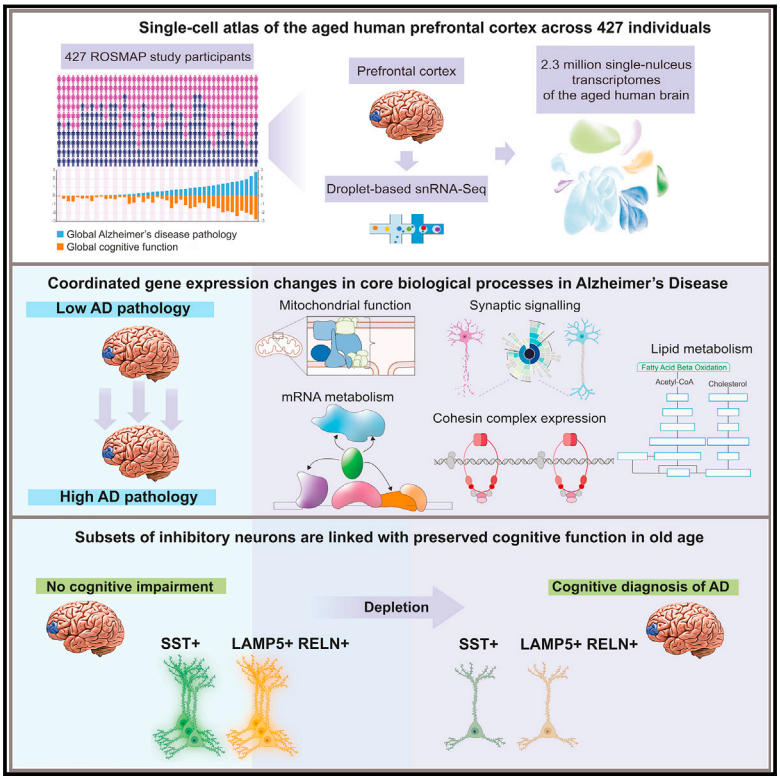
Cell

Article

Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology

Graphical abstract



Highlights

- Single-cell atlas of the aged human prefrontal cortex across 427 individuals
- Coordinated increase of the cohesin complex and DNA damage response factors in AD
- Selectively vulnerable somatostatin inhibitory neuron subtypes are depleted in AD
- Subsets of inhibitory neurons are linked with preserved cognitive function in old age

Authors

Hansruedi Mathys, Zhuyu Peng, Carles A. Boix, ..., David A. Bennett, Manolis Kellis, Li-Huei Tsai

Correspondence

mathysh@pitt.edu (H.M.), manoli@mit.edu (M.K.), lhtsai@mit.edu (L.-H.T.)

In brief

A single-cell atlas of the aged human prefrontal cortex across 2.3 million nuclei isolated from 427 ROSMAP study participants reveals cellular and molecular correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology.

Mathys et al., 2023, Cell *186*, 4365–4385
September 28, 2023
https://doi.org/10.1016/j.cell.2023.08.039

CellPress

Cell

CellPress
OPEN ACCESS

Article

Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology

Hansruedi Mathys,[1,2,3,7,8,*] Zhuyu Peng,[1,2,8] Carles A. Boix,[4,5,8] Matheus B. Victor,[1,2] Noelle Leary,[1,2] Sudhagar Babu,[3] Ghada Abdelhady,[3] Xueqiao Jiang,[1,2] Ayesha P. Ng,[1,2] Kimia Ghafari,[3] Alexander K. Kunisky,[3] Julio Mantero,[4,5] Kyriaki Galani,[4,5] Vanshika N. Lohia,[3] Gabrielle E. Fortier,[3] Yasmine Lotfi,[3] Jason Ivey,[3] Hannah P. Brown,[3] Pratham R. Patel,[3] Nehal Chakraborty,[3] Jacob I. Beaudway,[3] Elizabeth J. Imhoff,[3] Cameron F. Keeler,[3] Maren M. McChesney,[3] Haishal H. Patel,[3] Sahil P. Patel,[3] Megan T. Thai,[3] David A. Bennett,[6] Manolis Kellis,[4,5,7,*] and Li-Huei Tsai[1,2,5,7,9,*]

[1]Picower Institute for Learning and Memory, MIT, Cambridge, MA 02139, USA
[2]Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139, USA
[3]University of Pittsburgh Brain Institute and Department of Neurobiology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA
[4]Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA 02139, USA
[5]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[6]Rush Alzheimer's Disease Center, Chicago, IL 60612, USA
[7]Senior author
[8]These authors contributed equally
[9]Lead contact
*Correspondence: mathysh@pitt.edu (H.M.), manoli@mit.edu (M.K.), lhtsai@mit.edu (L.-H.T.)
https://doi.org/10.1016/j.cell.2023.08.039

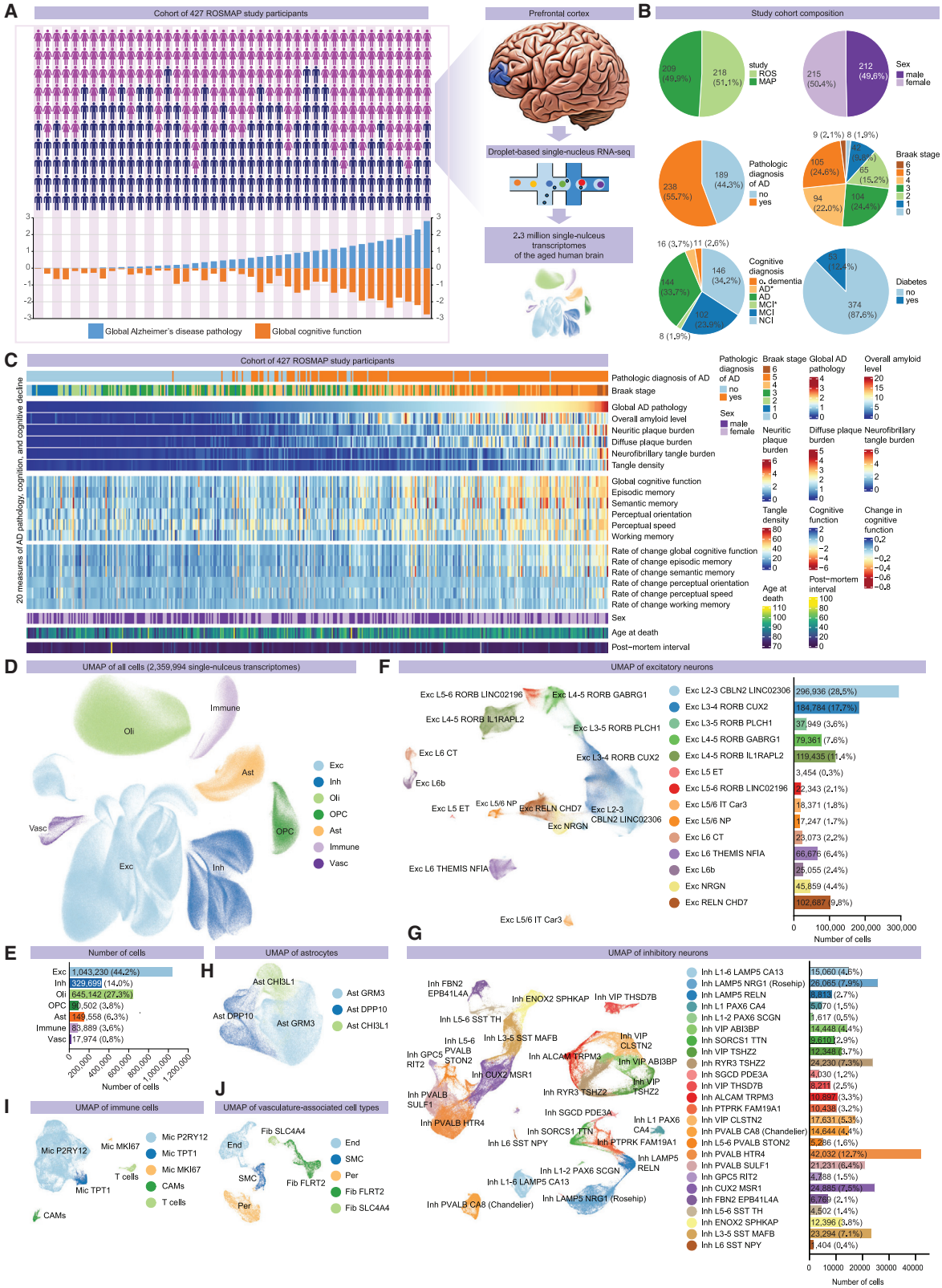
SUMMARY

Alzheimer's disease (AD) is the most common cause of dementia worldwide, but the molecular and cellular mechanisms underlying cognitive impairment remain poorly understood. To address this, we generated a single-cell transcriptomic atlas of the aged human prefrontal cortex covering 2.3 million cells from postmortem human brain samples of 427 individuals with varying degrees of AD pathology and cognitive impairment. Our analyses identified AD-pathology-associated alterations shared between excitatory neuron subtypes, revealed a coordinated increase of the cohesin complex and DNA damage response factors in excitatory neurons and in oligodendrocytes, and uncovered genes and pathways associated with high cognitive function, dementia, and resilience to AD pathology. Furthermore, we identified selectively vulnerable somatostatin inhibitory neuron subtypes depleted in AD, discovered two distinct groups of inhibitory neurons that were more abundant in individuals with preserved high cognitive function late in life, and uncovered a link between inhibitory neurons and resilience to AD pathology.

INTRODUCTION

Alzheimer's disease (AD) is the most common form of dementia.[1] It is a slowly progressing neurodegenerative disorder, starting with mild memory loss and culminating in severe impairment of broad executive and cognitive functions.[2,3] The pathological hallmarks of AD encompass a range of clinical presentations and cellular pathologies, including β-amyloid accumulation, paired helical filament (PHF) tau neurofibrillary tangles, glial cell activation, and neuronal and synaptic loss.[4–9]

Single-cell atlases have proven to be very informative for unraveling the biology of aging[10] and the etiology of diseases like COVID-19[11] by permitting a systematic investigation of how different cell types are impacted. For AD, however, single-cell studies have only been conducted on small groups of 10–21 individuals and thus have not captured the full spectrum of distinct types of AD pathology, cognitive decline, and end-stage cognition observed in the general older population.[12–16] Analyses of the aged human brain across larger unbiased cohorts are needed to capture this variability and answer central questions such as: What distinguishes individuals who age in a healthy way and do not show signs of cognitive impairment from individuals who develop dementia? Which cell types are particularly vulnerable to AD pathology burden, and can their loss explain cognitive impairment? Which genes and pathways are particularly active in individuals who maintain high cognitive function late in life, sometimes despite high levels of Alzheimer's disease pathology?[17,18] Understanding the biology behind cognitive decline versus preserved cognitive function late in life is crucial because it may lead to opportunities for therapeutic development.



(legend on next page)

Here, we address this challenge by reporting the insights gained from a comprehensive single-cell molecular atlas of the aged human prefrontal cortex across 2.3 million nuclei isolated from 427 study participants of two longitudinal cohort studies of aging and dementia (ROSMAP).[19] We capitalize on ROSMAP's rich clinical data and detailed *postmortem* pathological evaluations to determine cellular and molecular correlates of high cognitive function, dementia, and resilience to AD pathology. We report transcriptomic differences associated with multiple measures of AD pathology, identify AD-associated alterations conserved across excitatory neuron subtypes, and uncover a coordinated increase of the cohesin complex and the cellular response to DNA damage in excitatory neurons and in oligodendrocytes. Tracking compositional differences in neuronal subtypes across the spectrum of AD pathology and cognitive impairment, we identify selectively vulnerable inhibitory neuron subtypes depleted in AD as well as two distinct groups of inhibitory neurons that are more abundant in individuals with preserved high cognitive function late in life and uncover a link between inhibitory neurons and resilience to AD pathology.

RESULTS

A 2.3-million-nuclei transcriptional atlas of the aged human prefrontal cortex

We used single-nucleus RNA-sequencing (snRNA-seq) to profile more than 2.3 million nuclei isolated from the prefrontal cortex of *postmortem* brain tissue of 427 participants in the Religious Order Study (ROS) or the Rush Memory and Aging Project (MAP) (Figures 1A and 1B). The sample of 427 study participants included individuals across a range of AD progression, from non-AD and early-AD (N = 115; Braak Stages 0, I, and II) to intermediate- (N = 198; Stages III–IV) and late-AD (N = 114; Stages V–VI) (Figures 1A and 1B). Based on the clinical consensus diagnosis of cognitive status at the time of death, 146 study participants were diagnosed with no cognitive impairment, 102 with mild cognitive impairment, and 144 with AD dementia (Figure 1B). The distribution of categorical and quantitative measures of AD and other pathologies, as well as quantitative clinical phenotypes, in the cohort of 427 study participants (Figure 1C) closely reflected that of the entire ROSMAP cohort (Data S1, pages 1A–1Q).

After removal of doublets (using Scrublet) and low-quality cells, a total of 2,359,994 cells remained in the final dataset (Figure 1D). Based on Leiden clustering, marker gene analysis, and comparisons with previously published data,[20] we annotated 54 high-resolution cell types in 12 major groups (Figure S1; Data S1, pages 1R and 1S). These included 14 excitatory neuron subtypes (Exc, 1,043,230 cells, 44.2% of total); 25 inhibitory neuron subtypes (Inh, 329,699, 14.0%); oligodendrocytes (645,142 cells, 27.5%); oligodendrocyte precursor cells (OPCs, 90,502, 3.8%); 3 astrocyte subtypes (149,558, 6.3%); 5 immune cell types, including microglia, central nervous system (CNS)-associated macrophages (CAMs), and T cells (83,889, 3.6%); and several vascular cell types (17,974, 0.8%), including endothelial (End), smooth muscle cells (SMCs), fibroblast (Fib), and pericyte (Per) subpopulations (Figures 1F–1J) (see Sun et al.[21] in this issue for a more detailed analysis of microglial cellular states). Transcriptomic cell types were largely conserved across individuals, as most of the identified cellular subsets contained nuclei from the large majority of donors (Data S1, page 1R and 1S).

Systematic differential analysis of gene expression

To determine how individual cell types are affected by different clinical and lifestyle variables, we analyzed the differential gene expression in our 54 cell types with regard to each variable. We considered multiple measures of AD pathology, Lewy body and TDP-43 pathology, vascular pathology, medical conditions, and cognitive, physical, and social lifestyle variables (Figure 2A). After correcting for the level of global AD pathology, few genes were significantly associated with non-AD-pathology variables in any cell type (Figure 2A). One notable exception was diabetes, which exhibited a remarkably cell-type-specific association with gene expression alterations in oligodendrocytes (Figure 2A; Data S1, page 2). We also observed sex-specific responses to AD pathology (Data S1, page 2H), consistent with previous observations.[22]

We subsequently focused our analysis on the AD pathology variables because of the large number of genes associated with these measures (Figure 2A). Differentially expressed genes (DEGs; adjusted p value < 0.05) significantly associated with AD pathology were identified in nearly all cell types (Figure 2A), with the largest numbers of DEGs detected for measures of neuritic plaque and neurofibrillary tangle burden (Figure 2A). When comparing DEGs associated with different types of AD pathology, we observed a significant overlap of genes positively or negatively associated with global AD pathology, neuritic plaque burden, neurofibrillary tangle burden, and tangle density in many neuronal subtypes, as well as in astrocytes and oligodendrocytes (Figures 2B–2F and S2A), and the effect sizes of gene expression changes were strongly correlated between variables (Figures S2B and S2C). Thus, our results highlight large AD-pathology-associated consensus signatures that are reproducibly

Figure 1. Single-nucleus RNA-seq interrogation of prefrontal cortex tissue from 427 ROSMAP study participants

(A) Cohort and snRNA-seq profiling summary, covering PFC samples from 427 ROSMAP study participants along a spectrum of AD progression.

(B) Study cohort composition.

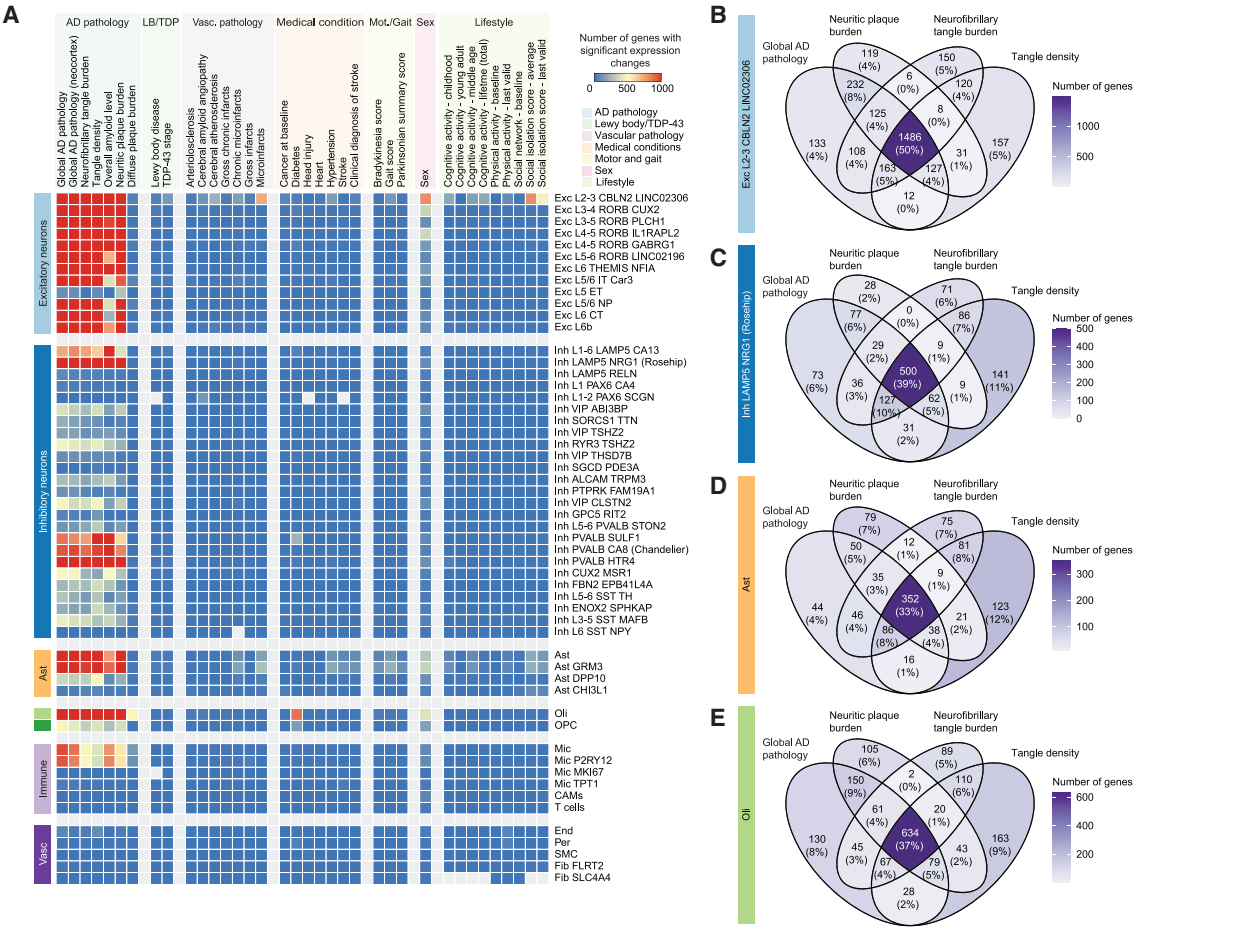
(C) Cohort metadata overview. The heatmap shows measures of AD pathology, cognitive function, sex, age at death, and *postmortem* interval (rows) for the 427 study participants (columns), ordered by global AD pathology.

