, we investigated the gene signature of disease-associated astrocytes (DAAs), described in a recent snRNA-seq study of the hippocampus in 5XFAD mice (*79*) as an AD-specific *GFAP*^{high} astrocyte subpopulation that is distinct from another *GFAP*^{high} astrocyte subpopulation found in aged wild-type and 5XFAD (*GFAP*-high). Based on DAA gene signature analysis, we reasoned that this trajectory follows a trend from a *GFAP*-low state to *GFAP*-high and DAA-like states (Fig. 2.6i, see Supplementary Note[i] for gene signature lists).

RVAE modeling of 12,487 astrocyte gl-cCREs and 1,797 astrocyte t-DEGs revealed rich gene-regulatory dynamics across the trajectory (Fig. 2.6j-k). We investigated the relationship between astrocyte t-DEGs and two TFs: CCCTC-binding factor (CTCF) and FOSL2, whose motif variability we have found to be downregulated and upregulated in late-stage AD, respectively (Bonferroni adjusted p-values 6.45 $\times10^{-17}$, 5.65 $\times10^{-99}$ respectively). CTCF is known as a master chromatin regulator (*156*, *157*), and we observed that the CTCF motif variability trajectory is anti-correlated with the DAA and *GFAP*-high signatures (end of the trajectory, Extended Data Fig. 2.6) and positively correlated with t-DEGs in the *GFAP*-low phase of the trajectory (Fig. 2.6l). Alternatively, we found a positive correlation between the motif variability trajectory of FOSL2 with the *GFAP*-high and DAA gene signatures and a positive correlation with genes at the end of the trajectory (Fig. 2.6l, Supplementary Note[i]). These findings suggest that FOSL2 may be an activator of the disease-associated astrocyte signature, whereas CTCF may promote a more homeostatic or non-diseased astrocyte state. By relating gene expression with TF motif enrichment, TF binding site accessibility, and using the temporal information learned by the RVAE, we begin to unravel the role of TFs in regulating cell states, such as disease-associated astrocytes.

Cell-type-specific cis-regulation at AD genetic risk loci

To further our understanding of AD genetic risk signals, we performed cell-type-specific linkage-disequilibrium score regression (LDSC) (*131*) analysis in our snATAC-seq clusters using GWAS summary statistics in AD (*9*, *15*) and other relevant traits (*132*, *133*, *135*–*138*, *158*) (Methods, Supplementary Table 6, Supplementary Note[i]). Microglia clusters MG.b and MG.c showed a significant enrichment ($FDR < 0.05$) for AD GWAS SNPs from the Kunkle *et al.* study, and all five microglia clusters showed a significant enrichment (MG.a, MG.e $FDR < 0.005$; MG.b, MG.c, MG.d $FDR < 0.0005$) for GWAS SNPs from the Jansen *et al.* study, which included familial AD-by-proxy samples in addition to AD patient data (Fig. 2.7a). The results of this GWAS heritability analysis supports previous findings in non-diseased human (*72*) and mouse (*73*) snATAC-seq data. We further investigated AD risk signals in microglia using gchromVAR (*159*) to compute the enrichment of fine-mapped AD-associated polymorphisms from Jansen *et al.* along the microglia pseudotime trajectory and observed a significant increase (Pearson correlation, p-value = 0.0048, Methods, Fig. 2.7b-c) in the gchromVAR deviation score in distal peaks throughout the microglia trajectory, in stark contrast with a significant decrease in the deviation score for the analogous gene-proximal peak analysis (Pearson correlation, p-value = 0.0053), highlighting AD-associated SNPs at distal enhancers in disease-associated microglia. By overlaying the co-accessibility map with chromatin accessibility signal and GWAS statistics along the genomic axis, we unraveled the potential cis-regulatory relationships disrupted by causal disease variants in GWAS genes, such as *BIN1*, *ADAM10*, *APOE* and *SCL24A4* (Fig. 2.7d-i). We found that the *APOE* locus, which harbors the main determinants of AD heritability and is one of the best studied AD risk loci, has cis-regulatory chromatin networks altered in disease in microglia and astrocytes, highlighting cCREs that are prime candidates for further study using genome editing technologies.

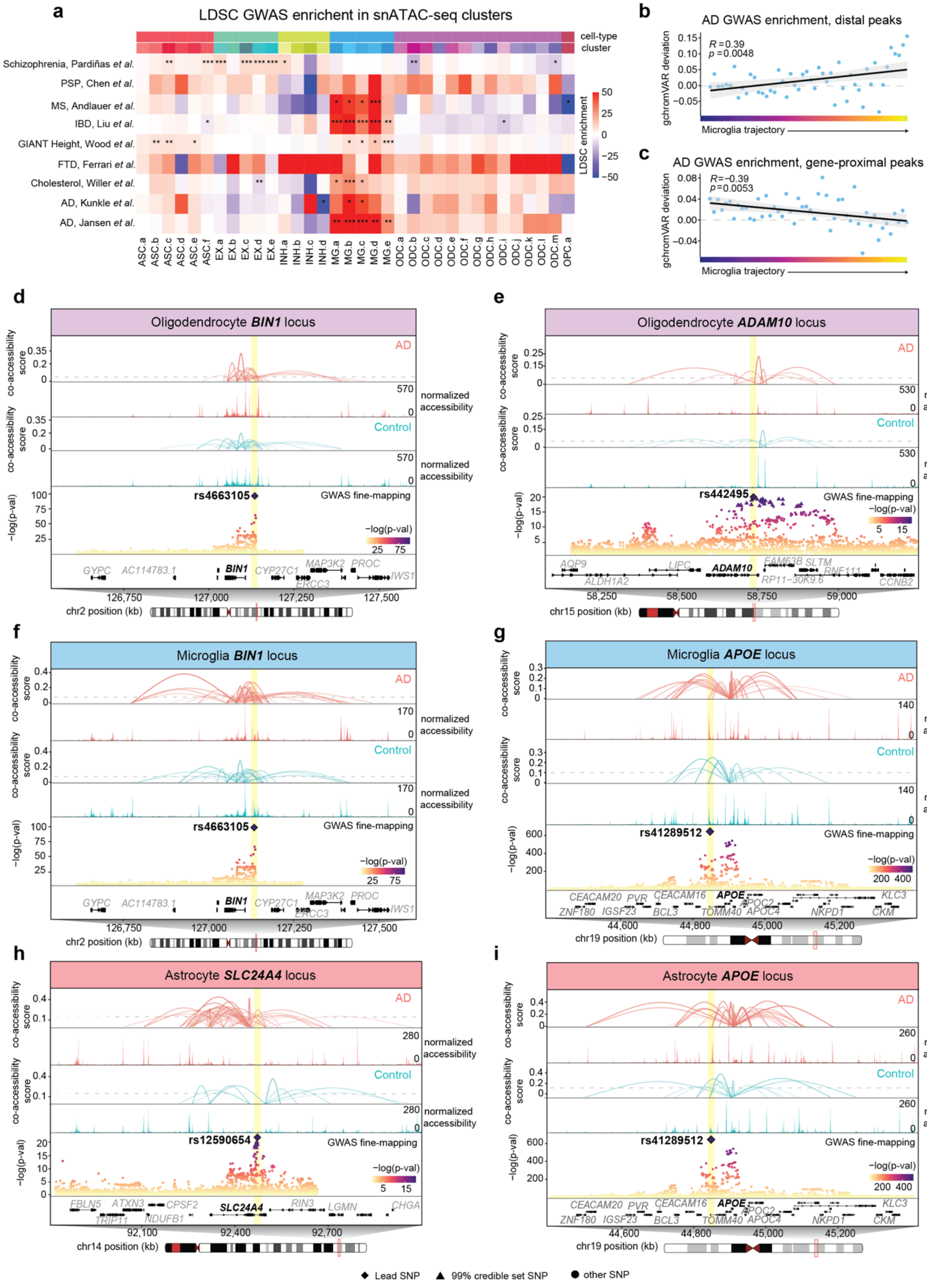

a
LDSC GWAS enrichent in snATAC-seq clusters
cell-type
cluster
Schizophrenia, Pardiñas et al.
PSP, Chen et al.
MS, Andlauer et al.
IBD, Liu et al.
GIANT Height, Wood et al.
FTD, Ferrari et al.
Cholesterol, Willer et al.
AD, Kunkle et al.
AD, Jansen et al.
LDSC enrichment
50
25
0
−25
−50
ASC.a ASC.b ASC.c ASC.d ASC.e ASC.f EX.a EX.b EX.c EX.d EX.e INH.a INH.b INH.c INH.d MG.a MG.b MG.c MG.d MG.e ODC.a ODC.b ODC.c ODC.d ODC.e ODC.f ODC.g ODC.h ODC.i ODC.j ODC.k ODC.l ODC.m OPC.a
b
AD GWAS enrichment, distal peaks
R=0.39
p=0.0048
gchromVAR deviation
Microglia trajectory
c
AD GWAS enrichment, gene-proximal peaks
R=−0.39
p=0.0053
gchromVAR deviation
Microglia trajectory
d
Oligodendrocyte BIN1 locus
e
Oligodendrocyte ADAM10 locus
f
Microglia BIN1 locus
g
Microglia APOE locus
h
Astrocyte SLC24A4 locus
i
Astrocyte APOE locus
co-accessibility score
normalized accessibility
AD
Control
GWAS fine-mapping
−log(p-val)
rs4663105
rs442495
rs41289512
rs12590654
chr2 position (kb)
chr15 position (kb)
chr19 position (kb)
chr14 position (kb)
GYPC AC114783.1 BIN1 CYP27C1 MAP3K2 PROC ERCC3 IWS1
AQP9 ALDH1A2 LIPC ADAM10 FAM63B SLTM RP11-30K9.6 RNF111 CCNB2
CEACAM20 PVR CEACAM16 APOE APOC2 KLC3 ZNF180 IGSF23 BCL3 TOMM40 APOC4 NKPD1 CKM
FBLN5 ATXN3 CPSF2 TRIP11 NDUFB1 SLC24A4 RIN3 LGMN CHGA
Lead SNP
99% credible set SNP
other SNP

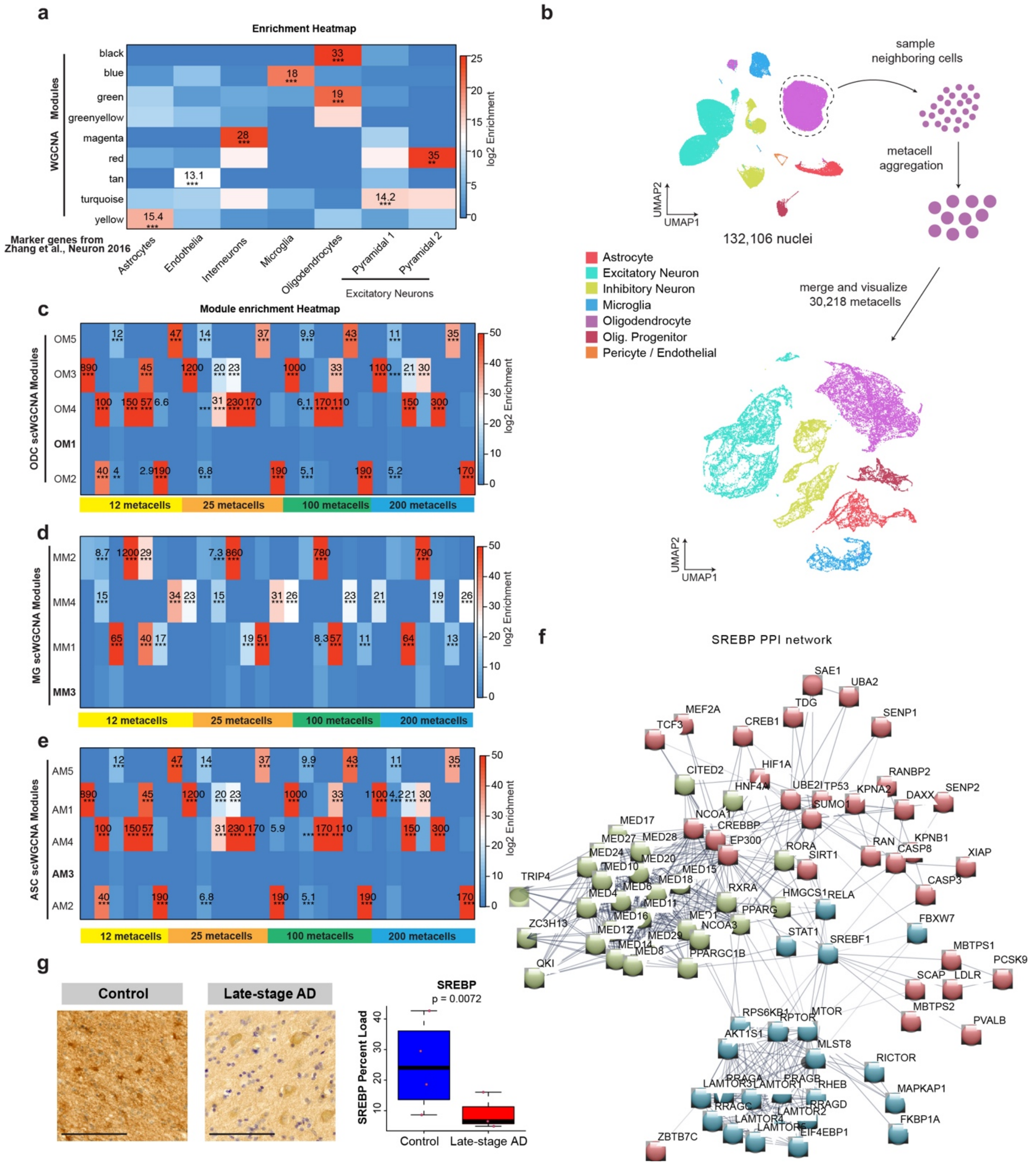
Figure 2.7. Cell-type specific regulatory landscapes of GWAS loci in the AD brain

a, Heatmap showing LDSC enrichment of GWAS traits and disorders in snATAC-seq clusters. P-values are derived from LDSC enrichment tests, and FDR corrected p-values are overlaid on the heatmap (*: FDR < 0.05, **: FDR < 0.005, ***: FDR < 0.0005). **b**, **c**, Scatter plots showing gchromVAR enrichments along the microglia pseudotime trajectory in distal peaks (**b**) and gene-proximal peaks (**c**) averaged for nuclei in each of the 50 trajectory bins. The black line shows a linear regression, and the gray outline represents the 95% confidence interval. Pearson correlation coefficient and p-value are shown. **d-i**, Cis-regulatory architecture at the following GWAS loci and cell-types: *BIN1* (**d**) and *ADAM10* (**e**) in oligodendrocytes; *BIN1* (**f**) and *APOE* (**g**) in microglia; *SLC24A4* (**h**) and *APOE* (**i**) in astrocytes. Co-accessible links for late-stage AD and control are shown separately, with the line height and opacity corresponding to the co-accessibility score; links with a score below the gray dotted line are removed for visualization purposes. Genomic coverage plots for AD and control are shown separately. Jansen *et al.* AD GWAS statistics for SNPs at each locus are shown. Lead SNPs are shown as diamonds, and SNPs in 99% credible set are shown as triangles. Chromosome ideogram indicates genomic coordinates in a 500 kilobase radius centered at each GWAS gene. Chromosome coordinates are the following: *BIN1* chr2:127047027-127110355; *ADAM10* chr15:58587807-58752978; *APOE* chr19:44902754-44910393; *SLC24A4* chr14:92319581-92502483.

Single-cell co-expression networks using scWGNCA

To recontextualize snRNA-seq data in systems-level framework, we sought to develop a gene co-expression network analysis approach for single-cell data based on weighted gene co-expression analysis (WGCNA) (*139*, *140*) a powerful analytical approach for identifying disease-associated gene modules (*30*, *160*) originally designed for bulk gene expression data. Our revised approach uses aggregated expression profiles in place of potentially sparse single cells, where metacells are constructed from specific cell populations by computing the mean expression from 50 neighboring cells using k-nearest neighbors (Methods, Extended Data Fig. 2.7, Supplementary Note^{i}). We re-processed published AD snRNA-seq data from Mathys *et al.* (*75*) and used iNMF to integrate with our snRNA-seq data (Methods, Extended Data Fig. 2.8). Additionally, we performed bulk RNA-seq in early- and late-stage AD cases, as well as pathological controls and curated additional AD bulk-tissue RNA-seq samples from ROSMAP (*46*). Finally, we used consensus WGCNA (*45*), a meta-analytical approach, to jointly form co-expression networks in metacells constructed from the integrated snRNA-seq dataset and bulk-tissue RNA-seq data of the human PFC from two distinct cohorts. We call this approach Single-nucleus Consensus WGCNA (scWGCNA; Extended Data Fig. 2.1, 2.7, 2.9, 2.10; Supplementary Data 7[i]), performed iteratively for each cell-type, where each edge in a co-expressed module is supported by both bulk-tissue RNA-seq data (this study and ROSMAP (*46*)) and aggregated snRNA-seq data (this study and Mathys *et al.*).

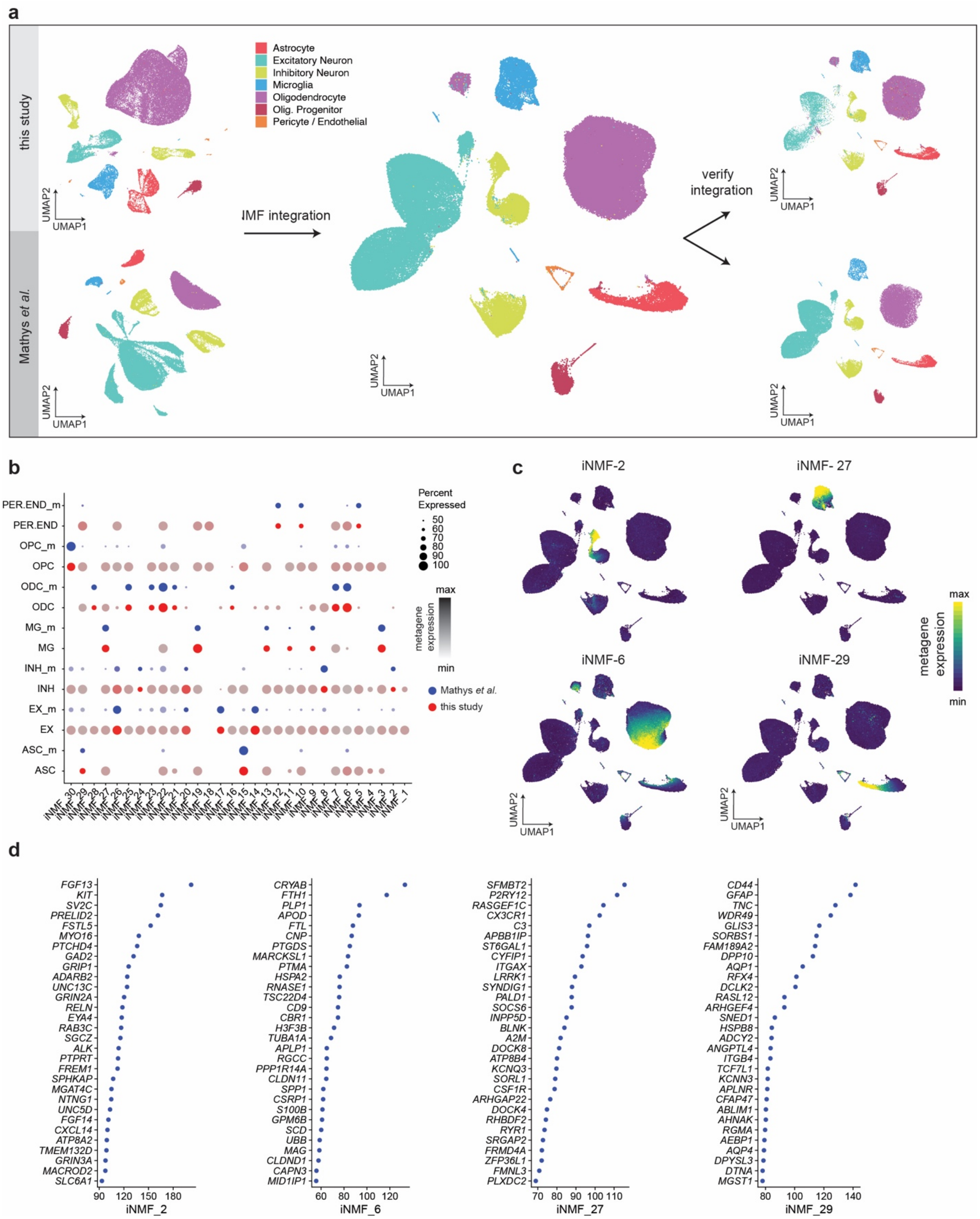
We specifically highlight our scWGCNA analysis for oligodendrocytes; we found four co-expression modules significantly correlated with AD diagnosis—OM1, OM2, OM4, and OM5 (Fig. 2.8a-b, Supplementary Data 7[i]). For example, hub genes of the AD-downregulated module OM1 encode ribosomal subunits (*RPS15A*, *RPL30*, *RPL23A*, etc.), consistent with its enrichment of GO terms related to protein synthesis and sorting (Supplementary Fig. 11[i]). OM2 gene members *MAG*, *CNP*, and *PLP1* are known to be involved in myelination, and unsurprisingly we found OM2 downregulated with disease.



Extended Data Figure 2.7. Metacell aggregation and SREBP.

a, Heatmap showing the enrichment of cell-type marker genes in standard WGCNA modules constructed from our snRNA-seq data. **b**, Schematic showing generation of 30,218 *metacells* from the integrated transcriptomic dataset of 132,106 nuclei from our snRNA-seq and Mathys *et al.* **c-e,** Heatmap showing enrichment of oligodendrocyte **(c)**, microglia **(d)**, and astrocyte **(e)** scWGCNA modules constructed with 12 metacells, 25 metacells, 100 metacells, and

200 metacells in the scWGCNA modules constructed with 50 metacells, as shown in Fig. 2.7 and Supplementary Fig. 15-16[i]. **f,** SREBP protein-protein interaction (PPI) network. Green circle denotes proteins involved in ribosome processing and transcription pathway, cyan circle for mTOR pathway, and red circle for lipid processing pathway. **g,** Left: Representative immunohistochemistry images from postmortem human brain tissue for SREBP with nuclear counterstain. Right: Quantification of SREBP staining. n = 4 pathological controls, 3 late-stage AD. Data is represented as the mean of four equally sized regions per sample. Scale bar represents 100 µm. Linear mixed-effects model ** $p < 0.01$. Box boundaries and line correspond to the interquartile range (IQR) and median respectively. Whiskers extend to the lowest or highest data points that are no further than 1.5 times the IQR from the box boundaries.



Extended Data Figure 2.8. iNMF integration of snRNA-seq with Mathys et al. snRNA-seq.

a, Schematic representation of iNMF integration of snRNA-seq with Mathys et al. snRNA-seq. UMAP plots are colored by cell-type assignments. **b**, Dot plot of iNMF metagene expression in each cell-type, split by dataset of origin. **c**, UMAP plots of the integrated dataset colored by selected iNMF metagenes. **d**, Dot plots showing the iNMF loading for the top 30 genes for the same metagenes in **c**.

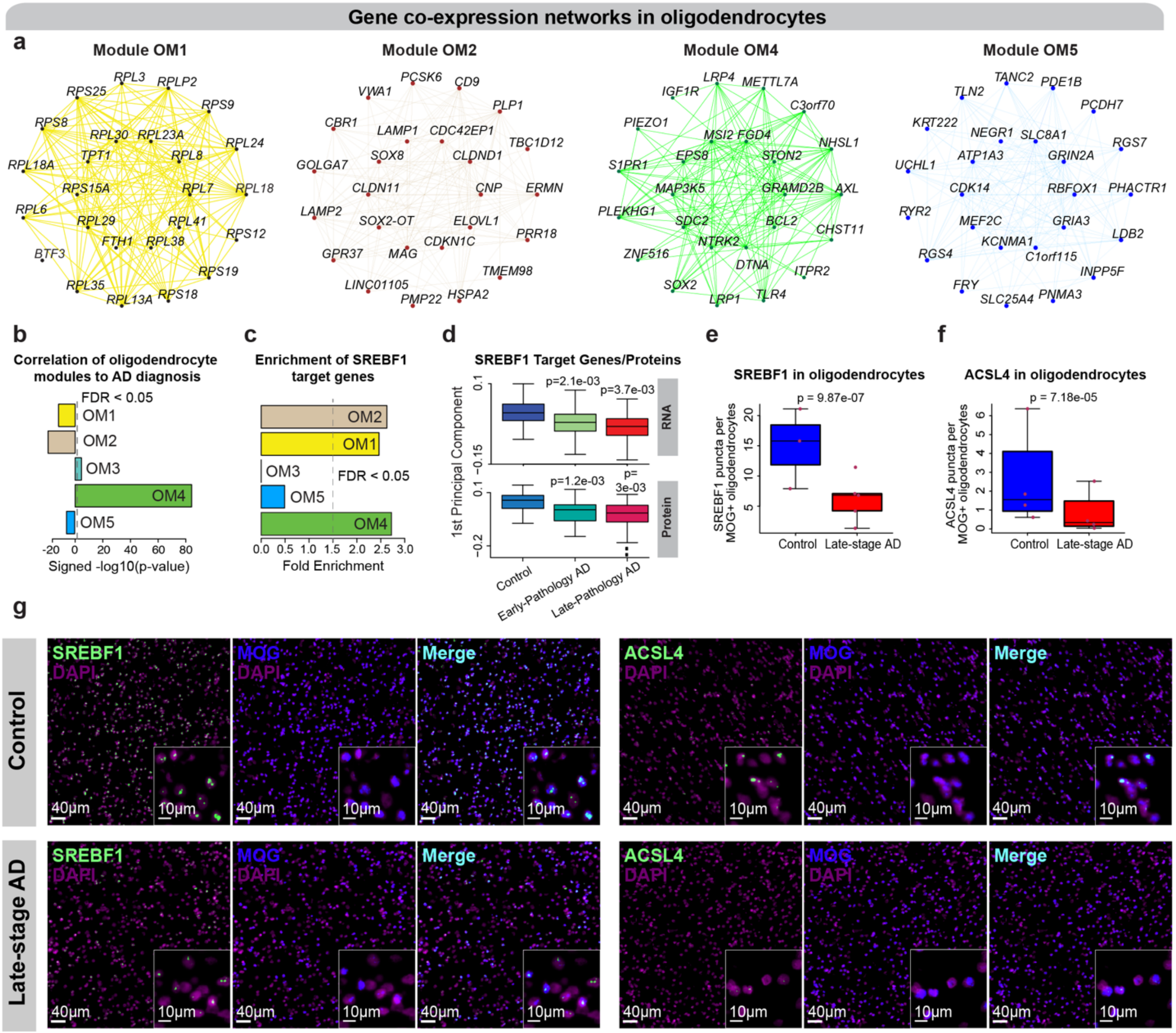
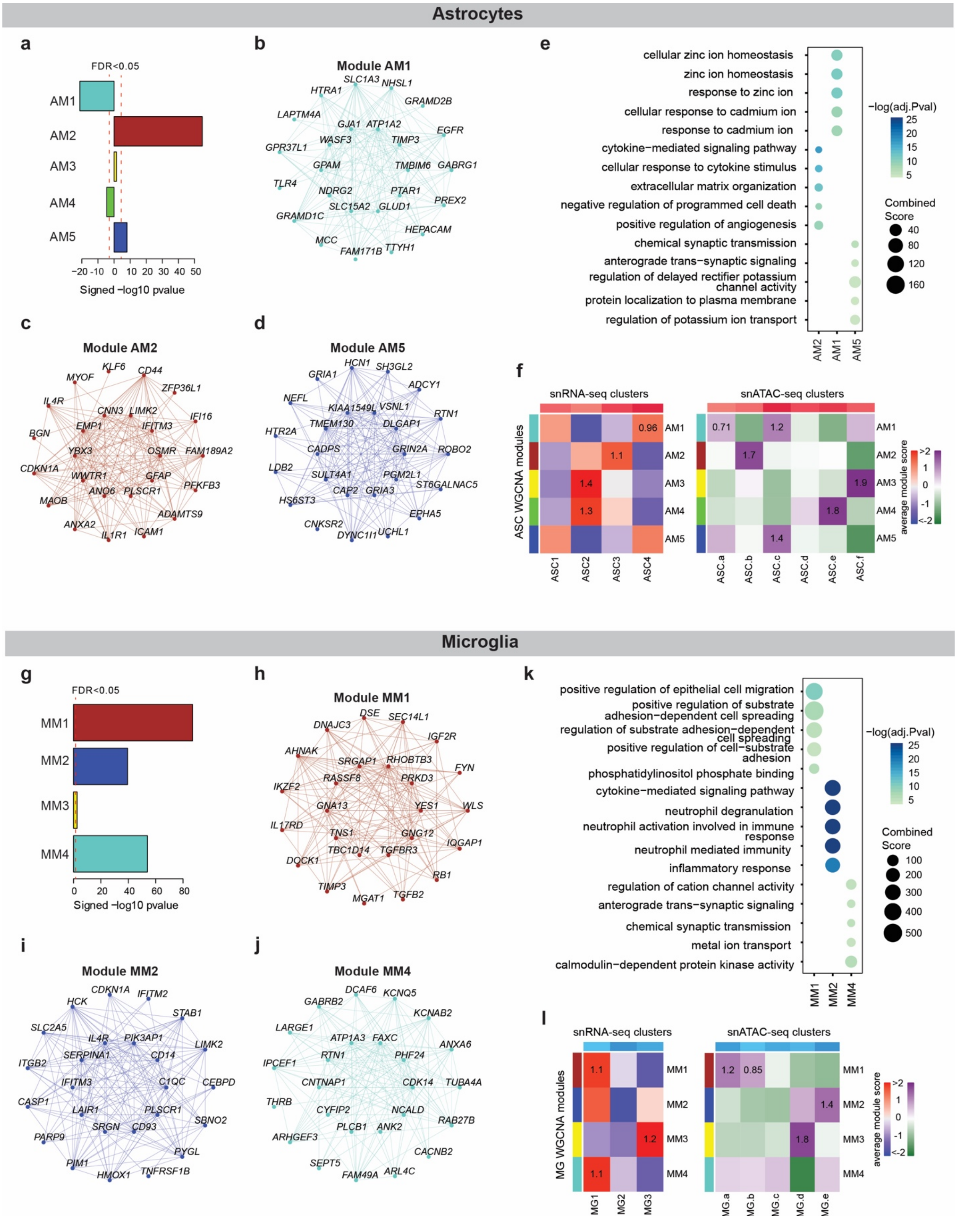


Figure 2.8. Robust co-expression modules revealed using integrated bulk and single-cell co-expression network analysis

a, Co-expression plots for modules OM1, OM2, OM4, and OM5. **b,** Signed correlation oligodendrocyte co-expression modules with AD diagnosis. **c,** Enrichment of SREBF1 target genes in oligodendrocyte co-expression modules. **d,** Boxplots showing RNA (top) and protein expression (*144*) (bottom; n = 98 controls, 76 early-pathology, 101 late-pathology) of SREBF1's targets with AD pathological staging. Two-sided Wilcoxon test. **e,** Boxplots showing quantification of *SREBF1* puncta per *MOG*$^{+}$ oligodendrocyte. n = 3 cognitively healthy controls, 5 late-stage AD. Data is represented as the mean of four equally sized regions per sample. Linear mixed-effects model. **f,** Boxplots showing quantification of *ACSL4* puncta per *MOG*$^{+}$ oligodendrocyte. n = 4 cognitively healthy controls, 4 late-stage AD. Data is represented as the mean of four equally sized regions per sample. Linear mixed-effects model. **g,** Representative RNA fluorescence *in situ* hybridization (RNAscope) images from postmortem human brain tissue for combined *SREBF1* and *MOG* staining as in **e** (left) and *ACSL4* and *MOG* staining as in **f** (right) with DAPI nuclear counterstain. For box and whisker plots, box boundaries and line correspond to the interquartile range (IQR) and median respectively. Whiskers extend to the lowest or highest data points that are no further than 1.5 times the IQR from the box boundaries.

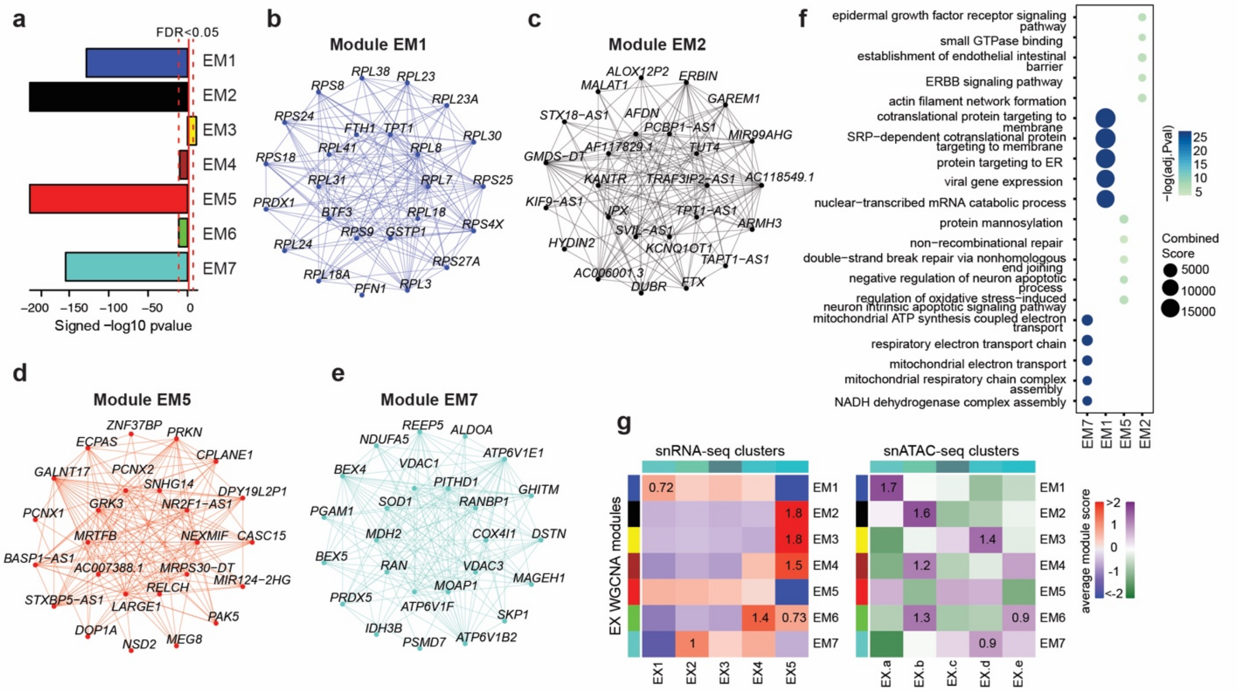
Additionally, we examined SREBF1's downstream regulatory targets in the context of co-expression networks (Methods). Notably, we found that three of the oligodendrocyte modules were significantly enriched for targets of SREBF1, indicating the importance of SREBF1 in regulating gene expression in these modules (Fig. 2.8c). Using a multi-scale dataset of bulk-tissue RNA-seq, high-throughput proteomics (*144*), and SREBF1 ChIP-seq data (ENCODE), we defined a protein-protein interaction (PPI) network of SREBF1 target genes (Extended Data Fig. 2.7). Additionally, we found module eigengene expression of SREBF1 targets downregulated in early- and late-pathology AD cases at the level of proteins (*144*) and RNA (Fig. 2.8d), corroborated by downregulation of SREBF1 motif variability in snATAC-seq data (Extended Data Fig. 2.4). We also validated the downregulation of SREBF1 in late-stage AD through RNA in situ hybridization and immunohistochemistry and found a decrease in *ACSL4* expression, one of SREBF1's targets identified in ENCODE ChIP-seq data, in late-stage AD (Fig. 2.8e-g, Extended Data Fig. 2.7). Overall, our co-expression network analysis approach facilitates the identification of cell-type-specific disease biology, and we have highlighted TF SREBF1, largely unstudied in the context of AD, in oligodendrocytes to demonstrate our approach's ability to yield novel disease insights.