(D) Joint UMAP of 2.3 million cells across 7 major cell classes including excitatory neurons (Exc), inhibitory neurons (Inh), oligodendrocytes (Oli), oligodendrocyte precursor cells (OPC), astrocytes (Ast), immune cells (Immune), and vascular and epithelial cells (Vasc).

(E) Barplot showing the number of cells per major cell class.

(F–J) UMAPs of excitatory neuron subtypes (F), inhibitory neuron subtypes (G), astrocyte subtypes (H), immune cell subsets (I), and vascular and epithelial cell subtypes (J). Barplots show the number of cells for excitatory (F) and inhibitory (G) neuronal subtypes.

See also Figure S1 and Data S1, page 1.



(legend on next page)

detectable across multiple distinct measures of AD pathology within many different cell types. We also found a significant overlap of DEGs associated with AD pathology between different excitatory neuron subtypes (Figures 2G, 2H, and S2D–S2G). The overlap of AD-associated genes between excitatory neuron subtypes was also indicated by self-organizing map (SOM) analyses of gene-trait correlations (Methods S1, page 1A), broadly observed across the expression level spectrum (Data S1, page 3A), and further confirmed using a permutation test (Data S1, pages 3B and 3C).

To validate our differential expression results, we compared them to those of previously published snRNA-seq studies of Alzheimer's disease[14,15,23,24] and found a significant positive correlation for gene expression changes across multiple cell types, with particularly high concordance between our results and those of other studies analyzing the prefrontal cortex (PFC; Data S1, pages 3D–3G). To further test the robustness and generalizability of our results, we divided the full cohort into two non-overlapping random subsets and found a highly significant overlap of DEGs between the results of the original analysis across the entire cohort and both random sub-cohorts (Data S1, page 4A). Furthermore, the association scores and effect sizes of significant DEGs were strongly correlated between the full cohort and both sub-cohorts (Data S1, pages 4B and 4C). We also confirmed that the observed associations between gene expression and global AD pathology remained significant after accounting for additional potential confounding factors such as technical variability, sex, and diabetes and after removing ambient RNA using DecontX[25] (Data S1, pages 4D–4F).

The results of the differential expression analysis, both from the original analysis and the analysis involving the removal of ambient RNA and adjustment for sex and unwanted sources of variation, are available on GitHub (https://github.com/mathyslab7/ROSMAP_snRNAseq_PFC).

To further validate the results of our differential gene expression analysis, we analyzed a separate large snRNA-seq dataset derived from the dorsolateral prefrontal cortex (DLPFC) of 424 ROSMAP study participants (referred to as the De Jager dataset).[26,27] A permutation test revealed a highly significant overlap of global AD-pathology-associated DEGs between the two datasets (Data S1, page 5A). We extended our validation to a different brain region using a snRNA-seq dataset derived from the middle temporal gyrus (MTG) (referred to as the SEA-AD dataset).[28] Interestingly, we observed that a large fraction of genes confirmed to be altered in the prefrontal cortex by both our dataset and the De Jager dataset were also significantly associated with the continuous pseudo-progression score of AD severity in the MTG dataset (Data S1, pages 5B and 5C). This compelling agreement across different cohorts and brain regions serves as a testament to the robustness and generalizability of our findings.

AD-pathology-associated changes conserved across excitatory neuron subtypes

To identify the genes, pathways, and biological processes recurrently altered across multiple excitatory neuron subtypes, we defined a consensus signature of genes significantly associated with global AD pathology in at least three different excitatory neuron subtypes. The consensus signature positively associated with global AD pathology was enriched for genes involved in neuron projection morphogenesis, mRNA metabolic processes, chromatin organization, and synaptic signaling (Figure 3A). Consensus signature genes associated with synaptic signaling were further examined in the context of synaptic location using the SynGO database,[29] and the results suggested that gene expression alterations in AD involve both the pre-and postsynaptic compartments (Figures S3A–S3E). We confirmed that many of the synaptic genes positively associated with AD pathology were also upregulated in a published dataset of gene expression from neurofibrillary tangle (NFT)-bearing neurons compared to NFT-free somas (Figure S3F).[30] However, analysis of expression changes based on quantitative proteomics revealed a decrease of many of the synaptic genes in AD brains, potentially suggesting that the greater expression of these genes in neuronal nuclei reflects a homeostatic response, possibly as a reaction to synaptic degeneration in the neuronal periphery (Figure S3G).[31] Consistent with this hypothesis, genes negatively associated with multiple measures of AD pathology across multiple excitatory neuron subtypes were confirmed to be decreased in AD brains based on quantitative proteomics (Figures S3H–S3K),[31] a finding that was further supported by cell-type-specific reverse transcription quantitative PCR (RT-qPCR) (Data S1, page 6A) and multiplex RNA *in situ* hybridization (Data S1, pages 7A–7C). Other consensus signature genes positively associated with global AD pathology annotated to be involved in mRNA metabolic processes centered around the spliceosome and included the heterogeneous nuclear ribonucleoproteins HNRNPA2B1 and HNRNPH3 (Figure 3C). A significant upregulation of HNRNPA2B1 in individuals with a pathologic diagnosis of AD was corroborated using RT-qPCR (Data S1, pages 6B and 6C) and multiplex RNA *in situ* hybridization (Data S1, pages 7D and 7E).

Figure 2. Gene expression changes associated with AD pathology are shared across excitatory neuron subtypes

(A) Heatmap showing the number of significantly differentially expressed genes (adjusted p value < 0.05) across 54 cell types (rows) and 36 variables (columns).

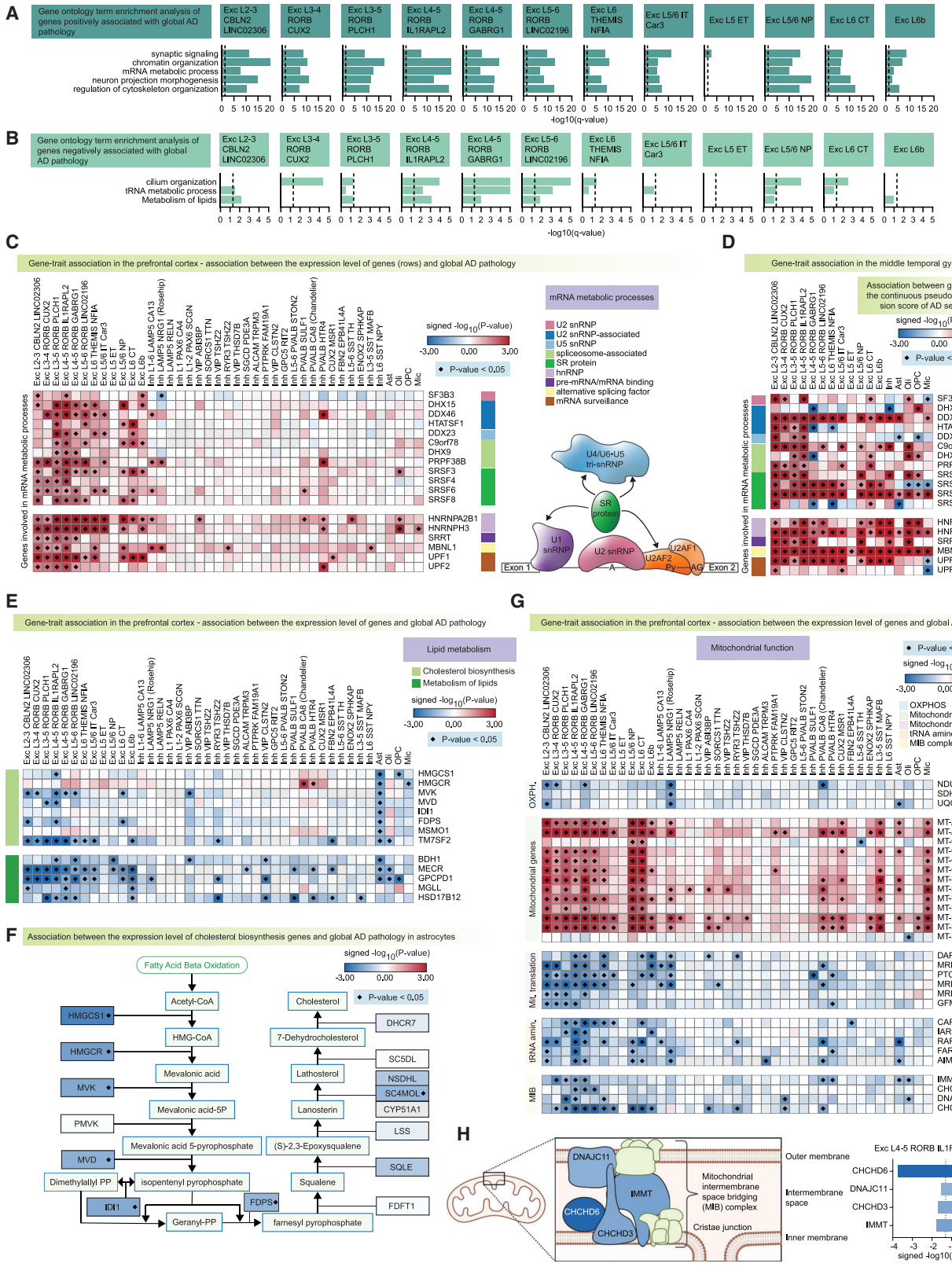
(B–E) Venn diagrams showing the number of genes significantly positively associated with global AD pathology, neuritic plaque burden, NFT burden, and tangle density in the excitatory neuron subtype Exc L2–3 CBLN2 LINC02306 (B), the inhibitory neuron subtype Inh LAMP5 NRG1 (Rosehip) (C), astrocytes (D), and oligodendrocytes (E).

(F) Heatmaps showing the overlap (odds ratio [OR], one-sided Fisher's exact test) of genes significantly positively associated with the variables indicated. **−log10(p value) > 200, *−log10(p value) > 100 (one-sided Fisher's exact test).

(G) Circos plot showing the overlap of genes positively associated with global AD pathology between excitatory neuron subtypes.

(H) Heatmaps showing the overlap (OR, one-sided Fisher's exact test) of genes significantly positively (left panel) or negatively (right panel) associated with global AD pathology between the excitatory neuron subtypes indicated. **−log10(p value) > 200, *−log10(p value) > 100 (one-sided Fisher's exact test).

See also Figure S2; Data S1, pages 2–5; and Methods S1, pages 1, 6, and 7.



(legend on next page)

The consensus signature negatively associated with global AD pathology was enriched for genes involved in tRNA metabolic processes and metabolism of lipids (Figure 3B). Genes negatively associated with multiple measures of AD pathology across multiple cell types included the lipid metabolism genes BDH1, GPCPD1, MECR, and TM7SF2, which is involved in cholesterol biosynthesis (Figure 3E). Consistent with these results, RT-qPCR on neuronal nuclei confirmed a significant downregulation of BDH1 and GPCPD1 in individuals with a pathologic diagnosis of AD (Data S1, pages 6D and 6E). While the consensus signature is based on gene expression differences in excitatory neurons, dysregulation of lipid metabolism and cholesterol biosynthesis genes was also evident in astrocytes (Figure 3F), an observation that was further supported by immunohistochemical analysis of *postmortem* prefrontal cortex tissue (Data S1, pages 8A–8C). Alterations in lipid metabolism were accompanied by widespread changes of genes encoding the mitochondrial electron transport chain components, and individuals with a high level of AD pathology exhibited a decreased expression of several genes encoding components of the mitochondrial intermembrane space bridging complex (MIB), a complex that is of central importance for the maintenance of mitochondrial inner membrane architecture (Figures 3G and 3H; Data S1, page 8D).[32] Our findings collectively point to several metabolic processes and facets of mitochondrial biology that may be compromised in AD and may contribute to cellular dysfunction in astrocytes and multiple neuronal subtypes.

The gene expression differences related to synaptic signaling, RNA splicing, lipid metabolism, and mitochondrial function observed in the PFC were largely confirmed in a separate dataset derived from the middle temporal gyrus[28] (Figure 3D; Data S1, pages 6F–6I) and were further supported by bulk RNA-sequencing data from the ROSMAP cohorts (n = 638), quantitative proteomics data[31] (Data S1, pages 6J–6O), and a group-based differential expression analysis (Methods S1, pages 2 and 3).

We performed an additional analysis focused on identifying AD-associated gene expression alterations that exhibit more cell-type specificity. We defined cell-type-restricted DEGs as genes significantly altered in no more than two cell types and found that a considerable fraction of DEGs fell within this category (Methods S1, pages 4A–4C). We successfully identified cell-type-restricted AD-associated gene expression changes not only at the level of major cell types but also among excitatory subtypes and inhibitory neuron subclasses (Methods S1, pages 4D–4F) and confirmed a large fraction of these alterations in the SEA-AD dataset (Methods S1, pages 4G and 4H).

Coordinated increase of cohesin complex expression and DNA damage response in AD

To evaluate whether AD affects additional protein complexes besides the MIB complex, we computed module scores for 2,417 human protein complexes from CORUM[33] and tested their association with global AD pathology. Cohesin, a protein complex that regulates sister chromatid cohesion, homologous recombination, and DNA looping, was most consistently and significantly elevated across excitatory neuron subtypes from the PFCs of individuals with high levels of AD pathology (Figures 4A and S4A–S4C) and was also positively associated with global AD pathology in PFC oligodendrocytes (Figure 4A). Increased expression of the cohesin complex in individuals with high levels of AD pathology was also observed in a subset of excitatory neuron subtypes in other brain regions (Figure 4A) and further confirmed based on bulk RNA-sequencing data from the ROSMAP cohorts (n = 638) (Figure 4B), RT-qPCR analysis (Figure S4D), and quantitative proteomics data (Figure 4C).[31]

To explore the context and possible function of increased cohesin expression in AD, we computed the correlation between the cohesin complex and all expressed genes across excitatory neuron cell types. Interestingly, many genes positively correlated with the cohesin complex, including the chromatin regulators ATRX, ZBTB1, SIRT1, CHD1, BPTF, and ZBTB7A, have been reported to play a critical role in the cellular response to DNA damage (Figures 4D and S4E).[34–42] Other genes positively correlated with the cohesin complex are recruited to sites of DNA damage (NIPBL),[43] regulate base excision repair (USP47),[44] promote DNA damage recovery (BAZ1B),[45] regulate the DNA damage response (CDKN2AIP),[46] and play a critical role in DNA damage repair (MACROD1)[47] (Figures 4D and S4E). Many of these genes were themselves positively associated with multiple measures of AD pathology across multiple excitatory neuron subtypes and were also positively correlated with cohesin complex expression in excitatory neurons in other brain regions (Figures S4F–S4H; Data S1, pages 9A–9C). A coordinated increase in the expression of the cohesin complex and DNA damage response genes in AD was also observed in oligodendrocytes (Figures 4E–4G; Data S1, page 9D). These results indicate that the AD-associated increase in expression of the cohesin complex is accompanied by elevated expression of a plethora of chromatin regulators

Figure 3. Gene expression changes associated with AD pathology

(A and B) Selected gene ontology (GO) terms enriched among genes significantly positively (A) or negatively (B) associated with global AD pathology.

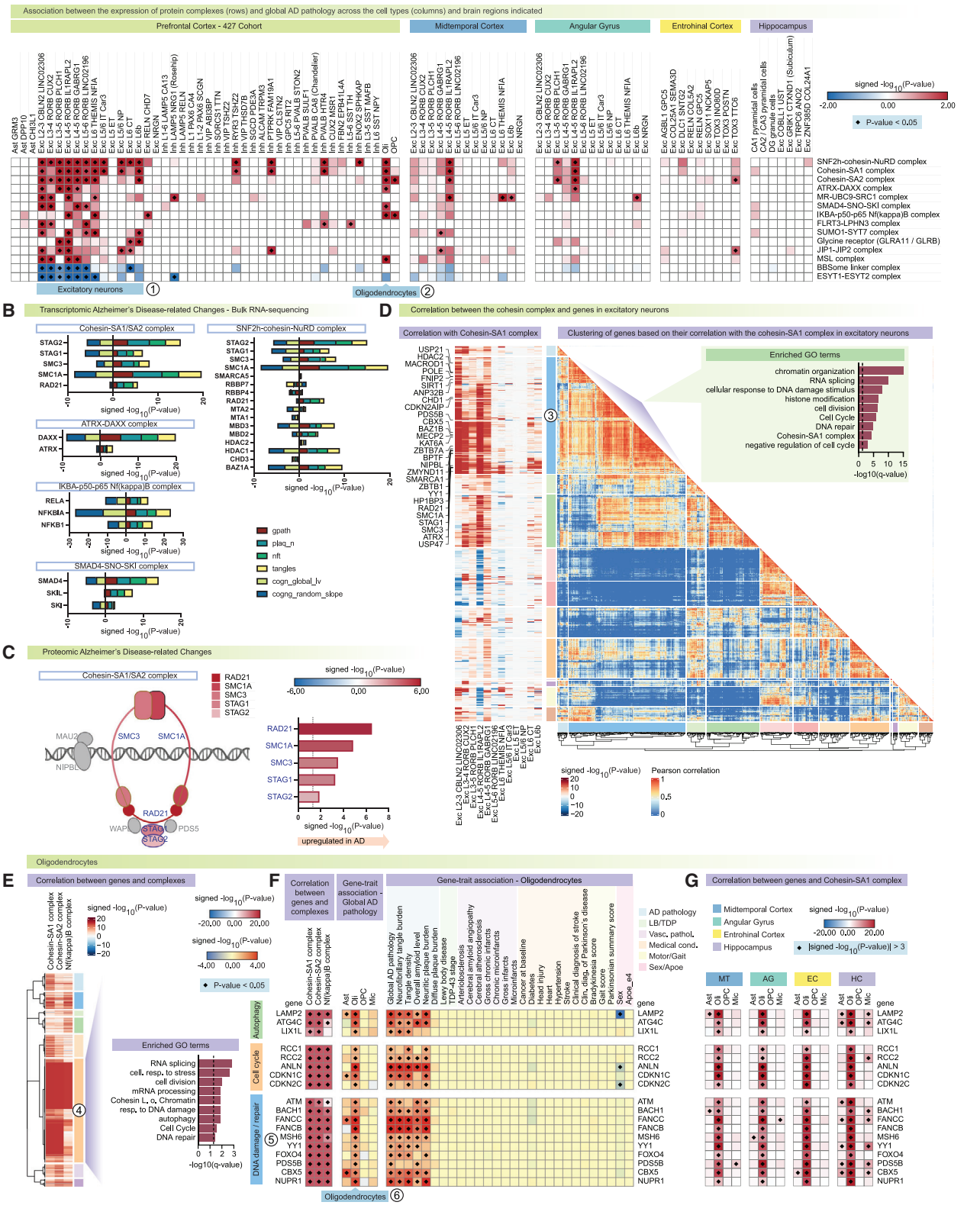
(C, E, and G) Association between the expression level of genes (rows) and global AD pathology across the cell types indicated (columns). Differential gene expression analysis results for genes involved in mRNA metabolic processes (C), lipid metabolism (E), and mitochondrial function (G) are displayed. The heatmap shows association scores (signed −log10 false discovery rate [FDR]-adjusted p value, where the sign was determined by the direction [positive or negative] of the association), and significant scores (FDR-adjusted p value <0.05) are indicated by a rhombus shape.

(D) Confirmation of differential gene expression (DGE) results in a separate dataset derived from the middle temporal gyrus. The heatmap shows association scores (signed −log10 FDR-adjusted p value, where the sign was determined by the direction [positive or negative] of the association), and significant scores (FDR-adjusted p value < 0.05) are indicated by a rhombus shape.

(F) Negative association between the expression level of cholesterol biosynthesis genes and global AD pathology in astrocytes (boxes are colored by the association scores).

(H) Association between the expression level of MIB complex genes and global AD pathology in layer 4/5 excitatory neurons (Exc L4–5 RORB IL1RAPL2). The bar plot shows association scores.

See also Figure S3; Data S1, pages 6–8; and Methods S1, pages 2–4.



(legend on next page)

and DNA damage response factors that promote genome integrity, an observation that is consistent with the results presented in the accompanying manuscripts by Dileep and colleagues[48] and Xiong and colleagues.[49] Collectively, our findings indicate that the deterioration of the genome and epigenome integrity may constitute important pathological events in the pathogenesis of AD.

Temporal changes in gene expression across multiple stages of Alzheimer's disease progression

To deepen our analysis of molecular changes associated with AD progression, we split the entire cohort into three groups based on the AD pathology variable of interest and determined DEGs comparing the second group to the first group of individuals with the lowest level of AD pathology and the third group with the highest level of AD pathology to the second group (Figure 5A). We observed a convergence in gene expression changes associated with different types of AD pathology as well as a significant overlap of DEGs between excitatory neuron subtypes at both the early and late stages of disease progression (Data S1, pages 10A–10C), consistent with the results based on the entire cohort of individuals.

To examine AD progression at the molecular level, we intersected genes linked to AD pathology across the entire cohort with the DEGs determined based on the disease-stage-specific differential expression analysis (Figure 5A). Many DEGs identified across the cohort also showed significant changes at early or late disease stages (Figure S5A). Across multiple cell types, early changes positively associated with AD pathology were enriched for genes involved in synapse organization, regulation of cell projection organization, chromatin organization, and DNA damage response (Figures 5B and S5B), whereas genes involved in cell activation were enriched only in microglia (Figure S5C). Genes involved in autophagy were enriched among early changes negatively associated with AD pathology in astrocytes, excitatory and inhibitory neurons, oligodendrocytes, and OPCs (Figures 5C and S5E). Late changes positively associated with AD pathology were enriched for genes involved in mRNA metabolic processes, chromatin organization, DNA damage response, cell projection organization, synaptic signaling, and protein catabolic processes in neurons (Figures 5B and S5B); cholesterol metabolism in oligodendrocytes; and cell activation, regulation of defense response, and the TYROBP causal network in microglia (Figures S5C and S5D). Conversely, genes negatively associated with AD pathology at a late stage of disease progression were enriched for lipid metabolism genes in astrocytes, excitatory and inhibitory neurons, and oligodendrocytes (Figures 5C and S5F).

We also performed a similar analysis in the SEA-AD dataset, splitting the cohort into three groups based on the continuous pseudo-progression score of each individual.[28] We then looked at a subset of genes that were significant across multiple cell types in both PFC and MTG and found that many of them were significantly associated with global AD pathology across multiple neuronal and glial cell types at the late stage of disease progression (Figure 5E). Consistent with these observations, we found that the cohesin complex expression tended to increase with AD progression and was significantly elevated in the individuals with the highest level of global AD pathology (Figure 5D). Finally, we explored disease progression in AD by comparing gene expression alterations between the PFC and the MTG datasets. Early changes considerably overlapped between the datasets (Figure 5G), and intriguingly, a sizable fraction of the genes that were significant at a late stage of disease progression in the PFC were also significantly altered at an early stage in the MTG (Figures 5E and 5F). These observations indicate that gene expression alterations in Alzheimer's disease are at least partially shared between different brain regions and may indicate that similar molecular alterations happen with different temporal patterns in different regions of the brain.

Molecular correlates of cognitive impairment in the aged human brain

Next, we explore what distinguishes individuals who age in a healthy way and do not show signs of cognitive impairment from those who develop dementia. To determine the molecular correlates of cognitive impairment in the aged human brain, we analyzed differential gene expression in individual cell types with regard to a summary measure of global cognitive function as well as measures of individual cognitive domains and the

Figure 4. Coordinated elevation of cohesin complex expression and DNA damage response in AD

(A) Association (association scores) between the expression of 14 protein complexes (rows) and global AD pathology across the cell types (columns) and brain regions indicated.

(B and C) Confirmation of increased cohesin complex expression in AD based on (B) bulk RNA-sequencing data from the dorsal lateral prefrontal cortex of 638 ROSMAP study participants and (C) quantitative proteomics.[31] Bar plots show association scores.

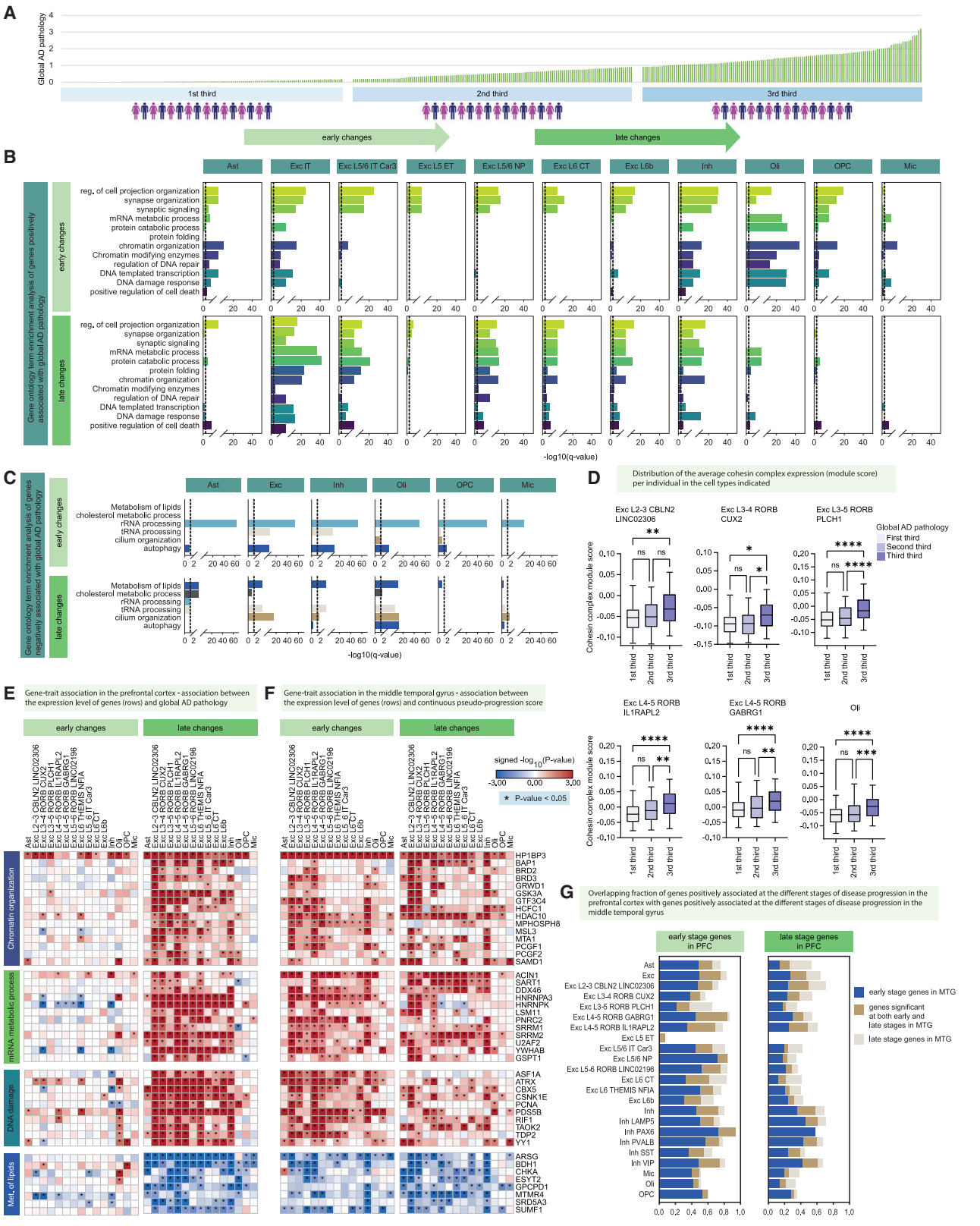
(D) Left: Correlation (association scores) between the cohesin complex and consensus signature genes positively associated with global AD pathology across excitatory neuron subtypes. Right: k-means clustering of genes based on the Pearson correlation of the association scores across excitatory neurons shown in the left panel. (3) Module of genes exhibiting a positive correlation with the cohesin complex across multiple excitatory neuron subtypes.

(E) Correlation (association scores) between genes positively associated with global AD pathology (in oligodendrocytes) and the protein complexes indicated (columns) in oligodendrocytes. (4) Module of genes positively correlated with the cohesin-SA1/SA2 complex in oligodendrocytes.

(F) Left: Correlation between the expression level of the protein complexes (columns) and genes (rows) indicated in oligodendrocytes. (5) Genes involved in cellular response to DNA damage and DNA repair are positively correlated with cohesin complex expression in oligodendrocytes. Middle: Association between global AD pathology and the expression level of genes co-regulated with the cohesin complex (rows) across the cell types indicated (columns). (6) DNA damage response genes are positively associated with global AD pathology in oligodendrocytes. Right: Association between the expression level of genes co-regulated with the cohesin complex (rows) and the variables indicated (columns) in oligodendrocytes.

(G) Correlation between the expression level of the cohesin-SA1 complex and the genes indicated (rows) across four glial cell types and four distinct regions of the human brain.

See also Figure S4 and Data S1, page 9.



(legend on next page)

person-specific rate of change in these variables over time. DEGs significantly associated with cognition were identified in nearly all cell types (Figure 6A), with the largest number of DEGs associated with measures of global cognitive function and episodic memory. Genes associated with distinct cognition variables overlapped significantly in most excitatory neuron cell types, a subset of inhibitory neuron subtypes, astrocytes, and oligodendrocytes. These genes also overlapped significantly with AD-pathology-associated genes, suggesting that within a given cell type, the gene expression alterations associated with cognitive impairment closely reflect AD-pathology-associated changes (Figures 6B and S6A). Accordingly, DEGs associated with global cognitive function significantly overlapped between different excitatory neuron subtypes (Figure S6B).

Just as we had done for genes associated with AD pathology, we also determined a consensus signature of genes significantly associated with global cognitive function in at least three different excitatory neuron subtypes. The consensus signature positively associated with global cognitive function was enriched for genes involved in lipid metabolism, whereas the negatively associated consensus signature included genes involved in synaptic signaling and neuron projection development (Figures S6C and S6D). Genes involved in lipid metabolism, particularly cholesterol biosynthesis, were also associated with global cognitive function in astrocytes and oligodendrocytes, with astrocytes showing positive associations while oligodendrocytes exhibited the opposite trend (Figures S6D and S6E). To further characterize synaptic genes significantly associated with global cognitive function, we examined consensus signature genes in the context of synaptic location using the SynGO database.[29] Negatively associated genes exhibited the most significant enrichment for postsynaptic terms (Figure S6F), while genes positively associated with global cognitive function, including CACNG3, PAK1, NPTX2, RPH3A, SVOP, and BDNF, mapped to both the pre- and postsynaptic compartments (Figures 6C–6E). Examination of PFC tissue using RNA *in situ* hybridization (Figure 6F) and immunohistochemistry (Figure 6G) confirmed a significant reduction in NPTX2 expression among individuals with cognitive impairment. Additional synaptic genes positively associated with global cognitive function included NUMB and VGF (Figures 6D and 6E), a finding that was further corroborated through RT-qPCR (Data S1, page 11A) and immunohistochemical analysis (Data S1, pages 11B and 11C). Finally, we observed that several genes involved in fibroblast growth factor (FGF) signaling were positively associated with multiple measures of cognition across multiple excitatory neuron subtypes (Figures 6D and 6E). In summary, our results highlight common and cell-type-specific genes, pathways, and cellular components associated with high cognitive function and dementia in the aged human brain.

Molecular correlates of cognitive resilience to AD pathology

Cognitive resilience (CR) has been defined as the continuum of better (or worse) than expected cognition, given the degree of neuropathology.[18,50,51] To identify potential molecular mediators that confer resilience to AD pathology, we next investigated what distinguishes the brains of individuals with high CR from individuals who are more susceptible to AD pathology. To determine the genes linked to CR in the aged human brain, we analyzed differential gene expression in excitatory neuron subtypes with regard to the CR and cognitive decline resilience (CDR) measures (see STAR Methods; Methods S1, pages 5A and 5B). DEGs significantly associated with CR were identified in multiple excitatory neuron subtypes (Methods S1, page 5C). We identified 5 genes (HES4, PDE10A, RPH3A, ST6GAL2, and UST) that were differentially expressed relative to at least one CR measure in at least 8 excitatory neuron cell types (Methods S1, pages 5D and 5E). Interestingly, UST is a predicted target of the MEF2 family of transcription factors, which we have recently identified as key regulators of CR.[52] These genes were also robustly differentially expressed relative to all the CR measures in at least 1 excitatory cell type (Methods S1, pages 5D and 5E). Greater expression of these genes in resilient individuals was further confirmed based on an overrepresentation analysis (Methods S1, pages 5F–5K), bulk RNA-sequencing data from the ROSMAP cohorts (Methods S1, pages 5L–5O), RT-qPCR analysis of RNA isolated from PFC *postmortem* tissue (Methods S1, page 5P), and a group-based differential expression analysis comparing individuals with both pathologic and cognitive diagnoses of AD versus individuals with no cognitive impairment despite a pathologic diagnosis of AD (Methods S1, page 5Q). In conclusion, these analyses highlight a small set of genes whose expression is linked to cognitive resilience in multiple excitatory neuron subtypes.

Neuronal and glial compositional changes with the accumulation of AD pathology

Cell death is the ultimate form of cellular dysfunction and is a hallmark of neurodegenerative disorders, including Alzheimer's

Figure 5. Temporal changes in gene expression across multiple stages of Alzheimer's disease progression

(A) Schematic illustration of differential expression analysis approach.