Astrocytes
a
FDR<0.05
AM1
AM2
AM3
AM4
AM5
−20 -10 0 10 20 30 40 50
Signed −log10 pvalue
b
Module AM1
SLC1A3 NHSL1 HTRA1 GRAMD2B LAPTM4A GJA1 ATP1A2 EGFR WASF3 TIMP3 GPR37L1 GPAM TMBIM6 GABRG1 TLR4 NDRG2 PTAR1 PREX2 SLC15A2 GLUD1 GRAMD1C HEPACAM MCC TTYH1 FAM171B
c
Module AM2
KLF6 CD44 MYOF ZFP36L1 IL4R CNN3 LIMK2 IFI16 EMP1 IFITM3 BGN YBX3 OSMR FAM189A2 CDKN1A WWTR1 GFAP ANO6 PLSCR1 PFKFB3 MAOB ADAMTS9 ANXA2 IL1R1 ICAM1
d
Module AM5
HCN1 SH3GL2 GRIA1 ADCY1 NEFL KIAA1549L VSNL1 RTN1 HTR2A TMEM130 DLGAP1 CADPS GRIN2A ROBO2 LDB2 SULT4A1 PGM2L1 CAP2 GRIA3 ST6GALNAC5 HS6ST3 EPHA5 CNKSR2 DYNC1I1 UCHL1
e
cellular zinc ion homeostasis
zinc ion homeostasis
response to zinc ion
cellular response to cadmium ion
response to cadmium ion
cytokine−mediated signaling pathway
cellular response to cytokine stimulus
extracellular matrix organization
negative regulation of programmed cell death
positive regulation of angiogenesis
chemical synaptic transmission
anterograde trans−synaptic signaling
regulation of delayed rectifier potassium channel activity
protein localization to plasma membrane
regulation of potassium ion transport
AM2 AM1 AM5
−log(adj.Pval)
25 20 15 10 5
Combined Score
40 80 120 160
f
snRNA-seq clusters
snATAC-seq clusters
ASC WGCNA modules
0.96 1.1 1.4 1.3
0.71 1.2 1.7 1.9 1.8 1.4
AM1 AM2 AM3 AM4 AM5
ASC1 ASC2 ASC3 ASC4
ASC.a ASC.b ASC.c ASC.d ASC.e ASC.f
average module score
>2 1 0 -1 <-2
Microglia
g
FDR<0.05
MM1
MM2
MM3
MM4
0 20 40 60 80
Signed −log10 pvalue
h
Module MM1
DSE SEC14L1 DNAJC3 IGF2R AHNAK SRGAP1 RHOBTB3 FYN RASSF8 PRKD3 IKZF2 GNA13 YES1 WLS IL17RD TNS1 GNG12 IQGAP1 TBC1D14 TGFBR3 DOCK1 RB1 TIMP3 MGAT1 TGFB2
i
Module MM2
CDKN1A IFITM2 HCK STAB1 SLC2A5 IL4R PIK3AP1 LIMK2 SERPINA1 CD14 ITGB2 IFITM3 C1QC CEBPD CASP1 LAIR1 PLSCR1 SBNO2 SRGN CD93 PARP9 PYGL PIM1 HMOX1 TNFRSF1B
j
Module MM4
DCAF6 KCNQ5 GABRB2 KCNAB2 LARGE1 ATP1A3 FAXC ANXA6 RTN1 PHF24 IPCEF1 CNTNAP1 CDK14 TUBA4A THRB CYFIP2 NCALD RAB27B PLCB1 ANK2 ARHGEF3 CACNB2 SEPT5 ARL4C FAM49A
k
positive regulation of epithelial cell migration
positive regulation of substrate adhesion−dependent cell spreading
regulation of substrate adhesion−dependent cell spreading
positive regulation of cell−substrate adhesion
phosphatidylinositol phosphate binding
cytokine−mediated signaling pathway
neutrophil degranulation
neutrophil activation involved in immune response
neutrophil mediated immunity
inflammatory response
regulation of cation channel activity
anterograde trans−synaptic signaling
chemical synaptic transmission
metal ion transport
calmodulin−dependent protein kinase activity
MM1 MM2 MM4
−log(adj.Pval)
25 20 15 10 5
Combined Score
100 200 300 400 500
l
snRNA-seq clusters
snATAC-seq clusters
MG WGCNA modules
1.1 1.2 1.1
1.2 0.85 1.4 1.8
MM1 MM2 MM3 MM4
MG1 MG2 MG3
MG.a MG.b MG.c MG.d MG.e
average module score
>2 1 0 -1 <-2

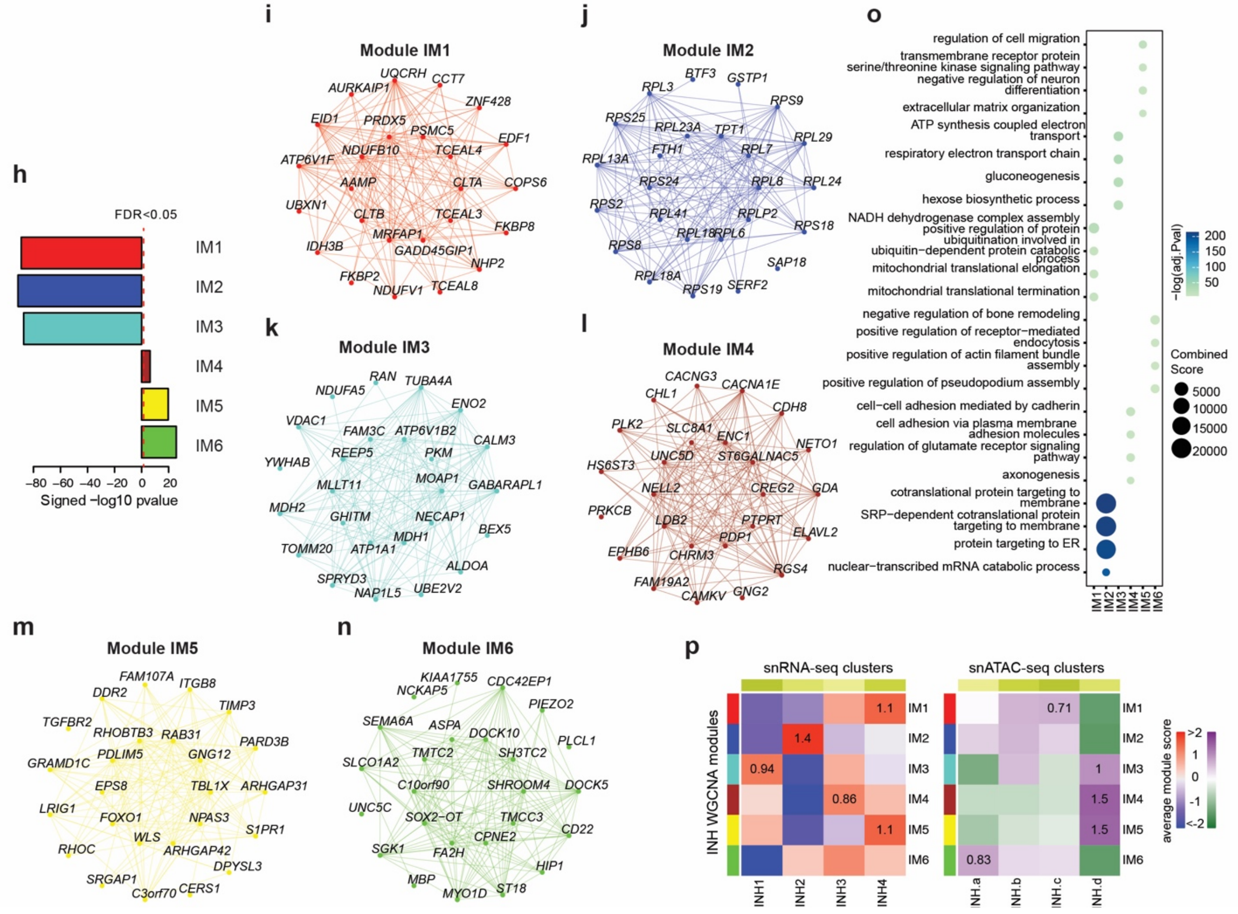
Extended Data Figure 2.9. scWGCNA in microglia and astrocytes.

a, Signed correlation of astrocyte modules to AD diagnosis. **b-d**, Co-expression plots for modules AM1 **(b)**, AM2 **(c)**, and AM5 **(d)**. **e,** GO term enrichment of astrocyte modules. **f,** Heatmaps showing row-normalized Seurat module scores of astrocyte modules in snRNA-seq (left) and snATAC-seq (right) astrocyte clusters. **g,** Signed correlation of microglia co-expression modules with AD diagnosis. **h-j,** Co-expression plots for modules MM1 **(h)**, MM2 **(i)**, and MM4 **(j)**. **k,** GO term enrichment of microglia modules. **l,** Heatmaps showing row-normalized Seurat module scores of microglia modules in snRNA-seq (left) and snATAC-seq (right) microglia clusters.

Excitatory Neurons



Inhibitory Neurons



Extended Data Figure 2.10. scWGCNA in neurons.

a, Signed correlation of excitatory neuron modules to AD diagnosis. **b-e,** Co-expression plots for modules EM1 **(b)**, EM2 **(c)**, EM5 **(d)**, and EM7 **(e)**. **f,** GO term enrichment of excitatory neuron modules. **g,** Heatmaps showing row-normalized Seurat module scores of excitatory neuron modules in snRNA-seq (left) and snATAC-seq (right) excitatory neuron clusters. **h,**Signed correlation of inhibitory neuron modules to AD diagnosis. **i-n,** Co-expression plots for modules IM1 **(i)**, IM2 **(j)**, IM3 **(k)**, IM4 **(l)**, IM5 **(m)**, and IM6 **(n)**. **o**, GO term enrichment of inhibitory neuron modules. **p,** Heatmaps showing row-normalized Seurat module scores of inhibitory neuron modules in snRNA-seq (left) and snATAC-seq (right) inhibitory neuron clusters.

Discussion

Our integrated multi-omic analysis of late-stage AD provides a unique lens into the continuum of cellular heterogeneity underlying disease pathogenesis. Pinpointing causal mechanisms of complex diseases requires a rigorous understanding of cell population specific gene regulatory systems at both the epigenomic and transcriptomic level. While single-cell chromatin accessibility can provide important insights into disease, it is a challenging data modality to work with due to its inherent sparsity. We circumvented the issue of sparsity by integrating single-nucleus open-chromatin and single-nucleus transcriptomes from the same samples, in addition to using aggregation methods for pseudo-bulk accessibility profiling and co-accessibility analysis. Taking these considerations into account, our multi-omic analysis enabled us to analyze cell-type-specific epigenomic dysregulation in neurodegeneration and expands on previous work to decipher the transcriptomes of single nuclei in human AD.

A major contribution of our study is that we identified cell-type specific gl-cCREs, which may be mediating gene regulatory changes in late-stage AD, along with TFs that may be binding to these gl-cCREs within the given cell-type. While cCREs can be identified with epigenetic data alone, our analysis is substantiated by integrating single-nucleus transcriptomic data, as we link the gene expression of candidate target genes with cCRE chromatin accessibility. Previous studies of AD have not explored cis-gene regulation at a cell-type or cell subpopulation level. We have highlighted both cis- and trans-gene regulation disrupted in late-stage AD, providing potential targets for further study into AD, like NRF1 in oligodendrocytes and FOSL2 in astrocytes and their corresponding gl-cCREs. Further, we examined cis-regulatory interactions in our multi-omic dataset to elucidate cell-type and disease specific patterns of genes implicated in inherited AD risk by GWAS, which are of particular interest as candidate therapeutic targets. For a subset of AD GWAS loci, we compared cis-regulatory networks between AD and control cell populations to identify interactions that are uniquely found in disease. Thus, this study serves as a resource for the broader AD community to explore cell-type and cell-state-specific regulatory landscapes of genes and genomic regions that may be of particular interest, such as AD GWAS loci.

Moreover, independent and joint analyses of the transcriptome and chromatin profiles of oligodendrocytes revealed disrupted gene regulation and biological pathways in AD (Supplementary Note[i]). We described a multi-omic oligodendrocyte trajectory and evaluated gene expression signatures in the transition from newly formed to mature oligodendrocytes, observing that the trajectory seemed to follow oligodendrocyte maturation. Notably, we analyzed the trajectory dynamics of SREBF1, a TF involved in regulation of cholesterol and lipid metabolism that has been shown to be involved in Aβ-related processes (*155*). We found that SREBF1 motif variability was decreased in late-stage AD, indicating that fewer SREBF1 binding sites are accessible in disease, and SREBF1 gene expression is also downregulated in AD oligodendrocytes. Trajectory analysis revealed that SREBF1 motif variability is positively correlated with t-DEGs throughout the trajectory, suggesting that it acts as a transcriptional activator in oligodendrocytes.

Co-expression network analysis methods like WGCNA have been widely used for discovery of disease-associated gene modules in bulk gene expression data (*39*, *47*); however, these approaches are rarely used in single-cell transcriptomics, with some exceptions (*161*) due to challenges in network construction from noisy data. Here we introduced scWGCNA, a method for interrogation of cell population-specific co-expression networks that leverages aggregated metacells to combat the sparsity of single-cell gene expression. Using scWGCNA, we found gene co-expression networks in human AD by jointly analyzing our snRNA-seq and bulk RNA-seq with additional snRNA-seq and bulk RNA-seq samples from the ROSMAP cohort (*46*, *75*). This meta-analytical approach ensured robustness of our network analysis and allowed us to evaluate the resulting gene modules in early-stage AD (Supplementary Note[i]). Notably, scWGCNA identified three oligodendrocyte modules that were enriched for target genes of SREBF1 and showed that the gene and protein expression of these targets were decreased in late-stage AD. With our co-expression and trajectory analysis of SREBF1 in oligodendrocytes, SREBF1 is clearly a gene to prioritize for follow-up studies as a candidate target for AD therapeutics, demonstrating the utility of our approach in identifying novel gene targets for disease.

While the causative molecular mechanisms of sporadic AD remain unknown, our work offers new insights which assist in unraveling the nature of gene regulation in AD, especially in regard to genomic loci with well-described heritable disease risk. Additional work is needed to spatially resolve the complexity of gene expression and epigenomics in AD and neurodegeneration in general. The data presented here are a valuable resource for understanding regulatory relationships in the diseased brain, and our analysis framework serves as a blueprint for making discoveries in complex traits using single-cell multi-omic data. Finally, our intuitive web portal for exploring single nuclei in the human brain allows for the accessibility of our results to anyone with an internet-equipped device.

---

[i] Please access at https://doi.org/10.1038/s41588-021-00894-z

Chapter Three: A comparative study of the Alzheimer's disease transcriptome with spatial and cellular resolution

Introduction

Dating back to Santiago Ramón y Cajal and Korbinian Brodmann, it was revealed that the human brain is highly spatially organized at both macro- and microscopic levels, where both brain circuitry and function underlie this structural organization. Cajal's drawings depicted the laminar organization of the brain and great morphological diversity across neurons. Multiple single-cell (scRNA-seq) and single-nucleus RNA-sequencing (snRNA-seq) studies over the past few years, however, discovered that brain cell populations are even more heterogeneous at the molecular level (*51*, *54–62*). In the Alzheimer's disease (AD) brain, specific cell subpopulations have been identified as under- or overrepresented relative to the cognitively healthy brain (*75*, *76*, *78*, *84*, *85*). Determining the functional significance of this multitude of cell subpopulations remains a large challenge; these single-cell studies majorly relied on approaches that result in a loss of potentially critical spatial information that could help us decipher the roles of these cells. Although earlier studies used laser capture microdissection to carefully profile cells by layer, this is a low-throughput method compared to the now "gold-standard" sc- or snRNA-seq approaches, and large numbers of cells are required to detect rarer cell populations. There is also a pressing need to increase sample numbers to robustly define disease-associated changes.

Recently, several spatial profiling methods were developed with varying levels of resolution, numbers of genes, and throughput (*94–97*, *162*, *163*). Spatial transcriptomics relies on spatial spots with primers to uniquely barcode transcripts based upon their spatial location and does not require pre-determined gene targets, thus allowing an unbiased assessment of gene expression changes. Further, this technique allows us to profile a whole coronally sectioned mouse hemisphere, compared to only single brain regions. The caveat, however, is a lack of single cell resolution. Therefore to circumvent this lack of single cell resolution, we generated both spatial transcriptomic and snRNA-seq data to perform an integrated analysis of AD with cellular and spatial resolution.

Here we examined the spatial and single-cell transcriptome of clinical AD samples including both early- and late-stage pathology cases, as well as AD in Down syndrome (AD in DS). Although individuals with DS aged >65 years old have an 80% risk of dementia (*22*), only one previous study has profiled cell-type specific gene expression changes in DS brains (*164*). Moreover, despite shared features between AD in the general population and AD in DS (*23*), there are no published single-cell or spatial transcriptomic studies examining both populations, and DS is a potentially advantageous group for clinical studies of AD. We discovered spatial and cellular AD transcriptomic changes in both DS and the general population and additionally extended our analyses to a commonly used amyloid mouse model of AD, 5XFAD, by generating an additional spatial transcriptomic dataset. Our analyses highlight the value of integrating not only different models of investigation, but also different data modalities. We identified a set of amyloid-associated genes shared between human and mouse, by integrating imaging data. Furthermore, we revealed spatially restricted cellular communication pathways dysregulated in disease, and we assessed the protein expression of genes identified in our analyses with a spatial proteomic approach.

Materials and Methods

Postmortem human brain tissue

Human brain tissue from prefrontal cortex and posterior cingulate cortex was obtained from UC Irvine's Alzheimer's Disease Research Center and the NIH NeuroBioBank under UCI's Institutional Review Board (IRB). Samples were assigned to groups based on both NFT and plaque staging, in addition to clinical diagnoses. Samples were also selected based upon several covariates, including age, sex, race, postmortem interval (PMI), RNA integrity number (RIN), and disease comorbidity. RIN values were obtained by isolating total RNA with the Zymo Direct-zol RNA Isolation kit and assessing with the Agilent TapeStation 4200. Sample information is available in Supplementary Table 1[ii].

Mouse brain tissue

All mouse work was approved by the Institutional Animal Care and Use (IACUC) committee at UCI. 5XFAD hemizygous (C57BL16) and wildtype littermates were bred and housed until sacrifice at 4, 6, 8, and 12 months. Sample information is available in Supplementary Table 1[ii]. For genotyping, we used the following primers (for *PSEN1*): 5' – AAT AGA GAA CGG CAG GAG CA – 3' (Forward), 5' – GCC ATG AGG GCA CTA ATC AT – 3' (Reverse). Mice were euthanized by carbon dioxide inhalation. After PBS transcardiac perfusion, one brain hemisphere was flash frozen in isopentane chilled with dry ice for spatial transcriptomics.

Single-nucleus RNA-sequencing

Single-nucleus isolations were performed in randomized groups of 12 samples. ~50mg fresh frozen postmortem human brain tissue was homogenized in Nuclei EZ Lysis buffer (NUC101-1KT, Sigma-Aldrich) and incubated for 5 min before being passed through a 70μm filter. Samples were then centrifuged at 500 *g* for 5 min at 4°C and resuspended in additional lysis buffer for 5 min. After another centrifugation at 500 *g* for 5 min at 4°C, samples were incubated in Nuclei Wash and Resuspension buffer (NWR, 1xPBS, 1% BSA, 0.2U/μl RNase inhibitor) for 5 min. To remove myelin contaminants, we prepared sucrose gradients with Nuclei PURE Sucrose Buffer and Nuclei PURE 2M Sucrose Cushion Solution from the Nuclei PURE Nuclei Isolation Kit (NUC-

201, Sigma-Aldrich), and samples were carefully overlaid and centrifuged at 13,000 *g* for 45 min at 4°C. Samples were then washed in NWR before processing with the Nuclei Fixation Kit (SB1003, Parse Biosciences). After nuclei fixation and permeabilization, samples were cryopreserved with DMSO until day of library preparation.

Library preparations were performed as 4 batches of 24 samples, with an additional batch to increase numbers of nuclei/sample for 16 samples. We generated single-nucleus libraries with the WTK Whole Transcriptome Kit (SB2001, Parse Biosciences). cDNA library quantification and quality were assessed with Qubit dsDNA HS assay kit (Q32851, Invitrogen) and D5000 HS kit (5067-5592, 5067-5593; Agilent) or D1000 HS kit (5067-5584, 5067-5585; Agilent) for the Agilent TapeStation 4200. Libraries were sequenced using Illumina Novaseq 6000 S4 platform using 100bp paired-end sequencing for a sequencing depth of 50,000 read pairs/cell.

Spatial transcriptomics

Fresh frozen tissue samples were sectioned on a HM525NX cryostat (Fisher) at -15°C for 10µm thick sections that are immediately mounted onto 10x Genomics Visium slides. Slides were individually stored in slide mailers (sealed airtight in a plastic bag) at -80°C until staining. We followed 10x Genomics Methanol Fixation, Immunofluorescence Staining & Imaging for Visium Spatial Protocols (RevC), except after tissue sections were fixed in methanol and blocked, the sections were incubated with Amylo-glo (TR-300, Biosensis) for 20 min. Sections were then incubated with the primary antibody OC (polyclonal, AB2286, Millipore) and respective secondary antibody (goat anti-rabbit secondary antibody Alexa Fluor 488, Life Tech or Alexa Fluor 647, Life Tech). Immediately after immunostaining, capture areas were imaged on a widefield Nikon Ti2-E microscope at 20X magnification. Spatial transcriptomic libraries were then generated from the tissue sections according to the 10x Genomics Visium User Guide. Library quantification, quality check, and sequencing were performed as previously described, but sequencing depth was based on an estimated 60% tissue area coverage per sample for 50,000 read pairs per covered spot.

Spatial proteomics

Fixed and cryoprotected tissue was sectioned on a HM525NX cryostat (Fisher) at -15°C for 14µm thick sections onto Fisher Superfrost Plus slides. Slides were stored at -80°C until staining, sealed airtight in a plastic bag. We followed the fresh frozen staining protocol from Standard BioTools (Fluidigm); however, since the tissue was previously fixed, we skipped the fixation step. Slides were transferred on dry ice to incubate at 37°C for 5 min on a PCR machine, similar to the 10x Genomics Visium protocol. Sections were washed in PBS 3 times for 5 min before drawing a hydrophobic barrier. After the hydrophobic barrier dried, we incubated the sections with 3% BSA in PBS with 0.2% Triton X-100 for 45 min at room temperature. We then incubated the sections with the primary antibody cocktail diluted in 0.5% BSA/PBS with 0.2% Triton X-100 overnight at 4°C. The antibodies and dilutions used in the primary antibody cocktail are in Supplementary Table 7[ii]. Sections were then washed in PBS with 0.2% Triton X-100 twice for 8 min before incubating with the iridium intercalator (1:100 in PBS) for 30 min at room temperature. We then washed the sections in water twice for 5 min before allowing them to air dry before ablation. Ablations were performed in 1000x1000µm regions of interest, with the exception of one due to unexpected consumption of Argon gas that resulted in a 1000x922 acquisition.

Preprocessing gene expression data

For the snRNA-seq dataset, we aligned sequencing reads to the reference transcriptome (GRCH38) and quantified gene expression using splitpipe (version X, Parse Biosciences) in each of the five snRNA-seq experiments. We quantified and corrected ambient RNA signal present in our samples using Cellbender (*165*) remove-background (v 0.2.0). Heterotypic barcode collisions were inferred in each snRNA-seq experiment using Scrublet (*166*) (v 0.2.3) with default settings. We merged the individual snRNA-seq experiments into a single anndata (v 0.8.0) object, totaling 611,999 barcodes and 29,889 genes before quality control (QC) filtering. For each snRNA-seq experiment, we removed barcodes in the 95th percentile for number of UMIs detected, doublet score from Scrublet, and percentage of mitochondrial reads. We also applied dataset-wide cutoffs to remove barcodes with less than or equal to 250 UMIs, greater than or equal to 50,000 total UMIs, and greater than or equal to 10% mitochondrial reads. For one of the snRNA-seq experiments, we applied a more stringent filter to remove cells with less than or equal to 500 UMIs, and greater than or equal to 5% mitochondrial reads. We retained 431,534 barcodes for downstream analysis.

The 10X Genomics Loupe Browser image alignment tool was used to select Visium ST spots that intersected the tissue based on the fluorescent images. Sequencing reads from the human and mouse Visium experiments were processed using the 10X Genomics Spaceranger (v 1.2.1) pipeline, with GRch38 and MM10 as the respective reference transcriptomes. Spaceranger count was used to align sequencing reads to the reference, quantify gene expression, and perform a preliminary clustering analysis for each sample. Unlike the snRNA-seq dataset, we did not filter out additional spots based on sequencing QC metrics. The UMI counts matrices and fluorescent images for the human and mouse samples were combined into merged Seurat (*167–169*) objects for the respective species.

Initial snRNA-seq data analysis

Following QC filtering, we processed the snRNA-seq dataset using SCANPY (*170*) and scVI (*171*). The UMI counts matrix was first normalized using the functions sc.pp.normalize_total and sc.ppl.log1p. We set up the anndata object to train the scVI model using snRNA-seq as the batch key and the following additional continuous and categorical covariates: sample ID, diagnosis, brain region, age at death, percentage of mitochondrial counts, number of UMI, postmortem interval (PMI) and RNA integrity number (RIN). We set up the scVI model with two hidden layers, 128 nodes per layer, a 30-dimensional latent embedding after the encoder phase, and a dropout rate of 0.1. We trained the model over 50 epochs and noted a flattened loss curve by the end of the training procedure. The latent embedding learned from the scVI model accounts for the batch effects and additional covariates specified in the model setup step, and we used this embedding for Leiden clustering and UMAP (*172*) dimensionality reduction in SCANPY. With a resolution parameter of 1.5 we identified 43 clusters. We inspected gene expression patterns in these clusters for a panel of canonical CNS cell-type marker genes to assign major cell-type labels to each cluster. We also checked the distribution of QC metrics in each cluster to identify outlier clusters. Six clusters (7, 29, 33, 35, 50, 51) were removed from downstream analysis as QC outliers, or due to presence of potential doublets. We recomputed the UMAP and Leiden clustering (resolution = 1.2) after filtering these clusters, yielding 29 clusters. Glutamatergic neuron clusters were annotated based on expression of known cortical layer marker genes, and GABAergic neuron clusters were annotated based on expression of known markers (*VIP, SST, PVALB, LAMP5*). At this stage, non-neuronal cell clusters were simply labeled by their major cell-type (astrocytes, microglia, oligodendrocytes, oligodendrocyte progenitors, vascular cells). To identify subpopulations in non-neuronal cells, we performed subclustering analysis in each of the major non-neuronal cell populations (microglia, astrocytes, oligodendrocytes, vascular cells).

Each group was isolated in its own anndata object, and Leiden clustering was performed (see GitHub repository for subclustering parameters).

Reprocessing publicly available single-nucleus gene expression datasets

We obtained sequencing data from obtained three published snRNA-seq studies (*173–175*) of AD. Sequencing data from the Mathys *et al.* 2019 and Zhou *et al.* 2020 datasets were downloaded from Synapse (syn18485175 and syn21670836), and the data from the Morabito *et al*. 2021 study generated by our own group was not re-downloaded. We used a uniform pipeline to process each of these datasets, with slightly varying parameters which are noted in our GitHub repository. This pipeline and the resulting anndata objects are identical to those used in another study from our group (*176*), and we reiterate the main analysis steps here. Sequencing reads were pseudoaligned to the reference transcriptome (GRch38) and gene expression was quantified using the count function from kallisto bustools (*177*). Ambient RNA signal was corrected in UMI counts matrices for each sample using Cellbender (*165*) remove-background, and we used Scrublet (*166*) to identify barcodes attributed to more than one cell. Individual samples were then merged into one anndata object for each of the three studies. Analogous to the snRNA-seq data we generated in this study, we performed percentile filtering based on the following QC metrics: doublet score, number of UMI per cell, and percentage of mitochondrial reads per cell. The downstream processing was performed using SCANPY (*170*). Gene expression was normalized using the functions sc.pp.normalize_total and sc.pp.log1p, resulting in a ln(CPM) transformation of the input UMI counts data. Highly variable features were identified using sc.pp.highly_variable_genes, which were then scaled to unit variance and centered at zero using sc.pp.scale. Linear dimensionality reduction was performed on the scaled expression matrix using PCA with the function sc.tl.pca. Harmony (*178*) was used to batch correct the PCA representation with the function sc.external.pp.harmony_integrate. A cell neighborhood graph was computed based on the harmony representation using sc.pp.neighbors, followed by Leiden (*179*) clustering and non-linear UMAP dimensionality reduction with sc.tl.leiden and sc.tl.umap respectively. Canonical CNS cell-type marker genes were used to assign coarse-grain identities to each cluster, and to identify additional doublet clusters that passed our previous filtering steps. We inspected the distribution of the QC metrics in each cluster and removed outlier clusters. After filtering additional low-quality clusters, we ran UMAP and Leiden clustering again to result in the final processed anndata object for each dataset.

Spatial transcriptomics clustering analysis

In the human and mouse ST datasets, we grouped spots into biologically relevant clusters by accounting for transcriptome measurements and spatial coordinates. The BayesSpace (*180*) clustering algorithm uses a low-dimensional representation of the transcriptome with a spatial prior to encourage assignment of neighboring spots in the same cluster. Critically, BayesSpace produces a single unified clustering across many different ST experiments, rather than separate clustering and annotation for each ST slide. Seurat objects were converted to the SingleCellExperiment format using the function as.SingleCellExperiment. Absolute and relative spatial coordinates were stored in the meta-data compartment of the SingleCellExperiment objects to inform the BayesSpace model of the spatial information, ensuring to offset each sample such that there was no overlap. Each dataset was log normalized and linear dimensionality reduction was performed with PCA using the function spatialPreprocess from the BayesSpace R package. Harmony batch correction was applied on the basis of individual samples using the RunHarmony function. For the human dataset, we ran BayesSpace clustering using the spatialCluster function in the BayesSpace R package, varying the q parameter (number of resulting clusters) from 5 through 10. We inspected the output of each clustering and found that q=9 produced results that were most consistent with the underlying anatomy of the cortex, allowing us to annotate clusters based on cortical layers and white matter. Similarly, we ran BayesSpace clustering on the mouse dataset varying the q parameter between 10 and 20, and we selected q=15 for downstream analysis.

Reference-based integration of snRNA-seq datasets

We performed reference-based integration of the snRNA-seq dataset from the current study with the three published AD snRNA-seq datasets. Using our new snRNA-seq dataset as the reference, we projected the three published datasets into the reference latent space using scANVI (*181*), and we performed transfer learning to predict cell identities using scArches (*182*). While scANVI shares similarities with the scVI model that we previously used to process our snRNA-seq data, it is a semi-supervised model that leverages cell annotations in the reference dataset to inform the latent representation of the query dataset. We trained the scANVI model separately for each of the query datasets using the class scvi.model.SCANVI, training for 100 epochs in each case. For each query dataset, this process resulted in a low-dimensional representation of the transcriptome in the latent space originally learned from the reference snRNA-seq dataset with the model.get_latent_representation function, and predicted cell annotation labels from the model.predict function. We merged the reference dataset with the three query datasets, and we ran UMAP on the scANVI latent representation to visually represent the unified dataset in two dimensions.

Differential expression analysis

We used MAST (*183*) with the function FindMarkers from Seurat to identify differentially expressed genes for each single-nucleus or spatial cluster. For both human datasets, we compared the disease group (early-stage AD, late-stage AD, or AD in DS) to control, and for the 5XFAD dataset, we compared 5XFAD to WT for each timepoint. The following latent variables were used for each dataset: Study and nCount_RNA (late-stage AD snRNA-seq); Batch, nCount_RNA, and PMI (AD in DS snRNA-seq); Sequencing batch, Library batch, RIN, PMI, and nCount_Spatial (Human Visium); Sequencing batch, Library batch, Sacrifice batch, and nCount_Spatial (5XFAD Visium). Gene ontology term enrichment was performed using enrichR (*129*, *130*) (v 3.0), and we used ComplexHeatmap (*124*) (v 2.10.0) for visualization and hierarchical clustering. Human genes were converted into mouse orthologs with biomaRt (*184*, *185*) (v 2.50.0).

Spatial mapping of snRNA-seq data

We mapped our snRNA-seq dataset into spatial coordinates using the R package CellTrek (*186*) (v 0.0.94). Briefly, the CellTrek pipeline enables spatial mapping of single-cell transcriptomes by creating an integrated co-embedding of ST and single-cell data, followed by a multivariate random forest model to predict the biological coordinates from the shared feature space. In our testing, we found that this algorithm was limited in that it could not scale to large datasets comprising hundreds of thousands of single cells. Additionally, this algorithm only maps data to a single ST slide at a time. We also found that the CellTrek algorithm only provided predicted coordinates for a subset of the input single cell transcriptomes. For these reasons, we mapped our snRNA-seq frontal cortex data to the human ST dataset in a pairwise fashion for each snRNA-seq sample and each ST sample. For a given pair of ST and snRNA-seq samples, we constructed an integrated co-embedding using the CellTrek function traint with default parameters. We then iteratively mapped the single-cell transcriptomes into the ST coordinates using the celltrek function over three iterations. The second iteration only included cells that were not mapped in the first iteration, and the third iteration only included cells that were not mapped in the first or second iterations. We then computed the Euclidean distance between each mapped cell and each of the ST spots, and we labeled each cell with a spatial annotation based on the most frequently observed annotation among the labels of the ten closest spots. After running the pairwise CellTrek mappings, we compiled the results into a single table. In sum, this process yielded multiple spatial coordinates and multiple annotations for each cell across the 39 human ST samples in this study. Given that these tissue samples varied in their grey and white matter content, the CellTrek mappings and inferred spatial annotations are generally not consistent across the ST samples. To come up with a consensus regional annotation across the different spatial mappings, we excluded the mappings from ST samples which were excessively high in white matter or grey matter content. We computed a metric summarizing the grey to white matter ratio in each ST sample by counting the number of grey matter spots and white matter spots, taking the difference, and dividing by the total number of spots. Positive values indicate higher grey matter content, while negative values indicate higher white matter content. We excluded samples with greater than 0.9 and less than -0.3, thereby retaining mappings from 34 of the ST samples. For each cell, we counted the number of times it was mapped to each spatial region, and labeled the cell based on the most frequently mapped region across the different samples. We further simplified these spatial annotations by upper cortical, lower cortical, or white matter regions.

Cell-cell signaling analysis

We performed cell-cell signaling analysis in our snRNA-seq frontal cortex dataset with CellChat (*187*) (v 1.1.3), using the predicted spatial annotations in addition to cell-type labels. The human CellChatDB ligand-receptor interaction database was used for this analysis. To facilitate downstream comparisons of the signaling networks in DSAD versus control samples, we ran the CellChat workflow separately based on disease status. The CellChat object was created using the normalized gene expression matrix and the cell-type annotations with the predicted spatial regions from CellTrek, removing any cell groups with fewer than 30 cells. We then ran the recommended CellChat workflow using the following functions: identifyOverExpressedGenes, identifyOverExpressedInteractions, projectData, computeCommunProb, filterCommunication, subsetCommunication, computeCommunProbPathway, aggregateNet, and netAnalysis_computeCentrality. The DSAD and control CellChat objects were merged into one object using the mergeCellChat function. We compared the signaling networks across conditions both functionally and structurally using the computeNetSimilarityPairwise function. Further, we used the rankNet function to compute the relative information flow changes between DSAD and control across all signaling pathways. We identified differentially expressed ligands and receptors as well as their signaling pathways using the identifyOverExpressedGenes function, visualizing selected results with the netVisual_bubble function.

Integration of amyloid imaging data and spatial transcriptomic data

We used custom automated imaging analysis protocols (General Analysis protocols on NIS-Elements) to obtain Amylo-glo+ and OC+ binaries, thresholding by intensity and size, as well as accounting for autofluorescence/non-specific staining by negative thresholding based on an empty channel. We exported the following values for each binary: Area (μm^2), EqDiameter, center

X and Y coordinates. Only 5XFAD samples were used for the mouse samples, as there is no amyloid pathology in WT mice. Samples with high background were excluded. We calculated the number of Amylo-glo or OC+ binaries per spot by testing for an intersection between a spot and binary with the following equation:

$$\sqrt{\left(X_{spot} - X_{binary}\right)^2 + \left(Y_{spot} - Y_{binary}\right)^2} \leq r_{spot} + r_{binary}$$

X_{spot} = X coordinate of spatial spot, Y_{spot} = Y coordinate of spatial spot

r_{spot} = radius of spatial spot

X_{binary} = X coordinate of binary, Y_{binary} = Y coordinate of binary

r_{binary} = radius of binary

The radius of a spatial spot was calculated according to values provided by 10x Genomics, where a spot is 55µm with a 100µm distance between spot centers, and expanded the radius of a spot to account for the gap between spots. The radius of a binary was calculated with EqDiameter. The number of binaries intersecting a given spot were summed, as well as their respective area and diameter, and the metrics were added as metadata values for the spatial spot in the SeuratObject. Amyloid-associated genes were identified by differential testing as previously described for diagnosis comparisons, but spots were grouped Amylo-glo+/- or OC+/-. Latent variables used were as follows: Library batch, RIN, and nCount_Spatial (human); Library batch, Sacrifice batch, and nCount_Spatial (mouse). We likewise used biomaRt to convert human genes to mouse orthologs to detect intersecting genes between human and mouse. Mouse genes were limited to those identified in the cortex and white matter.

Analysis of spatial proteomic data

Image preprocessing, segmentation, and single-cell measurements were performed using the steinbock pipeline (*188*) (v 0.15.0) with the Ilastik pixel classification and CellProfiler object segmentation workflow. Briefly, we trained an Ilastik classifier to identify nuclei based on the Ir191 and Ir193 channels, and we adjusted the CellProfiler pipeline to identify cells by dilating the nucleus objects with the DilateObjects module (star, size = 5). We also used steinbock to calculate object measurements, like mean intensity. We then imported the steinbock results as a SpatialExperiment object with imcRtools (v 1.0.0). Counts were transformed as an inverse hyperbolic sine, and we accounted for batch effects (different slides) with Harmony. We then used scran (v 1.22.1) to perform Louvain clustering (k = 20, rank-based edge weighting) and annotated clusters into neurons, microglia, and astrocytes based on marker gene expression. We used dittoSeq (*189*) (v 1.6.0) and Seurat, in which the SpatialExperiment object was converted into a SeuratObject with the as.Seurat function, for data visualization.

Results

Studying the Alzheimer's disease transcriptome with spatial and cellular resolution

We performed a cross-species study of spatially resolved gene expression changes in AD by generating spatial transcriptomic data (10x Genomics Visium) from both postmortem human prefrontal cortical tissue (n = 10 cognitively healthy controls, 9 early-stage AD, 10 late-stage AD, and 10 AD in DS; median 1,316 genes per spatial spot) and 5XFAD and wildtype (WT) mouse brains (n = 8-12 per group, 4 separate timepoints; median 2,438 genes per spatial spot) (Fig. 3.1a, Supplementary Fig. 3.1, Methods). We used BayesSpace, a clustering algorithm leveraging neighboring spot information, (*103*) to identify 9 transcriptionally distinct clusters in our human dataset—3 white matter (WM) clusters and 6 grey matter clusters encompassing the cortical layers—and 15 brain-region specific clusters in our mouse dataset (Fig. 3.1b-c, Methods). We annotated these clusters *post hoc* based on the expression of known marker genes and their localization within the tissue sections (Supplementary Fig. 3.1a-b, Supplementary Table 2 [ii]). Due to clear heterogeneity across the human samples, we examined the distribution of the spatial spots by diagnosis for each cluster and found that most clusters are well represented across the diagnoses, except for WM3 in early-stage AD cases (Supplementary Fig. 3.1c). In addition, two mouse clusters were sparsely distributed throughout tissue sections with no clear localization. One we identified as a mixture of erythrocytes and neurons based on marker gene expression, while we determined the other to be low quality spots based on the number of UMIs and did not include it in the rest of our analyses (Supplementary Fig. 3.1d).

We additionally generated snRNA-seq data (Parse Biosciences, SPLiT-seq) from cognitively healthy controls (n = 27 prefrontal cortex/FCX, 27 posterior cingulate cortex/PCC) and AD in DS (n = 21 FCX, 21 PCC; 55 individuals total) and integrated this dataset with 3 previously published datasets (*75*, *78*, *85*) (FCX; total n = 27 cognitively healthy controls, 23 early-stage AD, and 48 late-stage AD) for a total of 585,042 nuclei (Fig. 3.1d, Supplementary Fig. 3.2, Methods). In this integrated dataset, we identified not only layer-specific excitatory neuron subpopulations, but also vascular subpopulations as described previously (*74*) (Fig. 3.1e). Furthermore, we examined the expression of an "*ex vivo* activation" signature, previously defined as an artifact of sample processing (*190*). Although we did find expression of select *ex vivo* activation genes, like *HSPA1A*, *HSPA1B*, *FOS*, and *JUN*, in our datasets, the overall expression, calculated as a score (*191*), of this gene expression signature was relatively low (Supplementary Fig. 3.1h-i, 3.2g).