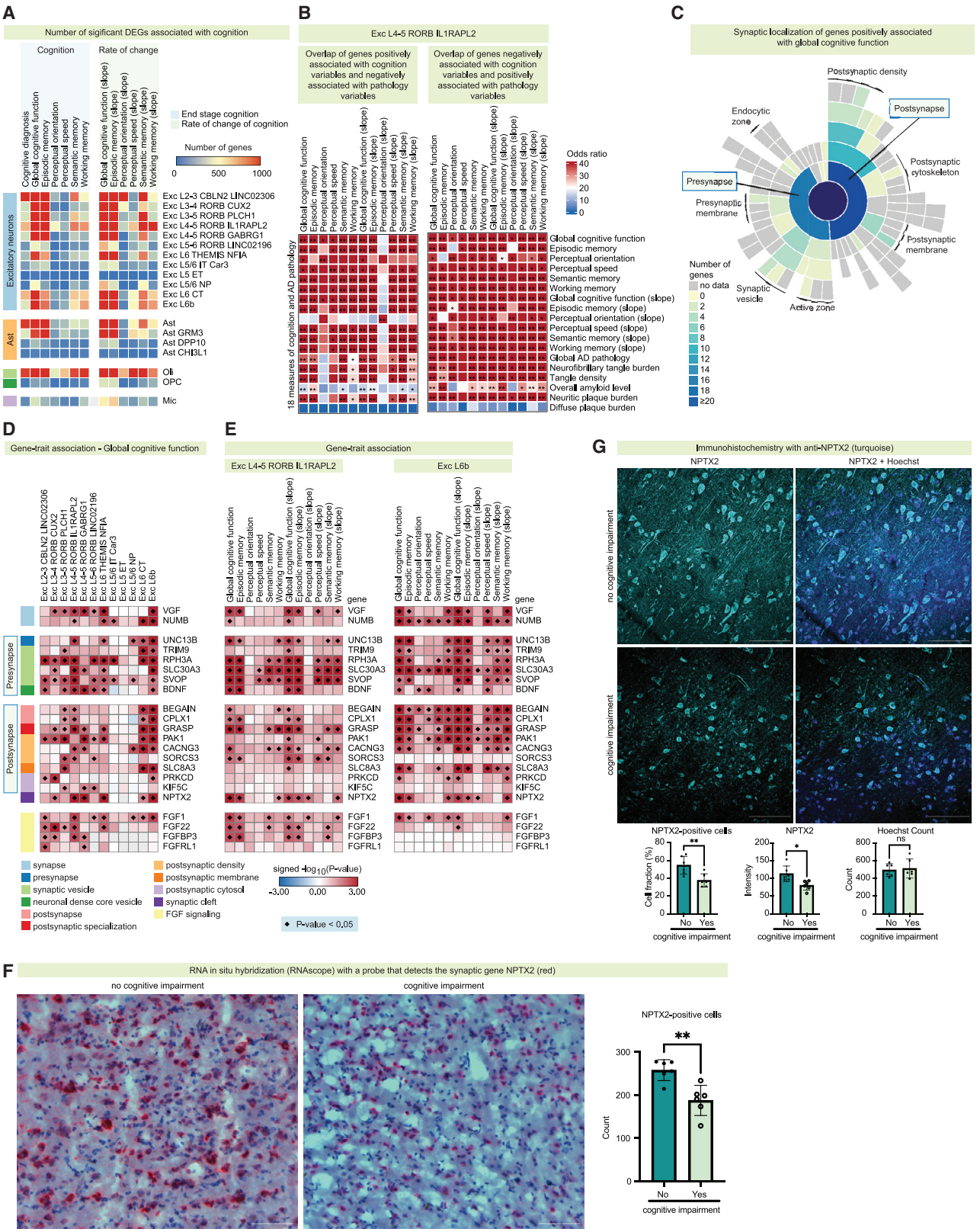
(B and C) Selected gene ontology (GO) terms (rows) enriched among genes significantly (B) positively or (C) negatively associated with global AD pathology across the cell types indicated. Upper panels: early changes (second versus first group), lower panels: late changes (third versus second group).

(D) Boxplots showing the distribution of the average cohesin complex expression (module score) per individual in the cell types indicated. Within each box, horizontal lines denote median values; boxes extend from the 25th to the 75th percentile of each group's distribution of values; whiskers extend from the 5th to the 95th percentile. ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$; ns, $p > 0.05$ (ordinary one-way ANOVA corrected for multiple comparisons using Tukey's multiple comparisons test).

(E and F) Association between the expression level of selected genes involved in the biological processes indicated (rows) and global AD pathology and continuous pseudo-progression score in the prefrontal cortex (E) and middle temporal cortex (F), respectively. Left column: early changes; right column: late changes. Significant scores (FDR-adjusted p value < 0.05) are indicated by an asterisk symbol.

(G) Overlapping fraction of genes positively associated at the different stages of disease progression in the PFC with genes positively associated at the different stages of disease progression in the MTG. Left: early-stage genes in the PFC; right: late-stage changes in the PFC.

See also Figure S5 and Data S1, page 10.



(legend on next page)

disease. We investigated how cell type composition is affected with the progression of AD pathology to identify particularly vulnerable cell types. First, we assessed whether the proportions of major cell types change with the progression of AD pathology and found that the relative abundance of most major cell types did not significantly change (Figures 7A and 7B; Data S1, pages 12A and 12B). One exception was the population of microglial cells that exhibited a significant increase in its relative abundance in individuals with the highest levels of global AD pathology, neuritic plaque burden, and overall amyloid level (Figures 7A and 7B; Data S1, pages 12A–12D). By contrast, excitatory neurons trended toward a subtle decrease in their relative abundance with the progression of AD pathology, reaching significance when considering the neocortex-specific measure of global AD pathology (Figures 7A and 7B; Data S1, page 12B).

To identify particularly vulnerable and resilient cell types, we next evaluated changes in cell type composition with the progression of AD pathology at the level of high-resolution cell types. We identified three selectively vulnerable inhibitory neuron subtypes of the SST subclass (Inh CUX2 MSR1, Inh ENOX2 SPHKAP, and Inh L3–5 SST MAFB) that were significantly overrepresented in the subset of individuals with the lowest levels of global AD pathology, neuritic plaque burden, diffuse plaque burden, overall amyloid level, NFT burden, and tangle density and in individuals without a pathologic diagnosis of AD (Figure 7C). Interestingly, Uniform Manifold Approximation and Projection (UMAP) representation of inhibitory neurons revealed that the three vulnerable SST subtypes clustered next to each other, suggesting that these subtypes are characterized by a high degree of transcriptomic similarity, potentially explaining the shared vulnerability to AD pathology (Figure 7D). Consistent with the results of the overrepresentation analysis, the relative abundance of the entire subclass of SST inhibitory neurons, as well as the relative abundance of individual subtypes of the group of selectively vulnerable SST neurons, was significantly decreased in the subset of individuals with the highest levels of global AD pathology, NFT burden, and tangle density, a finding that was further confirmed using a quasi-binomial regression model (Figures 7E and 7F; Data S1, pages 13A–13D). The relative abundance of many excitatory neuron subtypes trended toward a decrease with the progression of AD pathology but did not reach significance (Data S1, pages 13E and 13F).

To further validate the changes in cell type composition, we integrated our data with the De Jager dataset, generating a snRNA-seq dataset of the human PFC encompassing 615 individuals. Analyzing the combined dataset using quasi-binomial regression, we confirmed a decrease in the relative abundance of the entire subclass of SST inhibitory neurons and all three vulnerable SST neuron subtypes in individuals with higher levels of AD pathology (Data S1, pages 13G and 13H). In summary, our results highlight the subclass of somatostatin inhibitory neurons, and most notably three of its subtypes, as particularly vulnerable to AD pathology.

Differences in cell type composition associated with cognitive impairment

We next utilized our large-scale dataset to examine cell type composition differences between individuals who maintain high cognitive function late in life and those with dementia. The relative abundance of inhibitory neurons tended to be higher in the subset of individuals with the highest level of global cognitive function and was significantly decreased in the subset of individuals with the lowest scores in the cognitive domains of perceptual orientation and semantic memory (Data S1, pages 14A–14D). Likewise, the relative abundance of excitatory neurons was significantly decreased in the subset of individuals with the lowest scores in the cognitive domain of perceptual orientation (Data S1, pages 14B and 14D).

Next, we explored differences in cell type composition associated with cognitive decline at the level of high-resolution cell types. We identified multiple inhibitory neuron subtypes, including the vulnerable SST neurons, that were significantly overrepresented in the subset of individuals with the highest level of global cognitive function, episodic memory, and semantic memory (Figure 7G). Additional inhibitory neuron subtypes overrepresented in individuals with high cognitive function included the RELN-positive inhibitory neurons of the LAMP5 subclass (Inh LAMP5 RELN), the subtype marked by the expression of PTPRK and FAM19A1 (Inh PTPRK FAM19A1), and the subtype marked by the expression of SORCS1 and TTN (Inh SORCS1 TTN) (Figure 7G). UMAP representation of inhibitory neurons

Figure 6. Molecular correlates of cognitive impairment in the aged human brain

(A) Heatmap showing the number of significantly differentially expressed genes (adjusted p value < 0.05).

(B) Heatmaps showing the overlap (OR, one-sided Fisher's exact test) of genes significantly associated with the variables indicated in L4–5 excitatory neurons (Exc L4–5 RORB IL1RAPL2). **−log10(p value) > 200, *−log10(p value) > 100 (one-sided Fisher's exact test).

(C) The sunburst plot showing the synaptic localization of genes positively associated with global cognitive function (number of genes).

(D) Association (association scores) between the expression level of synaptic genes (rows) and global cognitive function across the cell types indicated (columns).

(E) Association (association scores) between the synaptic genes (rows) and variables indicated (columns) in L4–5 excitatory neurons (Exc L4–5 RORB IL1RAPL2).

(F) RNA *in situ* hybridization with an RNAscope probe for NPTX2 (red) in the gray matter of the PFC of an individual without cognitive impairment (left) and an individual with cognitive impairment (right). The tissue was counterstained with hematoxylin. Scale bar: 100 μm. Right: quantification of the number of cells detectably expressing NPTX2. Data are mean ± s.e.m.; **p = 0.0031 (Student's two tailed t test). n = 6 individuals without cognitive impairment and n = 6 individuals with a cognitive diagnosis of AD (cognitive impairment); n = 4 images per individual.

(G) Immunohistochemistry with an anti-NPTX2 (turquoise) antibody in the gray matter of the PFC of an individual without cognitive impairment (top) and an individual with a cognitive diagnosis of AD (cognitive impairment) (bottom) (scale bar: 100μm). DNA (nuclei) was stained with Hoechst 33342. Bottom: Quantification of NPTX2 immunostaining. Bottom left: fraction of cells detectably expressing NPTX2. Bottom middle: immunostaining signal intensity in cells detectably expressing NPTX2. Bottom right: total number of Hoechst positive nuclei. Data are mean ± s.e.m.; *p = 0.0142 (Student's two tailed t test). n = 6 individuals without cognitive impairment and n = 6 individuals with a cognitive diagnosis of AD (cognitive impairment); n = 2 images per individual.

See also Figure S6; Data S1, page 11; and Methods S1, page 5.

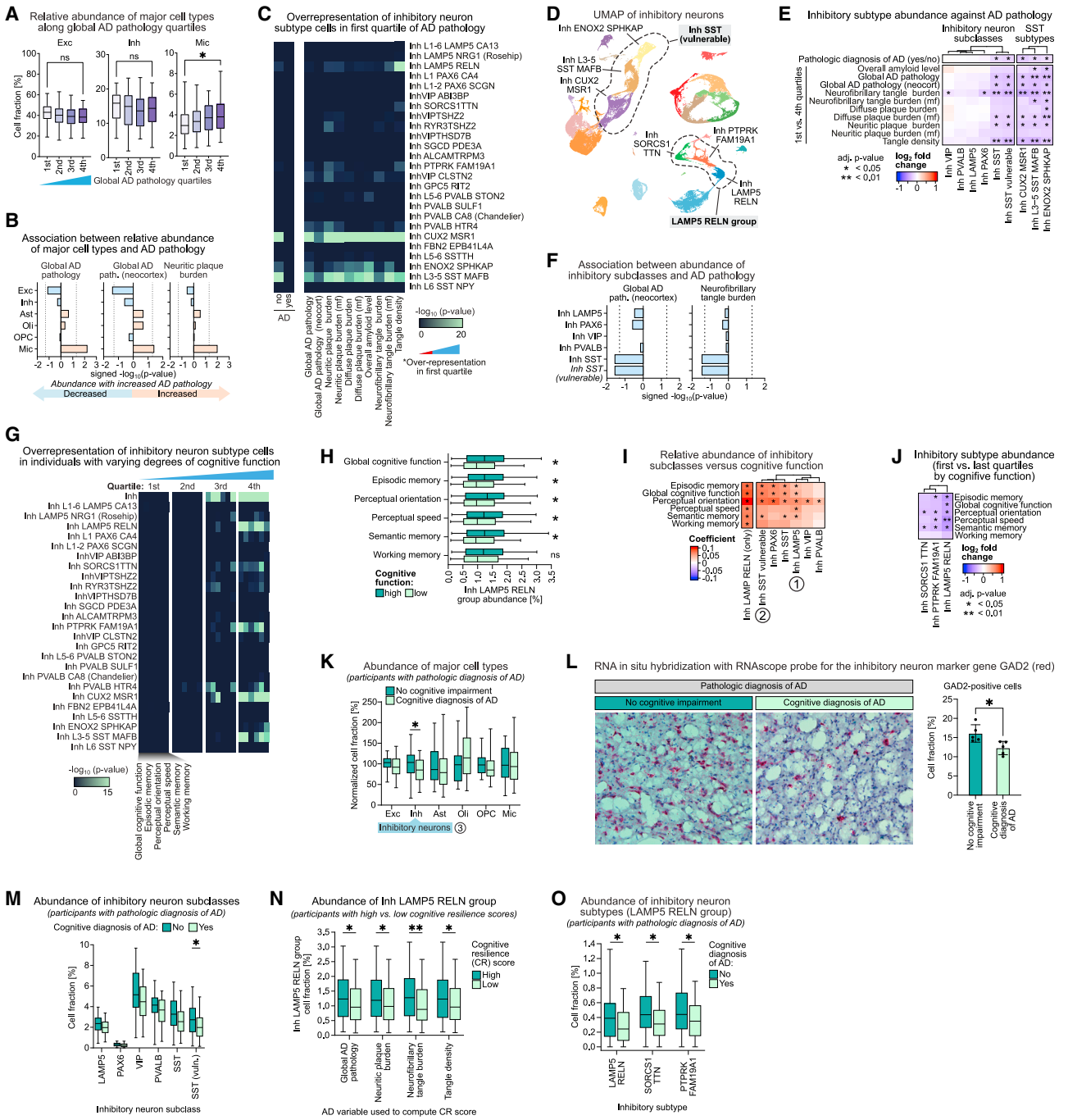


Figure 7. Cell type composition alterations associated with AD pathology and cognitive impairment

(A) Boxplots showing the relative abundance of major cell types along disease progression. Within each box, horizontal lines denote median values; boxes extend from the 25th to the 75th percentile of each group's distribution of values; whiskers extend from the 5th to the 95th percentile. *p < 0.05; n.s., p > 0.05 (Kruskal-Wallis test followed by Dunn's multiple comparison test).

(B) Association (association scores) between the relative abundance of major cell types and the measures of AD pathology indicated. The dotted lines indicate the significance level threshold of an FDR-corrected p value of 0.05.

(C and G) Overrepresentation within each inhibitory neuron subtype of cells isolated from individuals with varying degrees of (C) AD pathology or (G) cognitive function (hypergeometric test; the p values have been adjusted for multiple hypothesis testing; −log10[Bonferroni-corrected p values] are shown).

(D) UMAP of inhibitory neurons colored by subtype. Dashed line surrounds the group of vulnerable SST inhibitory neuron subtypes and the subtypes of the LAMP5 RELN group, respectively.

(legend continued on next page)

revealed that the three inhibitory neuron subtypes, Inh LAMP5 RELN, Inh PTPRK FAM19A1, and Inh SORCS1 TTN (hereafter referred to as LAMP5 RELN group), clustered next to each other, suggesting a high degree of transcriptomic similarity between these subtypes (Figure 7D). The relative abundance of the entire LAMP5 RELN group of inhibitory neurons and the entire subclass of SST inhibitory neurons was significantly decreased in individuals scoring low on tests of global cognitive function, episodic memory, perceptual orientation, and semantic memory (Figure 7H; Data S1, page 15A). These findings were further supported by additional group-based analyses comparing individuals with high versus low cognitive function and by quasi-binomial regression analysis (Figures 7I and 7J; Data S1, pages 15B–15D). In general, we did not find a significant association between the relative abundance of excitatory neuron subtypes and cognitive function except for layer 2/3 (L2–3) excitatory neurons (Exc L2–3 CBLN2 LINC02306), which were positively associated with the measure of perceptual orientation (Data S1, pages 16A and 16B).

In summary, our analysis uncovered two distinct groups of inhibitory neuron subtypes that are more abundant in the prefrontal cortex of individuals with preserved high cognitive function late in life. These findings were further supported by a quasi-binomial regression analysis of the combined snRNA-seq dataset encompassing 615 individuals (Data S1, pages 16C–16E).

Cell types associated with cognitive resilience to AD pathology

Finally, we explored what distinguishes the brains of individuals with CR to AD pathology in terms of cell type composition. We found that the relative abundance of inhibitory neurons was significantly higher in individuals with high CR to global AD pathology, neuritic plaque burden, and NFT burden (Figure S7A). Consistent with this observation, among individuals with a pathologic diagnosis of AD, the relative abundance of inhibitory neurons was significantly higher in individuals without cognitive impairment compared to individuals with a diagnosis of AD dementia (Figure 7K). This finding was further supported by quasi-binomial regression analysis (Figure S7B) and RNA *in situ* hybridization analysis on PFC tissue samples (Figure 7L).

We then asked whether cognitive resilience to AD pathology is associated with specific subclasses and subtypes of inhibitory neurons. RELN-positive inhibitory neurons of the LAMP5 subclass were significantly overrepresented in the subset of individuals with the highest level of CR to global AD pathology, neuritic plaque burden, NFT burden, and tangle density (Figure S7C). The remaining subtypes of the LAMP5 RELN group of inhibitory neurons and the vulnerable SST subtypes also tended to be overrepresented in individuals with high CR to AD pathology (Figure S7C). The relative abundance of the SST subclass of inhibitory neurons was significantly higher in individuals with high CR to global AD pathology, neuritic plaque burden, and NFT burden (Figures S7D–S7F). Among individuals with a pathologic diagnosis of AD, the relative abundance of the vulnerable SST neurons was significantly higher in individuals without cognitive impairment compared to individuals with a diagnosis of AD dementia (Figure 7M). Focusing on the LAMP5 RELN group of inhibitory neurons, we found that the relative abundance of the entire group and two of its subtypes (Inh LAMP5 RELN and Inh PTPRK FAM19A1) was significantly higher in individuals with high CR to global AD pathology, neuritic plaque burden, NFT burden, and tangle density (Figures 7N and S7G). Furthermore, among individuals with a pathologic diagnosis of AD, the relative abundance of the individual subtypes of the LAMP5 RELN group was significantly higher in individuals without cognitive impairment compared to individuals with a diagnosis of AD dementia (Figure 7O). These findings were further supported by quasi-binomial regression analysis (Data S1, pages 17A–17C), RT-qPCR (Data S1, page 17D), chromogenic RNA *in situ* hybridization (Data S1, pages 17E–17G), immunohistochemical analysis (Data S1, pages 18A–18C), a quasi-binomial regression analysis of the combined dataset across 615 individuals (Data S1, pages 18D–18F), and multiplex RNA

(E) Relative abundance of inhibitory neuron subclasses and SST inhibitory neuron subtypes comparing individuals with high versus low AD pathology. The cohort was split into quarters based on the pathology measures indicated. **$p < 0.01$, *$p < 0.05$; n.s., $p > 0.05$ (multiple Mann-Whitney tests corrected for multiple hypothesis testing using the Holm-Šídák method).

(F) Association between relative abundance of inhibitory neuron subclasses and the measures of AD pathology indicated.

(H) Relative abundance of the LAMP5 RELN group of inhibitory neurons comparing individuals with high versus low cognitive function. **$p < 0.01$, *$p < 0.05$; n.s., $p > 0.05$ (multiple Mann-Whitney tests corrected for multiple hypothesis testing using the Holm-Šídák method).

(I) Association (quasi-binomial regression) between the relative abundance of inhibitory neuron subclasses and measures of cognitive function. Significant positive association between the relative abundance of the LAMP5 (1) and SST (2) subclasses of inhibitory neurons and the measures of cognitive function.

(J) Relative abundance of inhibitory neuron subtypes (LAMP5 RELN group) comparing individuals with high versus low cognitive function. The cohort was split into quarters based on the cognitive domains indicated. **$p < 0.01$, *$p < 0.05$; n.s., $p > 0.05$ (multiple Mann-Whitney tests corrected for multiple hypothesis testing using the Holm-Šídák method).

(K, M, and O) Relative abundance of (K) major cell types, (M) inhibitory neuron subclasses, or (O) inhibitory neuron subtypes (LAMP5 RELN group) in study participants with a pathologic diagnosis of AD comparing individuals with or without a cognitive diagnosis of AD. *$p < 0.05$; n.s., $p > 0.05$ (multiple Mann-Whitney tests corrected for multiple hypothesis testing using the Holm-Šídák method). (3) Significantly higher relative abundance of inhibitory neurons in individuals without cognitive impairment compared to individuals with a diagnosis of AD dementia (in the subset of individuals with a pathologic diagnosis of AD).

(L) RNA *in situ* hybridization with an RNAscope probe for the inhibitory neuron marker gene GAD2 (red) in the gray matter of the PFC of an individual with no cognitive impairment (left) and an individual with a cognitive diagnosis of AD (cognitive impairment) (right). Scale bar: 100 μm. Right: Fraction of cells detectably expressing GAD2. Data are mean ± s.e.m.; *$p = 0.0178$ (Student's two tailed t test). n = 5 individuals with no cognitive impairment and n = 5 individuals with a cognitive diagnosis of AD; n = 4 images per individual. All the individuals considered in this analysis had a pathologic diagnosis of AD.

(N) Relative abundance of the LAMP5 RELN group of inhibitory neurons comparing individuals with high versus low cognitive resilience scores. **$p < 0.01$, *$p < 0.05$; n.s., $p > 0.05$ (multiple Mann-Whitney tests corrected for multiple hypothesis testing using the Holm-Šídák method).

See also Figure S7 and Data S1, pages 12–19.

in situ hybridization analysis of *postmortem* PFC tissue (Data S1, page 19).

In summary, our analysis revealed that inhibitory neurons, and in particular two distinct subgroups associated with high cognitive function (the LAMP5 RELN group and vulnerable SST subtypes), are more abundant in the prefrontal cortex of individuals with high cognitive resilience to AD pathology.

DISCUSSION

In this work, we present a transcriptomic atlas of the aged human prefrontal cortex from 427 ROSMAP study participants with varying degrees of Alzheimer's disease progression, cognitive impairment, and resilience to AD pathology. Our analysis of the single-cell resolution data from this large cohort enabled us to annotate the cellular diversity of the human PFC, identify AD-pathology-associated gene expression alterations shared between excitatory neuron subtypes, and pinpoint alterations in the expression of genes involved in critical and fundamental cellular processes, including synaptic signaling, chromatin organization, metabolism of lipids, mRNA and tRNA metabolic processes, and mitochondrial function. Systematic analysis of protein complex expression changes along the progression of AD pathology revealed a coordinated increase of the cohesin complex and the cellular response to DNA damage in excitatory neurons and in oligodendrocytes. Our large dataset enabled us to identify genes and pathways associated with high cognitive performance and discover factors linked to preserved cognitive function in the presence of AD pathology. Tracking compositional differences across the spectrum of AD pathology and cognitive impairment, we were able to pinpoint selectively vulnerable somatostatin inhibitory neuron subtypes depleted in AD, identify two distinct groups of inhibitory neurons that were more abundant in individuals with preserved high cognitive function late in life, and uncover a link between inhibitory neurons and resilience to AD pathology.

Systematic differential expression analysis revealed widespread gene expression changes linked to various measures of AD pathology across many cell types. Some of these changes, including the increased expression of genes involved in synaptic signaling, may represent an attempt of cells to reestablish homeostasis. Other alterations, including the decreased expression of electron transport chain genes and genes encoding mitochondrial ribosomal proteins and aminoacyl-tRNA synthetases, may contribute to cellular dysfunction in AD. Reduced expression of the mitochondrial intermembrane space bridging complex might compromise the architecture of the mitochondrial inner membrane and consequently contribute to mitochondrial dysfunction in AD.[53]

Greater expression of the cohesin complex in AD was another phenomenon recurrently observed across multiple cell types. In agreement with a well-established link between DNA damage and neurodegenerative diseases,[54–56] we found that greater cohesin complex expression was accompanied by higher expression of DNA damage response factors. Especially in excitatory neurons, the cohesin complex expression was also correlated with the expression of many chromatin regulators, warranting the examination of changes to the chromatin landscape in AD (see Dileep et al.[48] and Xiong et al.[49] in this issue). The cohesin complex has also been reported to prevent the end joining of distant DNA double-strand ends, thereby protecting against chromosomal re-arrangements.[57] In this issue of *Cell*, Dileep et al.[48] directly tested the prediction that DNA double-strand breaks (DSBs) will impact both the linear and the 3D genome organization in neurodegeneration. Indeed, Dileep et al.[48] found that excitatory neurons in AD are enriched for gene fusions indicating chromosomal structural variations and that neuronal DNA DSBs lead to mosaic chromosomal structural variations and 3D genome disruption in neurodegeneration.

We capitalized on the phenotypic variability between study participants in terms of cognitive function and cognitive decline to interrogate what characterizes individuals who maintain high cognitive function late in life. We identified genes positively associated with global cognitive function across multiple excitatory neuron subtypes, including the synaptic genes CACNG3, PAK1, NPTX2, RPH3A, SVOP, and BDNF. Another synaptic gene positively associated with global cognitive function was VGF, which is potently upregulated by BDNF[58] and has been shown to reduce amyloid plaque burden, disease-associated microglial activation, and memory impairment in a mouse model of AD.[59] We found multiple genes involved in FGF signaling, including FGF1 and FGF22, that were positively associated with global cognitive function across multiple excitatory neuron subtypes. Interestingly, cerebrospinal fluid of young mice has recently been shown to restore memory in aged mice via Fgf17.[60] Whether other members of the FGF family might have similar rejuvenating powers remains to be investigated.

Cell death is a defining feature of neurodegenerative disorders, including Alzheimer's disease. We analyzed how the composition of cell types changes as AD pathology advances to pinpoint especially vulnerable or resilient cell types. We identified a significant decrease in the relative abundance of somatostatin inhibitory neurons in individuals with high AD pathology, consistent with previous reports.[61–63] Going beyond the subclass level of inhibitory neurons, we identified a subset of three vulnerable somatostatin inhibitory neuron subtypes with transcriptomically similar identities that were significantly depleted in AD.

There is a 2.5-fold elevation in inhibitory neurons and an expanded inhibitory-to-inhibitory neuron network in humans compared with mice, highlighting the importance of this cell class for human brain function.[64–66] We tracked how the composition of cell types changes with the progression of cognitive impairment to pinpoint cell types linked to cognitive function in the aged human brain. We discovered two distinct groups of inhibitory neurons, including the vulnerable SST subtypes, that were more abundant in the PFC of individuals with preserved high cognitive function late in life. The second group of inhibitory neurons linked to cognitive function centered around a RELN-positive subtype of the LAMP5 subclass. Consistent with our findings, loss of LAMP5 interneurons has recently been reported to drive neuronal network dysfunction in AD mouse models.[67] Importantly, we found that the two groups of inhibitory neuron subtypes associated with cognition were also more abundant in the PFCs of individuals with high cognitive resilience to AD pathology. In summary, our results highlight a link between inhibitory neurons and preserved cognitive function late in life. The importance of inhibitory neurons in the context of preserved

high cognitive function is well in line with a large body of literature describing a crucial role of this cell class in controlling brain oscillations, which are hypothesized to enhance information processing.[68,69] Recent technological advancements in neuroscience, in particular non-invasive methods of brain stimulation, will help elucidate the function of inhibitory neurons in preserving high cognitive performance late in life.

Our work provides a comprehensive resource for the medical, scientific, and pharmaceutical communities that offers a detailed molecular and cell-type-specific portrait of Alzheimer's disease.

Limitations of the study

Despite a vast array of cells analyzed in this study, the limited number of cells acquired from each participant may impede certain analyses, particularly for less prevalent cell types. A simulation-based power analysis revealed considerably lower statistical power estimates for less abundant cell types, such as those associated with the vasculature and inhibitory neuron subtypes (Methods S1, page 6A). Also, our empirical investigation revealed a positive association between the number of cells evaluated per participant and the number of DEGs identified (Methods S1, pages 6B and 6C). This correlation indicates that the present study might have underestimated the number of DEGs in cell types with lower abundance, such as inhibitory neuron subtypes.

An additional important caveat to consider is the potential impact of ambient RNA contamination. To address this issue, we employed the DecontX method to estimate and remove ambient RNA from our snRNA-seq data. Overall, we found that the results of our original analysis were largely confirmed after removing ambient RNA contamination (Data S1, page 4F). However, it's worth noting that certain glial cell types (Oli, OPC, and Mic) showed a lower confirmation rate for downregulated genes after ambient RNA correction, suggesting the importance of considering this factor in their analysis.

It is also important to acknowledge certain limitations and complexities involved in analyzing data from a large number of individual human samples. One limitation is that gene expression variation can be influenced by multiple biological, demographic, and technical factors. To address this concern, we performed a comprehensive variance partitioning analysis, evaluating the contribution of known technical covariates and factors of unwanted variation estimated by Remove Unwanted Variation (RUV)-seq to the gene expression variance in our dataset. We found that the factors estimated by RUV-seq effectively captured the variance attributable to known technical covariates (Methods S1, page 7). While this analysis strengthens the study's robustness, it is important to note that the influence of unmeasured or unknown covariates remains a potential limitation in our analysis. Nevertheless, we believe our approach provides a solid foundation for interpreting the data and enhances the reliability of our findings.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Human subjects
- METHOD DETAILS
  - Clinical data
  - Pathologic AD and other pathologies
  - Cognitive resilience
  - Isolation of nuclei from frozen postmortem brain tissue
  - Single-nucleus RNA-seq
  - cDNA synthesis and RT-qPCR analysis
  - Immunohistochemistry
  - RNA *in situ* hybridization
  - Isolation of neuronal nuclei, cDNA synthesis and RT-qPCR analysis
  - RNAscope and HiPlex
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - snRNA-seq processing, QC, and annotation
  - Differential gene expression analysis
  - Comparison with previously published snRNA-seq studies of AD
  - Correlation analysis of gene expression and clinical and pathological traits using self-organizing maps
  - Protein complexes analysis
  - Cognitive resilience analysis
  - Cell type composition
  - External data sources
  - IMARIS colocalization quantification

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2023.08.039.

ACKNOWLEDGMENTS

We thank the study participants and staff of the Rush Alzheimer's Disease Center. We thank all members of the Tsai and Kellis labs at MIT for extensive discussions and feedback on computational methods and biological interpretation. This work was supported in part by the Cure Alzheimer's Fund; the JBP Foundation; the Robert A. and Renee E. Belfer Family Foundation; Eduardo Eurnekian; Joseph P. DiSabato; NIH grants AG058002, AG062377, NS110453, NS115064, AG062335, AG074003, NS127187, MH119509, HG008155, AG081017, NS129032, AG077227, and AG067151 (M.K.); NIH grants RF1AG062377, RF1 AG054321, and RO1 AG054012 (L.-H.T.); and the NIH training grant GM087237 (to C.A.B.). H.M. was supported by an Early Postdoc Mobility fellowship from the Swiss National Science Foundation (P2BSP3_151885). ROSMAP is supported by P30AG10161, P30AG72975, R01AG15819, R01AG17917, U01AG46152, and U01AG61356. M.B.V. is supported by the Howard Hughes Medical Institute Hannah Gray Postdoctoral Fellowship. ROSMAP resources can be requested at https://www.radc.rush.edu.

AUTHOR CONTRIBUTIONS

This study was designed by H.M., C.B., D.A.B., M.K., and L.-H.T. and directed and coordinated by H.M., M.K., and L.-H.T. H.M., A.P.N., X.J., J.M., and K.G. carried out sample preparations and single-cell RNA profiling. H.M. and C.B. led the computational analysis. Z.P., M.B.V., and N.L. carried out experimental

validation experiments. G.A., X.J., A.P.N., K.G., S.B., A.K.K., V.N.L., G.E.F., Y.L., J.B.I., H.P.B., P.R.P., N.C., J.I.B., E.J.I., C.F.K., M.M.M., H.H.P., S.P.P., and M.T.T. helped with the computational analysis. D.A.B. contributed samples and data. H.M., C.B., G.A., D.A.B., M.K., and L.-H.T. wrote the manuscript.

DECLARATION OF INTERESTS

L.-H.T. is a member of the Scientific Advisory Boards of Cognito Therapeutics, 4M Therapeutics, Cell Signaling Technology, and Souvien Therapeutics, which have no association to the work described in this manuscript.

Received: September 19, 2022
Revised: May 20, 2023
Accepted: August 29, 2023
Published: September 28, 2023

REFERENCES

1. Wilson, R.S., Segawa, E., Boyle, P.A., Anagnos, S.E., Hizel, L.P., and Bennett, D.A. (2012). The Natural History of Cognitive Decline in Alzheimer's Disease. Psychol. Aging *27*, 1008–1017. https://doi.org/10.1037/A0029857.
2. Scheltens, P., Blennow, K., Breteler, M.M.B., De Strooper, B., Frisoni, G.B., Salloway, S., and Van Der Flier, W.M. (2016). Alzheimer's disease. Lancet.. https://doi.org/10.1016/S0140-6736(15)01124-1.
3. Masters, C.L., Bateman, R., Blennow, K., Rowe, C.C., Sperling, R.A., and Cummings, J.L. (2015). Alzheimer's disease. Nat. Publ. Gr. *1*, 15056. https://doi.org/10.1038/nrdp.2015.56.
4. Serrano-Pozo, A., Frosch, M.P., Masliah, E., and Hyman, B.T. (2011). Neuropathological alterations in Alzheimer disease. Cold Spring Harb. Perspect. Med. *1*, a006189. https://doi.org/10.1101/cshperspect.a006189.
5. Hardy, J., and Selkoe, D.J. (2002). The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science *297*, 353–356. https://doi.org/10.1126/SCIENCE.1072994.
6. Braak, H., and Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol. *82*, 239–259. https://doi.org/10.1007/BF00308809.
7. Beach, T.G., Walker, R., and McGeer, E.G. (1989). Patterns of gliosis in Alzheimer's disease and aging cerebrum. Glia *2*, 420–436. https://doi.org/10.1002/GLIA.440020605.
8. Heneka, M.T., Carson, M.J., El Khoury, J., Landreth, G.E., Brosseron, F., Feinstein, D.L., Jacobs, A.H., Wyss-Coray, T., Vitorica, J., Ransohoff, R.M., et al. (2015). Neuroinflammation in Alzheimer's disease. Lancet Neurol. *14*, 388–405. https://doi.org/10.1016/S1474-4422(15)70016-5.
9. Wyss-Coray, T. (2016). Ageing, neurodegeneration and brain rejuvenation539. Nature *539*, 180–186. https://doi.org/10.1038/nature20411.
10. Almanzar, N., Antony, J., Baghel, A.S., Bakerman, I., Bansal, I., Barres, B.A., Beachy, P.A., Berdnik, D., Bilen, B., Brownfield, D., et al. (2020). A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. Nature *583*, 590–595. https://doi.org/10.1038/s41586-020-2496-1.
11. Delorey, T.M., Ziegler, C.G.K., Heimberg, G., Normand, R., Yang, Y., Segerstolpe, Å., Abbondanza, D., Fleming, S.J., Subramanian, A., Montoro, D.T., et al. (2021). COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. Nat *595*, 107–113. https://doi.org/10.1038/s41586-021-03570-8.
12. Grubman, A., Chew, G., Ouyang, J.F., Sun, G., Choo, X.Y., McLean, C., Simmons, R.K., Buckberry, S., Vargas-Landin, D.B., Poppe, D., et al. (2019). 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. https://doi.org/10.1038/s41593-019-0539-4.
13. Leng, K., Li, E., Eser, R., Piergies, A., Sit, R., Tan, M., Neff, N., Li, S.H., Rodriguez, R.D., Suemoto, C.K., et al. (2021). Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. Nat. Neurosci. *24*, 276–287. https://doi.org/10.1038/s41593-020-00764-7.
14. Morabito, S., Miyoshi, E., Michael, N., Shahin, S., Martini, A.C., Head, E., Silva, J., Leavy, K., Perez-Rosendahl, M., and Swarup, V. (2021). Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. Nat. Genet. *53*, 1143–1155. https://doi.org/10.1038/s41588-021-00894-z.
15. Sadick, J.S., O'Dea, M.R., Hasel, P., Dykstra, T., Faustin, A., and Liddelow, S.A. (2022). Astrocytes and oligodendrocytes undergo subtype-specific transcriptional changes in Alzheimer's disease. Neuron *110*, 1788–1805.e10. https://doi.org/10.1016/J.NEURON.2022.03.008.
16. Zhou, Y., Song, W.M., Andhey, P.S., Swain, A., Levy, T., Miller, K.R., Poliani, P.L., Cominelli, M., Grover, S., Gilfillan, S., et al. (2020). Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. Nat. Med. *26*, 131–142. https://doi.org/10.1038/S41591-019-0695-9.
17. Neuner, S.M., Telpoukhovskaia, M., Menon, V., O'Connell, K.M.S., Hohman, T.J., and Kaczorowski, C.C. (2022). Translational approaches to understanding resilience to Alzheimer's disease. Trends Neurosci. *45*, 369–383. https://doi.org/10.1016/J.TINS.2022.02.005.
18. Wagner, M., Wilson, R.S., Leurgans, S.E., Boyle, P.A., Bennett, D.A., Grodstein, F., and Capuano, A.W. (2022). Quantifying longitudinal cognitive resilience to Alzheimer's disease and other neuropathologies. Alzheimers Dement. *18*, 2252–2261. https://doi.org/10.1002/alz.12576.
19. Bennett, D.A., Buchman, A.S., Boyle, P.A., Barnes, L.L., Wilson, R.S., and Schneider, J.A. (2018). Religious Orders Study and Rush Memory and Aging Project. J. Alzheimers Dis. *64*, S161–S189. https://doi.org/10.3233/JAD-179939.
20. Hodge, R.D., Bakken, T.E., Miller, J.A., Smith, K.A., Barkan, E.R., Graybuck, L.T., Close, J.L., Long, B., Johansen, N., Penn, O., et al. (2019). Conserved cell types with divergent features in human versus mouse cortex. Nature *573*, 61–68. https://doi.org/10.1038/s41586-019-1506-7.
21. Sun, N., Victor, M.B., Park, Y., Xiong, X., Scannail, A.N., Leary, N., Prosper, S., Viswanathan, S., Luna, X., Boix, C.A., et al. (2023). Human Microglial State Dynamics in Alzheimer's Disease Progression. Cell *186*, 4386–4403. https://doi.org/10.1016/j.cell.2023.08.037.
22. Mathys, H., Davila-Velderrain, J., Peng, Z., Gao, F., Mohammadi, S., Young, J.Z., Menon, M., He, L., Abdurrob, F., Jiang, X., et al. (2019). Single-cell transcriptomic analysis of Alzheimer's disease. Nat *570*, 332–337. https://doi.org/10.1038/s41586-019-1195-2.
23. Lau, S.F., Cao, H., Fu, A.K.Y., and Ip, N.Y. (2020). Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and neuroprotective glia in Alzheimer's disease. Proc. Natl. Acad. Sci. USA. *117*, 25800–25809. https://doi.org/10.1073/PNAS.2008762117/-/DCSUPPLEMENTAL.
24. Su, Y., Zhou, Y., Bennett, M.L., Li, S., Carceles-Cordon, M., Lu, L., Huh, S., Jimenez-Cyrus, D., Kennedy, B.C., Kessler, S.K., et al. (2022). A single-cell transcriptome atlas of glial diversity in the human hippocampus across the postnatal lifespan. Cell Stem Cell *29*, 1594–1610.e8. https://doi.org/10.1016/j.stem.2022.09.010.
25. Yang, S., Corbett, S.E., Koga, Y., Wang, Z., Johnson, W.E., Yajima, M., and Campbell, J.D. (2020). Decontamination of ambient RNA in single-cell RNA-seq with DecontX. Genome Biol. *21*, 57. https://doi.org/10.1186/S13059-020-1950-6/FIGURES/6.
26. Green, G.S., Fujita, M., Yang, H.-S., Taga, M., McCabe, C., Cain, A., White, C.C., Schmidtner, A.K., Zeng, L., Wang, Y., et al. (2023). Cellular dynamics across aged human brains uncover a multicellular cascade leading to Alzheimer's disease. Preprint at bioRxiv.. https://doi.org/10.1101/2023.03.07.531493.
27. Fujita, M., Gao, Z., Zeng, L., McCabe, C., White, C.C., Ng, B., Green, G.S., Rozenblatt-Rosen, O., Phillips, D., Amir-Zilberstein, L., et al. (2022). Cell-subtype specific effects of genetic variation in the aging