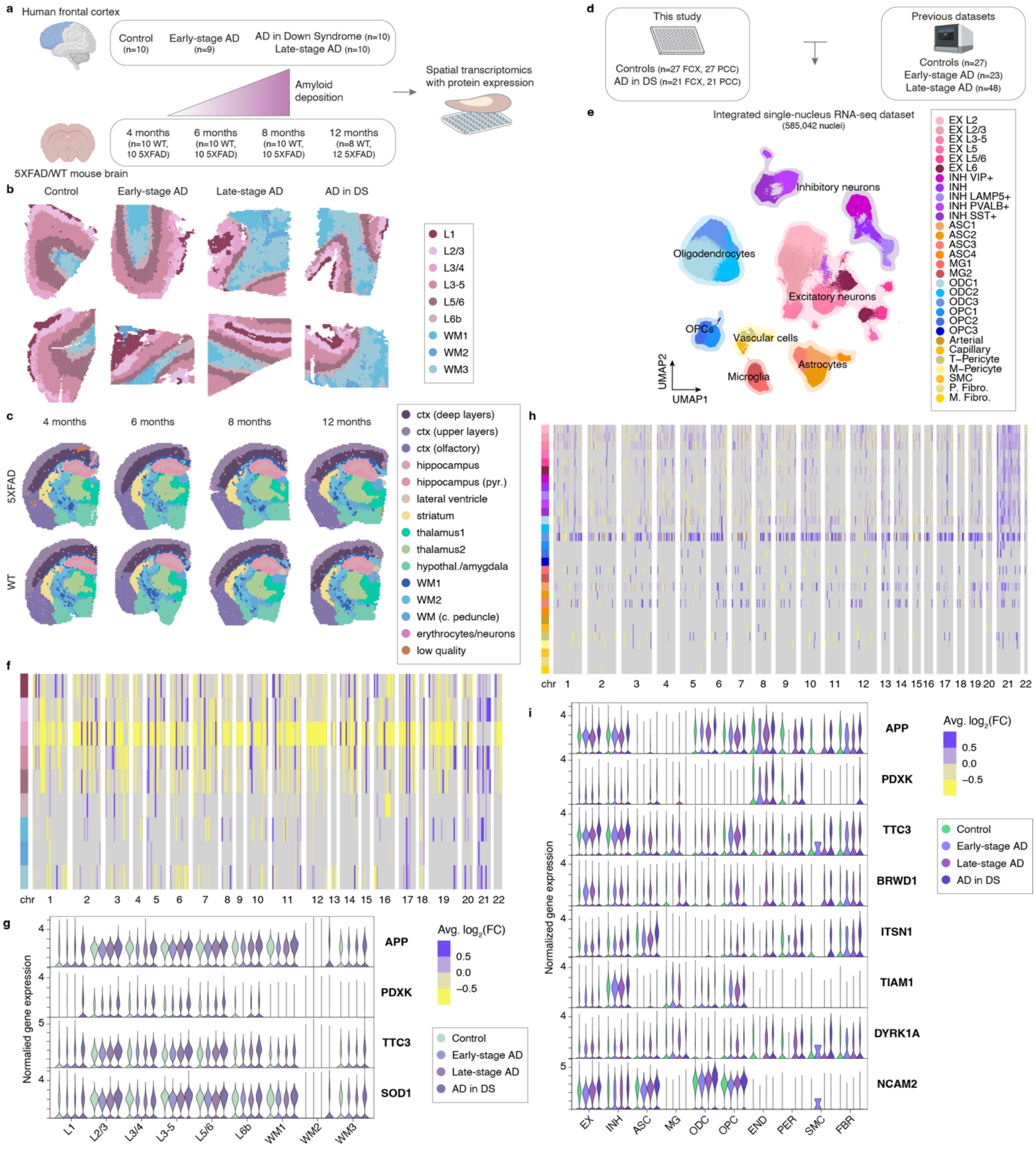
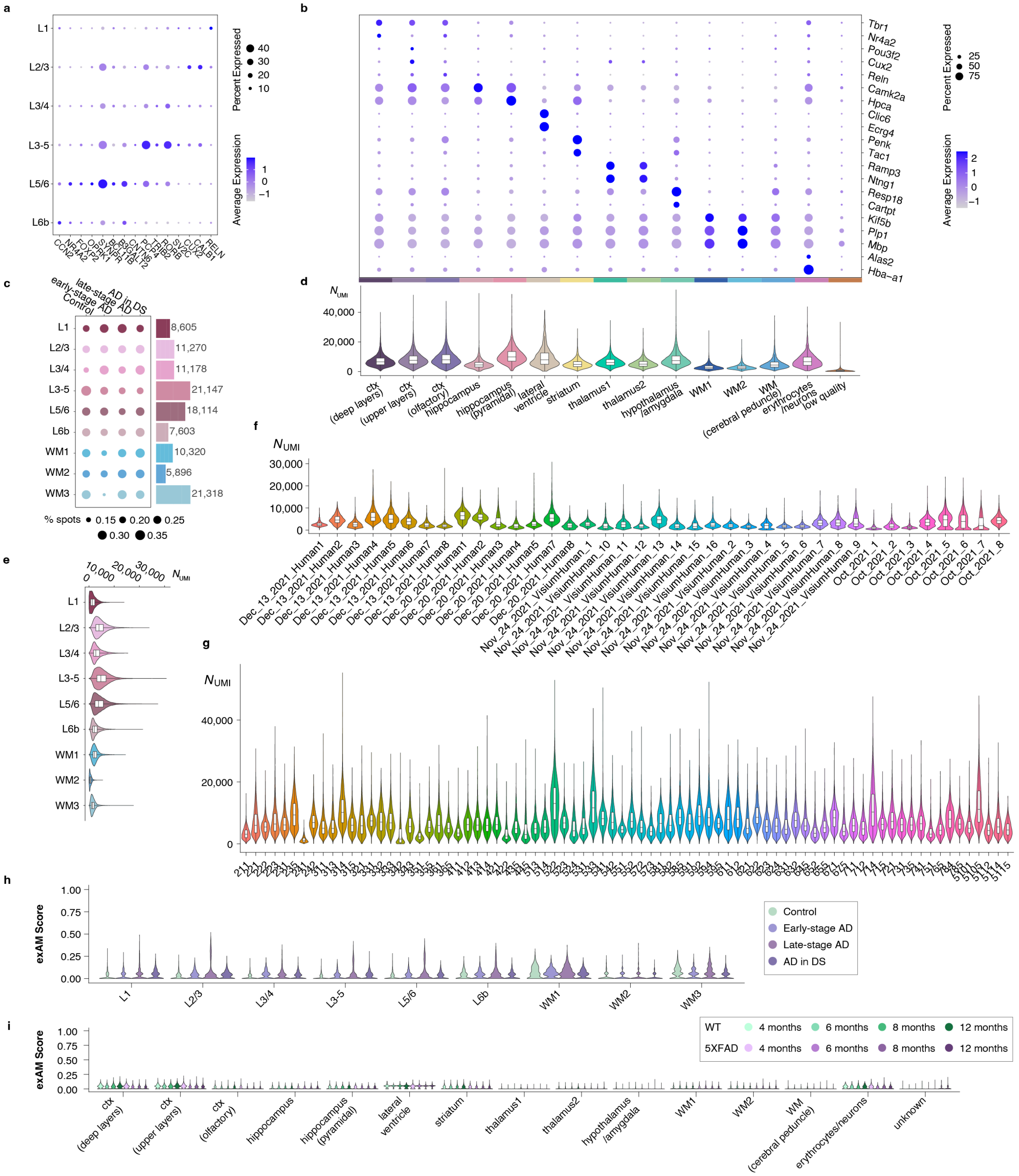


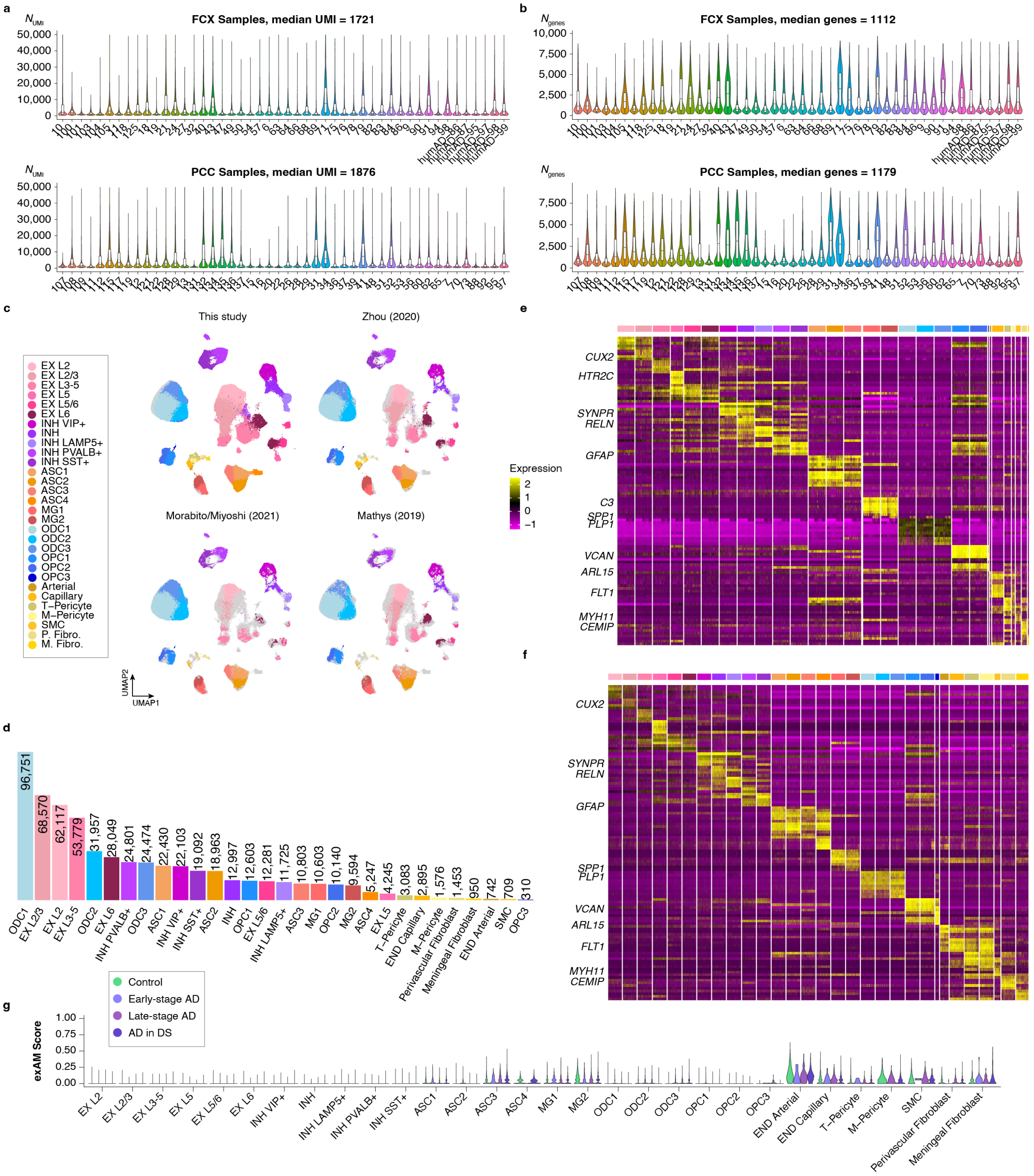
Figure 3.1. Spatial transcriptomics and snRNA-seq to study Alzheimer's disease

a, Schematic overview of the spatial transcriptomic datasets. **b**, Plots of spatial spots and their x, y coordinates from representative human samples, colored by BayesSpace cluster assignments. WM = white matter. **c**, Plots as in **(b)** from representative mouse samples. Ctx = cortex, pyr. = pyramidal, hypothal. = hypothalamus, c. peduncle = cerebral peduncle. **d**, Schematic overview of the snRNA-seq datasets. **e**, UMAP visualization of snRNA-seq clusters in the integrated dataset, colored by cluster assignments. ASC = astrocyte, Arterial = arteriole endothelial cell, Capillary = capillary endothelial cell, EX = excitatory neuron, INH = inhibitory neuron, MG = microglia, M. Fibro. = meningeal fibroblast, M-Pericyte = extracellular matrix pericyte, ODC = oligodendrocyte, OPC = oligodendrocyte precursor cell, P. Fibro. = perivascular fibroblast, SMC = smooth muscle cell, T-Pericyte = transport pericyte). **f**, Heatmap of differentially expressed genes (average $\log_2$(fold-change) cutoff = 0.5) by chromosome in AD in DS samples from the spatial transcriptomic dataset. **g**, Violin plots of genes residing in chromosome 21 by diagnosis and spatial cluster. **h**, Heatmap as in **(f)** but from the snRNA-seq dataset. **i**, Violin plots of genes residing in chromosome 21 by diagnosis and cell-type.



Supplementary Figure 3.1. Quality control for spatial transcriptomic datasets

a-b, Dot plots of marker gene expression in human spatial clusters **(a)** and mouse spatial clusters **(b)**. Size of the dot is the percent of spots expressing a gene. **c**, Left: Dot plot of percent of spots per human spatial cluster by diagnosis. Right: Bar plot of the total spatial spot number for each human spatial cluster. **d-e**, Violin plots of the number of unique molecular identifiers (UMIs) by mouse spatial cluster **(d)** and by human spatial cluster **(e)**. **f-g**, Violin plots as in **(d)** by individual human samples **(f)** and by individual mouse samples **(g)**. **h-i**, Violin plots of ex vivo activation score (*190*) by group across human **(h)** and mouse **(i)** spatial clusters.

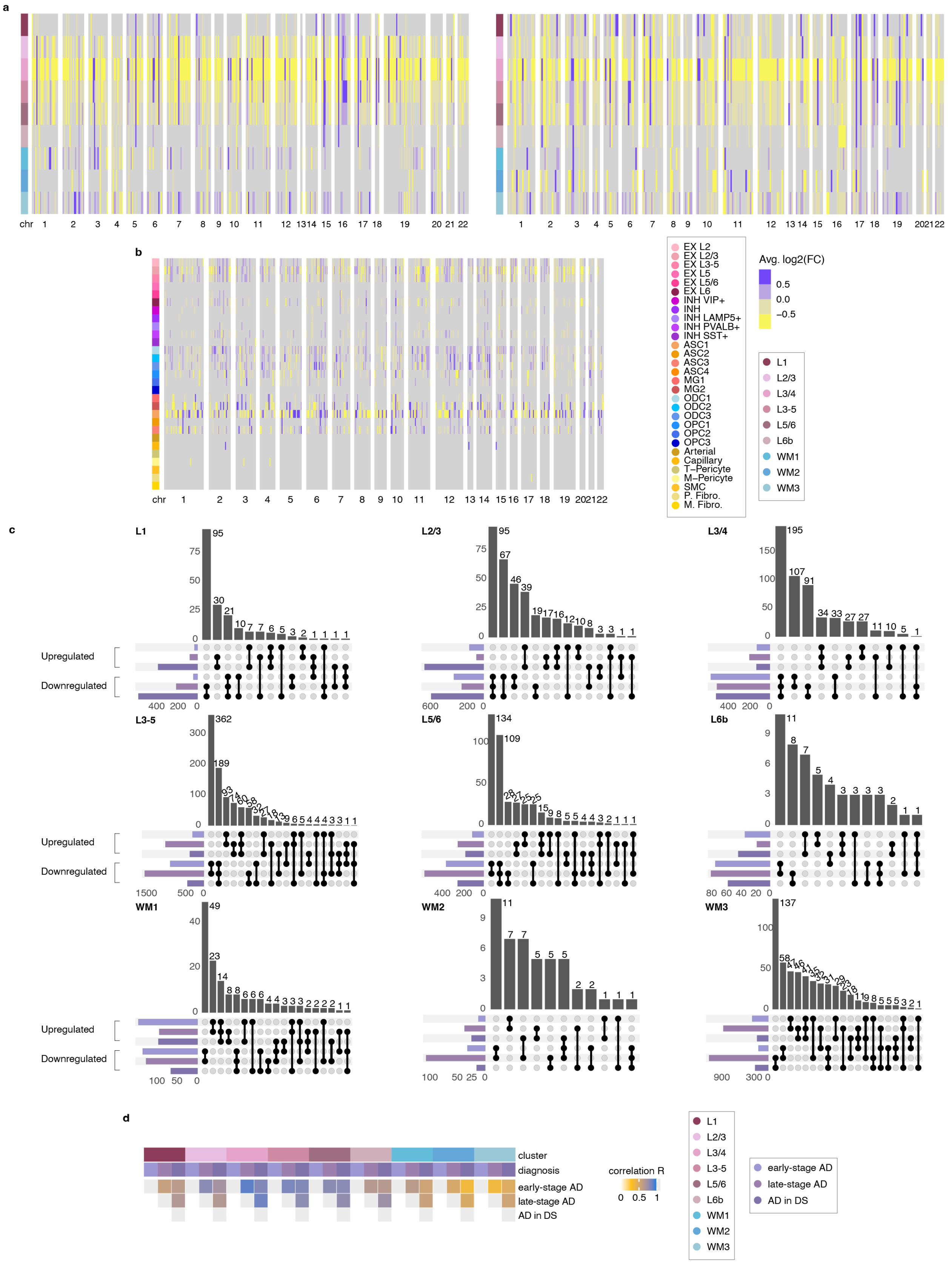


Supplementary Figure 3.2. Quality control for single-nucleus RNA-seq dataset

a, Violin plots of the number of unique molecular identifiers (UMIs) by sample, split by region frontal cortex (FCX, top) and posterior cingulate cortex (PCC, bottom). **b**, Violin plots as in **(a)** but showing the number of genes. **c**, UMAP visualizations of the individual snRNA-seq datasets comprising the integrated snRNA-seq dataset. Nuclei are colored by their cluster assignments in the integrated dataset. **d**, Bar plot of the number of nuclei in each cluster. **e-f**, Heatmaps of snRNA-seq cluster marker genes identified by differential expression analysis, split by the late-stage AD datasets **(e)** and AD in DS dataset **(f)**. **g**, Violin plots of ex vivo activation score (*190*) by group across the snRNA-seq clusters.

Regional and cell-type specific gene expression changes in clinical Alzheimer's disease

To identify disease-associated gene expression changes with cellular and spatial resolution, we performed differential expression analysis (with respect to controls) for our human spatial and single-nucleus datasets (Methods, Supplementary Table 3 [ii]). We focused our analyses on the FCX, since we had both spatial and single-nucleus data for this region. Considering the extra copy of chr21 suggests there will be overexpression of chr21 genes in our AD in DS samples, we first wanted to examine the differential expression of genes by chromosome. However, we found that not all chr21 genes are upregulated in both our spatial and single-nucleus datasets, and upregulation was dependent on region or cell-type (adjusted p-value < 0.05, Fig. 3.1f-h). For example, *APP* expectedly is upregulated in AD in DS samples but interestingly is not significantly different from control samples in spatial cluster L3/4. Our findings are in line with Palmer & Liu *et al.*'s snRNA-seq study of DS (*164*). Moreover, we found genes residing in other chromosomes are also significantly changing, similar to early- and late-stage AD (Supplementary Fig. 3.3a-b).


a
chr
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
b
chr
EX L2
EX L2/3
EX L3-5
EX L5
EX L5/6
EX L6
INH VIP+
INH
INH LAMP5+
INH PVALB+
INH SST+
ASC1
ASC2
ASC3
ASC4
MG1
MG2
ODC1
ODC2
ODC3
OPC1
OPC2
OPC3
Arterial
Capillary
T−Pericyte
M−Pericyte
SMC
P. Fibro.
M. Fibro.
Avg. log2(FC)
0.5
0.0
−0.5
L1
L2/3
L3/4
L3-5
L5/6
L6b
WM1
WM2
WM3
c
L1
L2/3
L3/4
L3-5
L5/6
L6b
WM1
WM2
WM3
Upregulated
Downregulated
d
cluster
diagnosis
early-stage AD
late-stage AD
AD in DS
correlation R
0 0.5 1

Supplementary Figure 3.3. Differential expressed genes by diagnosis in spatial and snRNA-seq data

a, Heatmaps of differentially expressed genes (DEGs, average $\log_2$(fold-change) cutoff = 0.5) by chromosome in early- (left) and late-stage AD (right) samples from the spatial transcriptomic dataset. **b**, Heatmap of DEGs as in **(a)** in late-stage AD samples from the snRNA-seq dataset. **c**, Upset plots of DEGs stratified by upregulated or downregulated in each diagnosis for each human spatial cluster. **d**, Heatmap of correlation coefficients from comparing average $\log_2$(fold-change) values. All comparisons were significant.

We next wanted to further assess the convergence or divergence of disease-associated gene expression changes between diagnoses for each cluster. In our spatial data, although the majority of DEGs were unique to a diagnosis, we still discovered many are conserved, including genes previously associated with AD, like *CST3*, *VIM*, *NEAT1*, and *CLU* in cluster L3-5 (Supplementary Fig. 3.3c). We also found significant and positive fold-change correlations across spatial and single-nucleus clusters, except in smaller vascular clusters and OPC2 (Supplementary Fig. 3.3d, 3.4). Notably correlations were stronger in grey matter clusters, compared to WM clusters, and this was matched in the single-nucleus data, where the correlations were also stronger in neuronal clusters, relative to glial clusters. We focused on the DEGs shared between late-stage AD and AD in DS, considering both conditions exhibit extensive amyloid and tau pathology. Disease-associated gene expression changes conserved between late-stage AD and AD in DS demonstrated strikingly region-specific patterns, separating by grey and WM (Fig. 3.2a). We additionally examined gene ontology (GO) term enrichment of these shared DEGs and found region-specific enrichment of AD-relevant biological pathways, such as upregulation of genes related to calcium signaling in WM3 and downregulation of those related to amyloid fibril formation in L3/4 and L3-5 (Fig. 3.2b). Furthermore, we found similar GO term enrichment in our cell subpopulations (Fig. 3.2c). Genes associated with NIK/NF-kappaB signaling were downregulated in spatial cluster L3/4 and excitatory neuron clusters EX L2/3 and L3-5. We also examined the differential expression of these same genes in the PCC and found that most are positively correlated with that in the FCX, with few unchanging in either region (Supplementary Fig. 3.5). In addition, effect sizes appeared to be larger in the PCC for select clusters ($\beta > 1$). Previously, we observed that gene expression changes mirror the anatomic progression of AD (*45*), and the PCC is a site of early changes in AD (*192*). Altogether we identified conserved regional and cellular transcriptional changes between late-stage AD and AD in DS, supporting the utility of studying AD in DS to understand AD molecular changes in both DS and the general population.

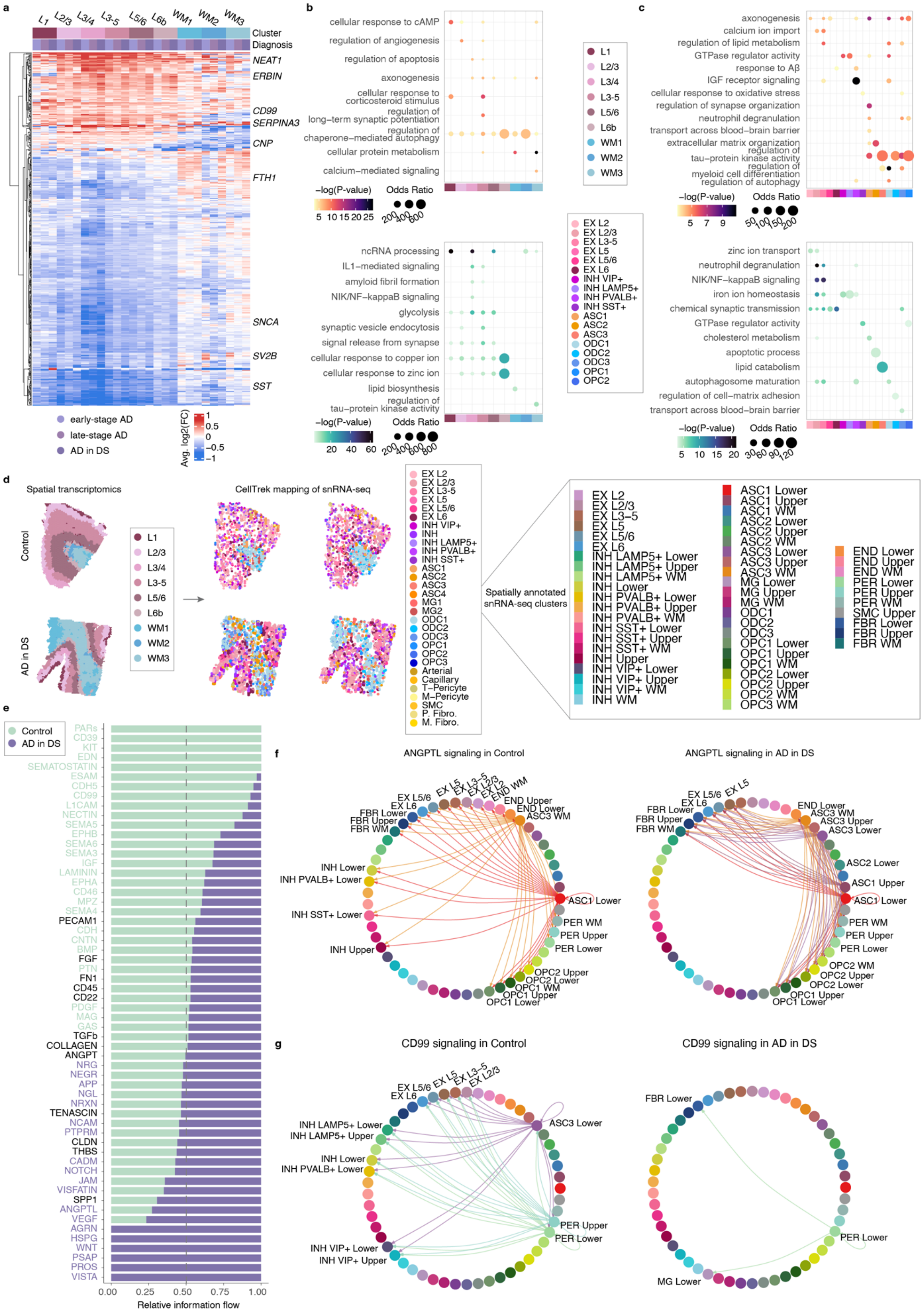
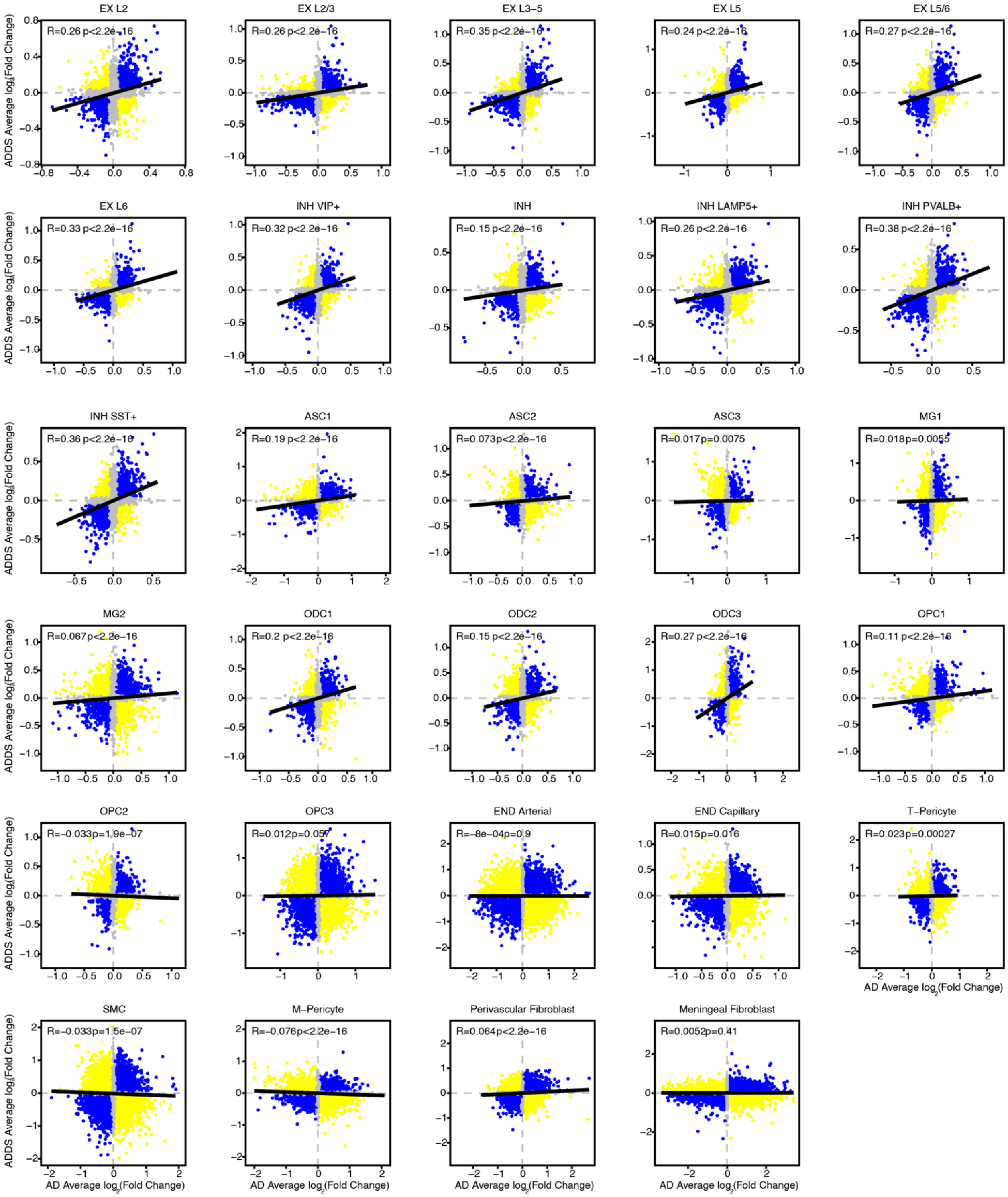

a
L1
L2/3
L3/4
L3-5
L5/6
L6b
WM1
WM2
WM3
Cluster
Diagnosis
NEAT1
ERBIN
CD99
SERPINA3
CNP
FTH1
SNCA
SV2B
SST
early-stage AD
late-stage AD
AD in DS
Avg. log2(FC)
1
0.5
0
−0.5
−1
b
cellular response to cAMP
regulation of angiogenesis
regulation of apoptosis
axonogenesis
cellular response to corticosteroid stimulus
regulation of long−term synaptic potentiation
regulation of chaperone−mediated autophagy
cellular protein metabolism
calcium−mediated signaling
−log(P-value)
5 10 15 20 25
Odds Ratio
200 400 600
ncRNA processing
IL1−mediated signaling
amyloid fibril formation
NIK/NF−kappaB signaling
glycolysis
synaptic vesicle endocytosis
signal release from synapse
cellular response to copper ion
cellular response to zinc ion
lipid biosynthesis
regulation of tau−protein kinase activity
−log(P-value)
20 40 60
Odds Ratio
200 400 600 800
L1
L2/3
L3/4
L3-5
L5/6
L6b
WM1
WM2
WM3
EX L2
EX L2/3
EX L3-5
EX L5
EX L5/6
EX L6
INH VIP+
INH LAMP5+
INH PVALB+
INH SST+
ASC1
ASC2
ASC3
ODC1
ODC2
ODC3
OPC1
OPC2
c
axonogenesis
calcium ion import
regulation of lipid metabolism
GTPase regulator activity
response to Aβ
IGF receptor signaling
cellular response to oxidative stress
regulation of synapse organization
neutrophil degranulation
transport across blood−brain barrier
extracellular matrix organization
regulation of tau−protein kinase activity
regulation of myeloid cell differentiation
regulation of autophagy
−log(P-value)
5 7 9
Odds Ratio
50 100 150 200
zinc ion transport
neutrophil degranulation
NIK/NF−kappaB signaling
iron ion homeostasis
chemical synaptic transmission
GTPase regulator activity
cholesterol metabolism
apoptotic process
lipid catabolism
autophagosome maturation
regulation of cell−matrix adhesion
transport across blood−brain barrier
−log(P-value)
5 10 15 20
Odds Ratio
30 60 90 120
d
Spatial transcriptomics
Control
AD in DS
L1
L2/3
L3/4
L3-5
L5/6
L6b
WM1
WM2
WM3
CellTrek mapping of snRNA-seq
EX L2
EX L2/3
EX L3-5
EX L5
EX L5/6
EX L6
INH VIP+
INH
INH LAMP5+
INH PVALB+
INH SST+
ASC1
ASC2
ASC3
ASC4
MG1
MG2
ODC1
ODC2
ODC3
OPC1
OPC2
OPC3
Arterial
Capillary
T−Pericyte
M−Pericyte
SMC
P. Fibro.
M. Fibro.
Spatially annotated snRNA-seq clusters
EX L2
EX L2/3
EX L3−5
EX L5
EX L5/6
EX L6
INH LAMP5+ Lower
INH LAMP5+ Upper
INH LAMP5+ WM
INH Lower
INH PVALB+ Lower
INH PVALB+ Upper
INH PVALB+ WM
INH SST+ Lower
INH SST+ Upper
INH SST+ WM
INH Upper
INH VIP+ Lower
INH VIP+ Upper
INH VIP+ WM
INH WM
ASC1 Lower
ASC1 Upper
ASC1 WM
ASC2 Lower
ASC2 Upper
ASC2 WM
ASC3 Lower
ASC3 Upper
ASC3 WM
MG Lower
MG Upper
MG WM
ODC1
ODC2
ODC3
OPC1 Lower
OPC1 Upper
OPC1 WM
OPC2 Lower
OPC2 Upper
OPC2 WM
OPC3 WM
END Lower
END Upper
END WM
PER Lower
PER Upper
PER WM
SMC Upper
FBR Lower
FBR Upper
FBR WM
e
Control
AD in DS
PARs
CD39
KIT
EDN
SEMATOSTATIN
ESAM
CDH5
CD99
L1CAM
NECTIN
SEMA5
EPHB
SEMA6
SEMA3
IGF
LAMININ
EPHA
CD46
MPZ
SEMA4
PECAM1
CDH
CNTN
BMP
FGF
PTN
FN1
CD45
CD22
PDGF
MAG
GAS
TGFb
COLLAGEN
ANGPT
NRG
NEGR
APP
NGL
NRXN
TENASCIN
NCAM
PTPRM
CLDN
THBS
CADM
NOTCH
JAM
VISFATIN
SPP1
ANGPTL
VEGF
AGRN
HSPG
WNT
PSAP
PROS
VISTA
0.00
0.25
0.50
0.75
1.00
Relative information flow
f
ANGPTL signaling in Control
EX L5/6
EX L6
EX L5
EX L3-5
EX L2/3
EX L2
END WM
END Upper
END Lower
ASC3 WM
FBR Lower
FBR Upper
FBR WM
INH Lower
INH PVALB+ Lower
INH SST+ Lower
INH Upper
ASC1 Lower
PER WM
PER Upper
PER Lower
OPC2 Upper
OPC2 Lower
OPC1 WM
OPC1 Upper
OPC1 Lower
ANGPTL signaling in AD in DS
EX L5/6
EX L5
EX L6
FBR Lower
FBR Upper
FBR WM
END Lower
ASC3 WM
ASC3 Upper
ASC3 Lower
ASC2 Lower
ASC1 Upper
ASC1 Lower
PER WM
PER Upper
PER Lower
OPC2 WM
OPC2 Upper
OPC2 Lower
OPC1 Upper
OPC1 Lower
g
CD99 signaling in Control
EX L5/6
EX L6
EX L5
EX L3-5
EX L2/3
ASC3 Lower
INH LAMP5+ Lower
INH LAMP5+ Upper
INH Lower
INH PVALB+ Lower
PER Upper
PER Lower
INH VIP+ Lower
INH VIP+ Upper
CD99 signaling in AD in DS
FBR Lower
PER Lower
MG Lower

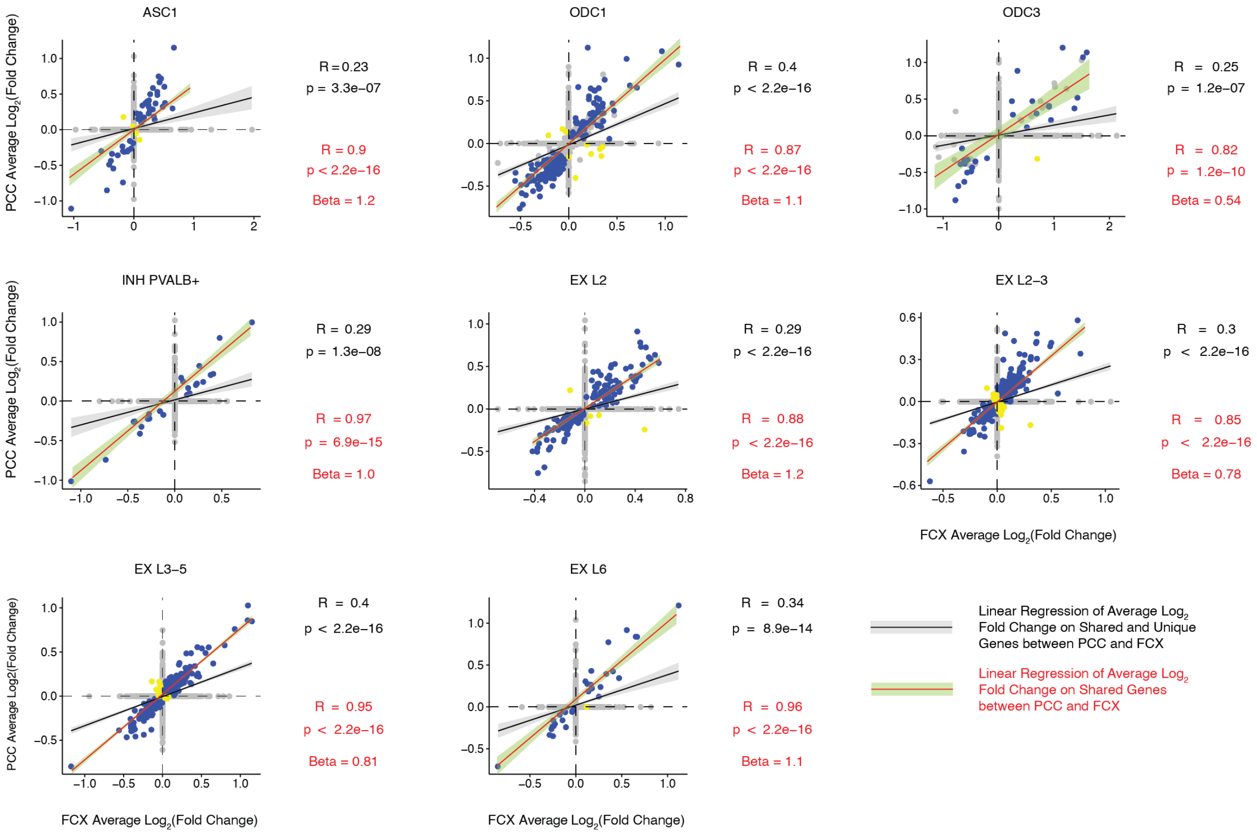
Figure 3.2. Spatially resolved transcriptomic changes in clinical Alzheimer's disease

a, Heatmap of average $\log_2$(fold-change) for differentially expressed genes (DEGs) shared between late-stage AD and AD in DS (adjusted p-value < 0.05 in both comparisons, average $\log_2$(fold-change) cutoff = 0.2) by diagnosis and spatial cluster. **b**, Dot plots of gene ontology (GO) term enrichment of late-stage AD and AD in DS shared DEGs (adjusted p-value < 0.05 in both comparisons) by spatial cluster, split by upregulated (top) and downregulated genes (bottom). **c**, Dot plots as in **(b)** but from snRNA-seq data, split by upregulated (top) and downregulated genes (bottom). **d**, Overview of our integration of spatial transcriptomics and snRNA-seq. Spatial transcriptomic samples were used to map nuclei profiled by snRNA-seq with CellTrek to spatial coordinates within tissue sections. We used our annotations of our spatial transcriptomic clusters to then assign spatial annotations to the snRNA-seq nuclei. **e**, Stacked bar plot of signaling pathways enriched in control (green) or AD in DS (purple) nuclei. **f-g**, Circle plots of ANGPTL **(f)** and CD99 **(g)** signaling pathway networks in control nuclei (left) and AD in DS nuclei (right) by spatially annotated snRNA-seq cluster. Arrows indicate direction (sender/receiver).