and Alzheimer cortex. Preprint at bioRxiv.. https://doi.org/10.1101/2022.11.07.515446.

28. Gabitto, M.I., Travaglini, K.J., Rachleff, V.M., Kaplan, E.S., Long, B., Ariza, J., Ding, Y., Mahoney, J.T., Dee, N., Goldy, J., et al. (2023). Integrated multimodal cell atlas of Alzheimer's disease. Preprint at bioRxiv.. https://doi.org/10.1101/2023.05.08.539485.
29. Koopmans, F., van Nierop, P., Andres-Alonso, M., Byrnes, A., Cijsouw, T., Coba, M.P., Cornelisse, L.N., Farrell, R.J., Goldschmidt, H.L., Howrigan, D.P., et al. (2019). SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. Neuron *103*, 217–234.e4. https://doi.org/10.1016/J.NEURON.2019.05.002/ATTACHMENT/80038C21-9CDC-40DD-AEB4-3DAC61339481/MMC8.XLSX.
30. Otero-Garcia, M., Xue, Y.-Q., Shakouri, T., Deng, Y., Morabito, S., Allison, T., Lowry, W.E., Kawaguchi, R., Swarup, V., and Cobos, I. (2020). Single-soma transcriptomics of tangle-bearing neurons in Alzheimer's disease reveals the signatures of tau-associated synaptic dysfunction. Preprint at bioRxiv.. https://doi.org/10.1101/2020.05.11.088591.
31. Johnson, E.C.B., Dammer, E.B., Duong, D.M., Ping, L., Zhou, M., Yin, L., Higginbotham, L.A., Guajardo, A., White, B., Troncoso, J.C., et al. (2020). Large-scale proteomic analysis of Alzheimer's disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. Nat. Med., 769–780. https://doi.org/10.1038/s41591-020-0815-6.
32. Pfanner, N., van der Laan, M., Amati, P., Capaldi, R.A., Caudy, A.A., Chacinska, A., Darshi, M., Deckers, M., Hoppins, S., Icho, T., et al. (2014). Uniform nomenclature for the mitochondrial contact site and cristae organizing system. J. Cell Biol. *204*, 1083–1086. https://doi.org/10.1083/JCB.201401006.
33. Giurgiu, M., Reinhard, J., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., Montrone, C., and Ruepp, A. (2019). CORUM: The comprehensive resource of mammalian protein complexes - 2019. Nucleic Acids Res. *47*, D559–D563. https://doi.org/10.1093/nar/gky973.
34. Kim, H., Dejsuphong, D., Adelmant, G., Ceccaldi, R., Yang, K., Marto, J.A., and D'Andrea, A.D. (2014). A Transcriptional Repressor ZBTB1 Promotes Chromatin Remodeling and Translesion DNA Synthesis. Mol. Cell *54*, 107–118. https://doi.org/10.1016/J.MOLCEL.2014.02.017.
35. Dobbin, M.M., Madabhushi, R., Pan, L., Chen, Y., Kim, D., Gao, J., Ahanonu, B., Pao, P.-C., Qiu, Y., Zhao, Y., and Tsai, L.H. (2013). SIRT1 collaborates with ATM and HDAC1 to maintain genomic stability in neurons. Nat. Neurosci. *16*, 1008–1015. https://doi.org/10.1038/nn.3460.
36. Oberdoerffer, P., Michan, S., McVay, M., Mostoslavsky, R., Vann, J., Park, S.K., Hartlerode, A., Stegmuller, J., Hafner, A., Loerch, P., et al. (2008). SIRT1 Redistribution on Chromatin Promotes Genomic Stability but Alters Gene Expression during Aging. Cell *135*, 907–918. https://doi.org/10.1016/J.CELL.2008.10.025/ATTACHMENT/6E5B2D1C-D04E-4449-9E3A-E3DA8A658F42/MMC1.PDF.
37. Farnung, L., Vos, S.M., Wigge, C., and Cramer, P. (2017). Nucleosome-Chd1 structure and implications for chromatin remodelling. Nature *550*, 539–542. https://doi.org/10.1038/NATURE24046.
38. Zhou, J., Li, J., Serafim, R.B., Ketchum, S., Ferreira, C.G., Liu, J.C., Coe, K.A., Price, B.D., and Yusufzai, T. (2018). Human CHD1 is required for early DNA-damage signaling and is uniquely regulated by its N terminus. Nucleic Acids Res. *46*, 3891–3905. https://doi.org/10.1093/NAR/GKY128.
39. Bulut-Karslioglu, A., Jin, H., Kim, Y.K., Cho, B., Guzman-Ayala, M., Williamson, A.J.K., Hejna, M., Stötzel, M., Whetton, A.D., Song, J.S., and Ramalho-Santos, M. (2021). Chd1 protects genome integrity at promoters to sustain hypertranscription in embryonic stem cells. Nat. Commun. *12*, 4859. https://doi.org/10.1038/s41467-021-25088-3.
40. Kari, V., Mansour, W.Y., Raul, S.K., Baumgart, S.J., Mund, A., Grade, M., Sirma, H., Simon, R., Will, H., Dobbelstein, M., et al. (2016). Loss of CHD1 causes DNA repair defects and enhances prostate cancer therapeutic responsiveness. EMBO Rep. *17*, 1609–1623. https://doi.org/10.15252/EMBR.201642352.
41. Chiu, L.Y., Gong, F., and Miller, K.M. (2017). Bromodomain proteins: repairing DNA damage within chromatin. Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*, 20160286. https://doi.org/10.1098/RSTB.2016.0286.
42. Liu, X.S., Chandramouly, G., Rass, E., Guan, Y., Wang, G., Hobbs, R.M., Rajendran, A., Xie, A., Shah, J.V., Davis, A.J., et al. (2015). LRF maintains genome integrity by regulating the non-homologous end joining pathway of DNA repair. Nat. Commun. *6*, 8325. https://doi.org/10.1038/ncomms9325.
43. Bot, C., Pfeiffer, A., Giordano, F., Manjeera, D.E., Dantuma, N.P., and Ström, L. (2017). Independent mechanisms recruit the cohesin loader protein NIPBL to sites of DNA damage. J. Cell Sci. *130*, 1134–1146. https://doi.org/10.1242/JCS.197236.
44. Parsons, J.L., Dianova, I.I., Khoronenkova, S.V., Edelmann, M.J., Kessler, B.M., and Dianov, G.L. (2011). USP47 is a deubiquitylating enzyme that regulates base excision repair by controlling steady-state levels of DNA polymerase β. Mol. Cell *41*, 609–615. https://doi.org/10.1016/J.MOLCEL.2011.02.016.
45. Oppikofer, M., Sagolla, M., Haley, B., Zhang, H.M., Kummerfeld, S.K., Sudhamsu, J., Flynn, E.M., Bai, T., Zhang, J., Ciferri, C., and Cochran, A.G. (2017). Non-canonical reader modules of BAZ1A promote recovery from DNA damage. Nat. Commun. *8*, 862. https://doi.org/10.1038/s41467-017-00866-0.
46. Singh, R., Kalra, R.S., Hasan, K., Kaul, Z., Cheung, C.T., Huschtscha, L., Reddel, R.R., Kaul, S.C., and Wadhwa, R. (2014). Molecular characterization of collaborator of ARF (CARF) as a DNA damage response and cell cycle checkpoint regulatory protein. Exp. Cell Res. *322*, 324–334. https://doi.org/10.1016/J.YEXCR.2014.01.022.
47. Yang, X., Ma, Y., Li, Y., Dong, Y., Yu, L.L., Wang, H., Guo, L., Wu, C., Yu, X., and Liu, X. (2020). Molecular basis for the MacroD1-mediated hydrolysis of ADP-ribosylation. DNA Repair *94*, 102899. https://doi.org/10.1016/J.DNAREP.2020.102899.
48. Dileep, V., Boix, C.A., Mathys, H., Marco, A., Welch, G.M., Meharena, H.S., Loon, A., Jeloka, R., Peng, Z., Bennett, D.A., et al. (2023). Neuronal DNA double-strand breaks lead to chromosomal structural variations and 3D genome disruption in neurodegeneration. Cell *186*, 4404–4421. https://doi.org/10.1016/j.cell.2023.08.038.
49. Xiong, X., James, B.T., Boix, C.A., Park, Y., Galani, K., Victor, M.B., Sun, N., Hou, L., Dileep, V., Ho, L.-L., et al. (2023). Epigenomic Dissection of Alzheimer's Disease Pinpoints Causal Variants and Reveals Epigenome Erosion. Cell *186*, 4422–4437. https://doi.org/10.1016/j.cell.2023.08.040.
50. Arenaza-Urquijo, E.M., and Vemuri, P. (2018). Resistance vs resilience to Alzheimer disease: Clarifying terminology for preclinical studies. Neurology *90*, 695–703.
51. Stern, Y., Arenaza-Urquijo, E.M., Bartrés-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., Franzmeier, N., Kempermann, G., Kremen, W.S., et al. (2020). Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. Alzheimer's Dement. *16*, 1305–1311. https://doi.org/10.1016/j.jalz.2018.07.219.
52. Barker, S.J., Raju, R.M., Milman, N.E.P., Wang, J., Davila-Velderrain, J., Gunter-Rahman, F., Parro, C.C., Bozzelli, P.L., Abdurrob, F., Abdelaal, K., et al. (2021). MEF2 is a key regulator of cognitive potential and confers resilience to neurodegeneration. Sci. Transl. Med. *13*, eabd7695. https://doi.org/10.1126/SCITRANSLMED.ABD7695/SUPPL_FILE/SCITRANSLMED.ABD7695_DATA_FILE_S1.ZIP.
53. Ashleigh, T., Swerdlow, R.H., and Beal, M.F. (2023). The role of mitochondrial dysfunction in Alzheimer's disease pathogenesis. Alzheimer's Dement. *19*, 333–342. https://doi.org/10.1002/ALZ.12683.
54. Welch, G., Boix, C., Schmauch, E., Davila-Velderrain, J., Victor, M.B., Dileep, V., Bozzelli, L., Su, Q., Cheng, J., Lee, A., et al. (2021). Neurons burdened by DNA double strand breaks incite microglia activation through antiviral-like signaling in neurodegeneration. Preprint at bioRxiv.. https://doi.org/10.1101/2021.12.23.474002.
55. Chow, H.M., and Herrup, K. (2015). Genomic integrity and the ageing brain. Nat. Rev. Neurosci. *16*, 672–684. https://doi.org/10.1038/nrn4020.

56. Welch, G., and Tsai, L.-H. (2022). Mechanisms of DNA damage-mediated neurotoxicity in neurodegenerative disease. EMBO Rep. *23*, e54217. https://doi.org/10.15252/EMBR.202154217.

57. Gelot, C., Guirouilh-Barbat, J., Le Guen, T., Dardillac, E., Chailleux, C., Canitrot, Y., and Lopez, B.S. (2016). The Cohesin Complex Prevents the End Joining of Distant DNA Double-Strand Ends. Mol. Cell *61*, 15–26. https://doi.org/10.1016/J.MOLCEL.2015.11.002.

58. Alder, J., Thakker-Varia, S., Bangasser, D.A., Kuroiwa, M., Plummer, M.R., Shors, T.J., and Black, I.B. (2003). Brain-Derived Neurotrophic Factor-Induced Gene Expression Reveals Novel Actions of VGF in Hippocampal Synaptic Plasticity. J. Neurosci. *23*, 10800–10808. https://doi.org/10.1523/JNEUROSCI.23-34-10800.2003.

59. Beckmann, N.D., Lin, W.J., Wang, M., Cohain, A.T., Charney, A.W., Wang, P., Ma, W., Wang, Y.C., Jiang, C., Audrain, M., et al. (2020). Multiscale causal networks identify VGF as a key regulator of Alzheimer's disease. Nat. Commun. *11*, 3942. https://doi.org/10.1038/S41467-020-17405-Z.

60. Xiao, M.F., Xu, D., Craig, M.T., Pelkey, K.A., Chien, C.C., Shi, Y., Zhang, J., Resnick, S., Pletnikova, O., Salmon, D., et al. (2017). NPTX2 and cognitive dysfunction in Alzheimer's Disease. Elife *6*, e23798. https://doi.org/10.7554/ELIFE.23798.

61. Cain, A., Taga, M., McCabe, C., Green, G., Hekselman, I., White, C.C., Lee, D.I., Gaur, P., Rozenblatt-Rosen, O., Zhang, F., et al. (2022). Multi-cellular communities are perturbed in the aging human brain and Alzheimer's disease. Preprint at bioRxiv.. https://doi.org/10.1101/2020.12.22.424084.

62. Consens, M.E., Chen, Y., Menon, V., Wang, Y., Schneider, J.A., De Jager, P.L., Bennett, D.A., Tripathy, S.J., and Felsky, D. (2022). Bulk and Single-Nucleus Transcriptomics Highlight Intra-Telencephalic and Somatostatin Neurons in Alzheimer's Disease. Front. Mol. Neurosci. *15*, 903175. https://doi.org/10.3389/FNMOL.2022.903175.

63. Davies, P., Katzman, R., and Terry, R.D. (1980). Reduced somatostatin-like immunoreactivity in cerebral cortex from cases of Alzheimer disease and Alzheimer senile dementa. Nature *288*, 279–280. https://doi.org/10.1038/288279A0.

64. Loomba, S., Straehle, J., Gangadharan, V., Heike, N., Khalifa, A., Motta, A., Ju, N., Sievers, M., Gempt, J., Meyer, H.S., and Helmstaedter, M. (2022). Connectomic comparison of mouse and human cortex. Science *377*, eabo0924. https://doi.org/10.1126/SCIENCE.ABO0924.

65. Krienen, F.M., Goldman, M., Zhang, Q., C H Del Rosario, R., Florio, M., Machold, R., Saunders, A., Levandowski, K., Zaniewski, H., Schuman, B., et al. (2020). Innovations present in the primate interneuron repertoire. Nature *586*, 262–269. https://doi.org/10.1038/S41586-020-2781-Z.

66. Bakken, T.E., Jorstad, N.L., Hu, Q., Lake, B.B., Tian, W., Kalmbach, B.E., Crow, M., Hodge, R.D., Krienen, F.M., Sorensen, S.A., et al. (2021). Comparative cellular analysis of motor cortex in human, marmoset and mouse. Nat *598*, 111–119. https://doi.org/10.1038/s41586-021-03465-8.

67. Deng, Y., Bi, M., Delerue, F., Forrest, S.L., Chan, G., Van Der Hoven, J., Van Hummel, A., Feiten, A.F., Lee, S., Martinez-Valbuena, I., et al. (2022). Loss of LAMP5 interneurons drives neuronal network dysfunction in Alzheimer's disease. Acta Neuropathol. *144*, 637–650. https://doi.org/10.1007/S00401-022-02457-W.

68. Sohal, V.S., Zhang, F., Yizhar, O., and Deisseroth, K. (2009). Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. Nat *459*, 698–702. https://doi.org/10.1038/nature07991.

69. Cardin, J.A., Carlén, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L.H., and Moore, C.I. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. Nature *459*, 663–667. https://doi.org/10.1038/NATURE08002.

70. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888–1902.e21. https://doi.org/10.1016/J.CELL.2019.05.031.

71. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: Large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15. https://doi.org/10.1186/S13059-017-1382-0/FIGURES/1.

72. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. *10*, 1523. https://doi.org/10.1038/s41467-019-09234-6.

73. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

74. Boyle, P.A., Yu, L., Leurgans, S.E., Wilson, R.S., Brookmeyer, R., Schneider, J.A., and Bennett, D.A. (2019). Attributable risk of Alzheimer's dementia attributed to age-related neuropathologies. Ann. Neurol. *85*, 114–124. https://doi.org/10.1002/ANA.25380.

75. Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell Syst. *8*, 281–291.e9. https://doi.org/10.1016/J.CELS.2018.11.005.

76. Risso, D., Ngai, J., Speed, T.P., and Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. Nat. Biotechnol. *32*, 896–902. https://doi.org/10.1038/nbt.2931.

77. Shen, L.; Sinai ISoMaM (2023). GeneOverlap: Test and visualize gene overlaps. R package version 1.36.0. http://shenlab-sinai.github.io/shenlab-sinai/.

78. Wehrens, R., and Buydens, L.M.C. (2007). Self- and Super-organizing Maps in R: The kohonen Package. J. Stat. Softw. *21*, 1–19. https://doi.org/10.18637/JSS.V021.I05.

79. Dewey, M. (2022). metap: meta-analysis of significance values. R package version 1.8.

80. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849. https://doi.org/10.1093/bioinformatics/btw313.

81. Bennett, D.A., Wilson, R.S., Schneider, J.A., Evans, D.A., Beckett, L.A., Aggarwal, N.T., Barnes, L.L., Fox, J.H., J, B., and Bach, J. (2002). Natural history of mild cognitive impairment in older persons. Neurology *59*, 198–205. https://doi.org/10.1212/WNL.59.2.198.

82. Bennett, D.A., Schneider, J.A., Wilson, R.S., Bienias, J.L., and Arnold, S.E. (2004). Neurofibrillary tangles mediate the association of amyloid load with clinical Alzheimer disease and level of cognitive function. Arch. Neurol. *61*, 378–384. https://doi.org/10.1001/ARCHNEUR.61.3.378.

83. Bennett, D.A., Schneider, J.A., Tang, Y., Arnold, S.E., and Wilson, R.S. (2006). The effect of social networks on the relation between Alzheimer's disease pathology and level of cognitive function in old people: a longitudinal cohort study. Lancet Neurol. *5*, 406–412. https://doi.org/10.1016/S1474-4422(06)70417-3.

84. Wilson, R.S., Boyle, P.A., Yu, L., Barnes, L.L., Sytsma, J., Buchman, A.S., Bennett, D.A., and Schneider, J.A. (2015). Temporal course and pathologic basis of unawareness of memory loss in dementia. Neurology *85*, 984–991. https://doi.org/10.1212/WNL.0000000000001935.

85. Boyle, P.A., Wang, T., Yu, L., Wilson, R.S., Dawe, R., Arfanakis, K., Schneider, J.A., Beck, T., Rajan, K.B., Evans, D., and Bennett, D.A. (2021). The "cognitive clock": A novel indicator of brain health. Alzheimers Dement. *17*, 1923–1937. https://doi.org/10.1002/ALZ.12351.

86. Buchman, A.S., Yu, L., Klein, H.-U., Zammit, A.R., Oveisgharan, S., Grodstein, F., Tasaki, S., Levey, A.I., Seyfried, N.T., and Bennett, D.A. (2022). Proteome-wide discovery of cortical proteins that may provide motor resilience to offset the negative effects of pathologies in older adults. J. Gerontol. A. Biol. Sci. Med. Sci. *78*, 494–503. https://doi.org/10.1093/GERONA/GLAC105.

87. Bai, Y., Wang, J., Song, R., Wang, Z., Qi, X., Buchman, A.S., Bennett, D.A., and Xu, W. (2022). Influence of cardiovascular risk burden on pulmonary function trajectory: role of physical and social activities. Aging (Albany. NY) *14*, 6081–6093. https://doi.org/10.18632/AGING.204201.

88. Grodstein, F., Wang, T., Leurgans, S.E., Wilson, R.S., and Bennett, D.A. (2022). Modifiable psychosocial risk factors and delayed onset of dementia in older populations: analysis of two prospective US cohorts. BMJ Open *12*, e059317. https://doi.org/10.1136/BMJOPEN-2021-059317.

89. Wilson, R.S., Wang, T., Yu, L., Grodstein, F., Bennett, D.A., and Boyle, P.A. (2021). Cognitive Activity and Onset Age of Incident Alzheimer Disease Dementia. Neurology *97*, e922–e929. https://doi.org/10.1212/WNL.0000000000012388.

90. Wilson, R.S., Segawa, E., Hizel, L.P., Boyle, P.A., and Bennett, D.A. (2012). Terminal dedifferentiation of cognitive abilities. Neurology *78*, 1116–1122. https://doi.org/10.1212/WNL.0B013E31824F7FF2.

91. Arvanitakis, Z., Schneider, J.A., Wilson, R.S., Li, Y., Arnold, S.E., Wang, Z., and Bennett, D.A. (2006). Diabetes is related to cerebral infarction but not to AD pathology in older persons. Neurology *67*, 1960–1965. https://doi.org/10.1212/01.WNL.0000247053.45483.4E.

92. Boyle, P.A., Wang, T., Yu, L., Wilson, R.S., Dawe, R., Arfanakis, K., Schneider, J.A., and Bennett, D.A. (2021). To what degree is late life cognitive decline driven by age-related neuropathologies? Brain *144*, 2166–2175. https://doi.org/10.1093/BRAIN/AWAB092.

93. Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. *8*, 14049. https://doi.org/10.1038/ncomms14049.

94. Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. *20*, 296. https://doi.org/10.1186/S13059-019-1874-1.

95. Lake, B.B., Chen, S., Sos, B.C., Fan, J., Kaeser, G.E., Yung, Y.C., Duong, T.E., Gao, D., Chun, J., Kharchenko, P.V., and Zhang, K. (2018). Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. Nat. Biotechnol. *36*, 70–80. https://doi.org/10.1038/nbt.4038.

96. Kierdorf, K., Masuda, T., Jordão, M.J.C., and Prinz, M. (2019). Macrophages at CNS interfaces: ontogeny and function in health and disease. Nat. Rev. Neurosci. *20*, 547–562. https://doi.org/10.1038/s41583-019-0201-x.

97. Habib, N., Avraham-Davidi, I., Basu, A., Burks, T., Shekhar, K., Hofree, M., Choudhury, S.R., Aguet, F., Gelfand, E., Ardlie, K., et al. (2017). Massively parallel single-nucleus RNA-seq with DroNc-seq. Nat. Methods *14*, 955–958. https://doi.org/10.1038/nmeth.4407.

98. Vanlandewijck, M., He, L., Mäe, M.A., Andrae, J., Ando, K., Del Gaudio, F., Nahar, K., Lebouvier, T., Laviña, B., Gouveia, L., et al. (2018). A molecular atlas of cell types and zonation in the brain vasculature. Nature *554*, 475–480. https://doi.org/10.1038/NATURE25739.

99. Velmeshev, D., Schirmer, L., Jung, D., Haeussler, M., Perez, Y., Mayer, S., Bhaduri, A., Goyal, N., Rowitch, D.H., and Kriegstein, A.R. (2019). Single-cell genomics identifies cell type–specific molecular changes in autism. Science *364*, 685–689. https://doi.org/10.1126/SCIENCE.AAV8130.

100. Crowell, H.L., Soneson, C., Germain, P.L., Calini, D., Collin, L., Raposo, C., Malhotra, D., and Robinson, M.D. (2020). muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. Nat. Commun. *11*, 6077. https://doi.org/10.1038/s41467-020-19894-4.

101. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140. https://doi.org/10.1093/BIOINFORMATICS/BTP616.

102. Zhao, S., Li, C.I., Guo, Y., Sheng, Q., and Shyr, Y. (2018). RnaSeqSampleSize: Real data based sample size estimation for RNA sequencing. BMC Bioinf. *19*, 191. https://doi.org/10.1186/S12859-018-2191-5/FIGURES/5.

103. Hoffman, G.E., Lee, D., Bendl, J., Fnu, P., Hong, A., Casey, C., Alvia, M., Shao, Z., Argyriou, S., Therrien, K., et al. (2023). Efficient differential expression analysis of large-scale single cell transcriptomics data using dreamlet. Preprint at bioRxiv.. https://doi.org/10.1101/2023.03.17.533005.

104. Kohonen, T. (1990). The Self-Organizing Map. Proc. IEEE *78*, 1464–1480. https://doi.org/10.1109/5.58325.

105. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

106. Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420. https://doi.org/10.1038/nbt.4096.

107. Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502. https://doi.org/10.1038/nbt.3192.

108. Liu, Y., and Xie, J. (2020). Cauchy combination test: a powerful test with analytic p-value calculation under arbitrary dependency structures. J. Am. Stat. Assoc. *115*, 393–402. https://doi.org/10.1080/01621459.2018.1554485.

109. Li, X., Li, Z., Zhou, H., Gaynor, S.M., Liu, Y., Chen, H., Sun, R., Dey, R., Arnett, D.K., Aslibekyan, S., et al. (2020). Dynamic incorporation of multiple in silico functional annotations empowers rare variant association analysis of large whole genome sequencing studies at scale. Nat. Genet. *52*, 969–983. https://doi.org/10.1038/S41588-020-0676-4.

110. Li, Z., Li, X., Zhou, H., Gaynor, S.M., Selvaraj, M.S., Arapoglou, T., Quick, C., Liu, Y., Chen, H., Sun, R., et al. (2022). A framework for detecting noncoding rare variant associations of large-scale whole-genome sequencing studies. Nat. Methods *19*, 1599–1611. https://doi.org/10.1038/S41592-022-01640-X.

111. Chen, Y., Lun, A.T.L., and Smyth, G.K. (2016). From reads to genes to pathways: Differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. F1000Research.. https://doi.org/10.12688/F1000RESEARCH.8987.2.

112. McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. *40*, 4288–4297. https://doi.org/10.1093/nar/gks042.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-NPTX2 antibody	Abcam	ab69858; RRID:AB_1524045
Anti-VGF antibody	Abcam	ab69989; RRID:AB_1271455
Anti-CHCHD6	Atlas Antibodies	HPA047673; RRID:AB_10961761
Anti-BDH1	Atlas Antibodies	HPA030947; RRID:AB_10669548
Anti-NeuN, clone A60, Alexa Fluor	EMD Millipore	MAB377X; RRID:AB_2149209
Anti-Somatostatin antibody	Abcam	ab108456; RRID:AB_11158517
Anti-GFAP	Abcam	Ab53554; RRID:AB_880202
Anti-NeuN	Synaptic Systems	266 004; RRID:AB_2619988
Donkey anti-Rabbit, Alexa Fluor 488	Invitrogen	A21206
Donkey anti-Mouse, Alexa Fluor 488	Invitrogen	A21202
Donkey anti-Rabbit, Alexa Fluor 594	Invitrogen	A11037
Donkey anti-Mouse, Alexa Fluor 594	Invitrogen	A11032
Donkey anti-Guinea pig, Alexa Fluor 647	Invitrogen	A21447
Biological samples		
Frozen postmortem prefrontal cortex brain tissue	Rush Alzheimer's Disease Center (RADC) (https://www.radc.rush.edu/)	N/A
Chemicals, peptides, and recombinant proteins		
D-Sucrose	Sigma-Aldrich	S7903-5KG
Calcium chloride solution	Sigma-Aldrich	21115-100ML
Magnesium acetate solution	Sigma-Aldrich	63052-100ML
Trizma® hydrochloride solution	Sigma-Aldrich	T2569-100ML
UltraPure™ 0.5M EDTA, pH 8.0	Thermo Fisher Scientific	15575–020
Igepal® CA-630	Sigma-Aldrich	I8896-50ML
2-Mercaptoethanol	Sigma-Aldrich	M6250-10ML
Recombinant RNase Inhibitor	Clontech	2313B
OptiPrep™ Density Gradient Medium	Sigma-Aldrich	D1556-250ML
Phosphate-Buffered Saline (PBS)	Fisher Scientific	SH3025601
BSA Fraction V, Protease-free	Sigma-Aldrich	03117332001
Tissue-Tek* O.C.T. Compound	VWR	25608–930
Gill's Hematoxylin	StatLab	HXGHE1LT
NucBlue™ Live ReadyProbes™ Reagent (Hoechst 33342)	ThermoFisher Scientific	R37605
eBioscience™ IHC Antigen Retrieval Solution - Low pH (10X)	ThermoFisher Scientific	00-4955-58
Fluromount-G	Electron microscopic Sciences	Cat#17984-25
Critical commercial assays		
Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3, 16 rxns	10X Genomics	PN-1000075
Chromium Single Cell B Chip Kit, 48 rxns	10X Genomics	PN-1000073
Chromium i7 Multiplex Kit	10X Genomics	PN-120262
Hoechst 33342, Trihydrochloride, Trihydrate - 10 mg/mL Solution in Water	Thermo Fisher Scientific	H3570

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RNAscope® 2.5 HD Duplex Reagent Kit, is an RNA *in situ* hybridization (ISH) chromogenic assay designed for multiplexing up to two targets.	Advanced Cell Diagnostics	322430
RNAscope® Probe - Hs-NPTX2-C2 - Homo sapiens neuronal pentraxin II (NPTX2) mRNA	Advanced Cell Diagnostics	464831-C2
RNAscope® Probe - Hs-SST - Homo sapiens somatostatin (SST), mRNA	Advanced Cell Diagnostics	310591
RNAscope® Probe - Hs-GAD2-C2 - Homo sapiens glutamate decarboxylase 2 (pancreatic islets and brain 65kDa) (GAD2) transcript variant 1 mRNA	Advanced Cell Diagnostics	415621-C2
RNAscope® Probe - Hs-LAMP5-C2 - Homo sapiens lysosomal associated membrane protein family member 5 (LAMP5) transcript variant 1 mRNA	Advanced Cell Diagnostics	487691-C2
RNAscope HiPlex Probe - Hs-SLC17A7-T1	Advanced Cell Diagnostics	415611-T1
RNAscope HiPlex Probe - Hs-RORB-T5	Advanced Cell Diagnostics	446061-T5
RNAscope HiPlex Probe - Hs-GAD2-T5	Advanced Cell Diagnostics	415621-T5
RNAscope HiPlex Probe - Hs-LAMP5-T8	Advanced Cell Diagnostics	487691-T8
RNAscope HiPlex Probe - Hs-CBLN2-T2	Advanced Cell Diagnostics	446051-T2
RNAscope HiPlex Probe - Hs-CUX2-T2	Advanced Cell Diagnostics	425581-T2
RNAscope HiPlex Probe - Hs-RELN-T7	Advanced Cell Diagnostics	413051-T7
RNAscope HiPlex Probe - Hs-SST-T1	Advanced Cell Diagnostics	310591-T1
RNAscope HiPlex Probe - Hs-HNRNPA2B1-T4	Advanced Cell Diagnostics	1270571-T4
RNAscope HiPlex Probe - Hs-RPH3A-T7	Advanced Cell Diagnostics	866721-T7
RNeasy Plus Mini Kit	QIAGEN	74134
iScript™ cDNA Synthesis Kit	Bio-Rad	1708891
SsoFast™ EvaGreen® Supermix	Bio-Rad	1725202
HiPlex12 Reagents Kit (488, 550; 650) Hs V2	Advanced Cell Diagnostics	324445
Deposited data		
snRNA-seq data	This paper	Synapse: syn52293417 (https://www.synapse.org/#!Synapse:syn52293417)
Differential expression analysis results	This paper	https://github.com/mathyslab7/ROSMAP_snRNAseq_PFC
De Jager dataset (snRNA-seq data)	Synapse	https://www.synapse.org/#!Synapse:syn51123521
SEA-AD (MTG) dataset (snRNA-seq data)	Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD)	https://sea-ad-single-cell-profiling.s3.amazonaws.com/index.html#MTG/RNAseq/
CORUM: the comprehensive resource of mammalian protein complexes	Giurgiu et al.[33]	https://www.uniprot.org/database/DB-0224
Oligonucleotides		
Primers for RT-qPCR, see STAR Methods	This paper	N/A
Software and algorithms		
Imairs	Imarisx64 8.1.2 (Bitplane, Zurich, Switzerland	
ImageJ	ImageJ 1.42q	
Prism 9	GraphPad	RRID:SCR_002798
R	N/A	RRID:SCR_001905
Seurat	Stuart et al.[70]	RRID:SCR_016341
scanpy	Wolf et al.[71]	RRID:SCR_018139

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BioRender	BioRender	www.biorender.com
Metascape	Zhou et al.[72]	RRID:SCR_016620
Cell Ranger	10x Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest
DESeq2	Love et al.[73]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html; RRID: SCR_015687
muscat	Crowell et al.[74]	https://www.bioconductor.org/packages/release/bioc/html/muscat.html
Scrublet	Wolock et al.[75]	https://github.com/AllonKleinLab/scrublet
RUVSeq	Risso et al.[76]	https://bioconductor.org/packages/release/bioc/html/RUVSeq.html
GeneOverlap	Shen L.[77]	https://bioconductor.org/packages/release/bioc/html/GeneOverlap.html
SynGO	Koopmans et al.[29]	https://www.syngoportal.org/
kohonen	Wehrens and Buydens[78]	https://cran.r-project.org/web/packages/kohonen/index.html
metap	Dewey[79]	https://cran.r-project.org/web/packages/metap/index.html
Hmisc	N/A	https://cran.r-project.org/web/packages/Hmisc/index.html
ComplexHeatmap	Gu et al.[80]	RRID:SCR_017270
Code used in this study	This paper	https://github.com/mathyslab7/ROSMAP_snRNAseq_PFC
Other		
LSM 880 confocal microscopes (Zeiss)	Zeiss	20× objective
LSM 900 confocal microscopes (Zeiss)	Zeiss	10× or 20× objectives
Leica VT1000S	Leica	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Li-Huei Tsai (lhtsai@mit.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Single-cell RNA-seq data will be available through the AD Knowledge Portal on Synapse upon publication of this manuscript. The data will be available here: https://www.synapse.org/#!Synapse:syn52293417. The data are available under controlled use conditions set by human privacy regulations. To access the data, a data use agreement is needed. This registration is in place solely to ensure the anonymity of the ROSMAP study participants. A data use agreement can be agreed with either Rush University Medical Center (RUMC) or with SAGE, which maintains Synapse, and can be downloaded from their websites (https://adknowledgeportal.synapse.org/). The single-cell RNA-seq data from our study can be explored on the UCSC Cell Browser portal (https://cells.ucsc.edu/?ds=ad-aging-brain). The results of this study are also available via the AD and aging brain atlas data repository (http://compbio.mit.edu/ad_aging_brain/).
- All original code has been deposited on GitHub and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human subjects

We selected 427 individuals from the Religious Orders Study and Rush Memory and Aging Project (ROSMAP), both ongoing longitudinal clinical-pathologic cohort studies of aging and dementia, in which all participants are brain donors. The studies include clinical data collected annually, detailed postmortem pathological evaluations, and extensive genetic, epigenomic, transcriptomic, proteomic, and metabolomic bulk-tissue profiling.[19] Individuals were balanced between sexes (male:female ratio 212:215). Among 427 study participants, 425 identified as White, and 2 identified as African American. To the question, "Are you of Spanish/Hispanic/Latino origin?", 423 participants answered "No," while 4 participants answered "Yes." Socioeconomic status of the participants was not determined for this study. Informed consent was obtained from each subject, and the Religious Orders Study and Rush Memory and Aging Project were each approved by an Institutional Review Board (IRB) of Rush University Medical Center. Participants also signed an Anatomic Gift Act, and a repository consent to allow their data to be repurposed.