EX L2
R=0.26 p<2.2e−16
EX L2/3
R=0.26 p<2.2e−16
EX L3−5
R=0.35 p<2.2e−16
EX L5
R=0.24 p<2.2e−16
EX L5/6
R=0.27 p<2.2e−16
EX L6
R=0.33 p<2.2e−16
INH VIP+
R=0.32 p<2.2e−16
INH
R=0.15 p<2.2e−16
INH LAMP5+
R=0.26 p<2.2e−16
INH PVALB+
R=0.38 p<2.2e−16
INH SST+
R=0.36 p<2.2e−16
ASC1
R=0.19 p<2.2e−16
ASC2
R=0.073p<2.2e−16
ASC3
R=0.017p=0.0075
MG1
R=0.018p=0.0055
MG2
R=0.067p<2.2e−16
ODC1
R=0.2 p<2.2e−16
ODC2
R=0.15 p<2.2e−16
ODC3
R=0.27 p<2.2e−16
OPC1
R=0.11 p<2.2e−16
OPC2
R=−0.033p=1.9e−07
OPC3
R=0.012p=0.057
END Arterial
R=−8e−04p=0.9
END Capillary
R=0.015p=0.016
T−Pericyte
R=0.023p=0.00027
SMC
R=−0.033p=1.5e−07
M−Pericyte
R=−0.076p<2.2e−16
Perivascular Fibroblast
R=0.064p<2.2e−16
Meningeal Fibroblast
R=0.0052p=0.41
ADDS Average log2(Fold Change)
AD Average log2(Fold Change)

Supplementary Figure 3.4. Correlations of snRNA-seq DEGs identified in late-stage AD and AD in DS

Scatterplots comparing average $\log_2$(fold-change) values in late-stage AD (X-axis) and AD in DS (Y-axis) for each snRNA-seq cluster. Points are colored blue if shared between the diagnoses, yellow if their directionalities are opposite, and grey if below average $\log_2$(fold-change) cutoff of 0.05.



Supplementary Figure 3.5. Regional correlations of snRNA-seq DEGs identified in late-stage AD and AD in DS

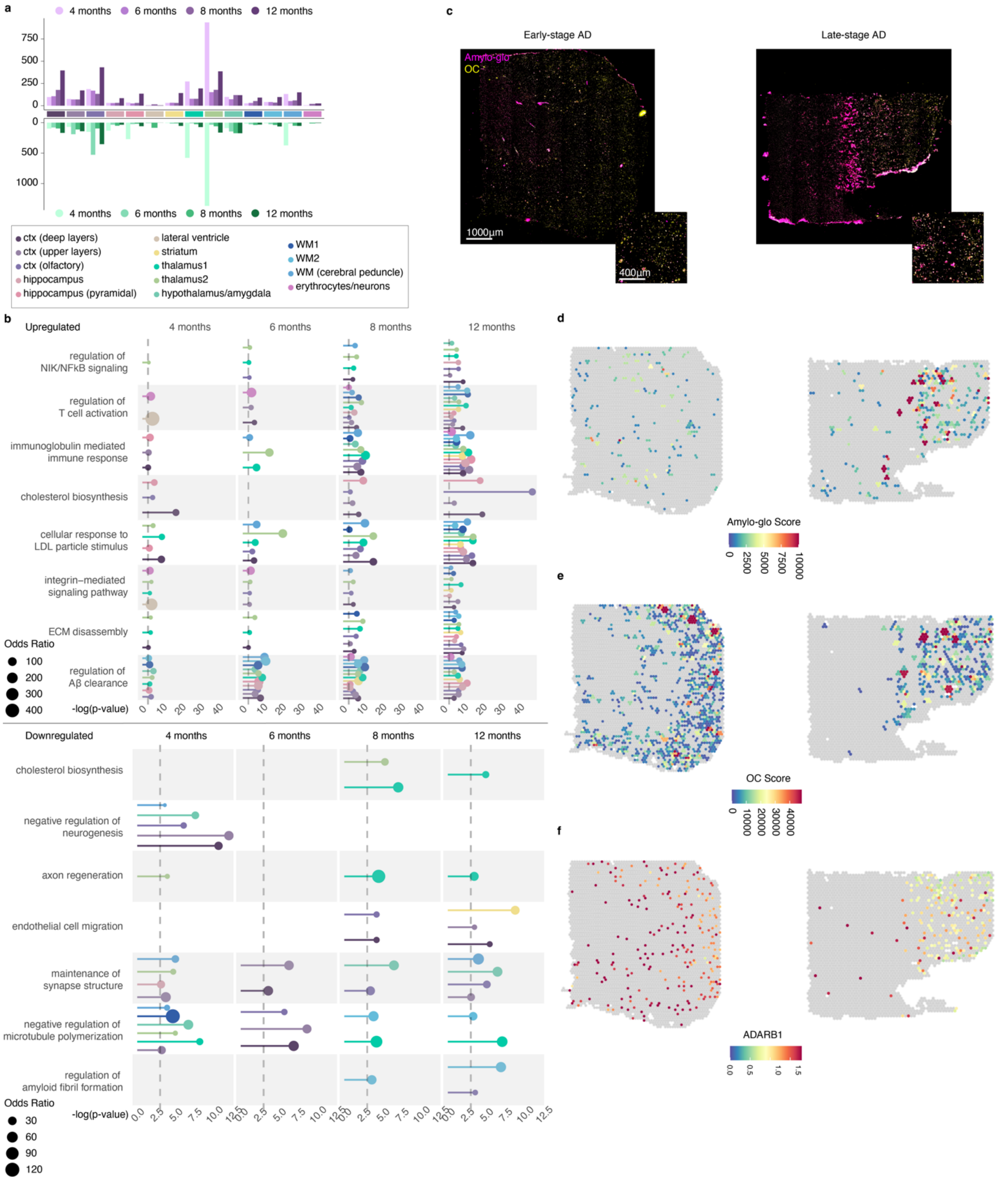
Scatterplots comparing average $\log_2$(fold-change) values in posterior cingulate cortex (PCC) (X-axis) and frontal cortex (FCX) (Y-axis) for each snRNA-seq cluster with more than 20 significant DEGs shared between late-stage AD and AD in DS. Points are colored blue if shared between the regions, yellow if their directionalities are opposite, and grey if not significant (adjusted p-value < 0.05).

Spatial annotation of cell clusters to identify cell-cell signaling dysregulated in disease

While excitatory neuron subpopulations clustered and easily could be annotated by cortical layer, in the case of inhibitory neurons and glia, we were not able to stratify these populations by region with snRNA-seq alone. Therefore, we sought to leverage our spatial transcriptomic data to infer the region localization of our snRNA-seq clusters. We used CellTrek, a method integrating spatial and single-cell data to predict the spatial coordinates of a cell in a tissue section (*193*) (Fig. 3.2d, Methods). With this expanded annotation of our snRNA-seq clusters, we wanted to assess how cell-cell communication may be dysregulated in AD with CellChat (*194*). CellChat identifies signaling pathways between cells based on the expression of a ligand and corresponding receptor. We rationalized that the additional spatial information would provide greater evidence for a communication pathway than snRNA-seq alone, since cells in proximity are more likely to be in communication with each other. We identified several signaling pathways up- or downregulated in disease (Fig. 3.2e, Supplementary Table 4 [ii]) and highlight ANGPTL and CD99 signaling pathways. We found ANGPTL signaling is upregulated in AD in DS, and in control samples, astrocyte clusters in cortical lower layers and the white matter (ASC1 and ASC3, respectively) communicate with a variety of cell-types, including neurons, pericytes, and oligodendrocyte progenitor cells (OPCs), by the ligand *ANGPTL4* (Fig. 3.2f). However, in AD in DS, additional astrocytes, like ASC1 and ASC3 in the cortical upper layers, also express *ANGPTL*4. Increased *ANGPTL4* expression has been previously observed in astrocytes from AD patients with vascular changes (*195*). Interestingly, we also see a loss of astrocyte-inhibitory neuron communication with disease. On the other hand, CD99 signaling is downregulated in AD in DS (Fig. 3.2g); we discovered pericyte and astrocyte (ASC3) communication with excitatory and inhibitory neuronal populations is absent, indicating that disrupted *CD99-CD99L2* signaling may underlie the neurovascular changes seen in AD. Notably, despite the overall downregulation of *CD99* signaling in AD in DS, we also found *CD99-PILRA* signaling between cortical lower layer pericytes and microglia exclusively in AD in DS. *PILRA*, a myeloid inhibitory signaling receptor, has been associated with AD genetic risk (*9*, *11*, *196*, *197*); our findings implicate pericytes in the modulation of microglial function in AD via *PILRA* activity.

Species-conserved disease-associated gene expression changes

In addition, we wanted to identify evolutionary-conserved AD transcriptional changes, since drug development largely relies on mouse models. However, mouse models of AD have been criticized for discrepancies with clinical AD. Considering previous literature shows baseline regional differences between human and mouse (*56*, *198*), we hypothesized that there may be also regional differences in disease. In addition, we believe that identification of the features shared between human and mouse will help forward *in* vivo research. We performed differential expression analysis on our mouse spatial transcriptomic dataset (5XFAD vs WT for each timepoint) and found an increasing number of upregulated genes over time, except in thalamic clusters, where the maximum was at the earliest timepoint, 4 months (Supplementary Fig. 3.6a-b, Supplementary Table 5 [ii]). Upregulated genes at 4 months in the thalamus included disease-associated microglia (DAM) genes (*77*), like *Cst7*, *Tyrobp*, *Ctsd*, and *Trem2*, suggesting an early response to plaques localized to the thalamus, and with increasing age, we found upregulation of these genes across brain regions.



Supplementary Figure 3.6. 5XFAD differentially expressed genes and human amyloid analysis

a, Bar plots of upregulated (top) and downregulated (bottom) genes identified in the 5XFAD by timepoint and spatial cluster. **b**, Lollipop plots of gene ontology (GO) term enrichment for upregulated (top) and downregulated (bottom) genes by timepoint and cluster. **c**, Representative whole-section images of the Amylo-glo and OC staining for amyloid pathology in early-stage (left) and late-stage AD (right) samples as in Fig. 3.3d. **d-f,** Plots of spatial spots and their x, y coordinates from the samples as in **(c)**, colored by Amylo-glo score (total number of intersecting Amylo-glo+ detections multiplied by the summed area) **(d)**, OC score (total number of intersecting OC+ detections multiplied by the summed area) **(e)**, and normalized *ADARB1* expression **(f)**.

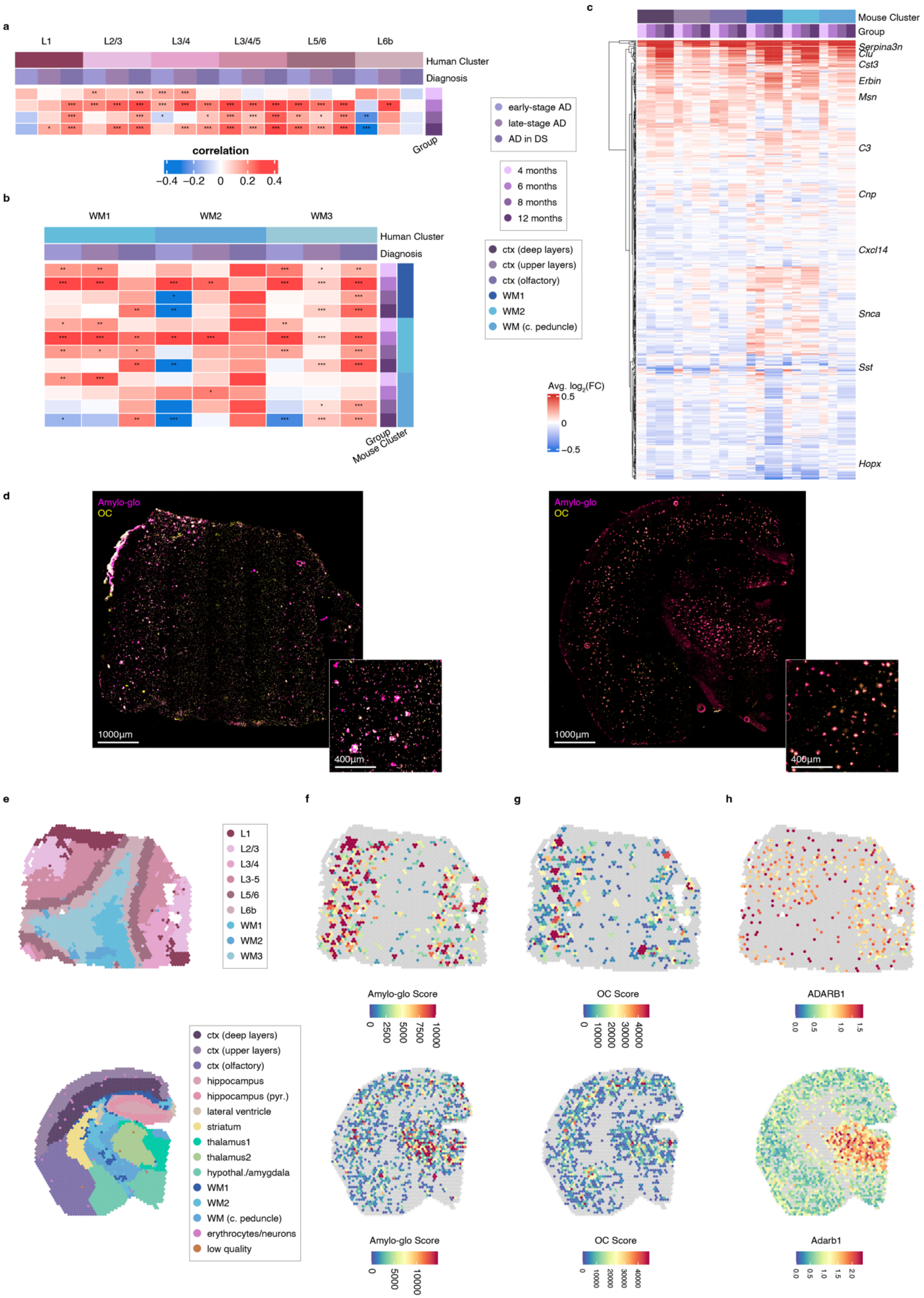
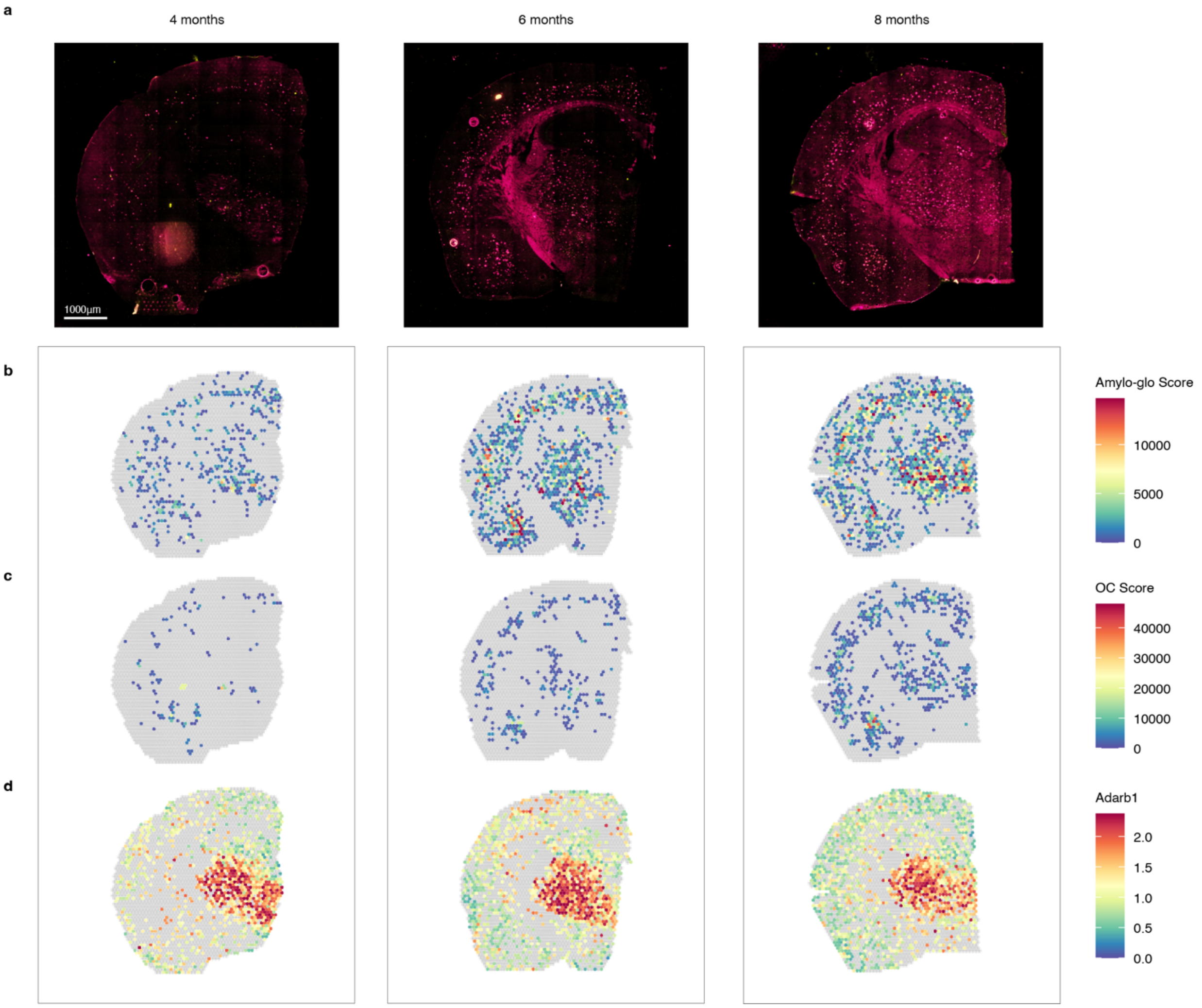

a
L1
L2/3
L3/4
L3/4/5
L5/6
L6b
Human Cluster
Diagnosis
Group
correlation
−0.4 −0.2 0 0.2 0.4
b
WM1
WM2
WM3
Human Cluster
Diagnosis
Group
Mouse Cluster
c
Mouse Cluster
Group
Serpina3n
Clu
Cst3
Erbin
Msn
C3
Cnp
Cxcl14
Snca
Sst
Hopx
early-stage AD
late-stage AD
AD in DS
4 months
6 months
8 months
12 months
ctx (deep layers)
ctx (upper layers)
ctx (olfactory)
WM1
WM2
WM (c. peduncle)
Avg. log$_2$(FC)
0.5
0
−0.5
d
Amylo-glo
OC
1000µm
400µm
Amylo-glo
OC
1000µm
400µm
e
L1
L2/3
L3/4
L3-5
L5/6
L6b
WM1
WM2
WM3
ctx (deep layers)
ctx (upper layers)
ctx (olfactory)
hippocampus
hippocampus (pyr.)
lateral ventricle
striatum
thalamus1
thalamus2
hypothal./amygdala
WM1
WM2
WM (c. peduncle)
erythrocytes/neurons
low quality
f
Amylo-glo Score
0 2500 5000 7500 10000
Amylo-glo Score
0 5000 10000
g
OC Score
0 10000 20000 30000 40000
OC Score
0 10000 20000 30000 40000
h
ADARB1
0.0 0.5 1.0 1.5
Adarb1
0.0 0.5 1.0 1.5 2.0

Figure 3.3. Species-conserved disease-associated gene expression changes

a, Heatmap of correlation coefficients from comparing average $\log_2$(fold-change) values between human and mouse spatial transcriptomic data for differentially expressed genes (DEGs) identified in human spatial transcriptomic data by cluster (adjusted p-value < 0.05). DEGs were identified individually for all diagnoses, and human L3-5, L5/6, and L6b DEGs were compared with mouse orthologs in mouse ctx (deep layers) at each timepoint. Human L1, L2/3, and L3/4 DEGs were compared with mouse orthologs in mouse ctx (upper layers) at each timepoint. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. **b**, Heatmap as in **(a)** but in human and mouse WM clusters. DEGs for each diagnosis from each human WM cluster was compared individually to a mouse WM cluster at each timepoint. **c**, Heatmap of average $\log_2$(fold-change) for DEGs identified in human spatial transcriptomic data that are also significantly changing in mouse spatial transcriptomic data (adjusted p-value < 0.05 in both human and mouse and with same directionality) by timepoint and mouse spatial cluster. **d,** Representative whole-section images of the Amylo-glo and OC staining for amyloid pathology in a human AD in DS sample (left) and a 12-month 5XFAD sample (right). The tissue sections were sectioned onto 10x Genomics Visium slides to perform immunofluorescence and generate cDNA libraries from the same tissue section. **e,** Plots of spatial spots and their x, y coordinates from the samples as in **(d)**, colored by cluster as in Fig. 3.1b-c. **f-h,** Plots of spatial spots as in **(e)**, colored by Amylo-glo score (total number of intersecting Amylo-glo+ detections multiplied by the summed area) **(f)**, OC score (total number of intersecting OC+ detections multiplied by the summed area) **(g)**, and normalized *ADARB1* expression **(h)**.

We then examined the expression of DEGs identified in our human spatial dataset in a matched comparison with our mouse spatial dataset (Fig. 3.3a-c). Mouse cortical upper and deep layer clusters were compared with the corresponding human cortical layer clusters (L1, L2/3, and L3/4; L3-5, L5/6, and L6b, respectively). However, white matter clusters were all individually compared with each other. We found significant and positive fold-change correlations in most cortical layer comparisons, except L1 and L6b. Like our comparison of AD in DS and late-stage AD, WM correlations were weaker, altogether indicating that the 5XFAD model recapitulates some, but not all, clinical AD changes. Our analyses compiled a list of species-conserved, regional DEGs comprising some of the genes previously identified in disease-associated glial signatures (*77*, *79*, *82*) (Supplementary Table 5 [ii]).

Furthermore, we investigated the preservation of amyloid-associated genes between mouse and human. The 5XFAD harbors five familial AD mutations inducing robust amyloid pathology. We performed immunostaining for amyloid pathology with Amylo-glo (*199*) and the conformation-specific antibody OC (*200*) prior to generating cDNA libraries for our mouse and human tissue sections, thus obtaining both protein and RNA expression data from the same tissue section (Fig. 3.3d-e). We integrated imaging analysis with the spatial transcriptomic data in order to obtain amyloid deposition scores for each spatial spot and identified genes upregulated in spatial spots with amyloid pathology (amyloid-associated genes, Fig. 3.3f-g, Supplementary Fig. 3.6-7, Supplementary Table 6 [ii], Methods). Similar to Chen *et al.* (*99*), we found limited overlap between mouse and human amyloid-associated genes, and our list of shared amyloid-associated genes contained 6 genes from their plaque-induced genes (PIGs)—*Apoe*, *Clu*, *Cst3*, *Gfap*, *Grn*, and *Vsir*. *Serpine2*, *Cpe*, and *Vim* were previously found in a proteomic study (*201*). We also identified novel genes, like *Adarb1*/*ADARB1*, which catalyzes adenosine to inosine RNA editing (Fig. 3.3h, Supplementary Fig. 3.6-7). Although a previous study found RNA editing decreases in AD (*202*), our results suggest that this may be spatially dependent, where differential RNA editing may be spatially restricted to amyloid pathology—possibly as a response to amyloid pathology.



Supplementary Figure 3.7. 5XFAD amyloid analysis

a, Representative whole-section images of the Amylo-glo and OC staining for amyloid pathology in 4, 6, and 8-month 5XFAD samples as in Fig. 3.3d. Amylo-glo intensity was increased to improve visualization of the tissue section. **b-d,** Plots of spatial spots and their x, y coordinates from the samples as in **(a)**, colored by Amylo-glo score (total number of intersecting Amylo-glo+ detections multiplied by the summed area) **(b)**, OC score (total number of intersecting OC+ detections multiplied by the summed area) **(c)**, and normalized *Adarb1* expression **(d)**.

Protein validation of transcriptomic analyses

Since gene expression measurements may not be reflective of protein expression, we used a spatial proteomic approach, imaging mass cytometry (IMC), in which antibodies are conjugated to metals, thereby avoiding autofluorescence issues inherent to aged postmortem human brain tissue. We generated IMC data from FCX tissue from 3 cognitively healthy controls, 4 late-stage AD, and 4 AD in DS and analyzed protein expression changes from a total of 76,178 cells (Fig. 3.4, Methods). After the removal of batch effects, we clustered the cells and annotated them based on the expression of GFAP, Map2, and Iba1 (Fig. 3.4c). We found that Cystatin C (*CST3*) is increased in microglia and astrocytes from late-stage AD and AD in DS (Fig. 3.4d). In our spatial transcriptomic data, *CST3* was upregulated in both groups in spatial clusters L3-5 and L5/6; however, we did not find the same congruence in our snRNA-seq data. *CST3* was upregulated in only ASC1 for AD in DS and MG2 for late-stage AD. We additionally discovered CD44 is upregulated in astrocytes from both diagnoses, but microglial upregulation of CD44 was confined to AD in DS (Fig. 3.4e). Similar to *CST3*, we found that *CD44* expression in the spatial transcriptomic data was more concordant than that in the snRNA-seq data. We expected *CD44* upregulation in astrocytes, but we did not detect microglial upregulation of *CD44* at the single-nucleus transcriptome. This discrepancy may be due to a variety of reasons, like lower resolution of microglia, but nevertheless points to the need for a comprehensive multi-omic approach.

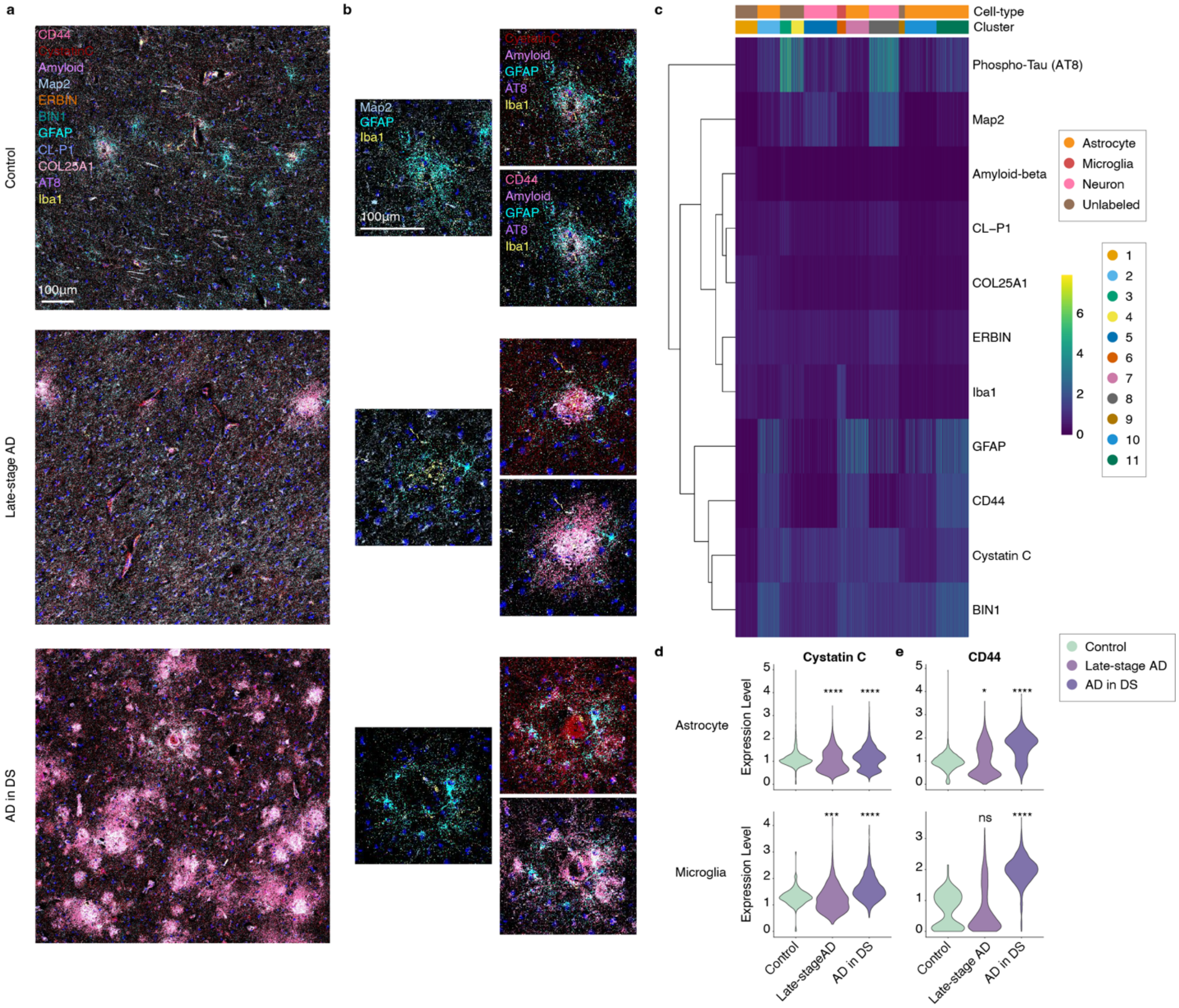


Figure 3.4. Spatial proteomics in clinical AD samples

a, Representative IMC images from control, late-stage AD, and AD in DS samples with select targets from the panel. **b**, Zoomed-in images as in **(a)** but showing only cell-type markers (left); amyloid and tau pathology localization with microglia and astrocytes, along with Cystatin C (right top) or CD44 (right bottom). **c**, Heatmap of transformed expression values for select targets in randomly sampled 5,000 cells. Louvain cluster assignments and cell-type annotations are shown above. **d**, Violin plots of Cystatin C expression in astrocytes (top) or microglia (bottom) for each diagnosis. **e**, Violin plots as in **(d)** but of CD44 expression. Two-sided Wilcoxon test (ns: $p > 0.05$, *: $p <= 0.05$, **: $p <= 0.01$, ***: $p <= 0.001$, ****: $p <= 0.0001$).

Discussion

Single-cell sequencing has uncovered cell-type and cell subpopulation-specific changes in disease, yielding novel insights into disease pathophysiology and gene targets for further study. However, a critical limitation of these methods is their inability to capture spatial information, and many disorders have spatially defined characteristics. AD pathology progresses in a predictably anatomical manner that may be linked to brain circuitry. We generated spatial transcriptomic data from postmortem human brain tissue samples of clinical AD, encompassing early- and late-stage AD in the general population, as well as AD in DS. Due to technical limitations, we could not profile gene expression at a macroscopic level of multiple brain regions, and our samples had varying amounts of cortical layers and white matter. We additionally generated spatial transcriptomic data from the 5XFAD mouse model of AD, in which we could assess both multiple brain regions and the temporal dynamics of disease progression, with the added benefit of lowered sample heterogeneity.

We identified regional transcriptomic changes shared between AD in the general population and DS, in line with previous accounts of shared genetic, clinical, and biomarker features (*23–25*). Further, we generated the first snRNA-seq study of AD in DS and integrated previously published AD studies to uncover cellular changes conserved between both populations. Both our human spatial and single-cell datasets are of the largest among currently published AD studies (n = 39 individuals; 55 new individuals in the integrated dataset, respectively). We leveraged a newly published method, CellTrek (*193*), to provide spatial coordinates for our snRNA-seq populations and the cell-cell communication analytical package CellChat (*194*) to reveal disease-associated cellular communication changes, which highlighted the role of astrocytes in AD pathophysiology and pinpointed downstream targets of AD astrocyte phenotype changes.

Our mouse spatial dataset is additionally the largest of our knowledge (n = 80 total), altogether representing a large resource for the AD community, and we investigated the expression of DEGs identified in clinical AD in the 5XFAD. Identification of cross-species disease changes are of particular interest for mouse model development and preclinical translation. We found regional evolutionary-conserved disease gene expression changes by analyzing both human and mouse spatial datasets. Furthermore, we integrated imaging analysis to reveal amyloid-associated genes shared between both species. Although previous 5XFAD studies identified amyloid-associated subpopulations and genes, not all may be recapitulated in clinical AD, and this has been previously highlighted in human AD snRNA-seq studies examining the DAM signature. While this may be due to single-cell vs single-nucleus comparisons, a previous study of snRNA-seq of human AD and 5XFAD (*78*), in addition to our previous re-analysis (*85*), demonstrated that this may be a species difference. In the present study, we present a list of species-conserved amyloid-associated genes, including previously identified and novel genes. However, we also note that the current resolution of spatial transcriptomics may have limited our findings.

Finally, we generated a spatial proteomic dataset to examine the protein expression of genes we identified in our transcriptomic datasets. Notably we found that findings in the spatial transcriptomic data were upheld at the protein level; however, we could not confirm layer-specific expression due to technical limitations. snRNA-seq results were less concordant with the proteomic data. Although this may be due to profiling only nuclear RNA, resultant protein expression may also be affected by posttranscriptional and posttranslational regulatory changes, as well as protein trafficking, which may not be captured in transcriptomic data. High-throughput methods to assess protein expression, however, are still limited, and the availability of effective antibodies remains a large impediment in the study of proteins. Altogether, this study clarifies the molecular changes occurring in clinical AD at both spatial and cellular levels and highlights those that are recapitulated in a commonly used AD mouse model.

[ii] Please contact for access.

Dissertation Concluding Remarks

AD remains a large societal issue, in which not only the patient, but also family and caregivers are affected. Years of research were poured into elucidating the role of Aβ in AD. Although the recent success with the anti-amyloid antibody lecanemab is uplifting, amyloid-targeting therapeutics have been largely underwhelming in clinical trials (*203*). Additionally, tau pathology was revealed to better correlate with cognitive decline than Aβ, thus causing attention to shift towards the mechanisms behind tau aggregation and propagation; however, anti-tau therapies likewise have been disappointing (*204*). Cognitively intact individuals with high amyloid and tau pathology (*205*), moreover, indicate that AD pathophysiology is more complex than its pathological hallmarks and impress a need to delineate the underlying biological pathways and regulators of neurodegeneration.

Early genetic studies of AD uncovered *APOE* as the strongest genetic risk factor for LOAD (*6*, *7*), and revolutionary technological advancements, in addition to the organization of several consortia, empowered us to perform large-scale AD GWAS. AD genetics is incredibly complex, and we have continued to identify novel loci associated with AD risk (*8–15*). However, there is also rising concern of the validity of these increasing numbers and whether they are truly AD risk factors, since recent cohorts have been comprised of individuals without pathologically confirmed diagnoses (*206*). Nevertheless, AD GWAS has expanded our knowledge of the disease by pinpointing new avenues for study, like the role of microglia and neuroinflammation in AD, and is a strong example of the beneficial outcomes of the ever-expanding "-omics" fields.

RNA and protein expression are more functional assessments of genomic changes, and multiple published studies have examined the transcriptome of AD (*26–35*), with increasing sample numbers as RNA-seq has become cheaper, as well as more widely accessible. Services, like Novogene, allow labs without RNA-seq expertise to generate transcriptomic data, and the increasing accessibility allows interdisciplinary research required for comprehensive investigation of complex disorders, like AD. We now realize, however, that “bulk” or tissue-level RNA-seq is limited; we obtain only the average of gene expression signals for a given tissue. The human brain is cellularly complex, and it is increasingly clear that we need to study the transcriptome at a cellular level. We and others have generated sc- or snRNA-seq data to study AD, revealing disease-associated gene expression changes are dependent on cell-type and cell subpopulation (*75–79*, *82*, *84*, *85*, *104*). In this dissertation, we build on previous literature describing disease-associated, transcriptionally-defined glial subpopulations in the 5XFAD mouse model and characterize similar subpopulations in human late-stage AD samples at both the transcriptome and epigenome (**Chapter Two**).

Much like the rapid emergence and adoption of single-cell profiling, we are now seeing a parallel in spatial profiling methods. The human brain is structurally organized into distinct functional regions, and neurons are spatially arranged in a laminar organization. AD pathology follows an anatomical progression (*207*), and cell-types, such as microglia and astrocytes, are morphologically distinct when in proximity or distal to Aβ pathology, suggesting that they may also be transcriptionally distinct. However, because snRNA-seq does not retain spatial information, we could not concretely associate AD molecular phenotypes with spatial localization. With spatial transcriptomics, we also cannot guarantee a 1:1 cell to spot ratio; one spatial barcode may represent the transcripts of a mixture of cells rather than one cell. Therefore, to overcome both assays’ limitations, we integrated spatial transcriptomic data with snRNA-seq data to define the spatial relationships between disease molecular signatures, and we expanded our analyses to the 5XFAD, as well as AD in DS, to perform a comprehensive, comparative study of the AD transcriptome (**Chapter Three**).

Although our analyses encompassed both neuron and glial populations, I will focus on our findings from microglia, oligodendrocytes, and astrocytes—the cell populations for which we constructed multi-omic trajectories. Microglia have been heavily implicated in AD; many AD GWAS risk genes are immune-related or expressed by microglia. With our snATAC-seq data, we confirmed the association of AD genetic risk factors and microglia. Our multi-omic trajectory analysis portrayed a transition from homeostatic to a disease-associated microglial state, similar to that identified in the 5XFAD (*77*). However, we found that the DAM signature identified in the 5XFAD is not fully recapitulated in clinical AD, and this is consistent with other snRNA-seq studies (*78*, *83*). We further revealed inconsistences between clinical AD and the 5XFAD mouse model in our spatial analyses but identified several microglial genes dysregulated in both species, including *C1qa*, *C1qb*, *Cd74, B2m, and Timp2*. We also highlighted the microglial TF PU.1, encoded by *SPI1*, in our single-nucleus multi-omic analysis, providing evidence of its role as a transcriptional repressor and its potential downstream targets. We note, though, several limitations that may have impeded our examination of microglia: artificial effects from sample processing, limited resolution or statistical power due to lower cell numbers, and a smaller cell size that impedes spatial profiling.

Oligodendrocytes are the dominant cell population of white matter regions and have been understudied in the context of AD. Dysregulation of oligodendrocytes may contribute to neurodegeneration through impairments in myelin formation and maintenance. We discovered substantial oligodendrocyte heterogeneity at both transcriptomic and epigenomic levels, mirroring that identified in other human and mouse studies (*60*, *68–70*). We found a decrease in both newly-formed and myelinating oligodendrocytes in late-stage AD, suggesting that there are changes in OPC differentiation and oligodendrocyte maturation. Additionally, we revealed SREBF1 as an oligodendrocyte-specific TF downregulated in AD, regulating genes that are members of AD-correlated oligodendrocyte gene co-expression networks. SREBF1 appears to act as a transcriptional activator, and we confirmed the downregulation of SREBF1 and its downstream target *ASCL4* with RNA *in situ* hybridization. Surprisingly we did not identify oligodendrocytes outside of the white matter in our integration of spatial and snRNA-seq data, but this may be a limitation of our integrative analysis that can be alleviated with a single-cell spatial profiling method.

Finally, astrocytes are morphologically and functionally diverse, with single-cell profiling expanding our understanding of astrocyte heterogeneity. Reactive astrocytes, marked by an increase in GFAP expression, around Aβ plaques have long been documented (*208*). Our multi-omic analysis similarly identified a disease-associated astrocyte state with increased *GFAP* expression, and these astrocytes had a similar gene expression signature to that found in the 5XFAD (*79*). We likewise discovered some of the same genes composing this signature (*Gfap*, *Ctsb*, *Vim*, *Serpina3n*) are upregulated in our human and mouse spatial data. Evolutionary-conserved amyloid-associated genes were also astrocytic (*Apoe*, *Clu*, *Cst3*, *Gfap*, *Vim*). Moreover, integration of snRNA-seq and spatial data allowed us to stratify our astrocyte subpopulations by WM and cortical layer localization, and cell-cell signaling analysis revealed regional changes in astrocyte communication with a broad number of cell-types, including neurons and pericytes. Astrocytes are known to be involved in synaptic transmission and blood-brain barrier maintenance; we discovered changes in astrocytic ANGPTL and CD99 signaling that highlight astrocyte modulation of brain vascular integrity in AD. Several AD dysregulated genes in astrocytes also include those related to the extracellular matrix (*CD44*, *CHI3L1*, *COLEC12*). Secreted and cell surface molecules, however, may not be fully captured with transcriptomics, and we can expect the rise of epitranscriptomics and proteomics, specifically secretomics, to further disentangle the molecular changes of AD.

Collectively this dissertation represents both technical and biological advancements in the study of AD. We developed protocols for technically challenging postmortem human brain tissue samples and generated precedent datasets for the AD community. In addition, the evaluation of both human and mouse samples is critical to forward our development of mouse models and therapeutics. We present six datasets available for researchers to further examine and highlight several areas for follow-up *in vitro* and *in vivo* mechanistic studies.

References

1. N. S. Ryan, M. N. Rossor, Correlating familial Alzheimers disease gene mutations with clinical phenotype. *Biomark Med*. 4, 99–112 (2010).

2. S. Forner, S. Kawauchi, G. Balderrama-Gutierrez, E. A. Kramár, D. P. Matheos, J. Phan, D. I. Javonillo, K. M. Tran, E. Hingco, C. da Cunha, N. Rezaie, J. A. Alcantara, D. Baglietto-Vargas, C. Jansen, J. Neumann, M. A. Wood, G. R. MacGregor, A. Mortazavi, A. J. Tenner, F. M. LaFerla, K. N. Green, Systematic phenotyping and characterization of the 5xFAD mouse model of Alzheimer's disease. *Sci Data*. 8, 270 (2021).

3. S. J. Webster, A. D. Bachstetter, L. J. V. Eldik, Comprehensive behavioral characterization of an APP/PS-1 double knock-in mouse model of Alzheimer's disease. *Alzheimer's Res Ther*. 5, 28 (2013).

4. T. Saito, Y. Matsuba, N. Mihira, J. Takano, P. Nilsson, S. Itohara, N. Iwata, T. C. Saido, Single App knock-in mouse models of Alzheimer's disease. *Nat Neurosci*. 17, 661–663 (2014).

5. M. Gatz, C. A. Reynolds, L. Fratiglioni, B. Johansson, J. A. Mortimer, S. Berg, A. Fiske, N. L. Pedersen, Role of Genes and Environments for Explaining Alzheimer Disease. *Arch Gen Psychiat*. 63, 168–174 (2006).

6. A. M. Saunders, W. J. Strittmatter, D. Schmechel, P. H. St. George-Hyslop, M. A. Pericak-Vance, S. H. Joo, B. L. Rosi, J. F. Gusella, D. R. Crapper-MacLachlan, M. J. Alberts, C. Hulette, B. Crain, D. Goldgaber, A. D. Roses, Association of apolipoprotein E allele ϵ4 with late-onset familial and sporadic Alzheimer's disease. *Neurology*. 43, 1467–1467 (1993).

7. E. Corder, A. Saunders, W. Strittmatter, D. Schmechel, P. Gaskell, G. Small, A. Roses, J. Haines, M. Pericak-Vance, Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*. 261, 921–923 (1993).

8. P. Hollingworth, D. Harold, R. Sims, A. Gerrish, J.-C. Lambert, M. M. Carrasquillo, R. Abraham, M. L. Hamshere, J. S. Pahwa, V. Moskvina, K. Dowzell, N. Jones, A. Stretton, C. Thomas, A. Richards, D. Ivanov, C. Widdowson, J. Chapman, S. Lovestone, J. Powell, P. Proitsi, M. K. Lupton, C. Brayne, D. C. Rubinsztein, M. Gill, B. Lawlor, A. Lynch, K. S. Brown, P. A. Passmore, D. Craig, B. McGuinness, S. Todd, C. Holmes, D. Mann, A. D. Smith, H. Beaumont, D. Warden, G. Wilcock, S. Love, P. G. Kehoe, N. M. Hooper, E. R. L. C. Vardy, J. Hardy, S. Mead, N. C. Fox, M. Rossor, J. Collinge, W. Maier, F. Jessen, E. Rüther, B. Schürmann, R. Heun, H. Kölsch, H. van den Bussche, I. Heuser, J. Kornhuber, J. Wiltfang, M. Dichgans, L. Frölich, H. Hampel, J. Gallacher, M. Hüll, D. Rujescu, I. Giegling, A. M. Goate, J. S. K. Kauwe, C. Cruchaga, P. Nowotny, J. C. Morris, K. Mayo, K. Sleegers, K. Bettens, S. Engelborghs, P. P. D. Deyn, C. V. Broeckhoven, G. Livingston, N. J. Bass, H. Gurling, A. McQuillin, R. Gwilliam, P. Deloukas, A. Al-Chalabi, C. E. Shaw, M. Tsolaki, A. B. Singleton, R. Guerreiro, T. W. Mühleisen, M. M. Nöthen, S. Moebus, K.-H. Jöckel, N. Klopp, H.-E. Wichmann, V. S. Pankratz, S. B. Sando, J. O. Aasly, M. Barcikowska, Z. K. Wszolek, D. W. Dickson, N. R. Graff-Radford, R. C. Petersen, A. D. N. Initiative, C. M. van Duijn, M. M. B. Breteler, M. A. Ikram, A. L. DeStefano, A. L. Fitzpatrick, O. Lopez, L. J. Launer, S. Seshadri, C. consortium, C.

Berr, D. Campion, J. Epelbaum, J.-F. Dartigues, C. Tzourio, A. Alpérovitch, M. Lathrop, E. consortium, T. M. Feulner, P. Friedrich, C. Riehle, M. Krawczak, S. Schreiber, M. Mayhaus, S. Nicolhaus, S. Wagenpfeil, S. Steinberg, H. Stefansson, K. Stefansson, J. Snaedal, S. Björnsson, P. V. Jonsson, V. Chouraki, B. Genier-Boley, M. Hiltunen, H. Soininen, O. Combarros, D. Zelenika, M. Delepine, M. J. Bullido, F. Pasquier, I. Mateo, A. Frank-Garcia, E. Porcellini, O. Hanon, E. Coto, V. Alvarez, P. Bosco, G. Siciliano, M. Mancuso, F. Panza, V. Solfrizzi, B. Nacmias, S. Sorbi, P. Bossù, P. Piccardi, B. Arosio, G. Annoni, D. Seripa, A. Pilotto, E. Scarpini, D. Galimberti, A. Brice, D. Hannequin, F. Licastro, L. Jones, P. A. Holmans, T. Jonsson, M. Riemenschneider, K. Morgan, S. G. Younkin, M. J. Owen, M. O'Donovan, P. Amouyel, J. Williams, Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet.* 43, 429–35 (2011).

9. B. W. Kunkle, B. Grenier-Boley, R. Sims, J. C. Bis, V. Damotte, A. C. Naj, A. Boland, M. Vronskaya, S. J. van der Lee, A. Amlie-Wolf, C. Bellenguez, A. Frizatti, V. Chouraki, E. R. Martin, K. Sleegers, N. Badarinarayan, J. Jakobsdottir, K. L. Hamilton-Nelson, S. Moreno-Grau, R. Olaso, R. Raybould, Y. Chen, A. B. Kuzma, M. Hiltunen, T. Morgan, S. Ahmad, B. N. Vardarajan, J. Epelbaum, P. Hoffmann, M. Boada, G. W. Beecham, J.-G. Garnier, D. Harold, A. L. Fitzpatrick, O. Valladares, M.-L. Moutet, A. Gerrish, A. V. Smith, L. Qu, D. Bacq, N. Denning, X. Jian, Y. Zhao, M. D. Zompo, N. C. Fox, S.-H. Choi, I. Mateo, J. T. Hughes, H. H. Adams, J. Malamon, F. Sanchez-Garcia, Y. Patel, J. A. Brody, B. A. Dombroski, M. C. D. Naranjo, M. Daniilidou, G. Eiriksdottir, S. Mukherjee, D. Wallon, J. Uphill, T. Aspelund, L. B. Cantwell, F. Garzia, D. Galimberti, E. Hofer, M. Butkiewicz, B. Fin, E. Scarpini, C. Sarnowski, W. S. Bush, S. Meslage, J. Kornhuber, C. C. White, Y. Song, R. C. Barber, S. Engelborghs, S. Sordon, D. Voijnovic, P. M. Adams, R. Vandenberghe, M. Mayhaus, L. A. Cupples, M. S. Albert, P. P. D. Deyn, W. Gu, J. J. Himali, D. Beekly, A. Squassina, A. M. Hartmann, A. Orellana, D. Blacker, E. Rodriguez-Rodriguez, S. Lovestone, M. E. Garcia, R. S. Doody, C. Munoz-Fernadez, R. Sussams, H. Lin, T. J. Fairchild, Y. A. Benito, C. Holmes, H. Karamujić-Čomić, M. P. Frosch, H. Thonberg, W. Maier, G. Roshchupkin, B. Ghetti, V. Giedraitis, A. Kawalia, S. Li, R. M. Huebinger, L. Kilander, S. Moebus, I. Hernández, M. I. Kamboh, R. Brundin, J. Turton, Q. Yang, M. J. Katz, L. Concari, J. Lord, A. S. Beiser, C. D. Keene, S. Helisalmi, I. Kloszewska, W. A. Kukull, A. M. Koivisto, A. Lynch, L. Tarraga, E. B. Larson, A. Haapasalo, B. Lawlor, T. H. Mosley, R. B. Lipton, V. Solfrizzi, M. Gill, W. T. Longstreth, T. J. Montine, V. Frisardi, M. Diez-Fairen, F. Rivadeneira, R. C. Petersen, V. Deramecourt, I. Alvarez, F. Salani, A. Ciaramella, E. Boerwinkle, E. M. Reiman, N. Fievet, J. I. Rotter, J. S. Reisch, O. Hanon, C. Cupidi, A. G. A. Uitterlinden, D. R. Royall, C. Dufouil, R. G. Maletta, I. de Rojas, M. Sano, A. Brice, R. Cecchetti, P. S. George-Hyslop, K. Ritchie, M. Tsolaki, D. W. Tsuang, B. Dubois, D. Craig, C.-K. Wu, H. Soininen, D. Avramidou, R. L. Albin, L. Fratiglioni, A. Germanou, L. G. Apostolova, L. Keller, M. Koutroumani, S. E. Arnold, F. Panza, O. Gkatzima, S. Asthana, D. Hannequin, P. Whitehead, C. S. Atwood, P. Caffarra, H. Hampel, I. Quintela, Á. Carracedo, L. Lannfelt, D. C. Rubinsztein, L. L. Barnes, F. Pasquier, L. Frölich, S. Barral, B. McGuinness, T. G. Beach, J. A. Johnston, J. T. Becker, P. Passmore, E. H. Bigio, J. M. Schott, T. D. Bird, J. D. Warren, B. F. Boeve, M. K. Lupton, J. D. Bowen, P. Proitsi, A. Boxer, J. F. Powell, J. R. Burke, J. S. K. Kauwe, J. M. Burns, M. Mancuso, J. D. Buxbaum, U. Bonuccelli, N. J. Cairns, A. McQuillin, C. Cao, G. Livingston, C. S. Carlson, N. J. Bass, C. M. Carlsson, J. Hardy, R. M. Carney, J. Bras, M. M. Carrasquillo, R. Guerreiro, M. Allen, H. C. Chui, E. Fisher, C. Masullo, E. A. Crocco, C. DeCarli, G. Bisceglio, M. Dick, L. Ma, R. Duara, N. R. Graff-Radford, D. A. Evans, A. Hodges, K. M. Faber, M. Scherer, K. B. Fallon, M. Riemenschneider, D. W. Fardo, R. Heun, M. R. Farlow, H. Kölsch, S. Ferris, M. Leber, T. M. Foroud, I. Heuser, D. R. Galasko, I. Giegling, M. Gearing, M. Hüll, D. H. Geschwind, J. R. Gilbert, J. Morris, R. C. Green, K. Mayo, J. H. Growdon, T. Feulner, R. L.

Hamilton, L. E. Harrell, D. Drichel, L. S. Honig, T. D. Cushion, M. J. Huentelman, P. Hollingworth, C. M. Hulette, B. T. Hyman, R. Marshall, G. P. Jarvik, A. Meggy, E. Abner, G. E. Menzies, L.-W. Jin, G. Leonenko, L. M. Real, G. R. Jun, C. T. Baldwin, D. Grozeva, A. Karydas, G. Russo, J. A. Kaye, R. Kim, F. Jessen, N. W. Kowall, B. Vellas, J. H. Kramer, E. Vardy, F. M. LaFerla, K.-H. Jöckel, J. J. Lah, M. Dichgans, J. B. Leverenz, D. Mann, A. I. Levey, S. Pickering-Brown, A. P. Lieberman, N. Klopp, K. L. Lunetta, H.-E. Wichmann, C. G. Lyketsos, K. Morgan, D. C. Marson, K. Brown, F. Martiniuk, C. Medway, D. C. Mash, M. M. Nöthen, E. Masliah, N. M. Hooper, W. C. McCormick, A. Daniele, S. M. McCurry, A. Bayer, A. N. McDavid, J. Gallacher, A. C. McKee, H. van den Bussche, M. Mesulam, C. Brayne, B. L. Miller, S. Riedel-Heller, C. A. Miller, J. W. Miller, A. Al-Chalabi, J. C. Morris, C. E. Shaw, A. J. Myers, J. Wiltfang, S. O'Bryant, J. M. Olichney, V. Alvarez, J. E. Parisi, A. B. Singleton, H. L. Paulson, J. Collinge, W. R. Perry, S. Mead, E. Peskind, D. H. Cribbs, M. Rossor, A. Pierce, N. S. Ryan, W. W. Poon, B. Nacmias, H. Potter, S. Sorbi, J. F. Quinn, E. Sacchinelli, A. Raj, G. Spalletta, M. Raskind, C. Caltagirone, P. Bossù, M. D. Orfei, B. Reisberg, R. Clarke, C. Reitz, A. D. Smith, J. M. Ringman, D. Warden, E. D. Roberson, G. Wilcock, E. Rogaeva, A. C. Bruni, H. J. Rosen, M. Gallo, R. N. Rosenberg, Y. Ben-Shlomo, M. A. Sager, P. Mecocci, A. J. Saykin, P. Pastor, M. L. Cuccaro, J. M. Vance, J. A. Schneider, L. S. Schneider, S. Slifer, W. W. Seeley, A. G. Smith, J. A. Sonnen, S. Spina, R. A. Stern, R. H. Swerdlow, M. Tang, R. E. Tanzi, J. Q. Trojanowski, J. C. Troncoso, V. M. V. Deerlin, L. J. V. Eldik, H. V. Vinters, J. P. Vonsattel, S. Weintraub, K. A. Welsh-Bohmer, K. C. Wilhelmsen, J. Williamson, T. S. Wingo, R. L. Woltjer, C. B. Wright, C.-E. Yu, L. Yu, Y. Saba, A. Pilotto, M. J. Bullido, O. Peters, P. K. Crane, D. Bennett, P. Bosco, E. Coto, V. Boccardi, P. L. D. Jager, A. Lleo, N. Warner, O. L. Lopez, M. Ingelsson, P. Deloukas, C. Cruchaga, C. Graff, R. Gwilliam, M. Fornage, A. M. Goate, P. Sanchez-Juan, P. G. Kehoe, N. Amin, N. Ertekin-Taner, C. Berr, S. Debette, S. Love, L. J. Launer, S. G. Younkin, J.-F. Dartigues, C. Corcoran, M. A. Ikram, D. W. Dickson, G. Nicolas, D. Campion, J. Tschanz, H. Schmidt, H. Hakonarson, J. Clarimon, R. Munger, R. Schmidt, L. A. Farrer, C. V. Broeckhoven, M. C. O'Donovan, A. L. DeStefano, L. Jones, J. L. Haines, J.-F. Deleuze, M. J. Owen, V. Gudnason, R. Mayeux, V. Escott-Price, B. M. Psaty, A. Ramirez, L.-S. Wang, A. Ruiz, C. M. van Duijn, P. A. Holmans, S. Seshadri, J. Williams, P. Amouyel, G. D. Schellenberg, J.-C. Lambert, M. A. Pericak-Vance, Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat Genet*. 51, 414–430 (2019).

10. J.-C. Lambert, S. Heath, G. Even, D. Campion, K. Sleegers, M. Hiltunen, O. Combarros, D. Zelenika, M. J. Bullido, B. Tavernier, L. Letenneur, K. Bettens, C. Berr, F. Pasquier, N. Fiévet, P. Barberger-Gateau, S. Engelborghs, P. D. Deyn, I. Mateo, A. Franck, S. Helisalmi, E. Porcellini, O. Hanon, E. A. D. I. Investigators, M. M. de Pancorbo, C. Lendon, C. Dufouil, C. Jaillard, T. Leveillard, V. Alvarez, P. Bosco, M. Mancuso, F. Panza, B. Nacmias, P. Bossù, P. Piccardi, G. Annoni, D. Seripa, D. Galimberti, D. Hannequin, F. Licastro, H. Soininen, K. Ritchie, H. Blanché, J.-F. Dartigues, C. Tzourio, I. Gut, C. V. Broeckhoven, A. Alpérovitch, M. Lathrop, P. Amouyel, Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet*. 41, 1094–1099 (2009).

11. J.-C. Lambert, C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, C. Bellenguez, G. Jun, A. L. DeStefano, J. C. Bis, G. W. Beecham, B. Grenier-Boley, G. Russo, T. A. Thornton-Wells, N. Jones, A. V. Smith, V. Chouraki, C. Thomas, M. A. Ikram, D. Zelenika, B. N. Vardarajan, Y. Kamatani, C.-F. Lin, A. Gerrish, H. Schmidt, B. Kunkle, M. L. Dunstan, A. Ruiz, M.-T. Bihoreau, S.-H. Choi, C. Reitz, F. Pasquier, P. Hollingworth, A. Ramirez, O. Hanon, A. L. Fitzpatrick, J. D. Buxbaum, D. Campion, P. K. Crane, C. Baldwin, T. Becker, V. Gudnason, C. Cruchaga, D.

Craig, N. Amin, C. Berr, O. L. Lopez, P. L. D. Jager, V. Deramecourt, J. A. Johnston, D. Evans, S. Lovestone, L. Letenneur, F. J. Morón, D. C. Rubinsztein, G. Eiriksdottir, K. Sleegers, A. M. Goate, N. Fiévet, M. J. Huentelman, M. Gill, K. Brown, M. I. Kamboh, L. Keller, P. Barberger-Gateau, B. McGuinness, E. B. Larson, R. Green, A. J. Myers, C. Dufouil, S. Todd, D. Wallon, S. Love, E. Rogaeva, J. Gallacher, P. S. George-Hyslop, J. Clarimon, A. Lleo, A. Bayer, D. W. Tsuang, L. Yu, M. Tsolaki, P. Bossù, G. Spalletta, P. Proitsi, J. Collinge, S. Sorbi, F. Sanchez-Garcia, N. C. Fox, J. Hardy, M. C. D. Naranjo, P. Bosco, R. Clarke, C. Brayne, D. Galimberti, M. Mancuso, F. Matthews, S. Moebus, P. Mecocci, M. D. Zompo, W. Maier, H. Hampel, A. Pilotto, M. Bullido, F. Panza, P. Caffarra, B. Nacmias, J. R. Gilbert, M. Mayhaus, L. Lannfelt, H. Hakonarson, S. Pichler, M. M. Carrasquillo, M. Ingelsson, D. Beekly, V. Alvarez, F. Zou, O. Valladares, S. G. Younkin, E. Coto, K. L. Hamilton-Nelson, W. Gu, C. Razquin, P. Pastor, I. Mateo, M. J. Owen, K. M. Faber, P. V. Jonsson, O. Combarros, M. C. O'Donovan, L. B. Cantwell, H. Soininen, D. Blacker, S. Mead, T. H. Mosley, D. A. Bennett, T. B. Harris, L. Fratiglioni, C. Holmes, R. F. A. G. de Bruijn, P. Passmore, T. J. Montine, K. Bettens, J. I. Rotter, A. Brice, K. Morgan, T. M. Foroud, W. A. Kukull, D. Hannequin, J. F. Powell, M. A. Nalls, K. Ritchie, K. L. Lunetta, J. S. K. Kauwe, E. Boerwinkle, M. Riemenschneider, M. Boada, M. Hiltunen, E. R. Martin, R. Schmidt, D. Rujescu, L.-S. Wang, J.-F. Dartigues, R. Mayeux, C. Tzourio, A. Hofman, M. M. Nöthen, C. Graff, B. M. Psaty, L. Jones, J. L. Haines, P. A. Holmans, M. Lathrop, M. A. Pericak-Vance, L. J. Launer, L. A. Farrer, C. M. van Duijn, C. V. Broeckhoven, V. Moskvina, S. Seshadri, J. Williams, G. D. Schellenberg, P. Amouyel, Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet*. 45, 1452–1458 (2013).

12. J. Schwartzentruber, S. Cooper, J. Z. Liu, I. Barrio-Hernandez, E. Bello, N. Kumasaka, A. M. H. Young, R. J. M. Franklin, T. Johnson, K. Estrada, D. J. Gaffney, P. Beltrao, A. Bassett, Genome-wide meta-analysis, fine-mapping and integrative prioritization implicate new Alzheimer's disease risk genes. *Nat Genet*. 53, 392–402 (2021).

13. C. Bellenguez, F. Küçükali, I. E. Jansen, L. Kleineidam, S. Moreno-Grau, N. Amin, A. C. Naj, R. Campos-Martin, B. Grenier-Boley, V. Andrade, P. A. Holmans, A. Boland, V. Damotte, S. J. van der Lee, M. R. Costa, T. Kuulasmaa, Q. Yang, I. de Rojas, J. C. Bis, A. Yaqub, I. Prokic, J. Chapuis, S. Ahmad, V. Giedraitis, D. Aarsland, P. Garcia-Gonzalez, C. Abdelnour, E. Alarcón-Martín, D. Alcolea, M. Alegret, I. Alvarez, V. Álvarez, N. J. Armstrong, A. Tsolaki, C. Antúnez, I. Appollonio, M. Arcaro, S. Archetti, A. A. Pastor, B. Arosio, L. Athanasiu, H. Bailly, N. Banaj, M. Baquero, S. Barral, A. Beiser, A. B. Pastor, J. E. Below, P. Benchek, L. Benussi, C. Berr, C. Besse, V. Bessi, G. Binetti, A. Bizarro, R. Blesa, M. Boada, E. Boerwinkle, B. Borroni, S. Boschi, P. Bossù, G. Bråthen, J. Bressler, C. Bresner, H. Brodaty, K. J. Brookes, L. I. Brusco, D. Buiza-Rueda, K. Bûrger, V. Burholt, W. S. Bush, M. Calero, L. B. Cantwell, G. Chene, J. Chung, M. L. Cuccaro, Á. Carracedo, R. Cecchetti, L. Cervera-Carles, C. Charbonnier, H.-H. Chen, C. Chillotti, S. Ciccone, J. A. H. R. Claassen, C. Clark, E. Conti, A. Corma-Gómez, E. Costantini, C. Custodero, D. Daian, M. C. Dalmasso, A. Daniele, E. Dardiotis, J.-F. Dartigues, P. P. de Deyn, K. de P. Lopes, L. D. de Witte, S. Debette, J. Deckert, T. del Ser, N. Denning, A. DeStefano, M. Dichgans, J. Diehl-Schmid, M. Diez-Fairen, P. D. Rossi, S. Djurovic, E. Duron, E. Düzel, C. Dufouil, G. Eiriksdottir, S. Engelborghs, V. Escott-Price, A. Espinosa, M. Ewers, K. M. Faber, T. Fabrizio, S. F. Nielsen, D. W. Fardo, L. Farotti, C. Fenoglio, M. Fernández-Fuertes, R. Ferrari, C. B. Ferreira, E. Ferri, B. Fin, P. Fischer, T. Fladby, K. Fließbach, B. Fongang, M. Fornage, J. Fortea, T. M. Foroud, S. Fostinelli, N. C. Fox, E. Franco-Macías, M. J. Bullido, A. Frank-García, L. Froelich, B. Fulton-Howard, D. Galimberti, J. M. García-Alberca, P. García-

González, S. Garcia-Madrona, G. Garcia-Ribas, R. Ghidoni, I. Giegling, G. Giorgio, A. M. Goate, O. Goldhardt, D. Gomez-Fonseca, A. González-Pérez, C. Graff, G. Grande, E. Green, T. Grimmer, E. Grünblatt, M. Grunin, V. Gudnason, T. Guetta-Baranes, A. Haapasalo, G. Hadjigeorgiou, J. L. Haines, K. L. Hamilton-Nelson, H. Hampel, O. Hanon, J. Hardy, A. M. Hartmann, L. Hausner, J. Harwood, S. Heilmann-Heimbach, S. Helisalmi, M. T. Heneka, I. Hernández, M. J. Herrmann, P. Hoffmann, C. Holmes, H. Holstege, R. H. Vilas, M. Hulsman, J. Humphrey, G. J. Biessels, X. Jian, C. Johansson, G. R. Jun, Y. Kastumata, J. Kauwe, P. G. Kehoe, L. Kilander, A. K. Ståhlbom, M. Kivipelto, A. Koivisto, J. Kornhuber, M. H. Kosmidis, W. A. Kukull, P. P. Kuksa, B. W. Kunkle, A. B. Kuzma, C. Lage, E. J. Laukka, L. Launer, A. Lauria, C.-Y. Lee, J. Lehtisalo, O. Lerch, A. Lleó, W. Longstreth, O. Lopez, A. L. de Munain, S. Love, M. Löwemark, L. Luckcuck, K. L. Lunetta, Y. Ma, J. Macías, C. A. MacLeod, W. Maier, F. Mangialasche, M. Spallazzi, M. Marquié, R. Marshall, E. R. Martin, A. M. Montes, C. M. Rodríguez, C. Masullo, R. Mayeux, S. Mead, P. Mecocci, M. Medina, A. Meggy, S. Mehrabian, S. Mendoza, M. Menéndez-González, P. Mir, S. Moebus, M. Mol, L. Molina-Porcel, L. Montrreal, L. Morelli, F. Moreno, K. Morgan, T. Mosley, M. M. Nöthen, C. Muchnik, S. Mukherjee, B. Nacmias, T. Ngandu, G. Nicolas, B. G. Nordestgaard, R. Olaso, A. Orellana, M. Orsini, G. Ortega, A. Padovani, C. Paolo, G. Papenberg, L. Parnetti, F. Pasquier, P. Pastor, G. Peloso, A. Pérez-Cordón, J. Pérez-Tur, P. Pericard, O. Peters, Y. A. L. Pijnenburg, J. A. Pineda, G. Piñol-Ripoll, C. Pisanu, T. Polak, J. Popp, D. Posthuma, J. Priller, R. Puerta, O. Quenez, I. Quintela, J. Q. Thomassen, A. Rábano, I. Rainero, F. Rajabli, I. Ramakers, L. M. Real, M. J. T. Reinders, C. Reitz, D. Reyes-Dumeyer, P. Ridge, S. Riedel-Heller, P. Riederer, N. Roberto, E. Rodriguez-Rodriguez, A. Rongve, I. R. Allende, M. Rosende-Roca, J. L. Royo, E. Rubino, D. Rujescu, M. E. Sáez, P. Sakka, I. Saltvedt, Á. Sanabria, M. B. Sánchez-Arjona, F. Sanchez-Garcia, P. S. Juan, R. Sánchez-Valle, S. B. Sando, C. Sarnowski, C. L. Satizabal, M. Scamosci, N. Scarmeas, E. Scarpini, P. Scheltens, N. Scherbaum, M. Scherer, M. Schmid, A. Schneider, J. M. Schott, G. Selbæk, D. Seripa, M. Serrano, J. Sha, A. A. Shadrin, O. Skrobot, S. Slifer, G. J. L. Snijders, H. Soininen, V. Solfrizzi, A. Solomon, Y. Song, S. Sorbi, O. Sotolongo-Grau, G. Spalletta, A. Spottke, A. Squassina, E. Stordal, J. P. Tartan, L. Tárraga, N. Tesí, A. Thalamuthu, T. Thomas, G. Tosto, L. Traykov, L. Tremolizzo, A. Tybjærg-Hansen, A. Uitterlinden, A. Ullgren, I. Ulstein, S. Valero, O. Valladares, C. V. Broeckhoven, J. Vance, B. N. Vardarajan, A. van der Lugt, J. V. Dongen, J. van Rooij, J. van Swieten, R. Vandenberghe, F. Verhey, J.-S. Vidal, J. Vogelgsang, M. Vyhnalek, M. Wagner, D. Wallon, L.-S. Wang, R. Wang, L. Weinhold, J. Wiltfang, G. Windle, B. Woods, M. Yannakoulia, H. Zare, Y. Zhao, X. Zhang, C. Zhu, M. Zulaica, J. Laczo, V. Matoska, M. Serpente, F. Assogna, F. Piras, F. Piras, V. Ciullo, J. Shofany, C. Ferrarese, S. Andreoni, G. Sala, C. P. Zoia, M. D. Zompo, A. Benussi, P. Bastiani, M. Takalo, T. Natunen, T. Laatikainen, J. Tuomilehto, R. Antikainen, T. Strandberg, J. Lindström, M. Peltonen, R. Abraham, A. Al-Chalabi, N. J. Bass, C. Brayne, K. S. Brown, J. Collinge, D. Craig, P. Deloukas, N. Fox, A. Gerrish, M. Gill, R. Gwilliam, D. Harold, P. Hollingworth, J. A. Johnston, L. Jones, B. Lawlor, G. Livingston, S. Lovestone, M. Lupton, A. Lynch, D. Mann, B. McGuinness, A. McQuillin, M. C. O'Donovan, M. J. Owen, P. Passmore, J. F. Powell, P. Proitsi, M. Rossor, C. E. Shaw, A. D. Smith, H. Gurling, S. Todd, C. Mummery, N. Ryan, G. Lacidogna, A. Adarmes-Gómez, A. Mauleón, A. Pancho, A. Gailhajenet, A. Lafuente, D. Macias-García, E. Martín, E. Pelejà, F. Carrillo, I. S. Merlín, L. Garrote-Espina, L. Vargas, M. Carrion-Claro, M. Marín, M. Labrador, M. Buendia, M. D. Alonso, M. Guitart, M. Moreno, M. Ibarria, M. Periñán, N. Aguilera, P. Gómez-Garre, P. Cañabate, R. Escuela, R. Pineda-Sánchez, R. Vigo-Ortega, S. Jesús, S. Preckler, S. Rodrigo-Herrero, S. Diego, A. Vacca, F. Roveta, N. Salvadori, E. Chipi, H. Boecker, C. Laske, R. Perneczky, C. Anastasiou, D. Janowitz, R. Malik, A. Anastasiou, K. Parveen, C. Lage, S. López-García, A. Antonell, K. Y. Mihova, D. Belezhanska, H. Weber, S. Kochen, P. Solis, N. Medel, J. Lisso, Z. Sevillano, D. G. Politis, V. Cores, C. Cuesta, C. Ortiz, J. I. Bacha,

M. Rios, A. Saenz, M. S. Abalos, E. Kohler, D. L. Palacio, I. Etchepareborda, M. Kohler, G. Novack, F. A. Prestia, P. Galeano, E. M. Castaño, S. Germani, C. R. Toso, M. Rojo, C. Ingino, C. Mangone, D. C. Rubinsztein, S. Teipel, N. Fievet, V. Deramerourt, C. Forsell, H. Thonberg, M. Bjerke, E. D. Roeck, M. T. Martínez-Larrad, N. Olivar, N. Aguilera, A. Cano, P. Cañabate, J. Macias, O. Maroñas, R. Nuñez-Llaves, C. Olivé, E. Pelejá, A. D. Adarmes-Gómez, M. D. Alonso, G. Amer-Ferrer, M. Antequera, J. A. Burguera, F. Carrillo, M. Carrión-Claro, M. J. Casajeros, M. M. de Pancorbo, R. Escuela, L. Garrote-Espina, P. Gómez-Garre, S. Hevilla, S. Jesús, M. A. L. Espinosa, A. Legaz, S. López-García, D. Macias-García, S. Manzanares, M. Marín, J. Marín-Muñoz, T. Marín, B. Martínez, V. Martínez, P. M.-L. Álvarez, M. M. Iriarte, M. T. Periñán-Tocino, R. Pineda-Sánchez, D. R. de Asúa, S. Rodrigo, I. Sastre, M. P. Vicente, R. Vigo-Ortega, L. Vivancos, J. Epelbaum, D. Hannequin, D. campion, V. Deramecourt, C. Tzourio, A. Brice, B. Dubois, A. Williams, C. Thomas, C. Davies, W. Nash, K. Dowzell, A. C. Morales, M. Bernardo-Harrington, J. Turton, J. Lord, K. Brown, E. Vardy, E. Fisher, J. D. Warren, M. Rossor, N. S. Ryan, R. Guerreiro, J. Uphill, N. Bass, R. Heun, H. Kölsch, B. Schürmann, A. Lacour, C. Herold, J. A. Johnston, P. Passmore, J. Powell, Y. Patel, A. Hodges, T. Becker, D. Warden, G. Wilcock, R. Clarke, P. Deloukas, Y. Ben-Shlomo, N. M. Hooper, S. Pickering-Brown, R. Sussams, N. Warner, A. Bayer, I. Heuser, D. Drichel, N. Klopp, M. Mayhaus, M. Riemenschneider, S. Pinchler, T. Feulner, W. Gu, H. van den Bussche, M. Hüll, L. Frölich, H.-E. Wichmann, K.-H. Jöckel, M. O'Donovan, M. Owen, S. Bahrami, I. Bosnes, P. Selnes, S. Bergh, A. Palotie, M. Daly, H. Jacob, A. Matakidou, H. Runz, S. John, R. Plenge, M. McCarthy, J. Hunkapiller, M. Ehm, D. Waterworth, C. Fox, A. Malarstig, K. Klinger, K. Call, T. Behrens, P. Loerch, T. Mäkelä, J. Kaprio, P. Virolainen, K. Pulkki, T. Kilpi, M. Perola, J. Partanen, A. Pitkäranta, R. Kaarteenaho, S. Vainio, M. Turpeinen, R. Serpi, T. Laitinen, J. Mäkelä, V.-M. Kosma, U. Kujala, O. Tuovila, M. Hendolin, R. Pakkanen, J. Waring, B. Riley-Gillis, J. Liu, S. Biswas, D. Diogo, C. Marshall, X. Hu, M. Gossel, R. Graham, B. Cummings, S. Ripatti, J. Schleutker, M. Arvas, O. Carpén, R. Hinttala, J. Kettunen, A. Mannermaa, J. Laukkanen, V. Julkunen, A. Remes, R. Kälviäinen, J. Peltola, P. Tienari, J. Rinne, A. Ziemann, J. Waring, S. Esmaeeli, N. Smaoui, A. Lehtonen, S. Eaton, S. Lahdenperä, J. van Adelsberg, J. Michon, G. Kerchner, N. Bowers, E. Teng, J. Eicher, V. Mehta, P. Gormley, K. Linden, C. Whelan, F. Xu, D. Pulford, M. Färkkilä, S. Pikkarainen, A. Jussila, T. Blomster, M. Kiviniemi, M. Voutilainen, B. Georgantas, G. Heap, F. Rahimov, K. Usiskin, T. Lu, D. Oh, K. Kalpala, M. Miller, L. McCarthy, K. Eklund, A. Palomäki, P. Isomäki, L. Pirilä, O. Kaipiainen-Seppänen, J. Huhtakangas, A. Lertratanakul, M. Hochfeld, N. Bing, J. E. Gordillo, N. Mars, M. Pelkonen, P. Kauppi, H. Kankaanranta, T. Harju, D. Close, S. Greenberg, H. Chen, J. Betts, S. Ghosh, V. Salomaa, T. Niiranen, M. Juonala, K. Metsärinne, M. Kähönen, J. Junttila, M. Laakso, J. Pihlajamäki, J. Sinisalo, M.-R. Taskinen, T. Tuomi, B. Challis, A. Peterson, A. Chu, J. Parkkinen, A. Muslin, H. Joensuu, T. Meretoja, L. Aaltonen, J. Mattson, A. Auranen, P. Karihtala, S. Kauppila, P. Auvinen, K. Elenius, R. Popovic, J. Schutzman, A. Loboda, A. Chhibber, H. Lehtonen, S. McDonough, M. Crohns, D. Kulkarni, K. Kaarniranta, J. A. Turunen, T. Ollila, S. Seitsonen, H. Uusitalo, V. Aaltonen, H. Uusitalo-Järvinen, M. Luodonpää, N. Hautala, S. Loomis, E. Strauss, H. Chen, A. Podgornaia, J. Hoffman, K. Tasanen, L. Huilaja, K. Hannula-Jouppi, T. Salmi, S. Peltonen, L. Koulu, I. Harvima, Y. Wu, D. Choy, P. Pussinen, A. Salminen, T. Salo, D. Rice, P. Nieminen, U. Palotie, M. Siponen, L. Suominen, P. Mäntylä, U. Gursoy, V. Anttonen, K. Sipilä, J. W. Davis, D. Quarless, S. Petrovski, E. Wigmore, C.-Y. Chen, P. Bronson, E. Tsai, Y. Huang, J. Maranville, E. Shaikho, E. Mohammed, S. Wadhawan, E. Kvikstad, M. Caliskan, D. Chang, T. Bhangale, S. Pendergrass, E. Holzinger, X. Chen, Å. Hedman, K. S. King, C. Wang, E. Xu, F. Auge, C. Chatelain, D. Rajpal, D. Liu, K. Call, T. Xia, M. Brauer, M. Kurki, J. Karjalainen, A. Havulinna, A. Jalanko, P. Palta, P. della B. Parolo, W. Zhou, S. Lemmelä, M. Rivas, J. Harju, A. Lehisto, A. Ganna, V. Llorens, H. Laivuori, S. Rüeger, M. E. Niemi, T. Tukiainen, M. P. Reeve,

H. Heyne, K. Palin, J. Garcia-Tabuenca, H. Siirtola, T. Kiiskinen, J. Lee, K. Tsuo, A. Elliott, K. Kristiansson, K. Hyvärinen, J. Ritari, M. Koskinen, K. Pylkäs, M. Kalaoja, M. Karjalainen, T. Mantere, E. Kangasniemi, S. Heikkinen, E. Laakkonen, C. Sipeky, S. Heron, A. Karlsson, D. Jambulingam, V. S. Rathinakannan, R. Kajanne, M. Aavikko, M. G. Jiménez, P. della B. Parola, A. Lehistö, M. Kanai, M. Kaunisto, E. Kilpeläinen, T. P. Sipilä, G. Brein, G. Awaisa, A. Shcherban, K. Donner, A. Loukola, P. Laiho, T. Sistonen, E. Kaiharju, M. Laukkanen, E. Järvensivu, S. Lähteenmäki, L. Männikkö, R. Wong, H. Mattsson, T. Hiekkalinna, T. Paajanen, K. Pärn, J. Gracia-Tabuenca, E. Abner, P. M. Adams, A. Aguirre, M. S. Albert, R. L. Albin, M. Allen, L. Alvarez, L. G. Apostolova, S. E. Arnold, S. Asthana, C. S. Atwood, G. Ayres, C. T. Baldwin, R. C. Barber, L. L. Barnes, S. Barral, T. G. Beach, J. T. Becker, G. W. Beecham, D. Beekly, J. E. Below, P. Benchek, B. A. Benitez, D. Bennett, J. Bertelson, F. E. Margaret, T. D. Bird, D. Blacker, B. F. Boeve, J. D. Bowen, A. Boxer, J. Brewer, J. R. Burke, J. M. Burns, W. S. Bush, J. D. Buxbaum, N. J. Cairns, C. Cao, C. S. Carlson, C. M. Carlsson, R. M. Carney, M. M. Carrasquillo, S. Chasse, M.-F. Chesselet, A. Chesi, N. A. Chin, H. C. Chui, J. Chung, S. Craft, P. K. Crane, D. H. Cribbs, E. A. Crocco, C. Cruchaga, M. L. Cuccaro, M. Cullum, E. Darby, B. Davis, P. L. D. Jager, C. DeCarli, J. DeToledo, M. Dick, D. W. Dickson, B. A. Dombroski, R. S. Doody, R. Duara, N. Ertekin-Taner, D. A. Evans, T. J. Fairchild, K. B. Fallon, M. R. Farlow, J. J. Farrell, V. Fernandez-Hernandez, S. Ferris, M. P. Frosch, B. Fulton-Howard, D. R. Galasko, A. Gamboa, M. Gearing, D. H. Geschwind, B. Ghetti, J. R. Gilbert, T. J. Grabowski, N. R. Graff-Radford, S. F. A. Grant, R. C. Green, J. H. Growdon, J. L. Haines, H. Hakonarson, J. Hall, R. L. Hamilton, O. Harari, L. E. Harrell, J. Haut, E. Head, V. W. Henderson, M. Hernandez, T. Hohman, L. S. Honig, R. M. Huebinger, M. J. Huentelman, C. M. Hulette, B. T. Hyman, L. S. Hynan, L. Ibanez, G. P. Jarvik, S. Jayadev, L.-W. Jin, K. Johnson, L. Johnson, M. I. Kamboh, A. M. Karydas, M. J. Katz, J. A. Kaye, C. D. Keene, A. Khaleeq, R. Kim, J. Knebl, N. W. Kowall, J. H. Kramer, P. P. Kuksa, F. M. LaFerla, J. J. Lah, E. B. Larson, C.-Y. Lee, E. B. Lee, A. Lerner, Y. Y. Leung, J. B. Leverenz, A. I. Levey, M. Li, A. P. Lieberman, R. B. Lipton, M. Logue, C. G. Lyketsos, J. Malamon, D. Mains, D. C. Marson, F. Martiniuk, D. C. Mash, E. Masliah, P. Massman, A. Masurkar, W. C. McCormick, S. M. McCurry, A. N. McDavid, S. McDonough, A. C. McKee, M. Mesulam, J. Mez, B. L. Miller, C. A. Miller, J. W. Miller, T. J. Montine, E. S. Monuki, J. C. Morris, A. J. Myers, T. Nguyen, S. O'Bryant, J. M. Olichney, M. Ory, R. Palmer, J. E. Parisi, H. L. Paulson, V. Pavlik, D. Paydarfar, V. Perez, E. Peskind, R. C. Petersen, J. E. Phillips-Cremins, A. Pierce, M. Polk, W. W. Poon, H. Potter, L. Qu, M. Quiceno, J. F. Quinn, A. Raj, M. Raskind, E. M. Reiman, B. Reisberg, J. S. Reisch, J. M. Ringman, E. D. Roberson, M. Rodriguear, E. Rogaeva, H. J. Rosen, R. N. Rosenberg, D. R. Royall, M. A. Sager, M. Sano, A. J. Saykin, J. A. Schneider, L. S. Schneider, W. W. Seeley, S. H. Slifer, S. Small, A. G. Smith, J. P. Smith, Y. E. Song, J. A. Sonnen, S. Spina, P. S. George-Hyslop, R. A. Stern, A. B. Stevens, S. M. Strittmatter, D. Sultzer, R. H. Swerdlow, R. E. Tanzi, J. L. Tilson, J. Q. Trojanowski, J. C. Troncoso, D. W. Tsuang, O. Valladares, V. M. V. Deerlin, L. J. van Eldik, R. Vassar, H. V. Vinters, J.-P. Vonsattel, S. Weintraub, K. A. Welsh-Bohmer, P. L. Whitehead, E. M. Wijsman, K. C. Wilhelmsen, B. Williams, J. Williamson, H. Wilms, T. S. Wingo, T. Wisniewski, R. L. Woltjer, M. Woon, C. B. Wright, C.-K. Wu, S. G. Younkin, C.-E. Yu, L. Yu, Y. Zhang, Y. Zhao, X. Zhu, H. Adams, R. O. Akinyemi, M. Ali, N. Armstrong, H. J. Aparicio, M. Bahadori, J. T. Becker, M. Breteler, D. Chasman, G. Chauhan, H. Comic, S. Cox, A. L. Cupples, G. Davies, C. S. DeCarli, M.-G. Duperron, J. Dupuis, T. Evans, F. Fan, A. Fitzpatrick, A. E. Fohner, M. Ganguli, M. Geerlings, S. J. Glatt, H. M. Gonzalez, M. Goss, H. Grabe, M. Habes, S. R. Heckbert, E. Hofer, E. Hong, T. Hughes, T. F. Kautz, M. Knol, W. Kremen, P. Lacaze, J. Lahti, Q. L. Grand, E. Litkowski, S. Li, D. Liu, X. Liu, M. Loitfelder, A. Manning, P. Maillard, R. Marioni, B. Mazoyer, D. M. van Lent, H. Mei, A. Mishra, P. Nyquist, J. O'Connell, Y. Patel, T. Paus, Z. Pausova, K. Raikkonen-Talvitie, M. Riaz, S. Rich, J. Rotter, J. Romero, G. Roshchupkin, Y. Saba, M.

Sargurupremraj, H. Schmidt, R. Schmidt, J. M. Shulman, J. Smith, H. Sekhar, R. Rajula, J. Shin, J. Simino, E. Sliz, A. Teumer, A. Thomas, A. Tin, E. Tucker-Drob, D. Vojinovic, Y. Wang, G. Weinstein, D. Williams, K. Wittfeld, L. Yanek, Y. Yang, L. A. Farrer, B. M. Psaty, M. Ghanbari, T. Raj, P. Sachdev, K. Mather, F. Jessen, M. A. Ikram, A. de Mendonça, J. Hort, M. Tsolaki, M. A. Pericak-Vance, P. Amouyel, J. Williams, R. Frikke-Schmidt, J. Clarimon, J.-F. Deleuze, G. Rossi, S. Seshadri, O. A. Andreassen, M. Ingelsson, M. Hiltunen, K. Sleegers, G. D. Schellenberg, C. M. van Duijn, R. Sims, W. M. van der Flier, A. Ruiz, A. Ramirez, J.-C. Lambert, New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet*. 54, 412–436 (2022).

14. R. E. Marioni, S. E. Harris, Q. Zhang, A. F. McRae, S. P. Hagenaars, W. D. Hill, G. Davies, C. W. Ritchie, C. R. Gale, J. M. Starr, A. M. Goate, D. J. Porteous, J. Yang, K. L. Evans, I. J. Deary, N. R. Wray, P. M. Visscher, GWAS on family history of Alzheimer's disease. *Transl Psychiat*. 8, 99 (2018).

15. I. E. Jansen, J. E. Savage, K. Watanabe, J. Bryois, D. M. Williams, S. Steinberg, J. Sealock, I. K. Karlsson, S. Hägg, L. Athanasiu, N. Voyle, P. Proitsi, A. Witoelar, S. Stringer, D. Aarsland, I. S. Almdahl, F. Andersen, S. Bergh, F. Bettella, S. Bjornsson, A. Brækhus, G. Bråthen, C. de Leeuw, R. S. Desikan, S. Djurovic, L. Dumitrescu, T. Fladby, T. J. Hohman, P. V. Jonsson, S. J. Kiddle, A. Rongve, I. Saltvedt, S. B. Sando, G. Selbæk, M. Shoai, N. G. Skene, J. Snaedal, E. Stordal, I. D. Ulstein, Y. Wang, L. R. White, J. Hardy, J. Hjerling-Leffler, P. F. Sullivan, W. M. van der Flier, R. Dobson, L. K. Davis, H. Stefansson, K. Stefansson, N. L. Pedersen, S. Ripke, O. A. Andreassen, D. Posthuma, Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet*. 51, 404–413 (2019).

16. J. F. Arboleda-Velasquez, F. Lopera, M. O'Hare, S. Delgado-Tirado, C. Marino, N. Chmielewska, K. L. Saez-Torres, D. Amarnani, A. P. Schultz, R. A. Sperling, D. Leyton-Cifuentes, K. Chen, A. Baena, D. Aguillon, S. Rios-Romenets, M. Giraldo, E. Guzmán-Vélez, D. J. Norton, E. Pardilla-Delgado, A. Artola, J. S. Sanchez, J. Acosta-Uribe, M. Lalli, K. S. Kosik, M. J. Huentelman, H. Zetterberg, K. Blennow, R. A. Reiman, J. Luo, Y. Chen, P. Thiyyagura, Y. Su, G. R. Jun, M. Naymik, X. Gai, M. Bootwalla, J. Ji, L. Shen, J. B. Miller, L. A. Kim, P. N. Tariot, K. A. Johnson, E. M. Reiman, Y. T. Quiroz, Resistance to autosomal dominant Alzheimer's disease in an APOE3 Christchurch homozygote: a case report. *Nat Med*. 25, 1680–1683 (2019).

17. T. Jonsson, J. K. Atwal, S. Steinberg, J. Snaedal, P. V. Jonsson, S. Bjornsson, H. Stefansson, P. Sulem, D. Gudbjartsson, J. Maloney, K. Hoyte, A. Gustafson, Y. Liu, Y. Lu, T. Bhangale, R. R. Graham, J. Huttenlocher, G. Bjornsdottir, O. A. Andreassen, E. G. Jönsson, A. Palotie, T. W. Behrens, O. T. Magnusson, A. Kong, U. Thorsteinsdottir, R. J. Watts, K. Stefansson, A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature*. 488, 96–99 (2012).

18. C. Sassi, M. A. Nalls, P. G. Ridge, J. R. Gibbs, J. Ding, M. K. Lupton, C. Troakes, K. Lunnon, S. Al-Sarraj, K. S. Brown, C. Medway, N. Clement, J. Lord, J. Turton, J. Bras, M. R. Almeida, A. Consortium, P. Passmore, D. Craig, J. Johnston, B. McGuinness, S. Todd, R. Heun, H. Kölsch, P. G. Kehoe, E. R. L. C. Vardy, N. M. Hooper, D. M. Mann, S. Pickering-Brown, K. Brown, J. Lowe, K. Morgan, A. D. Smith, G. Wilcock, D. Warden, C. Holmes, H. Holstege, E. Louwersheimer, W. M. van der Flier, P. Scheltens, J. C. V. Swieten, I. Santana, C.

Oliveira, K. Morgan, J. F. Powell, J. S. Kauwe, C. Cruchaga, A. M. Goate, A. B. Singleton, R. Guerreiro, J. Hardy, *Neurobiol Aging*, in press, doi:10.1016/j.neurobiolaging.2016.04.004.

19. Á. Fehér, Z. Giricz, A. Juhász, M. Pákáski, Z. Janka, J. Kálmán, ABCA1 rs2230805 and rs2230806 common gene variants are associated with Alzheimer's disease. *Neurosci Lett*. 664, 79–83 (2018).

20. Y. S. Davidson, A. Robinson, V. P. Prasher, D. M. A. Mann, The age of onset and evolution of Braak tangle stage and Thal amyloid pathology of Alzheimer's disease in individuals with Down syndrome. *Acta Neuropathologica Commun*. 6, 56 (2018).

21. D. M. A. Mann, The pathological association between down syndrome and Alzheimer disease. *Mech Ageing Dev*. 43, 99–136 (1988).

22. M. McCarron, P. McCallion, E. Reilly, P. Dunne, R. Carroll, N. Mulryan, A prospective 20-year longitudinal follow-up of dementia in persons with Down syndrome. *J Intell Disabil Res*. 61, 843–852 (2017).

23. J. Fortea, E. Vilaplana, M. Carmona-Iragui, B. Benejam, L. Videla, I. Barroeta, S. Fernández, M. Altuna, J. Pegueroles, V. Montal, S. Valldeneu, S. Giménez, S. González-Ortiz, L. Muñoz, T. Estellés, I. Illán-Gala, O. Belbin, V. Camacho, L. R. Wilson, T. Annus, R. S. Osorio, S. Videla, S. Lehmann, A. J. Holland, D. Alcolea, J. Clarimón, S. H. Zaman, R. Blesa, A. Lleó, Clinical and biomarker changes of Alzheimer's disease in adults with Down syndrome: a cross-sectional study. *Lancet*. 395, 1988–1997 (2020).

24. V. P. Prasher, S. G. Sajith, S. D. Rees, A. Patel, S. Tewari, N. Schupf, W. B. Zigman, Significant effect of APOE epsilon 4 genotype on the risk of dementia in Alzheimer's disease and mortality in persons with Down Syndrome. *Int J Geriatr Psych*. 23, 1134–1140 (2008).

25. J. H. Lee, A. J. Lee, L.-H. Dang, D. Pang, S. Kisselev, S. J. Krinsky-McHale, W. B. Zigman, J. A. Luchsinger, W. Silverman, B. Tycko, L. N. Clark, N. Schupf, Candidate gene analysis for Alzheimer's disease in adults with Down syndrome. *Neurobiol Aging*. 56, 150–158 (2017).

26. K. Bossers, K. T. S. Wirz, G. F. Meerhoff, A. H. W. Essing, J. W. van Dongen, P. Houba, C. G. Kruse, J. Verhaagen, D. F. Swaab, Concerted changes in transcripts in the prefrontal cortex precede neuropathology in Alzheimer's disease. *Brain*. 133, 3699–3723 (2010).

27. E. M. Blalock, H. M. Buechel, J. Popovic, J. W. Geddes, P. W. Landfield, Microarray analyses of laser-captured hippocampus reveal distinct gray and white matter signatures associated with incipient Alzheimer's disease. *J Chem Neuroanat*. 42, 118–126 (2011).

28. N. C. Berchtold, M. N. Sabbagh, T. G. Beach, R. C. Kim, D. H. Cribbs, C. W. Cotman, Brain gene expression patterns differentiate mild cognitive impairment from normal aged and Alzheimer's disease. *Neurobiol Aging*. 35, 1961–72 (2014).

29. N. C. Berchtold, P. D. Coleman, D. H. Cribbs, J. Rogers, D. L. Gillen, C. W. Cotman, Synaptic genes are extensively downregulated across multiple brain regions in normal human aging and Alzheimer's disease. *Neurobiol Aging*. 34, 1653–1661 (2013).

30. B. Zhang, C. Gaiteri, L.-G. Bodea, Z. Wang, J. McElwee, A. A. Podtelezhnikov, C. Zhang, T. Xie, L. Tran, R. Dobrin, E. Fluder, B. Clurman, S. Melquist, M. Narayanan, C. Suver, H. Shah, M. Mahajan, T. Gillis, J. Mysore, M. E. MacDonald, J. R. Lamb, D. A. Bennett, C. Molony, D. J. Stone, V. Gudnason, A. J. Myers, E. E. Schadt, H. Neumann, J. Zhu, V. Emilsson, Integrated Systems Approach Identifies Genetic Nodes and Networks in Late-Onset Alzheimer's Disease. *Cell*. 153, 707–720 (2013).

31. A. Annese, C. Manzari, C. Lionetti, E. Picardi, D. S. Horner, M. Chiara, M. F. Caratozzolo, A. Tullo, B. Fosso, G. Pesole, A. M. D'Erchia, Whole transcriptome profiling of Late-Onset Alzheimer's Disease patients provides insights into the molecular changes involved in the disease. *Sci Rep-uk*. 8, 4282 (2018).

32. B. Guennewig, J. Lim, L. Marshall, A. N. McCorkindale, P. J. Paasila, E. Patrick, J. J. Kril, G. M. Halliday, A. A. Cooper, G. T. Sutherland, Defining early changes in Alzheimer's disease from RNA sequencing of brain regions differentially affected by pathology. *Sci Rep-uk*. 11, 4865 (2021).

33. M. Wang, P. Roussos, A. McKenzie, X. Zhou, Y. Kajiwara, K. J. Brennand, G. C. D. Luca, J. F. Crary, P. Casaccia, J. D. Buxbaum, M. Ehrlich, S. Gandy, A. Goate, P. Katsel, E. Schadt, V. Haroutunian, B. Zhang, Integrative network analysis of nineteen brain regions identifies molecular signatures and networks underlying selective regional vulnerability to Alzheimer's disease. *Genome Med*. 8, 104 (2016).

34. R. A. Neff, M. Wang, S. Vatansever, L. Guo, C. Ming, Q. Wang, E. Wang, E. Horgusluoglu-Moloch, W. Song, A. Li, E. L. Castranio, J. TCW, L. Ho, A. Goate, V. Fossati, S. Noggle, S. Gandy, M. E. Ehrlich, P. Katsel, E. Schadt, D. Cai, K. J. Brennand, V. Haroutunian, B. Zhang, Molecular subtyping of Alzheimer's disease using RNA sequencing data reveals novel mechanisms and targets. *Sci Adv*. 7, eabb5398 (2021).

35. N. A. Bihlmeyer, E. Merrill, Y. Lambert, G. P. Srivastava, T. W. Clark, B. T. Hyman, S. Das, Novel methods for integration and visualization of genomics and genetics data in Alzheimer's disease. *Alzheimer's Dementia*. 15, 788–798 (2019).

36. J. A. Webster, J. R. Gibbs, J. Clarke, M. Ray, W. Zhang, P. Holmans, K. Rohrer, A. Zhao, L. Marlowe, M. Kaleem, D. S. McCorquodale, C. Cuello, D. Leung, L. Bryden, P. Nath, V. L. Zismann, K. Joshipura, M. J. Huentelman, D. Hu-Lince, K. D. Coon, D. W. Craig, J. V. Pearson, N.-N. Group, C. B. Heward, E. M. Reiman, D. Stephan, J. Hardy, A. J. Myers, Genetic Control of Human Brain Transcript Expression in Alzheimer Disease. *Am J Hum Genetics*. 84, 445–458 (2009).

37. B. Ng, C. C. White, H.-U. Klein, S. K. Sieberts, C. McCabe, E. Patrick, J. Xu, L. Yu, C. Gaiteri, D. A. Bennett, S. Mostafavi, P. L. D. Jager, An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. *Nat Neurosci*. 20, 1418–1426 (2017).

38. D. Patel, X. Zhang, J. J. Farrell, J. Chung, T. D. Stein, K. L. Lunetta, L. A. Farrer, Cell-type-specific expression quantitative trait loci associated with Alzheimer disease in blood and brain tissue. *Transl Psychiat*. 11, 250 (2021).

39. V. Swarup, F. I. Hinz, J. E. Rexach, K. Noguchi, H. Toyoshiba, A. Oda, K. Hirai, A. Sarkar, N. T. Seyfried, C. Cheng, S. J. Haggarty, R. Ferrari, J. D. Rohrer, A. Ramasamy, J. Hardy, D. G. Hernandez, M. A. Nalls, A. B. Singleton, J. B. J. Kwok, C. Dobson-Stone, W. S. Brooks, P. R. Schofield, G. M. Halliday, J. R. Hodges, O. Piguet, L. Bartley, E. Thompson, E. Haan, I. Hernández, A. Ruiz, M. Boada, B. Borroni, A. Padovani, N. J. Cairns, C. Cruchaga, G. Binetti, R. Ghidoni, L. Benussi, G. Forloni, D. Albani, D. Galimberti, C. Fenoglio, M. Serpente, E. Scarpini, J. Clarimón, A. Lleó, R. Blesa, M. L. Waldö, K. Nilsson, C. Nilsson, I. R. A. Mackenzie, G.-Y. R. Hsiung, D. M. A. Mann, J. Grafman, C. M. Morris, J. Attems, T. D. Griffiths, I. G. McKeith, A. J. Thomas, E. Jaros, P. Pietrini, E. D. Huey, E. M. Wassermann, M. C. Tierney, A. Baborie, P. Pastor, S. Ortega-Cubero, C. Razquin, E. Alonso, R. Perneczky, J. Diehl-Schmid, P. Alexopoulos, A. Kurz, I. Rainero, E. Rubino, L. Pinessi, E. Rogaeva, P. St. George-Hyslop, G. Rossi, F. Tagliavini, G. Giaccone, J. B. Rowe, J. C. M. Schlachetzki, J. Uphill, J. Collinge, S. Mead, A. Danek, V. M. V. Deerlin, M. Grossman, J. Q. Trojanowski, S. Pickering-Brown, P. Momeni, J. van der Zee, M. Cruts, C. V. Broeckhoven, S. F. Cappa, I. Leber, A. Brice, D. Hannequin, V. Golfier, M. Vercelletto, B. Nacmias, S. Sorbi, S. Bagnoli, I. Piaceri, J. E. Nielsen, L. E. Hjermind, M. Riemenschneider, M. Mayhaus, G. Gasparoni, S. Pichler, B. Ibach, M. N. Rossor, N. C. Fox, J. D. Warren, M. G. Spillantini, H. R. Morris, P. Rizzu, P. Heutink, J. S. Snowden, S. Rollinson, A. Gerhard, A. Richardson, A. C. Bruni, R. Maletta, F. Frangipane, C. Cupidi, L. Bernardi, M. Anfossi, M. Gallo, M. E. Conidi, N. Smirne, R. Rademakers, M. Baker, D. W. Dickson, N. R. Graff-Radford, R. C. Petersen, D. Knopman, K. A. Josephs, B. F. Boeve, J. E. Parisi, B. L. Miller, A. M. Karydas, H. Rosen, W. W. Seeley, J. C. van Swieten, E. G. P. Dopper, H. Seelaar, Y. A. L. Pijnenburg, P. Scheltens, G. Logroscino, R. Capozzo, V. Novelli, A. A. Puca, M. Franceschi, A. Postiglione, G. Milan, P. Sorrentino, M. Kristiansen, H.-H. Chiang, C. Graff, F. Pasquier, A. Rollin, V. Deramecourt, T. Lebouvier, L. Ferrucci, D. Kapogiannis, M. Grossman, V. M. V. Deerlin, J. Q. Trojanowski, J. J. Lah, A. I. Levey, S. Kondou, D. H. Geschwind, Identification of evolutionarily conserved gene networks mediating neurodegenerative dementia. *Nat Med*. 25, 152–164 (2019).

40. P. Lau, K. Bossers, R. Janky, E. Salta, C. S. Frigerio, S. Barbash, R. Rothman, A. S. R. Sierksma, A. Thathiah, D. Greenberg, A. S. Papadopoulou, T. Achsel, T. Ayoubi, H. Soreq, J. Verhaagen, D. F. Swaab, S. Aerts, B. D. Strooper, Alteration of the microRNA network during the progression of Alzheimer's disease. *Embo Mol Med*. 5, 1613–1634 (2013).

41. T. Raj, Y. I. Li, G. Wong, J. Humphrey, M. Wang, S. Ramdhani, Y.-C. Wang, B. Ng, I. Gupta, V. Haroutunian, E. E. Schadt, T. Young-Pearse, S. Mostafavi, B. Zhang, P. Sklar, D. A. Bennett, P. L. D. Jager, Integrative transcriptome analyses of the aging brain implicate altered splicing in Alzheimer's disease susceptibility. *Nat Genet*. 50, 1584–1592 (2018).

42. U. Dube, J. L. Del-Aguila, Z. Li, J. P. Budde, S. Jiang, S. Hsu, L. Ibanez, M. V. Fernandez, F. Farias, J. Norton, J. Gentsch, F. Wang, R. Allegri, F. Amtashar, T. Benzinger, S. Berman, C. Bodge, S. Brandon, W. Brooks, J. Buck, V. Buckles, S. Chea, P. Chrem, H. Chui, J. Cinco, J. Clifford, M. D'Mello, T. Donahue, J. Douglas, N. Edigo, N. Erekin-Taner, A. Fagan, M. Farlow, A. Farrar, H. Feldman, G. Flynn, N. Fox, E. Franklin, H. Fujii, C. Gant, S. Gardener, B. Ghetti, A. Goate, J. Goldman, B. Gordon, J. Gray, J. Gurney, J. Hassenstab, M. Hirohara, D. Holtzman, R. Hornbeck, S. H. DiBari, T. Ikeuchi, S. Ikonomovic, G. Jerome, M. Jucker, K. Kasuga, T. Kawarabayashi, W. Klunk, R. Koeppe, E. Kuder-Buletta, C. Laske, J. Levin, D. Marcus, R. Martins, N. S. Mason, D. Maue-Dreyfus, E. McDade, L. Montoya, H. Mori, A. Nagamatsu, K. Neimeyer, J. Noble, J. Norton, R. Perrin, M. Raichle, J. Ringman, J. H. Roh, P. Schofield, H.

Shimada, T. Shiroto, M. Shoji, W. Sigurdson, H. Sohrabi, P. Sparks, K. Suzuki, L. Swisher, K. Taddei, J. Wang, P. Wang, M. Weiner, M. Wolfsberger, C. Xiong, X. Xu, S. Salloway, C. L. Masters, J.-H. Lee, N. R. Graff-Radford, J. P. Chhatwal, R. J. Bateman, J. C. Morris, C. M. Karch, O. Harari, C. Cruchaga, An atlas of cortical circular RNA expression in Alzheimer disease brains demonstrates clinical and pathological associations. *Nat Neurosci*. 22, 1903–1912 (2019).

43. Y. Cheng, L. Saville, B. Gollen, A. A. Veronesi, M. Mohajerani, J. T. Joseph, A. Zovoilis, Increased Alu RNA processing in Alzheimer brains is linked to gene expression changes. *Embo Rep*. 22, e52255 (2021).

44. C. Guo, H.-H. Jeong, Y.-C. Hsieh, H.-U. Klein, D. A. Bennett, P. L. D. Jager, Z. Liu, J. M. Shulman, Tau Activates Transposable Elements in Alzheimer's Disease. *Cell Reports*. 23, 2874–2880 (2018).

45. S. Morabito, E. Miyoshi, N. Michael, V. Swarup, Integrative genomics approach identifies conserved transcriptomic networks in Alzheimer's disease. *Hum Mol Genet*. 29, 2899–2919 (2020).

46. S. Mostafavi, C. Gaiteri, S. E. Sullivan, C. C. White, S. Tasaki, J. Xu, M. Taga, H.-U. Klein, E. Patrick, V. Komashko, C. McCabe, R. Smith, E. M. Bradshaw, D. E. Root, A. Regev, L. Yu, L. B. Chibnik, J. A. Schneider, T. L. Young-Pearse, D. A. Bennett, P. L. D. Jager, A molecular network of the aging human brain provides insights into the pathology and cognitive decline of Alzheimer's disease. *Nat Neurosci*. 21, 811–819 (2018).

47. M. Allen, X. Wang, J. D. Burgess, J. Watzlawik, D. J. Serie, C. S. Younkin, T. Nguyen, K. G. Malphrus, S. Lincoln, M. M. Carrasquillo, C. Ho, P. Chakrabarty, S. Strickland, M. E. Murray, V. Swarup, D. H. Geschwind, N. T. Seyfried, E. B. Dammer, J. J. Lah, A. I. Levey, T. E. Golde, C. Funk, H. Li, N. D. Price, R. C. Petersen, N. R. Graff-Radford, S. G. Younkin, D. W. Dickson, J. R. Crook, Y. W. Asmann, N. Ertekin-Taner, Conserved brain myelination networks are altered in Alzheimer's and other neurodegenerative diseases. *Alzheimer's Dementia*. 14, 352–366 (2018).

48. J. A. Miller, M. C. Oldham, D. H. Geschwind, A Systems Level Analysis of Transcriptional Changes in Alzheimer's Disease and Normal Aging. *J Neurosci*. 28, 1410–1420 (2008).

49. P. Forabosco, A. Ramasamy, D. Trabzuni, R. Walker, C. Smith, J. Bras, A. P. Levine, J. Hardy, J. M. Pocock, R. Guerreiro, M. E. Weale, M. Ryten, Insights into TREM2 biology by network analysis of human brain gene expression data. *Neurobiol Aging*. 34, 2699–2714 (2013).

50. F. Tang, C. Barbacioru, Y. Wang, E. Nordman, C. Lee, N. Xu, X. Wang, J. Bodeau, B. B. Tuch, A. Siddiqui, K. Lao, M. A. Surani, mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*. 6, 377–382 (2009).

51. A. B. Rosenberg, C. M. Roco, R. A. Muscat, A. Kuchina, P. Sample, Z. Yao, L. Gray, D. J. Peeler, S. Mukherjee, W. Chen, S. H. Pun, D. L. Sellers, B. Tasic, G. Seelig, Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*. 360, eaam8999 (2018).

52. E. Z. Macosko, A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, J. J. Trombetta, D. A. Weitz, J. R. Sanes, A. K. Shalek, A. Regev, S. A. McCarroll, Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*. 161, 1202–1214 (2015).

53. N. Habib, I. Avraham-Davidi, A. Basu, T. Burks, K. Shekhar, M. Hofree, S. R. Choudhury, F. Aguet, E. Gelfand, K. Ardlie, D. A. Weitz, O. Rozenblatt-Rosen, F. Zhang, A. Regev, Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat Methods*. 14, 955–958 (2017).

54. B. B. Lake, R. Ai, G. E. Kaeser, N. S. Salathia, Y. C. Yung, R. Liu, A. Wildberg, D. Gao, H.-L. Fung, S. Chen, R. Vijayaraghavan, J. Wong, A. Chen, X. Sheng, F. Kaper, R. Shen, M. Ronaghi, J.-B. Fan, W. Wang, J. Chun, K. Zhang, Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science*. 352, 1586–1590 (2016).

55. S. Darmanis, S. A. Sloan, Y. Zhang, M. Enge, C. Caneda, L. M. Shuer, M. G. H. Gephart, B. A. Barres, S. R. Quake, A survey of human brain transcriptome diversity at the single cell level. *Proc National Acad Sci*. 112, 7285–7290 (2015).

56. R. D. Hodge, T. E. Bakken, J. A. Miller, K. A. Smith, E. R. Barkan, L. T. Graybuck, J. L. Close, B. Long, N. Johansen, O. Penn, Z. Yao, J. Eggermont, T. Höllt, B. P. Levi, S. I. Shehata, B. Aevermann, A. Beller, D. Bertagnolli, K. Brouner, T. Casper, C. Cobbs, R. Dalley, N. Dee, S.-L. Ding, R. G. Ellenbogen, O. Fong, E. Garren, J. Goldy, R. P. Gwinn, D. Hirschstein, C. D. Keene, M. Keshk, A. L. Ko, K. Lathia, A. Mahfouz, Z. Maltzer, M. McGraw, T. N. Nguyen, J. Nyhus, J. G. Ojemann, A. Oldre, S. Parry, S. Reynolds, C. Rimorin, N. V. Shapovalova, S. Somasundaram, A. Szafer, E. R. Thomsen, M. Tieu, G. Quon, R. H. Scheuermann, R. Yuste, S. M. Sunkin, B. Lelieveldt, D. Feng, L. Ng, A. Bernard, M. Hawrylycz, J. W. Phillips, B. Tasic, H. Zeng, A. R. Jones, C. Koch, E. S. Lein, Conserved cell types with divergent features in human versus mouse cortex. *Nature*. 573, 61–68 (2019).

57. B. Tasic, V. Menon, T. N. Nguyen, T. K. Kim, T. Jarsky, Z. Yao, B. Levi, L. T. Gray, S. A. Sorensen, T. Dolbeare, D. Bertagnolli, J. Goldy, N. Shapovalova, S. Parry, C. Lee, K. Smith, A. Bernard, L. Madisen, S. M. Sunkin, M. Hawrylycz, C. Koch, H. Zeng, Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat Neurosci*. 19, 335–346 (2016).

58. Z. Yao, C. T. J. van Velthoven, T. N. Nguyen, J. Goldy, A. E. Sedeno-Cortes, F. Baftizadeh, D. Bertagnolli, T. Casper, M. Chiang, K. Crichton, S.-L. Ding, O. Fong, E. Garren, A. Glandon, N. W. Gouwens, J. Gray, L. T. Graybuck, M. J. Hawrylycz, D. Hirschstein, M. Kroll, K. Lathia, C. Lee, B. Levi, D. McMillen, S. Mok, T. Pham, Q. Ren, C. Rimorin, N. Shapovalova, J. Sulc, S. M. Sunkin, M. Tieu, A. Torkelson, H. Tung, K. Ward, N. Dee, K. A. Smith, B. Tasic, H. Zeng, A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* (2021), doi:10.1016/j.cell.2021.04.021.

59. B. Tasic, Z. Yao, L. T. Graybuck, K. A. Smith, T. N. Nguyen, D. Bertagnolli, J. Goldy, E. Garren, M. N. Economo, S. Viswanathan, O. Penn, T. Bakken, V. Menon, J. Miller, O. Fong, K. E. Hirokawa, K. Lathia, C. Rimorin, M. Tieu, R. Larsen, T. Casper, E. Barkan, M. Kroll, S. Parry, N. V. Shapovalova, D. Hirschstein, J. Pendergraft, H. A. Sullivan, T. K. Kim, A. Szafer, N. Dee, P. Groblewski, I. Wickersham, A. Cetin, J. A. Harris, B. P. Levi, S. M. Sunkin, L. Madisen, T. L. Daigle, L. Looger, A. Bernard, J. Phillips, E. Lein, M. Hawrylycz, K. Svoboda, A. R. Jones, C.

Koch, H. Zeng, Shared and distinct transcriptomic cell types across neocortical areas. *Nature*. 563, 72–78 (2018).

60. A. Zeisel, A. B. Munoz-Manchado, S. Codeluppi, P. Lonnerberg, G. L. Manno, A. Jureus, S. Marques, H. Munguba, L. He, C. Betsholtz, C. Rolny, G. Castelo-Branco, J. Hjerling-Leffler, S. Linnarsson, Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science*. 347, 1138–1142 (2015).

61. R. C. Bandler, I. Vitali, R. N. Delgado, M. C. Ho, E. Dvoretskova, J. S. I. Molinas, P. W. Frazel, M. Mohammadkhani, R. Machold, S. Maedler, S. A. Liddelow, T. J. Nowakowski, G. Fishell, C. Mayer, Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature*. 601, 404–409 (2022).

62. T. E. Bakken, N. L. Jorstad, Q. Hu, B. B. Lake, W. Tian, B. E. Kalmbach, M. Crow, R. D. Hodge, F. M. Krienen, S. A. Sorensen, J. Eggermont, Z. Yao, B. D. Aevermann, A. I. Aldridge, A. Bartlett, D. Bertagnolli, T. Casper, R. G. Castanon, K. Crichton, T. L. Daigle, R. Dalley, N. Dee, N. Dembrow, D. Diep, S.-L. Ding, W. Dong, R. Fang, S. Fischer, M. Goldman, J. Goldy, L. T. Graybuck, B. R. Herb, X. Hou, J. Kancherla, M. Kroll, K. Lathia, B. van Lew, Y. E. Li, C. S. Liu, H. Liu, J. D. Lucero, A. Mahurkar, D. McMillen, J. A. Miller, M. Moussa, J. R. Nery, P. R. Nicovich, S.-Y. Niu, J. Orvis, J. K. Osteen, S. Owen, C. R. Palmer, T. Pham, N. Plongthongkum, O. Poirion, N. M. Reed, C. Rimorin, A. Rivkin, W. J. Romanow, A. E. Sedeño-Cortés, K. Siletti, S. Somasundaram, J. Sulc, M. Tieu, A. Torkelson, H. Tung, X. Wang, F. Xie, A. M. Yanny, R. Zhang, S. A. Ament, M. M. Behrens, H. C. Bravo, J. Chun, A. Dobin, J. Gillis, R. Hertzano, P. R. Hof, T. Höllt, G. D. Horwitz, C. D. Keene, P. V. Kharchenko, A. L. Ko, B. P. Lelieveldt, C. Luo, E. A. Mukamel, A. Pinto-Duarte, S. Preissl, A. Regev, B. Ren, R. H. Scheuermann, K. Smith, W. J. Spain, O. R. White, C. Koch, M. Hawrylycz, B. Tasic, E. Z. Macosko, S. A. McCarroll, J. T. Ting, H. Zeng, K. Zhang, G. Feng, J. R. Ecker, S. Linnarsson, E. S. Lein, Comparative cellular analysis of motor cortex in human, marmoset and mouse. *Nature*. 598, 111–119 (2021).

63. O. A. Bayraktar, T. Bartels, S. Holmqvist, V. Kleshchevnikov, A. Martirosyan, D. Polioudakis, L. B. Haim, A. M. H. Young, M. Y. Batiuk, K. Prakash, A. Brown, K. Roberts, M. F. Paredes, R. Kawaguchi, J. H. Stockley, K. Sabeur, S. M. Chang, E. Huang, P. Hutchinson, E. M. Ullian, M. Hemberg, G. Coppola, M. G. Holt, D. H. Geschwind, D. H. Rowitch, Astrocyte layers in the mammalian cerebral cortex revealed by a single-cell in situ transcriptomic map. *Nat Neurosci*. 23, 500–509 (2020).

64. M. Y. Batiuk, A. Martirosyan, J. Wahis, F. de Vin, C. Marneffe, C. Kusserow, J. Koeppen, J. F. Viana, J. F. Oliveira, T. Voet, C. P. Ponting, T. G. Belgard, M. G. Holt, Identification of region-specific astrocyte subtypes at single cell resolution. *Nat Commun*. 11, 1220 (2020).

65. R. Sankowski, C. Böttcher, T. Masuda, L. Geirsdottir, Sagar, E. Sindram, T. Seredenina, A. Muhs, C. Scheiwe, M. J. Shah, D. H. Heiland, O. Schnell, D. Grün, J. Priller, M. Prinz, Mapping microglia states in the human brain through the integration of high-dimensional techniques. *Nat Neurosci*. 22, 2098–2110 (2019).

66. T. Masuda, R. Sankowski, O. Staszewski, C. Böttcher, L. Amann, Sagar, C. Scheiwe, S. Nessler, P. Kunz, G. van Loo, V. A. Coenen, P. C. Reinacher, A. Michel, U. Sure, R. Gold, D.

Grün, J. Priller, C. Stadelmann, M. Prinz, Spatial and temporal heterogeneity of mouse and human microglia at single-cell resolution. *Nature*. 566 (2019), doi:10.1038/s41586-019-0924-x.

67. J. A. Stogsdill, K. Kim, L. Binan, S. L. Farhi, J. Z. Levin, P. Arlotta, Pyramidal neuron subtype diversity governs microglia states in the neocortex. *Nature*. 608, 750–756 (2022).

68. S. Jäkel, E. Agirre, A. M. Falcão, D. van Bruggen, K. W. Lee, I. Knuesel, D. Malhotra, C. ffrench-Constant, A. Williams, G. Castelo-Branco, Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature*. 566, 543–547 (2019).

69. A. M. Falcão, D. van Bruggen, S. Marques, M. Meijer, S. Jäkel, E. Agirre, Samudyata, E. M. Floriddia, D. P. Vanichkina, C. Ffrench-Constant, A. Williams, A. O. Guerreiro-Cacais, G. Castelo-Branco, Disease-specific oligodendrocyte lineage cells arise in multiple sclerosis. *Nat Med*. 24, 1837–1844 (2018).

70. S. Marques, A. Zeisel, S. Codeluppi, D. van Bruggen, A. M. Falcão, L. Xiao, H. Li, M. Häring, H. Hochgerner, R. A. Romanov, D. Gyllborg, A. B. Muñoz-Manchado, G. L. Manno, P. Lönnerberg, E. M. Floriddia, F. Rezayee, P. Ernfors, E. Arenas, J. Hjerling-Leffler, T. Harkany, W. D. Richardson, S. Linnarsson, G. Castelo-Branco, Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. *Sci New York N Y*. 352, 1326–9 (2016).

71. B. B. Lake, S. Chen, B. C. Sos, J. Fan, G. E. Kaeser, Y. C. Yung, T. E. Duong, D. Gao, J. Chun, P. V. Kharchenko, K. Zhang, Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat Biotechnol*. 36, 70–80 (2017).

72. M. R. Corces, A. Shcherbina, S. Kundu, M. J. Gloudemans, L. Frésard, J. M. Granja, B. H. Louie, T. Eulalio, S. Shams, S. T. Bagdatli, M. R. Mumbach, B. Liu, K. S. Montine, W. J. Greenleaf, A. Kundaje, S. B. Montgomery, H. Y. Chang, T. J. Montine, Single-cell epigenomic analyses implicate candidate causal variants at inherited risk loci for Alzheimer's and Parkinson's diseases. *Nat Genet*. 52, 1158–1168 (2020).

73. D. A. Cusanovich, A. J. Hill, D. Aghamirzaie, R. M. Daza, H. A. Pliner, J. B. Berletch, G. N. Filippova, X. Huang, L. Christiansen, W. S. DeWitt, C. Lee, S. G. Regalado, D. F. Read, F. J. Steemers, C. M. Disteche, C. Trapnell, J. Shendure, A Single-Cell Atlas of In Vivo Mammalian Chromatin Accessibility. *Cell*. 174, 1309-1324.e18 (2018).

74. A. C. Yang, R. T. Vest, F. Kern, D. P. Lee, M. Agam, C. A. Maat, P. M. Losada, M. B. Chen, N. Schaum, N. Khoury, A. Toland, K. Calcuttawala, H. Shin, R. Pálovics, A. Shin, E. Y. Wang, J. Luo, D. Gate, W. J. Schulz-Schaeffer, P. Chu, J. A. Siegenthaler, M. W. McNerney, A. Keller, T. Wyss-Coray, A human brain vascular atlas reveals diverse mediators of Alzheimer's risk. *Nature*. 603, 885–892 (2022).

75. H. Mathys, J. Davila-Velderrain, Z. Peng, F. Gao, S. Mohammadi, J. Z. Young, M. Menon, L. He, F. Abdurrob, X. Jiang, A. J. Martorell, R. M. Ransohoff, B. P. Hafler, D. A. Bennett, M. Kellis, L.-H. Tsai, Single-cell transcriptomic analysis of Alzheimer's disease. *Nature*. 570, 332–337 (2019).

76. K. Leng, E. Li, R. Eser, A. Piergies, R. Sit, M. Tan, N. Neff, S. H. Li, R. D. Rodriguez, C. K. Suemoto, R. E. P. Leite, A. J. Ehrenberg, C. A. Pasqualucci, W. W. Seeley, S. Spina, H. Heinsen, L. T. Grinberg, M. Kampmann, Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. *Nat Neurosci*. 24, 276–287 (2021).

77. H. Keren-Shaul, A. Spinrad, A. Weiner, O. Matcovitch-Natan, R. Dvir-Szternfeld, T. K. Ulland, E. David, K. Baruch, D. Lara-Astaiso, B. Toth, S. Itzkovitz, M. Colonna, M. Schwartz, I. Amit, A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell*. 169, 1276-1290.e17 (2017).

78. Y. Zhou, W. M. Song, P. S. Andhey, A. Swain, T. Levy, K. R. Miller, P. L. Poliani, M. Cominelli, S. Grover, S. Gilfillan, M. Cella, T. K. Ulland, K. Zaitsev, A. Miyashita, T. Ikeuchi, M. Sainouchi, A. Kakita, D. A. Bennett, J. A. Schneider, M. R. Nichols, S. A. Beausoleil, J. D. Ulrich, D. M. Holtzman, M. N. Artyomov, M. Colonna, Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. *Nat Med*. 26, 131–142 (2020).

79. N. Habib, C. McCabe, S. Medina, M. Varshavsky, D. Kitsberg, R. Dvir-Szternfeld, G. Green, D. Dionne, L. Nguyen, J. L. Marshall, F. Chen, F. Zhang, T. Kaplan, A. Regev, M. Schwartz, Disease-associated astrocytes in Alzheimer's disease and aging. *Nat Neurosci*. 23, 701–706 (2020).

80. C. S. Frigerio, L. Wolfs, N. Fattorelli, N. Thrupp, I. Voytyuk, I. Schmidt, R. Mancuso, W.-T. Chen, M. E. Woodbury, G. Srivastava, T. Möller, E. Hudry, S. Das, T. Saido, E. Karran, B. Hyman, V. H. Perry, M. Fiers, B. D. Strooper, The Major Risk Factors for Alzheimer's Disease: Age, Sex, and Genes Modulate the Microglia Response to Aβ Plaques. *Cell Reports*. 27, 1293-1306.e6 (2019).

81. A. M. Smith, K. Davey, S. Tsartsalis, C. Khozoie, N. Fancy, S. S. Tang, E. Liaptsi, M. Weinert, A. McGarry, R. C. J. Muirhead, S. Gentleman, D. R. Owen, P. M. Matthews, Diverse human astrocyte and microglial transcriptional responses to Alzheimer's pathology. *Acta Neuropathol*. 143, 75–91 (2022).

82. M. Kenigsbuch, P. Bost, S. Halevi, Y. Chang, S. Chen, Q. Ma, R. Hajbi, B. Schwikowski, B. Bodenmiller, H. Fu, M. Schwartz, I. Amit, A shared disease-associated oligodendrocyte signature among multiple CNS pathologies. *Nat Neurosci*. 25, 876–886 (2022).

83. A. T. Nguyen, K. Wang, G. Hu, X. Wang, Z. Miao, J. A. Azevedo, E. Suh, V. M. V. Deerlin, D. Choi, K. Roeder, M. Li, E. B. Lee, APOE and TREM2 regulate amyloid-responsive microglia in Alzheimer's disease. *Acta Neuropathol*. 140, 477–493 (2020).

84. S.-F. Lau, H. Cao, A. K. Y. Fu, N. Y. Ip, Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and neuroprotective glia in Alzheimer's disease. *Proc National Acad Sci*. 117, 25800–25809 (2020).

85. S. Morabito, E. Miyoshi, N. Michael, S. Shahin, A. C. Martini, E. Head, J. Silva, K. Leavy, M. Perez-Rosendahl, V. Swarup, Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nat Genet*. 53, 1143–1155 (2021).

86. H.-U. Klein, C. McCabe, E. Gjoneska, S. E. Sullivan, B. J. Kaskow, A. Tang, R. V. Smith, J. Xu, A. R. Pfenning, B. E. Bernstein, A. Meissner, J. A. Schneider, S. Mostafavi, L.-H. Tsai, T. L. Young-Pearse, D. A. Bennett, P. L. D. Jager, Epigenome-wide study uncovers large-scale changes in histone acetylation driven by tau pathology in aging and Alzheimer's human brains. *Nat Neurosci*. 22, 37–46 (2018).

87. P. L. D. Jager, G. Srivastava, K. Lunnon, J. Burgess, L. C. Schalkwyk, L. Yu, M. L. Eaton, B. T. Keenan, J. Ernst, C. McCabe, A. Tang, T. Raj, J. Replogle, W. Brodeur, S. Gabriel, H. S. Chai, C. Younkin, S. G. Younkin, F. Zou, M. Szyf, C. B. Epstein, J. A. Schneider, B. E. Bernstein, A. Meissner, N. Ertekin-Taner, L. B. Chibnik, M. Kellis, J. Mill, D. A. Bennett, Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat Neurosci*. 17, 1156–1163 (2014).

88. K. Lunnon, R. Smith, E. Hannon, P. L. D. Jager, G. Srivastava, M. Volta, C. Troakes, S. Al-Sarraj, J. Burrage, R. Macdonald, D. Condliffe, L. W. Harries, P. Katsel, V. Haroutunian, Z. Kaminsky, C. Joachim, J. Powell, S. Lovestone, D. A. Bennett, L. C. Schalkwyk, J. Mill, Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nat Neurosci*. 17, 1164–1170 (2014).

89. R. Nativio, G. Donahue, A. Berson, Y. Lan, A. Amlie-Wolf, F. Tuzer, J. B. Toledo, S. J. Gosai, B. D. Gregory, C. Torres, J. Q. Trojanowski, L.-S. Wang, F. B. Johnson, N. M. Bonini, S. L. Berger, Dysregulation of the epigenetic landscape of normal aging in Alzheimer's disease. *Nat Neurosci*. 21, 497–505 (2018).

90. A. Nott, I. R. Holtman, N. G. Coufal, J. C. M. Schlachetzki, M. Yu, R. Hu, C. Z. Han, M. Pena, J. Xiao, Y. Wu, Z. Keulen, M. P. Pasillas, C. O'Connor, C. K. Nickl, S. T. Schafer, Z. Shen, R. A. Rissman, J. B. Brewer, D. Gosselin, D. D. Gonda, M. L. Levy, M. G. Rosenfeld, G. McVicker, F. H. Gage, B. Ren, C. K. Glass, Brain cell type–specific enhancer–promoter interactome maps and disease-risk association. *Science*. 366, 1134–1139 (2019).

91. J. D. Buenrostro, B. Wu, U. M. Litzenburger, D. Ruff, M. L. Gonzales, M. P. Snyder, H. Y. Chang, W. J. Greenleaf, Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*. 523, 486–90 (2015).

92. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 10, 1213–1218 (2013).

93. C. Zhu, Y. Zhang, Y. E. Li, J. Lucero, M. M. Behrens, B. Ren, Joint profiling of histone modifications and transcriptome in single cells from mouse brain. *Nat Methods*. 18, 283–292 (2021).

94. K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, X. Zhuang, Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*. 348, aaa6090 (2015).

95. R. Ke, M. Mignardi, A. Pacureanu, J. Svedlund, J. Botling, C. Wählby, M. Nilsson, In situ sequencing for RNA analysis in preserved tissue and cells. *Nat Methods*. 10, 857–860 (2013).

96. P. L. Ståhl, F. Salmén, S. Vickovic, A. Lundmark, J. F. Navarro, J. Magnusson, S. Giacomello, M. Asp, J. O. Westholm, M. Huss, A. Mollbrink, S. Linnarsson, S. Codeluppi, Å. Borg, F. Pontén, P. I. Costea, P. Sahlén, J. Mulder, O. Bergmann, J. Lundeberg, J. Frisén, Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*. 353, 78–82 (2016).

97. S. G. Rodriques, R. R. Stickels, A. Goeva, C. A. Martin, E. Murray, C. R. Vanderburg, J. Welch, L. M. Chen, F. Chen, E. Z. Macosko, Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science*. 363, 1463–1467 (2019).

98. K. R. Maynard, L. Collado-Torres, L. M. Weber, C. Uytingco, B. K. Barry, S. R. Williams, J. L. Catallini, M. N. Tran, Z. Besich, M. Tippani, J. Chew, Y. Yin, J. E. Kleinman, T. M. Hyde, N. Rao, S. C. Hicks, K. Martinowich, A. E. Jaffe, Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat Neurosci*. 24, 425–436 (2021).

99. W.-T. Chen, A. Lu, K. Craessaerts, B. Pavie, C. S. Frigerio, N. Corthout, X. Qian, J. Laláková, M. Kühnemund, I. Voytyuk, L. Wolfs, R. Mancuso, E. Salta, S. Balusu, A. Snellinx, S. Munck, A. Jurek, J. F. Navarro, T. C. Saido, I. Huitinga, J. Lundeberg, M. Fiers, B. D. Strooper, Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. *Cell*. 182, 976-991.e19 (2020).

100. J. F. Navarro, D. L. Croteau, A. Jurek, Z. Andrusivova, B. Yang, Y. Wang, B. Ogedegbe, T. Riaz, M. Støen, C. Desler, L. J. Rasmussen, T. Tønjum, M.-C. Galas, J. Lundeberg, V. A. Bohr, Spatial Transcriptomics Reveals Genes Associated with Dysregulated Mitochondrial Functions and Stress Signaling in Alzheimer Disease. *Iscience*. 23, 101556 (2020).

101. S. Maniatis, T. Äijö, S. Vickovic, C. Braine, K. Kang, A. Mollbrink, D. Fagegaltier, Ž. Andrusivová, S. Saarenpää, G. Saiz-Castro, M. Cuevas, A. Watters, J. Lundeberg, R. Bonneau, H. Phatnani, Spatiotemporal dynamics of molecular pathology in amyotrophic lateral sclerosis. *Sci New York N Y*. 364, 89–93 (2019).

102. M. Elosua-Bayes, P. Nieto, E. Mereu, I. Gut, H. Heyn, SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res*. 49, e50–e50 (2021).

103. E. Zhao, M. R. Stone, X. Ren, J. Guenthoer, K. S. Smythe, T. Pulliam, S. R. Williams, C. R. Uytingco, S. E. B. Taylor, P. Nghiem, J. H. Bielas, R. Gottardo, Spatial transcriptomics at subspot resolution with BayesSpace. *Nat Biotechnol*. 39, 1375–1384 (2021).

104. J. S. Sadick, M. R. O'Dea, P. Hasel, T. Dykstra, A. Faustin, S. A. Liddelow, Astrocytes and oligodendrocytes undergo subtype-specific transcriptional changes in Alzheimer's disease. *Neuron*. 110, 1788-1805.e10 (2022).

105. P. Hasel, I. V. L. Rose, J. S. Sadick, R. D. Kim, S. A. Liddelow, Neuroinflammatory astrocyte subtypes in the mouse brain. *Nat Neurosci*. 24, 1475–1487 (2021).

106. T. Kamath, A. Abdulraouf, S. J. Burris, J. Langlieb, V. Gazestani, N. M. Nadaf, K. Balderrama, C. Vanderburg, E. Z. Macosko, Single-cell genomic profiling of human dopamine

neurons identifies a population that selectively degenerates in Parkinson’s disease. *Nat Neurosci*. 25, 588–595 (2022).

107. L. Garcia-Alonso, L.-F. Handfield, K. Roberts, K. Nikolakopoulou, R. C. Fernando, L. Gardner, B. Woodhams, A. Arutyunyan, K. Polanski, R. Hoo, C. Sancho-Serra, T. Li, K. Kwakwa, E. Tuck, V. Lorenzi, H. Massalha, M. Prete, V. Kleshchevnikov, A. Tarkowska, T. Porter, C. I. Mazzeo, S. van Dongen, M. Dabrowska, V. Vaskivskyi, K. T. Mahbubani, J. Park, M. Jimenez-Linan, L. Campos, V. Yu. Kiselev, C. Lindskog, P. Ayuk, E. Prigmore, M. R. Stratton, K. Saeb-Parsy, A. Moffett, L. Moore, O. A. Bayraktar, S. A. Teichmann, M. Y. Turco, R. Vento-Tormo, Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro. *Nat Genet*. 53, 1698–1711 (2021).

108. S. Z. Wu, G. Al-Eryani, D. L. Roden, S. Junankar, K. Harvey, A. Andersson, A. Thennavan, C. Wang, J. R. Torpy, N. Bartonicek, T. Wang, L. Larsson, D. Kaczorowski, N. I. Weisenfeld, C. R. Uytingco, J. G. Chew, Z. W. Bent, C.-L. Chan, V. Gnanasambandapillai, C.-A. Dutertre, L. Gluch, M. N. Hui, J. Beith, A. Parker, E. Robbins, D. Segara, C. Cooper, C. Mak, B. Chan, S. Warrier, F. Ginhoux, E. Millar, J. E. Powell, S. R. Williams, X. S. Liu, S. O’Toole, E. Lim, J. Lundeberg, C. M. Perou, A. Swarbrick, A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet*. 53, 1334–1347 (2021).

109. W. L. Hwang, K. A. Jagadeesh, J. A. Guo, H. I. Hoffman, P. Yadollahpour, J. W. Reeves, R. Mohan, E. Drokhlyansky, N. V. Wittenberghe, O. Ashenberg, S. L. Farhi, D. Schapiro, P. Divakar, E. Miller, D. R. Zollinger, G. Eng, J. M. Schenkel, J. Su, C. Shiau, P. Yu, W. A. Freed-Pastor, D. Abbondanza, A. Mehta, J. Gould, C. Lambden, C. B. M. Porter, A. Tsankov, D. Dionne, J. Waldman, M. S. Cuoco, L. Nguyen, T. Delorey, D. Phillips, J. L. Barth, M. Kem, C. Rodrigues, D. Ciprani, J. Roldan, P. Zelga, V. Jorgji, J. H. Chen, Z. Ely, D. Zhao, K. Fuhrman, R. Fropf, J. M. Beechem, J. S. Loeffler, D. P. Ryan, C. D. Weekes, C. R. Ferrone, M. Qadan, M. J. Aryee, R. K. Jain, D. S. Neuberg, J. Y. Wo, T. S. Hong, R. Xavier, A. J. Aguirre, O. Rozenblatt-Rosen, M. Mino-Kenudson, C. F. Castillo, A. S. Liss, D. T. Ting, T. Jacks, A. Regev, Single-nucleus and spatial transcriptome profiling of pancreatic cancer identifies multicellular dynamics associated with neoadjuvant treatment. *Nat Genet*. 54, 1178–1191 (2022).

110. Y. Deng, M. Bartosovic, S. Ma, D. Zhang, P. Kukanja, Y. Xiao, G. Su, Y. Liu, X. Qin, G. B. Rosoklija, A. J. Dwork, J. J. Mann, M. L. Xu, S. Halene, J. E. Craft, K. W. Leong, M. Boldrini, G. Castelo-Branco, R. Fan, Spatial profiling of chromatin accessibility in mouse and human tissues. *Nature*. 609, 375–383 (2022).

111. C. A. Thornton, R. M. Mulqueen, K. A. Torkenczy, A. Nishida, E. G. Lowenstein, A. J. Fields, F. J. Steemers, W. Zhang, H. L. McConnell, R. L. Woltjer, A. Mishra, K. M. Wright, A. C. Adey, Spatially mapped single-cell chromatin accessibility. *Nat Commun*. 12, 1274 (2021).

112. Y. Deng, M. Bartosovic, P. Kukanja, D. Zhang, Y. Liu, G. Su, A. Enninful, Z. Bai, G. Castelo-Branco, R. Fan, Spatial-CUT&Tag: Spatially resolved chromatin modification profiling at the cellular level. *Science*. 375, 681–686 (2022).

113. A. Zeisel, H. Hochgerner, P. Lönnerberg, A. Johnsson, F. Memic, J. van der Zwan, M. Häring, E. Braun, L. E. Borm, G. L. Manno, S. Codeluppi, A. Furlan, K. Lee, N. Skene, K. D.

Harris, J. Hjerling-Leffler, E. Arenas, P. Ernfors, U. Marklund, S. Linnarsson, Molecular Architecture of the Mouse Nervous System. *Cell*. 174, 999-1014.e22 (2018).

114. A. Grubman, G. Chew, J. F. Ouyang, G. Sun, X. Y. Choo, C. McLean, R. K. Simmons, S. Buckberry, D. B. Vargas-Landin, D. Poppe, J. Pflueger, R. Lister, O. J. L. Rackham, E. Petretto, J. M. Polo, A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. *Nat Neurosci*. 22, 2087–2097 (2019).

115. J. L. Del-Aguila, Z. Li, U. Dube, K. A. Mihindukulasuriya, J. P. Budde, M. V. Fernandez, L. Ibanez, J. Bradley, F. Wang, K. Bergmann, R. Davenport, J. C. Morris, D. M. Holtzman, R. J. Perrin, B. A. Benitez, J. Dougherty, C. Cruchaga, O. Harari, Single- nuclei RNA sequencing from human brain to study for Mendelian and sporadic AD. *Biorxiv*, 593756 (2019).

116. M. M. Carrasquillo, F. Zou, V. S. Pankratz, S. L. Wilcox, L. Ma, L. P. Walker, S. G. Younkin, C. S. Younkin, L. H. Younkin, G. D. Bisceglio, N. Ertekin-Taner, J. E. Crook, D. W. Dickson, R. C. Petersen, N. R. Graff-Radford, S. G. Younkin, Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. *Nat Genet*. 41, 192–8 (2009).

117. D. P. Hibar, H. H. H. Adams, N. Jahanshad, G. Chauhan, J. L. Stein, E. Hofer, M. E. Renteria, J. C. Bis, A. Arias-Vasquez, M. K. Ikram, S. Desrivières, M. W. Vernooij, L. Abramovic, S. Alhusaini, N. Amin, M. Andersson, K. Arfanakis, B. S. Aribisala, N. J. Armstrong, L. Athanasiu, T. Axelsson, A. H. Beecham, A. Beiser, M. Bernard, S. H. Blanton, M. M. Bohlken, M. P. Boks, J. Bralten, A. M. Brickman, O. Carmichael, M. M. Chakravarty, Q. Chen, C. R. K. Ching, V. Chouraki, G. Cuellar-Partida, F. Crivello, A. D. Braber, N. T. Doan, S. Ehrlich, S. Giddaluru, A. L. Goldman, R. F. Gottesman, O. Grimm, M. E. Griswold, T. Guadalupe, B. A. Gutman, J. Hass, U. K. Haukvik, D. Hoehn, A. J. Holmes, M. Hoogman, D. Janowitz, T. Jia, K. N. Jørgensen, N. Karbalai, D. Kasperaviciute, S. Kim, M. Klein, B. Kraemer, P. H. Lee, D. C. M. Liewald, L. M. Lopez, M. Luciano, C. Macare, A. F. Marquand, M. Matarin, K. A. Mather, M. Mattheisen, D. R. McKay, Y. Milaneschi, S. M. Maniega, K. Nho, A. C. Nugent, P. Nyquist, L. M. O. Loohuis, J. Oosterlaan, M. Papmeyer, L. Pirpamer, B. Pütz, A. Ramasamy, J. S. Richards, S. L. Risacher, R. Roiz-Santiañez, N. Rommelse, S. Ropele, E. J. Rose, N. A. Royle, T. Rundek, P. G. Sämann, A. Saremi, C. L. Satizabal, L. Schmaal, A. J. Schork, L. Shen, J. Shin, E. Shumskaya, A. V. Smith, E. Sprooten, L. T. Strike, A. Teumer, D. Tordesillas-Gutierrez, R. Toro, D. Trabzuni, S. Trompet, D. Vaidya, J. V. der Grond, S. J. V. der Lee, D. V. der Meer, M. M. J. V. Donkelaar, K. R. V. Eijk, T. G. M. V. Erp, D. V. Rooij, E. Walton, L. T. Westlye, C. D. Whelan, B. G. Windham, A. M. Winkler, K. Wittfeld, G. Woldehawariat, C. Wolf, T. Wolfers, L. R. Yanek, J. Yang, A. Zijdenbos, M. P. Zwiers, I. Agartz, L. Almasy, D. Ames, P. Amouyel, O. A. Andreassen, S. Arepalli, A. A. Assareh, S. Barral, M. E. Bastin, D. M. Becker, J. T. Becker, D. A. Bennett, J. Blangero, H. van Bokhoven, D. I. Boomsma, H. Brodaty, R. M. Brouwer, H. G. Brunner, R. L. Buckner, J. K. Buitelaar, K. B. Bulayeva, W. Cahn, V. D. Calhoun, D. M. Cannon, G. L. Cavalleri, C.-Y. Cheng, S. Cichon, M. R. Cookson, A. Corvin, B. Crespo-Facorro, J. E. Curran, M. Czisch, A. M. Dale, G. E. Davies, A. J. M. D. Craen, E. J. C. D. Geus, P. L. D. Jager, G. I. D. Zubicaray, I. J. Deary, S. Debette, C. DeCarli, N. Delanty, C. Depondt, A. DeStefano, A. Dillman, S. Djurovic, G. Donohoe, W. C. Drevets, R. Duggirala, T. D. Dyer, C. Enzinger, S. Erk, T. Espeseth, I. O. Fedko, G. Fernández, L. Ferrucci, S. E. Fisher, D. A. Fleischman, I. Ford, M. Fornage, T. M. Foroud, P. T. Fox, C. Francks, M. Fukunaga, J. R. Gibbs, D. C. Glahn, R. L. Gollub, H. H. H. Göring, R. C. Green, O. Gruber, V. Gudnason, S. Guelfi, A. K. Håberg, N. K. Hansell, J. Hardy, C. A. Hartman, R. Hashimoto, K. Hegenscheid, A. Heinz, S. L. Hellard, D. G.

Hernandez, D. J. Heslenfeld, B.-C. Ho, P. J. Hoekstra, W. Hoffmann, A. Hofman, F. Holsboer, G. Homuth, N. Hosten, J.-J. Hottenga, M. Huentelman, H. E. H. Pol, M. Ikeda, C. R. J. Jr, M. Jenkinson, R. Johnson, E. G. Jönsson, J. W. Jukema, R. S. Kahn, R. Kanai, I. Kloszewska, D. S. Knopman, P. Kochunov, J. B. Kwok, S. M. Lawrie, H. Lemaître, X. Liu, D. L. Longo, O. L. Lopez, S. Lovestone, O. Martinez, J.-L. Martinot, V. S. Mattay, C. McDonald, A. M. McIntosh, F. J. McMahon, K. L. McMahon, P. Mecocci, I. Melle, A. Meyer-Lindenberg, S. Mohnke, G. W. Montgomery, D. W. Morris, T. H. Mosley, T. W. Mühleisen, B. Müller-Myhsok, M. A. Nalls, M. Nauck, T. E. Nichols, W. J. Niessen, M. M. Nöthen, L. Nyberg, K. Ohi, R. L. Olvera, R. A. Ophoff, M. Pandolfo, T. Paus, Z. Pausova, B. W. J. H. Penninx, G. B. Pike, S. G. Potkin, B. M. Psaty, S. Reppermund, M. Rietschel, J. L. Roffman, N. Romanczuk-Seiferth, J. I. Rotter, M. Ryten, R. L. Sacco, P. S. Sachdev, A. J. Saykin, R. Schmidt, H. Schmidt, P. R. Schofield, S. Sigursson, A. Simmons, A. Singleton, S. M. Sisodiya, C. Smith, J. W. Smoller, H. Soininen, V. M. Steen, D. J. Stott, J. E. Sussmann, A. Thalamuthu, A. W. Toga, B. J. Traynor, J. Troncoso, M. Tsolaki, C. Tzourio, A. G. Uitterlinden, M. C. V. Hernández, M. V. der Brug, A. van der Lugt, N. J. A. van der Wee, N. E. M. V. Haren, D. van 't Ent, M.-J. V. Tol, B. N. Vardarajan, B. Vellas, D. J. Veltman, H. Völzke, H. Walter, J. M. Wardlaw, T. H. Wassink, M. E. Weale, D. R. Weinberger, M. W. Weiner, W. Wen, E. Westman, T. White, T. Y. Wong, C. B. Wright, R. H. Zielke, A. B. Zonderman, N. G. Martin, C. M. V. Duijn, M. J. Wright, W. T. Longstreth, G. Schumann, H. J. Grabe, B. Franke, L. J. Launer, S. E. Medland, S. Seshadri, P. M. Thompson, M. A. Ikram, Novel genetic loci associated with hippocampal volume. *Nat Commun*. 8, 13624 (2017).

118. D. Harold, R. Abraham, P. Hollingworth, R. Sims, A. Gerrish, M. L. Hamshere, J. S. Pahwa, V. Moskvina, K. Dowzell, A. Williams, N. Jones, C. Thomas, A. Stretton, A. R. Morgan, S. Lovestone, J. Powell, P. Proitsi, M. K. Lupton, C. Brayne, D. C. Rubinsztein, M. Gill, B. Lawlor, A. Lynch, K. Morgan, K. S. Brown, P. A. Passmore, D. Craig, B. McGuinness, S. Todd, C. Holmes, D. Mann, A. D. Smith, S. Love, P. G. Kehoe, J. Hardy, S. Mead, N. Fox, M. Rossor, J. Collinge, W. Maier, F. Jessen, B. Schürmann, R. Heun, H. van den Bussche, I. Heuser, J. Kornhuber, J. Wiltfang, M. Dichgans, L. Frölich, H. Hampel, M. Hüll, D. Rujescu, A. M. Goate, J. S. K. Kauwe, C. Cruchaga, P. Nowotny, J. C. Morris, K. Mayo, K. Sleegers, K. Bettens, S. Engelborghs, P. P. D. Deyn, C. V. Broeckhoven, G. Livingston, N. J. Bass, H. Gurling, A. McQuillin, R. Gwilliam, P. Deloukas, A. Al-Chalabi, C. E. Shaw, M. Tsolaki, A. B. Singleton, R. Guerreiro, T. W. Mühleisen, M. M. Nöthen, S. Moebus, K.-H. Jöckel, N. Klopp, H.-E. Wichmann, M. M. Carrasquillo, V. S. Pankratz, S. G. Younkin, P. A. Holmans, M. O'Donovan, M. J. Owen, J. Williams, Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet*. 41, 1088–93 (2009).

119. Y. Sun, P. Ip, A. Chakrabartty, Simple Elimination of Background Fluorescence in Formalin-Fixed Human Brain Tissue for Immunofluorescence Microscopy. *J Vis Exp* (2017), doi:10.3791/56188.

120. T. Stuart, A. Srivastava, S. Madad, C. A. Lareau, R. Satija, Single-cell chromatin state analysis with Signac. *Nat Methods*. 18, 1333–1341 (2021).

121. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive Integration of Single-Cell Data. *Cell*. 177, 1888-1902.e21 (2019).

122. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 36, 411–420 (2018).

123. J. M. Granja, M. R. Corces, S. E. Pierce, S. T. Bagdatli, H. Choudhry, H. Y. Chang, W. J. Greenleaf, ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet*. 53, 403–411 (2021).

124. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 32, 2847–2849 (2016).

125. A. Khan, O. Fornes, A. Stigliani, M. Gheorghe, J. A. Castro-Mondragon, R. van der Lee, A. Bessy, J. Chèneby, S. R. Kulkarni, G. Tan, D. Baranasic, D. J. Arenillas, A. Sandelin, K. Vandepoele, B. Lenhard, B. Ballester, W. W. Wasserman, F. Parcy, A. Mathelier, JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res*. 46, gkx1126- (2017).

126. A. N. Schep, B. Wu, J. D. Buenrostro, W. J. Greenleaf, chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods*. 14, 975–978 (2017).

127. H. A. Pliner, J. S. Packer, J. L. McFaline-Figueroa, D. A. Cusanovich, R. M. Daza, D. Aghamirzaie, S. Srivatsan, X. Qiu, D. Jackson, A. Minkina, A. C. Adey, F. J. Steemers, J. Shendure, C. Trapnell, Cicero Predicts cis-Regulatory DNA Interactions from Single-Cell Chromatin Accessibility Data. *Mol Cell*. 71, 858-871.e8 (2018).

128. R. Gaujoux, C. Seoighe, A flexible R package for nonnegative matrix factorization. *Bmc Bioinformatics*. 11, 367 (2010).

129. M. V. Kuleshov, M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik, A. Lachmann, M. G. McDermott, C. D. Monteiro, G. W. Gundersen, A. Ma'ayan, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 44, W90–W97 (2016).

130. E. Y. Chen, C. M. Tan, Y. Kou, Q. Duan, Z. Wang, G. V. Meirelles, N. R. Clark, A. Ma'ayan, Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *Bmc Bioinformatics*. 14, 128 (2013).

131. H. K. Finucane, Y. A. Reshef, V. Anttila, K. Slowikowski, A. Gusev, A. Byrnes, S. Gazal, P.-R. Loh, C. Lareau, N. Shoresh, G. Genovese, A. Saunders, E. Macosko, S. Pollack, T. B. Consortium, J. R. B. Perry, J. D. Buenrostro, B. E. Bernstein, S. Raychaudhuri, S. McCarroll, B. M. Neale, A. L. Price, Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet*. 50, 621–629 (2018).

132. A. F. Pardiñas, P. Holmans, A. J. Pocklington, V. Escott-Price, S. Ripke, N. Carrera, S. E. Legge, S. Bishop, D. Cameron, M. L. Hamshere, J. Han, L. Hubbard, A. Lynham, K. Mantripragada, E. Rees, J. H. MacCabe, S. A. McCarroll, B. T. Baune, G. Breen, E. M. Byrne, U. Dannlowski, T. C. Eley, C. Hayward, N. G. Martin, A. M. McIntosh, R. Plomin, D. J. Porteous,

N. R. Wray, A. Caballero, D. H. Geschwind, L. M. Huckins, D. M. Ruderfer, E. Santiago, P. Sklar, E. A. Stahl, H. Won, E. Agerbo, T. D. Als, O. A. Andreassen, M. Bækvad-Hansen, P. B. Mortensen, C. B. Pedersen, A. D. Børglum, J. Bybjerg-Grauholm, S. Djurovic, N. Durmishi, M. G. Pedersen, V. Golimbet, J. Grove, D. M. Hougaard, M. Mattheisen, E. Molden, O. Mors, M. Nordentoft, M. Pejovic-Milovancevic, E. Sigurdsson, T. Silagadze, C. S. Hansen, K. Stefansson, H. Stefansson, S. Steinberg, S. Tosato, T. Werge, G. Consortium:, C. Consortium:, D. A. Collier, D. Rujescu, G. Kirov, M. J. Owen, M. C. O'Donovan, J. T. R. Walters, G. Consortium, C. Consortium, G. Consortium, C. Consortium, Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet*. 50, 381–389 (2018).

133. R. Ferrari, D. G. Hernandez, M. A. Nalls, J. D. Rohrer, A. Ramasamy, J. B. J. Kwok, C. Dobson-Stone, W. S. Brooks, P. R. Schofield, G. M. Halliday, J. R. Hodges, O. Piguet, L. Bartley, E. Thompson, E. Haan, I. Hernández, A. Ruiz, M. Boada, B. Borroni, A. Padovani, C. Cruchaga, N. J. Cairns, L. Benussi, G. Binetti, R. Ghidoni, G. Forloni, D. Galimberti, C. Fenoglio, M. Serpente, E. Scarpini, J. Clarimón, A. Lleó, R. Blesa, M. L. Waldö, K. Nilsson, C. Nilsson, I. R. A. Mackenzie, G.-Y. R. Hsiung, D. M. A. Mann, J. Grafman, C. M. Morris, J. Attems, T. D. Griffiths, I. G. McKeith, A. J. Thomas, P. Pietrini, E. D. Huey, E. M. Wassermann, A. Baborie, E. Jaros, M. C. Tierney, P. Pastor, C. Razquin, S. Ortega-Cubero, E. Alonso, R. Perneczky, J. Diehl-Schmid, P. Alexopoulos, A. Kurz, I. Rainero, E. Rubino, L. Pinessi, E. Rogaeva, P. S. George-Hyslop, G. Rossi, F. Tagliavini, G. Giaccone, J. B. Rowe, J. C. M. Schlachetzki, J. Uphill, J. Collinge, S. Mead, A. Danek, V. M. V. Deerlin, M. Grossman, J. Q. Trojanowski, J. van der Zee, W. Deschamps, T. V. Langenhove, M. Cruts, C. V. Broeckhoven, S. F. Cappa, I. L. Ber, D. Hannequin, V. Golfier, M. Vercelletto, A. Brice, B. Nacmias, S. Sorbi, S. Bagnoli, I. Piaceri, J. E. Nielsen, L. E. Hjermind, M. Riemenschneider, M. Mayhaus, B. Ibach, G. Gasparoni, S. Pichler, W. Gu, M. N. Rossor, N. C. Fox, J. D. Warren, M. G. Spillantini, H. R. Morris, P. Rizzu, P. Heutink, J. S. Snowden, S. Rollinson, A. Richardson, A. Gerhard, A. C. Bruni, R. Maletta, F. Frangipane, C. Cupidi, L. Bernardi, M. Anfossi, M. Gallo, M. E. Conidi, N. Smirne, R. Rademakers, M. Baker, D. W. Dickson, N. R. Graff-Radford, R. C. Petersen, D. Knopman, K. A. Josephs, B. F. Boeve, J. E. Parisi, W. W. Seeley, B. L. Miller, A. M. Karydas, H. Rosen, J. C. van Swieten, E. G. P. Dopper, H. Seelaar, Y. A. L. Pijnenburg, P. Scheltens, G. Logroscino, R. Capozzo, V. Novelli, A. A. Puca, M. Franceschi, A. Postiglione, G. Milan, P. Sorrentino, M. Kristiansen, H.-H. Chiang, C. Graff, F. Pasquier, A. Rollin, V. Deramecourt, F. Lebert, D. Kapogiannis, L. Ferrucci, S. Pickering-Brown, A. B. Singleton, J. Hardy, P. Momeni, Frontotemporal dementia and its subtypes: a genome-wide association study. *Lancet Neurology*. 13, 686–699 (2014).

134. J. Chen, J.-T. Yu, K. Wojta, H.-F. Wang, H. Zetterberg, K. Blennow, J. S. Yokoyama, M. W. Weiner, J. H. Kramer, H. Rosen, B. L. Miller, G. Coppola, A. L. Boxer, F. the A. D. N. Initiative, Genome-wide association study identifies MAPT locus influencing human plasma tau levels. *Neurology*. 88, 669–676 (2017).

135. T. F. M. Andlauer, D. Buck, G. Antony, A. Bayas, L. Bechmann, A. Berthele, A. Chan, C. Gasperi, R. Gold, C. Graetz, J. Haas, M. Hecker, C. Infante-Duarte, M. Knop, T. Kümpfel, V. Limmroth, R. A. Linker, V. Loleit, F. Luessi, S. G. Meuth, M. Mühlau, S. Nischwitz, F. Paul, M. Pütz, T. Ruck, A. Salmen, M. Stangel, J.-P. Stellmann, K. H. Stürner, B. Tackenberg, F. T. Bergh, H. Tumani, C. Warnke, F. Weber, H. Wiendl, B. Wildemann, U. K. Zettl, U. Ziemann, F. Zipp, J. Arloth, P. Weber, M. Radivojkov-Blagojevic, M. O. Scheinhardt, T. Dankowski, T.

Bettecken, P. Lichtner, D. Czamara, T. Carrillo-Roa, E. B. Binder, K. Berger, L. Bertram, A. Franke, C. Gieger, S. Herms, G. Homuth, M. Ising, K.-H. Jöckel, T. Kacprowski, S. Kloiber, M. Laudes, W. Lieb, C. M. Lill, S. Lucae, T. Meitinger, S. Moebus, M. Müller-Nurasyid, M. M. Nöthen, A. Petersmann, R. Rawal, U. Schminke, K. Strauch, H. Völzke, M. Waldenberger, J. Wellmann, E. Porcu, A. Mulas, M. Pitzalis, C. Sidore, I. Zara, F. Cucca, M. Zoledziewska, A. Ziegler, B. Hemmer, B. Müller-Myhsok, Novel multiple sclerosis susceptibility loci implicated in epigenetic regulation. *Sci Adv*. 2, e1501678 (2016).

136. J. Z. Liu, S. van Sommeren, H. Huang, S. C. Ng, R. Alberts, A. Takahashi, S. Ripke, J. C. Lee, L. Jostins, T. Shah, S. Abedian, J. H. Cheon, J. Cho, N. E. Daryani, L. Franke, Y. Fuyuno, A. Hart, R. C. Juyal, G. Juyal, W. H. Kim, A. P. Morris, H. Poustchi, W. G. Newman, V. Midha, T. R. Orchard, H. Vahedi, A. Sood, J. J. Y. Sung, R. Malekzadeh, H.-J. Westra, K. Yamazaki, S.-K. Yang, I. M. S. G. Consortium, I. I. G. Consortium, J. C. Barrett, A. Franke, B. Z. Alizadeh, M. Parkes, T. B. K, M. J. Daly, M. Kubo, C. A. Anderson, R. K. Weersma, Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet*. 47, 979–986 (2015).

137. T. E. M. R. and G. (eMERGE) Consortium, T. Mig. Consortium, T. P. Consortium, T. L. C. Study, A. R. Wood, T. Esko, J. Yang, S. Vedantam, T. H. Pers, S. Gustafsson, A. Y. Chu, K. Estrada, J. Luan, Z. Kutalik, N. Amin, M. L. Buchkovich, D. C. Croteau-Chonka, F. R. Day, Y. Duan, T. Fall, R. Fehrmann, T. Ferreira, A. U. Jackson, J. Karjalainen, K. S. Lo, A. E. Locke, R. Mägi, E. Mihailov, E. Porcu, J. C. Randall, A. Scherag, A. A. E. Vinkhuyzen, H.-J. Westra, T. W. Winkler, T. Workalemahu, J. H. Zhao, D. Absher, E. Albrecht, D. Anderson, J. Baron, M. Beekman, A. Demirkan, G. B. Ehret, B. Feenstra, M. F. Feitosa, K. Fischer, R. M. Fraser, A. Goel, J. Gong, A. E. Justice, S. Kanoni, M. E. Kleber, K. Kristiansson, U. Lim, V. Lotay, J. C. Lui, M. Mangino, I. M. Leach, C. Medina-Gomez, M. A. Nalls, D. R. Nyholt, C. D. Palmer, D. Pasko, S. Pechlivanis, I. Prokopenko, J. S. Ried, S. Ripke, D. Shungin, A. Stancáková, R. J. Strawbridge, Y. J. Sung, T. Tanaka, A. Teumer, S. Trompet, S. W. van der Laan, J. van Setten, J. V. V. Vliet-Ostaptchouk, Z. Wang, L. Yengo, W. Zhang, U. Afzal, J. Ärnlöv, G. M. Arscott, S. Bandinelli, A. Barrett, C. Bellis, A. J. Bennett, C. Berne, M. Blüher, J. L. Bolton, Y. Böttcher, H. A. Boyd, M. Bruinenberg, B. M. Buckley, S. Buyske, I. H. Caspersen, P. S. Chines, R. Clarke, S. Claudi-Boehm, M. Cooper, E. W. Daw, P. A. D. Jong, J. Deelen, G. Delgado, J. C. Denny, R. Dhonukshe-Rutten, M. Dimitriou, A. S. F. Doney, M. Dörr, N. Eklund, E. Eury, L. Folkersen, M. E. Garcia, F. Geller, V. Giedraitis, A. S. Go, H. Grallert, T. B. Grammer, J. Gräßler, H. Grönberg, L. C. P. G. M. de Groot, C. J. Groves, J. Haessler, P. Hall, T. Haller, G. Hallmans, A. Hannemann, C. A. Hartman, M. Hassinen, C. Hayward, N. L. Heard-Costa, Q. Helmer, G. Hemani, A. K. Henders, H. L. Hillege, M. A. Hlatky, W. Hoffmann, P. Hoffmann, O. Holmen, J. J. Houwing-Duistermaat, T. Illig, A. Isaacs, A. L. James, J. Jeff, B. Johansen, Å. Johansson, J. Jolley, T. Juliusdottir, J. Junttila, A. N. Kho, L. Kinnunen, N. Klopp, T. Kocher, W. Kratzer, P. Lichtner, L. Lind, J. Lindström, S. Lobbens, M. Lorentzon, Y. Lu, V. Lyssenko, P. K. E. Magnusson, A. Mahajan, M. Maillard, W. L. McArdle, C. A. McKenzie, S. McLachlan, P. J. McLaren, C. Menni, S. Merger, L. Milani, A. Moayyeri, K. L. Monda, M. A. Morken, G. Müller, M. Müller-Nurasyid, A. W. Musk, N. Narisu, M. Nauck, I. M. Nolte, M. M. Nöthen, L. Oozageer, S. Pilz, N. W. Rayner, F. Renstrom, N. R. Robertson, L. M. Rose, R. Roussel, S. Sanna, H. Scharnagl, S. Scholtens, F. R. Schumacher, H. Schunkert, R. A. Scott, J. Sehmi, T. Seufferlein, J. Shi, K. Silventoinen, J. H. Smit, A. V. Smith, J. Smolonska, A. V. Stanton, K. Stirrups, D. J. Stott, H. M. Stringham, J. Sundström, M. A. Swertz, A.-C. Syvänen, B. O. Tayo, G. Thorleifsson, J. P. Tyrer, S. van Dijk, N. M. van Schoor, N. van der Velde, D. van Heemst, F. V. A. van Oort,

S. H. Vermeulen, N. Verweij, J. M. Vonk, L. L. Waite, M. Waldenberger, R. Wennauer, L. R. Wilkens, C. Willenborg, T. Wilsgaard, M. K. Wojczynski, A. Wong, A. F. Wright, Q. Zhang, D. Arveiler, S. J. L. Bakker, J. Beilby, R. N. Bergman, S. Bergmann, R. Biffar, J. Blangero, D. I. Boomsma, S. R. Bornstein, P. Bovet, P. Brambilla, M. J. Brown, H. Campbell, M. J. Caulfield, A. Chakravarti, R. Collins, F. S. Collins, D. C. Crawford, L. A. Cupples, J. Danesh, U. de Faire, H. M. den Ruijter, R. Erbel, J. Erdmann, J. G. Eriksson, M. Farrall, E. Ferrannini, J. Ferrières, I. Ford, N. G. Forouhi, T. Forrester, R. T. Gansevoort, P. V. Gejman, C. Gieger, A. Golay, O. Gottesman, V. Gudnason, U. Gyllensten, D. W. Haas, A. S. Hall, T. B. Harris, A. T. Hattersley, A. C. Heath, C. Hengstenberg, A. A. Hicks, L. A. Hindorff, A. D. Hingorani, A. Hofman, G. K. Hovingh, S. E. Humphries, S. C. Hunt, E. Hypponen, K. B. Jacobs, M.-R. Jarvelin, P. Jousilahti, A. M. Jula, J. Kaprio, J. J. P. Kastelein, M. Kayser, F. Kee, S. M. Keinanen-Kiukaanniemi, L. A. Kiemeney, J. S. Kooner, C. Kooperberg, S. Koskinen, P. Kovacs, A. T. Kraja, M. Kumari, J. Kuusisto, T. A. Lakka, C. Langenberg, L. L. Marchand, T. Lehtimäki, S. Lupoli, P. A. F. Madden, S. Männistö, P. Manunta, A. Marette, T. C. Matise, B. McKnight, T. Meitinger, F. L. Moll, G. W. Montgomery, A. D. Morris, A. P. Morris, J. C. Murray, M. Nelis, C. Ohlsson, A. J. Oldehinkel, K. K. Ong, W. H. Ouwehand, G. Pasterkamp, A. Peters, P. P. Pramstaller, J. F. Price, L. Qi, O. T. Raitakari, T. Rankinen, D. C. Rao, T. K. Rice, M. Ritchie, I. Rudan, V. Salomaa, N. J. Samani, J. Saramies, M. A. Sarzynski, P. E. H. Schwarz, S. Sebert, P. Sever, A. R. Shuldiner, J. Sinisalo, V. Steinthorsdottir, R. P. Stolk, J.-C. Tardif, A. Tönjes, A. Tremblay, E. Tremoli, J. Virtamo, M.-C. Vohl, P. Amouyel, F. W. Asselbergs, T. L. Assimes, M. Bochud, B. O. Boehm, E. Boerwinkle, E. P. Bottinger, C. Bouchard, S. Cauchi, J. C. Chambers, S. J. Chanock, R. S. Cooper, P. I. W. de Bakker, G. Dedoussis, L. Ferrucci, P. W. Franks, P. Froguel, L. C. Groop, C. A. Haiman, A. Hamsten, M. G. Hayes, J. Hui, D. J. Hunter, K. Hveem, J. W. Jukema, R. C. Kaplan, M. Kivimaki, D. Kuh, M. Laakso, Y. Liu, N. G. Martin, W. März, M. Melbye, S. Moebus, P. B. Munroe, I. Njølstad, B. A. Oostra, C. N. A. Palmer, N. L. Pedersen, M. Perola, L. Pérusse, U. Peters, J. E. Powell, C. Power, T. Quertermous, R. Rauramaa, E. Reinmaa, P. M. Ridker, F. Rivadeneira, J. I. Rotter, T. E. Saaristo, D. Saleheen, D. Schlessinger, P. E. Slagboom, H. Snieder, T. D. Spector, K. Strauch, M. Stumvoll, J. Tuomilehto, M. Uusitupa, P. van der Harst, H. Völzke, M. Walker, N. J. Wareham, H. Watkins, H.-E. Wichmann, J. F. Wilson, P. Zanen, P. Deloukas, I. M. Heid, C. M. Lindgren, K. L. Mohlke, E. K. Speliotes, U. Thorsteinsdottir, I. Barroso, C. S. Fox, K. E. North, D. P. Strachan, J. S. Beckmann, S. I. Berndt, M. Boehnke, I. B. Borecki, M. I. McCarthy, A. Metspalu, K. Stefansson, A. G. Uitterlinden, C. M. van Duijn, L. Franke, C. J. Willer, A. L. Price, G. Lettre, R. J. F. Loos, M. N. Weedon, E. Ingelsson, J. R. O'Connell, G. R. Abecasis, D. I. Chasman, M. E. Goddard, P. M. Visscher, J. N. Hirschhorn, T. M. Frayling, Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet.* 46, 1173–1186 (2014).

138. C. J. Willer, E. M. Schmidt, S. Sengupta, G. M. Peloso, S. Gustafsson, S. Kanoni, A. Ganna, J. Chen, M. L. Buchkovich, S. Mora, J. S. Beckmann, J. L. Bragg-Gresham, H.-Y. Chang, A. Demirkan, H. M. D. Hertog, R. Do, L. A. Donnelly, G. B. Ehret, T. Esko, M. F. Feitosa, T. Ferreira, K. Fischer, P. Fontanillas, R. M. Fraser, D. F. Freitag, D. Gurdasani, K. Heikkilä, E. Hyppönen, A. Isaacs, A. U. Jackson, Å. Johansson, T. Johnson, M. Kaakinen, J. Kettunen, M. E. Kleber, X. Li, J. Luan, L.-P. Lyytikäinen, P. K. E. Magnusson, M. Mangino, E. Mihailov, M. E. Montasser, M. Müller-Nurasyid, I. M. Nolte, J. R. O'Connell, C. D. Palmer, M. Perola, A.-K. Petersen, S. Sanna, R. Saxena, S. K. Service, S. Shah, D. Shungin, C. Sidore, C. Song, R. J. Strawbridge, I. Surakka, T. Tanaka, T. M. Teslovich, G. Thorleifsson, E. G. V. den Herik, B. F. Voight, K. A. Volcik, L. L. Waite, A. Wong, Y. Wu, W. Zhang, D. Absher, G. Asiki, I. Barroso, L. F. Been, J. L. Bolton, L. L. Bonnycastle, P. Brambilla, M. S. Burnett, G. Cesana, M. Dimitriou, A. S. F. Doney, A. Döring, P. Elliott, S. E. Epstein, G. I. Eyjolfsson, B. Gigante, M. O. Goodarzi, H.

Grallert, M. L. Gravito, C. J. Groves, G. Hallmans, A.-L. Hartikainen, C. Hayward, D. Hernandez, A. A. Hicks, H. Holm, Y.-J. Hung, T. Illig, M. R. Jones, P. Kaleebu, J. J. P. Kastelein, K.-T. Khaw, E. Kim, N. Klopp, P. Komulainen, M. Kumari, C. Langenberg, T. Lehtimäki, S.-Y. Lin, J. Lindström, R. J. F. Loos, F. Mach, W. L. McArdle, C. Meisinger, B. D. Mitchell, G. Müller, R. Nagaraja, N. Narisu, T. V. M. Nieminen, R. N. Nsubuga, I. Olafsson, K. K. Ong, A. Palotie, T. Papamarkou, C. Pomilla, A. Pouta, D. J. Rader, M. P. Reilly, P. M. Ridker, F. Rivadeneira, I. Rudan, A. Ruokonen, N. Samani, H. Scharnagl, J. Seeley, K. Silander, A. Stančáková, K. Stirrups, A. J. Swift, L. Tiret, A. G. Uitterlinden, L. J. van Pelt, S. Vedantam, N. Wainwright, C. Wijmenga, S. H. Wild, G. Willemsen, T. Wilsgaard, J. F. Wilson, E. H. Young, J. H. Zhao, L. S. Adair, D. Arveiler, T. L. Assimes, S. Bandinelli, F. Bennett, M. Bochud, B. O. Boehm, D. I. Boomsma, I. B. Borecki, S. R. Bornstein, P. Bovet, M. Burnier, H. Campbell, A. Chakravarti, J. C. Chambers, Y.-D. I. Chen, F. S. Collins, R. S. Cooper, J. Danesh, G. Dedoussis, U. de Faire, A. B. Feranil, J. Ferrières, L. Ferrucci, N. B. Freimer, C. Gieger, L. C. Groop, V. Gudnason, U. Gyllensten, A. Hamsten, T. B. Harris, A. Hingorani, J. N. Hirschhorn, A. Hofman, G. K. Hovingh, C. A. Hsiung, S. E. Humphries, S. C. Hunt, K. Hveem, C. Iribarren, M.-R. Järvelin, A. Jula, M. Kähönen, J. Kaprio, A. Kesäniemi, M. Kivimaki, J. S. Kooner, P. J. Koudstaal, R. M. Krauss, D. Kuh, J. Kuusisto, K. O. Kyvik, M. Laakso, T. A. Lakka, L. Lind, C. M. Lindgren, N. G. Martin, W. März, M. I. McCarthy, C. A. McKenzie, P. Meneton, A. Metspalu, L. Moilanen, A. D. Morris, P. B. Munroe, I. Njølstad, N. L. Pedersen, C. Power, P. P. Pramstaller, J. F. Price, B. M. Psaty, T. Quertermous, R. Rauramaa, D. Saleheen, V. Salomaa, D. K. Sanghera, J. Saramies, P. E. H. Schwarz, W. H.-H. Sheu, A. R. Shuldiner, A. Siegbahn, T. D. Spector, K. Stefansson, D. P. Strachan, B. O. Tayo, E. Tremoli, J. Tuomilehto, M. Uusitupa, C. M. van Duijn, P. Vollenweider, L. Wallentin, N. J. Wareham, J. B. Whitfield, B. H. R. Wolffenbuttel, J. M. Ordovas, E. Boerwinkle, C. N. A. Palmer, U. Thorsteinsdottir, D. I. Chasman, J. I. Rotter, P. W. Franks, S. Ripatti, L. A. Cupples, M. S. Sandhu, S. S. Rich, M. Boehnke, P. Deloukas, S. Kathiresan, K. L. Mohlke, E. Ingelsson, G. R. Abecasis, G. L. G. Consortium, Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 45, 1274–1283 (2013).

139. P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis. *Bmc Bioinformatics*. 9, 559 (2008).

140. B. Zhang, S. Horvath, A General Framework for Weighted Gene Co-Expression Network Analysis. *Stat Appl Genet Mol*. 4, Article17 (2005).

141. K. Lage, E. O. Karlberg, Z. M. Størling, P. Í. Ólason, A. G. Pedersen, O. Rigina, A. M. Hinsby, Z. Tümer, F. Pociot, N. Tommerup, Y. Moreau, S. Brunak, A human phenome-interactome network of protein complexes implicated in genetic disorders. *Nat Biotechnol*. 25, 309–316 (2007).

142. C. Stark, B.-J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, M. Tyers, BioGRID: a general repository for interaction datasets. *Nucleic Acids Res*. 34, D535–D539 (2006).

143. D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, L. J. Jensen, C. von Mering, STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 47, gky1131 (2018).

144. V. Swarup, T. S. Chang, D. M. Duong, E. B. Dammer, J. Dai, J. J. Lah, E. C. B. Johnson, N. T. Seyfried, A. I. Levey, D. H. Geschwind, Identification of Conserved Proteomic Networks in Neurodegenerative Dementia. *Cell Reports*. 31, 107807 (2020).

145. L. Haghverdi, A. T. L. Lun, M. D. Morgan, J. C. Marioni, Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol*. 36, 421–427 (2018).

146. J. D. Welch, V. Kozareva, A. Ferreira, C. Vanderburg, C. Martin, E. Z. Macosko, Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. *Cell*. 177, 1873-1887.e17 (2019).

147. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *Arxiv* (2018), doi:10.48550/arxiv.1802.03426.

148. V. A. Traag, L. Waltman, N. J. van Eck, From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep*. 9, 5233 (2019).

149. J. Satoh, N. Kawana, Y. Yamamoto, *Gene Regul Syst Biology*, in press, doi:10.4137/grsb.s13204.

150. C. Trapnell, D. Cacchiarelli, J. Grimsby, P. Pokharel, S. Li, M. Morse, N. J. Lennon, K. J. Livak, T. S. Mikkelsen, J. L. Rinn, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 32, 381–6 (2014).

151. J. Cao, M. Spielmann, X. Qiu, X. Huang, D. M. Ibrahim, A. J. Hill, F. Zhang, S. Mundlos, L. Christiansen, F. J. Steemers, C. Trapnell, J. Shendure, The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 566, 496–502 (2019).

152. X. Qiu, Q. Mao, Y. Tang, L. Wang, R. Chawla, H. A. Pliner, C. Trapnell, Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 14, 979–982 (2017).

153. R. Mitra, A. L. MacLean, RVAgene: generative modeling of gene expression time series data. *Bioinformatics*. 37, 3252–3262 (2021).

154. H. Shimano, R. Sato, SREBP-regulated lipid metabolism: convergent physiology — divergent pathophysiology. *Nat Rev Endocrinol*. 13, 710–730 (2017).

155. A. Mohamed, A. Viveiros, K. Williams, E. P. de Chaves, Aβ inhibits SREBP-2 activation through Akt inhibition. *J Lipid Res*. 59, 1–13 (2018).

156. E. M. Pugacheva, N. Kubo, D. Loukinov, M. Tajmul, S. Kang, A. L. Kovalchuk, A. V. Strunnikov, G. E. Zentner, B. Ren, V. V. Lobanenkov, CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention. *Proc National Acad Sci*. 117, 2020–2031 (2020).

157. S. Kim, N.-K. Yu, B.-K. Kaang, CTCF as a multifunctional protein in genome regulation and gene expression. *Exp Mol Medicine*. 47, e166–e166 (2015).

158. J. A. Chen, Z. Chen, H. Won, A. Y. Huang, J. K. Lowe, K. Wojta, J. S. Yokoyama, G. Bensimon, P. N. Leigh, C. Payan, A. Shatunov, A. R. Jones, C. M. Lewis, P. Deloukas, P. Amouyel, C. Tzourio, J.-F. Dartigues, A. Ludolph, A. L. Boxer, J. M. Bronstein, A. Al-Chalabi, D. H. Geschwind, G. Coppola, Joint genome-wide association study of progressive supranuclear palsy identifies novel susceptibility loci and genetic correlation to neurodegenerative diseases. *Mol Neurodegener*. 13, 41 (2018).

159. J. C. Ulirsch, C. A. Lareau, E. L. Bao, L. S. Ludwig, M. H. Guo, C. Benner, A. T. Satpathy, V. K. Kartha, R. M. Salem, J. N. Hirschhorn, H. K. Finucane, M. J. Aryee, J. D. Buenrostro, V. G. Sankaran, Interrogation of human hematopoiesis at single-cell and single-variant resolution. *Nat Genet*. 51, 683–693 (2019).

160. J. E. Rexach, D. Polioudakis, A. Yin, V. Swarup, T. S. Chang, T. Nguyen, A. Sarkar, L. Chen, J. Huang, L.-C. Lin, W. Seeley, J. Q. Trojanowski, D. Malhotra, D. H. Geschwind, Tau Pathology Drives Dementia Risk-Associated Gene Networks toward Chronic Inflammatory States and Immunosuppression. *Cell Reports*. 33, 108398 (2020).

161. Y. E. Wu, L. Pan, Y. Zuo, X. Li, W. Hong, Detecting Activated Cell Populations Using Single-Cell RNA-Seq. *Neuron*. 96, 313-329.e6 (2017).

162. S. R. Srivatsan, M. C. Regier, E. Barkan, J. M. Franks, J. S. Packer, P. Grosjean, M. Duran, S. Saxton, J. J. Ladd, M. Spielmann, C. Lois, P. D. Lampe, J. Shendure, K. R. Stevens, C. Trapnell, Embryo-scale, single-cell spatial transcriptomics. *Science*. 373, 111–117 (2021).

163. X. Wang, W. E. Allen, M. A. Wright, E. L. Sylwestrak, N. Samusik, S. Vesuna, K. Evans, C. Liu, C. Ramakrishnan, J. Liu, G. P. Nolan, F.-A. Bava, K. Deisseroth, Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science*. 361 (2018), doi:10.1126/science.aat5691.

164. C. R. Palmer, C. S. Liu, W. J. Romanow, M.-H. Lee, J. Chun, Altered cell and RNA isoform diversity in aging Down syndrome brains. *Proc National Acad Sci*. 118, e2114326118 (2021).

165. S. J. Fleming, J. C. Marioni, M. Babadi, CellBender remove-background: a deep generative model for unsupervised removal of background noise from scRNA-seq datasets. *Biorxiv*, 791699 (2019).

166. S. L. Wolock, R. Lopez, A. M. Klein, Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst*. 8, 281-291.e9 (2019).

167. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 36, 411–420 (2018).

168. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive Integration of Single-Cell Data. *Cell*. 177, 1888-1902.e21 (2019).

169. Y. Hao, S. Hao, E. Andersen-Nissen, W. M. Mauck, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data. *Cell*. 184, 3573-3587.e29 (2021).

170. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol*. 19, 15 (2018).

171. R. Lopez, J. Regier, M. B. Cole, M. I. Jordan, N. Yosef, Deep generative modeling for single-cell transcriptomics. *Nat Methods*. 15, 1053–1058 (2018).

172. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *Arxiv* (2018).

173. H. Mathys, J. Davila-Velderrain, Z. Peng, F. Gao, S. Mohammadi, J. Z. Young, M. Menon, L. He, F. Abdurrob, X. Jiang, A. J. Martorell, R. M. Ransohoff, B. P. Hafler, D. A. Bennett, M. Kellis, L.-H. Tsai, Single-cell transcriptomic analysis of Alzheimer's disease. *Nature*. 570, 332–337 (2019).

174. Y. Zhou, W. M. Song, P. S. Andhey, A. Swain, T. Levy, K. R. Miller, P. L. Poliani, M. Cominelli, S. Grover, S. Gilfillan, M. Cella, T. K. Ulland, K. Zaitsev, A. Miyashita, T. Ikeuchi, M. Sainouchi, A. Kakita, D. A. Bennett, J. A. Schneider, M. R. Nichols, S. A. Beausoleil, J. D. Ulrich, D. M. Holtzman, M. N. Artyomov, M. Colonna, Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. *Nat Med*. 26, 131–142 (2020).

175. S. Morabito, E. Miyoshi, N. Michael, S. Shahin, A. C. Martini, E. Head, J. Silva, K. Leavy, M. Perez-Rosendahl, V. Swarup, Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nat Genet*. 53, 1143–1155 (2021).

176. S. Morabito, F. Reese, N. Rahimzadeh, E. Miyoshi, V. Swarup, High dimensional co-expression networks enable discovery of transcriptomic drivers in complex biological systems. *bioRxiv* (2022), doi:10.1101/2022.09.22.509094.

177. P. Melsted, A. S. Booeshaghi, L. Liu, F. Gao, L. Lu, K. H. (Joseph) Min, E. da V. Beltrame, K. E. Hjörleifsson, J. Gehring, L. Pachter, Modular, efficient and constant-memory single-cell RNA-seq preprocessing. *Nat Biotechnol*. 39, 813–818 (2021).

178. I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods*. 16, 1289–1296 (2019).

179. V. A. Traag, L. Waltman, N. J. van Eck, From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep-uk*. 9, 5233 (2019).

180. E. Zhao, M. R. Stone, X. Ren, J. Guenthoer, K. S. Smythe, T. Pulliam, S. R. Williams, C. R. Uytingco, S. E. B. Taylor, P. Nghiem, J. H. Bielas, R. Gottardo, Spatial transcriptomics at subspot resolution with BayesSpace. *Nat Biotechnol*, 1–10 (2021).

181. C. Xu, R. Lopez, E. Mehlman, J. Regier, M. I. Jordan, N. Yosef, Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. *Mol Syst Biol*. 17, e9620 (2021).

182. M. Lotfollahi, M. Naghipourfar, M. D. Luecken, M. Khajavi, M. Büttner, M. Wagenstetter, Ž. Avsec, A. Gayoso, N. Yosef, M. Interlandi, S. Rybakov, A. V. Misharin, F. J. Theis, Mapping single-cell data to reference atlases by transfer learning. *Nat Biotechnol*. 40, 121–130 (2022).

183. G. Finak, A. McDavid, M. Yajima, J. Deng, V. Gersuk, A. K. Shalek, C. K. Slichter, H. W. Miller, M. J. McElrath, M. Prlic, P. S. Linsley, R. Gottardo, MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol*. 16, 278 (2015).

184. S. Durinck, P. T. Spellman, E. Birney, W. Huber, Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*. 4, 1184–1191 (2009).

185. S. Durinck, Y. Moreau, A. Kasprzyk, S. Davis, B. D. Moor, A. Brazma, W. Huber, BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*. 21, 3439–3440 (2005).

186. R. Wei, S. He, S. Bai, E. Sei, M. Hu, A. Thompson, K. Chen, S. Krishnamurthy, N. E. Navin, Spatial charting of single-cell transcriptomes in tissues. *Nat Biotechnol*, 1–10 (2022).

187. S. Jin, C. F. Guerrero-Juarez, L. Zhang, I. Chang, R. Ramos, C.-H. Kuan, P. Myung, M. V. Plikus, Q. Nie, Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. 12, 1088 (2021).

188. J. Windhager, B. Bodenmiller, N. Eling, *Biorxiv*, in press, doi:10.1101/2021.11.12.468357.

189. D. G. Bunis, J. Andrews, G. K. Fragiadakis, T. D. Burt, M. Sirota, dittoSeq: universal user-friendly single-cell and bulk RNA sequencing visualization toolkit. *Bioinformatics*. 36, 5535–5536 (2020).

190. S. E. Marsh, A. J. Walker, T. Kamath, L. Dissing-Olesen, T. R. Hammond, T. Y. de Soysa, A. M. H. Young, S. Murphy, A. Abdulraouf, N. Nadaf, C. Dufort, A. C. Walker, L. E. Lucca, V. Kozareva, C. Vanderburg, S. Hong, H. Bulstrode, P. J. Hutchinson, D. J. Gaffney, D. A. Hafler, R. J. M. Franklin, E. Z. Macosko, B. Stevens, Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. *Nat Neurosci*. 25, 306–316 (2022).

191. M. Andreatta, S. J. Carmona, UCell: Robust and scalable single-cell gene signature scoring. *Comput Struct Biotechnology J*. 19, 3796–3798 (2021).

192. S. Minoshima, B. Giordani, S. Berent, K. A. Frey, N. L. Foster, D. E. Kuhl, Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann Neurol*. 42, 85–94 (1997).

193. R. Wei, S. He, S. Bai, E. Sei, M. Hu, A. Thompson, K. Chen, S. Krishnamurthy, N. E. Navin, Spatial charting of single-cell transcriptomes in tissues. *Nat Biotechnol*, 1–10 (2022).

194. S. Jin, C. F. Guerrero-Juarez, L. Zhang, I. Chang, R. Ramos, C.-H. Kuan, P. Myung, M. V. Plikus, Q. Nie, Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. 12, 1088 (2021).

195. A. Chakraborty, A. Kamermans, B. van het Hof, K. Castricum, E. Aanhane, J. van Horssen, V. L. Thijssen, P. Scheltens, C. E. Teunissen, R. D. Fontijn, W. M. van der Flier, H. E. de Vries, Angiopoietin like-4 as a novel vascular mediator in capillary cerebral amyloid angiopathy. *Brain*. 141, 3377–3388 (2018).

196. N. Rathore, S. R. Ramani, H. Pantua, J. Payandeh, T. Bhangale, A. Wuster, M. Kapoor, Y. Sun, S. B. Kapadia, L. Gonzalez, A. A. Zarrin, A. Goate, D. V. Hansen, T. W. Behrens, R. R. Graham, Paired Immunoglobulin-like Type 2 Receptor Alpha G78R variant alters ligand binding and confers protection to Alzheimer's disease. *Plos Genet*. 14, e1007427 (2018).

197. K. L. Lindman, C. Jonsson, B. Weidung, J. Olsson, J. P. Pandey, D. Prokopenko, R. E. Tanzi, G. Hallmans, S. Eriksson, F. Elgh, H. Lövheim, PILRA polymorphism modifies the effect of APOE4 and GM17 on Alzheimer's disease risk. *Sci Rep-uk*. 12, 13264 (2022).

198. H. Zeng, E. H. Shen, J. G. Hohmann, S. W. Oh, A. Bernard, J. J. Royall, K. J. Glattfelder, S. M. Sunkin, J. A. Morris, A. L. Guillozet-Bongaarts, K. A. Smith, A. J. Ebbert, B. Swanson, L. Kuan, D. T. Page, C. C. Overly, E. S. Lein, M. J. Hawrylycz, P. R. Hof, T. M. Hyde, J. E. Kleinman, A. R. Jones, Large-Scale Cellular-Resolution Gene Profiling in Human Neocortex Reveals Species-Specific Molecular Signatures. *Cell*. 149, 483–496 (2012).

199. L. Schmued, J. Raymick, W. Tolleson, S. Sarkar, Y.-H. Zhang, A. Bell-Cohn, Introducing Amylo-Glo, a novel fluorescent amyloid specific histochemical tracer especially suited for multiple labeling and large scale quantification studies. *J Neurosci Meth*. 209, 120–126 (2012).

200. R. Kayed, E. Head, F. Sarsoza, T. Saing, C. W. Cotman, M. Necula, L. Margol, J. Wu, L. Breydo, J. L. Thompson, S. Rasool, T. Gurlo, P. Butler, C. G. Glabe, Fibril specific, conformation dependent antibodies recognize a generic epitope common to amyloid fibrils and fibrillar oligomers that is absent in prefibrillar oligomers. *Mol Neurodegener*. 2, 18 (2007).

201. E. Drummond, T. Kavanagh, G. Pires, M. Marta-Ariza, E. Kanshin, S. Nayak, A. Faustin, V. Berdah, B. Ueberheide, T. Wisniewski, The amyloid plaque proteome in early onset Alzheimer's disease and Down syndrome. *Acta Neuropathologica Commun*. 10, 53 (2022).

202. K. Khermesh, A. M. D'Erchia, M. Barak, A. Annese, C. Wachtel, E. Y. Levanon, E. Picardi, E. Eisenberg, Reduced levels of protein recoding by A-to-I RNA editing in Alzheimer's disease. *Rna*. 22, 290–302 (2016).

203. D. Mehta, R. Jackson, G. Paul, J. Shi, M. Sabbagh, Why do trials for Alzheimer's disease drugs keep failing? A discontinued drug perspective for 2010-2015. *Expert Opin Inv Drug*. 26, 735–739 (2017).

204. L. A. Sandusky-Beltran, E. M. Sigurdsson, Tau immunotherapies: Lessons learned, current status and future considerations. *Neuropharmacology*. 175, 108104 (2020).

205. E. J. A. Bowles, P. K. Crane, R. L. Walker, J. Chubak, A. Z. LaCroix, M. L. Anderson, D. Rosenberg, C. D. Keene, E. B. Larson, Cognitive Resilience to Alzheimer's Disease Pathology in the Human Brain. *J Alzheimer's Dis*. Preprint, 1–13 (2019).

206. V. Escott-Price, J. Hardy, Genome-wide association studies for Alzheimer's disease: bigger is not always better. *Brain Commun*. 4, fcac125 (2022).

207. H. Braak, E. Braak, Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol*. 82, 239–259 (1991).

208. M. V. Sofroniew, H. V. Vinters, Astrocytes: biology and pathology. *Acta Neuropathol*. 119, 7–35 (2010).