METHOD DETAILS

Clinical data

A global measure of cognition was based on the average *Z* score of 17 cognitive performance tests proximate to death using the mean and SD from the baseline of all participants. Subsets of these tests were averaged to yield individual measures of episodic memory, perceptual orientation, perceptual speed, semantic memory, and working memory. Medical conditions including diabetes, hypertension, and stroke are documented. Motor function and gait are assessed. Cognitive, physical, and social lifestyle variables are documented (Figure 2A). Methods for these variables have been previously reported.[19,81–91]

Pathologic AD and other pathologies

For categorical analyses comparing individuals with and without pathologic AD, the cohort was dichotomized based on the modified NIA-Reagan diagnosis of AD which reflects a combination of Braak Score and CERAD AD, with 238 individuals having AD and 189 individuals without. A global burden of AD was based on standardized counts of neuritic and diffuse plaques, and neurofibrillary tangles. Other pathologies included Lewy bodis, TDP-43 pathology, hippocampal sclerosis, and vascular pathologies, i.e., macro- and microscopic infarcts, athero- and arteriolar-sclerosis, amyloid angiopathy. Methods for these pathologies have been previously reported.[19,52,81–84,92,74]

Cognitive resilience

To quantify cognitive resilience, we computed a cognitive resilience score (CR score) as the difference between the observed cognition and the cognition predicted by a linear regression model, given the level of pathology (Methods S1, page 5A and 5B). Using this approach, we computed Cognitive Resilience (CR) scores based on the measure of global cognitive function and Cognitive Decline Resilience (CDR) scores based on the measure estimating the rate of change of global cognitive function over time (Methods S1, page 5A and 5B). Four distinct CR and CDR scores were derived based on four distinct measures of AD pathology, namely global AD pathology, and separately neuritic plaque burden, neurofibrillary tangle burden, and tangle density.

Isolation of nuclei from frozen postmortem brain tissue

The protocol for the isolation of nuclei from frozen postmortem brain tissue was adapted from a previous study.[22] All procedures were carried out on ice or at 4°C. In brief, postmortem brain tissue was homogenized in 700 μL homogenization buffer (320 mM sucrose, 5 mM $CaCl_2$, 3 mM $Mg(CH_3COO)_2$, 10 mM Tris HCl pH 7.8, 0.1 mM EDTA pH 8.0, 0.1% IGEPAL CA-630, 1 mM β-mercaptoethanol, and 0.4 U μl^{-1} recombinant RNase inhibitor (Clontech)) using a Wheaton Dounce tissue grinder (15 strokes with the loose pestle). Then the homogenized tissue was filtered through a 40-μm cell strainer, mixed with an equal volume of working solution (50% OptiPrep density gradient medium (Sigma-Aldrich), 5 mM $CaCl_2$, 3 mM $Mg(CH_3COO)_2$, 10 mM Tris HCl pH 7.8, 0.1 mM EDTA pH 8.0, and 1 mM β-mercaptoethanol) and loaded on top of an OptiPrep density gradient (750 μL 30% OptiPrep solution (30% OptiPrep density gradient medium,134 mM sucrose, 5 mM $CaCl_2$, 3 mM $Mg(CH_3COO)_2$, 10 mM Tris HCl pH 7.8, 0.1 mM EDTA pH 8.0, 1 mM β-mercaptoethanol, 0.04% IGEPAL CA-630, and 0.17 U μl^{-1} recombinant RNase inhibitor) on top of 300 μL 40% OptiPrep solution (40% OptiPrep density gradient medium, 96 mM sucrose, 5 mM $CaCl_2$, 3 mM $Mg(CH_3COO)_2$, 10 mM Tris HCl pH 7.8, 0.1 mM EDTA pH 8.0, 1 mM β-mercaptoethanol, 0.03% IGEPAL CA-630, and 0.12 U μl^{-1} recombinant RNase inhibitor). The nuclei were separated by centrifugation (5 min, 10,000 g, 4°C). A total of 100 μL of nuclei was collected from the 30%/40% interphase and washed with 1 mL of PBS containing 0.04% BSA. The nuclei were centrifuged at 300*g* for 3 min (4°C) and washed with 1 mL of PBS containing 0.04% BSA. Then the nuclei were centrifuged at 300*g* for 3 min (4°C) and re-suspended in 100 μL PBS containing 0.04% BSA. The nuclei were counted and diluted to a concentration of 1,000 nuclei per microliter in PBS containing 0.04% BSA.

Single-nucleus RNA-seq

Droplet-based snRNA-seq

For droplet-based snRNA-seq, libraries were prepared using the Chromium Single Cell 3′ Reagent Kits v3 according to the manufacturer's protocol (10x Genomics). The generated snRNA-seq libraries were sequenced using NextSeq 500/550 High Output v2 kits (150 cycles) or NovaSeq 6000 S2 Reagent Kits.

cDNA synthesis and RT-qPCR analysis

Postmortem brain tissue was homogenized in RNA lysis buffer (Qiagen, 74134) and total RNA was extracted using the RNeasy Plus Mini kit (Qiagen, 74134) according to manufacturer's instruction. The RNA was reverse transcribed using the iScript cDNA Synthesis Kit (Bio-Rad, 170–8891). For gene expression analysis cDNA was quantitatively amplified on a thermal cycler (BioRad) using SsoFast EvaGreen Supermix (BioRad, 1725202) and gene-specific primers (HES4 forward, CCACTCGAAGCTGGAGAAGG; HES4 reverse, CGGCCGTCACCTGCAC; PDE10A forward, GACTCCTTGGGTTTTCCGGG; PDE10A reverse, TGAGGAAACACTCGGTCAGC; ST6GAL2 forward, TGCGCATTCCTGCATTCCTG; ST6GAL2 reverse, AATGCAGATGGGTGGCAGAA; GAD2 forward, GTCTGCCTC GCATCCTCACG; GAD2 reverse TTAGGGACGTGGCAACCCTT; LAMP5 forward, TTTGTGGTGCGGGAAAATGG; LAMP5 reverse, ATGCGATATCGGCCTGTTCT; SST forward, GCCGCTTTAGGAGCGAGGTT; SST reverse, AACTGACGGAGTCTGGGGTC; RPL13 forward, CGGCATTCACAAGAAGGGAGA; RPL13 reverse, TACGGAGACTAGCGAAGGCT). The comparative Ct method was used to examine differences in gene expression and values were normalized to expression levels of RPL13.

Immunohistochemistry

Fresh frozen human brain tissue (prefrontal cortex) of approximately 0.1 cm^3 in volume was immersed in 50 mL 4% paraformaldehyde (4% PFA) in 1X PBS (phosphate buffered saline) for 36 h at 4°C on a shaker before being embedded into 3% low melting point agarose. The embedded brain tissue was then sectioned at 30 μm using a vibratome (Leica). To retrieve the antigens, sections were incubated at 95°C in immunohistochemistry (IHC) antigen retrieval solution (ThermoFisher Scientific; 00-4955-58) containing 0.05% Tween 20 for 45 min and then placed in PBS for 20 min at room temperature. After washing with PBS (3 × 10 min), the brain sections were incubated in quenching solution (50 mM ammonium acetate, 100 mM $CuSO_4$) at room temperature overnight. After washing with 1X PBS (3 × 15 min), the sections were permeabilized in PBS containing 0.3% Triton X-100 for 10 min and blocked in PBS containing 0.3% Triton X-100 and 5% normal donkey serum at room temperature for 1 h. The sections were incubated for 48 h at 4°C in primary antibody in PBS with 0.3% Triton X-100 and 5% normal donkey serum. Primary antibodies were anti-VGF antibody (1:200; Abcam; ab69989), anti-NPTX2 antibody (1:200; Abcam; ab69858), and anti-Somatostatin antibody (1:200, Abcam, ab108456). The sections were washed with PBS containing 0.1% Triton X-100 at room temperature (4 × 15 min), and then incubated with secondary antibodies (dilution 1:2,000) overnight at 4°C. Primary antibodies were visualized with Alexa Fluor 488, Alexa Fluor 594, and Alexa Fluor 647 antibodies (Molecular Probes), and cell nuclei were visualized with Hoechst 33342 (Sigma-Aldrich; 94403). The sections were washed with PBS containing 0.1% Triton X-100 at room temperature (4 × 15 min) and then mounted on Fisherbrand Superfrost Plus microscope slides in fluromount-G (Electron microscopic Sciences). Images were acquired using a confocal microscope (LSM 880; Zeiss) with a 20X objective. z stack images were taken with a step size of 0.15–0.3 mm, averaging of 2, resolution 2048x2048 suitable for 3D reconstruction. Imarisx64 8.3.1 (Bitplane, Zurich, Switzerland) software was used for 3-D rendering and quantification of fluorescence images.

RNA *in situ* hybridization

Frozen brain tissue was embedded in Tissue-Tek OCT compound (VWR; 25608-930), cryo-sectioned to 20-μm thickness, and placed onto Fisherbrand Superfrost Plus microscope slides (ThermoFisher Scientific; 12-550-15). The RNAscope 2.5 HD Duplex Detection Kit (Chromogenic) was used according to the manufacturer's instructions, with minor modifications to tissue preparation. In brief, sections frozen at −80°C were fixed in 4% PFA for 12 h at 4°C. Following fixation, the tissue was dehydrated using 50%, 70%, and 100% ethanol and then baked at 60°C for 30 min to avoid detachment. The tissue was then treated with H_2O_2 for 10 min at room temperature, followed by protease IV for 30 min. Brain sections were hybridized with the following RNAscope probes: GAD2 (RNAscope Probe - Hs-GAD2-C2; ACD; 415621-C2), SST (RNAscope Probe - Hs-SST; ACD; 310591), LAMP5 (RNAscope Probe - Hs-LAMP5-C2; ACD; 487691-C2), NPTX2 (RNAscope Probe - Hs-NPTX2-C2; ACD; 464831-C2). Probes in channel 1 were labeled with HRP enzyme and visualized with a green/cyan substrate, and probes in channel 2 were labeled with alkaline phosphatase enzyme and a red substrate. Haematoxylin was used to mark cell nuclei in blue. Slides were then imaged using a confocal microscope (LSM 900; Zeiss) to take 10X and 20X bright-field images, with four images acquired per individual. For quantifying the number of nuclei and the number of GAD2-positive cells, the 'analyze particles' function in ImageJ 1.42q was used. For quantifying the number of SST-positive and LAMP5-positive cells, a researcher manually counted the cells labeled with green/cyan (SST) or red (LAMP5) substrate, respectively.

Isolation of neuronal nuclei, cDNA synthesis and RT-qPCR analysis

Postmortem brain tissue was homogenized in 2 mL NF1 buffer (0.5% Triton X-100, 0.1 M sucrose, 5 mM MgCl2, 1 mM EDTA, 10 mM Tris HCl pH 8.0, 1 mM β-mercaptoethanol, 0.4 U μl^{-1} recombinant RNase inhibitor (Clontech)) using a Wheaton Dounce tissue grinder (15 strokes with the loose pestle). After homogenization, an additional 8 mL of NF1 buffer was added and the samples

were inverted 10 times to mix. The homogenate was passed through a 40-μm cell strainer (VWR; 21008-949) and centrifuged at 300g for 3 min at 4°C. Then the supernatant was removed, and the nuclei were incubated with Alexa-Fluor-488-conjugated anti-NeuN antibody (Millipore Sigma; MAB377X) in PBS containing 1% BSA for 15 min at 4°C on a shaker. The nuclei were washed by adding 35 mL PBS containing 1% BSA, spun down at 300g for 3 min at 4°C, resuspended in 0.5 mL PBS containing 0.04% BSA and 0.4 U μl−1 recombinant RNase inhibitor (Clontech), and stained with NucBlue Live ReadyProbes Reagent (ThermoFisher Scientific; R37605). NeuN-positive nuclei were then directly sorted into RNA lysis buffer (Qiagen, 74134) using fluorescence-activated cell sorting. Total RNA was extracted using the RNeasy Plus Mini Kit (Qiagen; 74134) according to the manufacturer's instructions, and reverse transcribed using the iScript cDNA Synthesis Kit (Bio-Rad; 170–8891). For gene-expression analysis, cDNA was quantitatively amplified on a thermal cycler (BioRad) using SsoFast EvaGreen Supermix (BioRad; 1725202) and gene-specific primers (Table S1).

RNAscope and HiPlex

Tissue blocks from prefrontal cortex were cryosection at 12um thickness on the cryostat (Leica CM3050S). Sections were directly mounted on Superfrost Plus Slides (Fisherbrand; Cat#12-550-15) and allowed to dry at −20 celsius for 1 to 2 h. Tissue was then processed following the recommended protocol by ACD Biotechne RNAscope assays. To quench autofluorescence, we applied TrueBlack (Biotium; Cat#23007) for 90 s after the final step (DAPI application) of the RNAscope assay procedure, and gently rinsed with PBS prior to coverslipping and imaging. For HiPlex multiplexing, we reapplied DAPI and TrueBlack at the end of every round. 2x2 tile scans with 12z steps were acquired with a confocal microscope (LSM 900; Zeiss) with 63× oil immersion objective. Maximum projection images were manually aligned on Photoshop (Adobe Photoshop 2021) with DAPI nuclei stain and tissue boundaries serving as guides for registration. Excitatory neurons were then identified based on probe detection for SLC17A7 (VGlut1), while inhibitory cells were labeled with the inhibitory marker GAD2. After manually assigning neuronal identity as excitatory or inhibitory (greater than 7 dots). Manual curation of images was performed to ensure accurate probe detection by an investigator blind to the sample group.

QUANTIFICATION AND STATISTICAL ANALYSIS

snRNA-seq processing, QC, and annotation

snRNA-seq data preprocessing

Gene counts were obtained by aligning reads to the GRCh38 genome using Cell Ranger software (v.3.0.2) (10x Genomics).[93] To account for unspliced nuclear transcripts, reads mapping to pre-mRNA were counted. After quantification of pre-mRNA using the Cell Ranger count pipeline, the Cell Ranger aggr pipeline was used to aggregate all libraries (without equalizing the read depth between groups) to generate a gene-count matrix. The Cell Ranger 3.0 default parameters were used to call cell barcodes. We used SCANPY[71] to process and cluster the expression profiles and infer cell identities of major cell classes. To remove doublets and poor-quality cells, cells were excluded from subsequent analysis if they were extreme outliers (observations outside the range [Q1 – k(Q3 – Q1),Q3 + k(Q3-Q1)], with k = 3 and Q1 and Q3 as the lower and upper quartiles) in terms of number of genes, number of unique molecular identifiers (UMIs), and percentage of mitochondrial genes. In addition, we called doublets using Scrublet[75] and flagged and removed cells that were labeled as doublets, leaving a final dataset of 2,359,994 cells.

For the UMAP visualization of individual major cell type classes (excitatory neurons, inhibitory neurons, astrocytes, oligodendrocytes, OPCs, immune cells), the SCTransform-based integration workflow of Seurat[70,94] was used to align data from individual batches, using default settings. We selected the set of relevant principal components based on Elbow plots. High resolution cell clusters were identified using the Seurat functions FindNeighbors and FindClusters (using the Leiden clustering algorithm).

We annotated cell types using previously published marker genes and single-cell RNA-sequencing data.[20,22,95–99] To annotate cell types based on previously published single-cell RNA-sequencing data (Allen Institute's cell types database; https://portal.brain-map.org/atlases-and-data/rnaseq/human-multiple-cortical-areas-smart-seq), we used three separate approaches. First, Spearman rank correlation coefficients between the average expression profiles of neuronal subpopulations previously defined by the Allen Brain Institute[20] and the neuronal subtypes identified in this study were computed using the cor function in R. Second, to project annotations of neuronal subpopulations previously defined by the Allen Brain Institute onto the neuronal cells analyzed in this study, we followed the integration and label transfer workflow of Seurat.[70] Third, we determined cell type marker genes based on data published by the Allen Brain Institute[20] using the FindAllMarkers function from Seurat (Wilcoxon Rank-Sum test with Bonferroni correction for multiple testing; adjusted $p < 0.05$) and computed module scores for each cell type marker gene set across all neuronal cells analyzed in this study using the AddModuleScore function of Seurat.

To further annotate cell types, we determined marker genes using the FindAllMarkers function from Seurat (Wilcoxon Rank-Sum test with Bonferroni correction for multiple testing; adjusted $p < 0.05$). We tested only genes that were detected in a minimum of 25% of the cells within the cell type (min.pct = 0.25) and that showed, on average, at least a 0.25-fold difference (log-scale) between the cells of the cell type and all remaining cells (logfc.threshold = 0.25). Marker genes of the high-resolution cell types or states were determined separately for each major cell type class.

Differential gene expression analysis

Differential gene expression analysis using muscat

We performed differential expression analyses using the R package muscat.[100] Lowly expressed genes were excluded and only genes with more than one count in at least 10 cells were considered. To take advantage of robust bulk RNA-sequencing differential expression frameworks, such as edgeR,[101] in a first step muscat aggregates measurements for each sample (in each cluster) to obtain pseudobulk data. Using this approach single-cell measurements were aggregated per study participant and cell type using the sum of raw counts option. Differential expression analysis was run using the edgeR method as implemented in muscat. For the analysis of variables measuring AD pathology, cognition, or cognitive resilience, we included as covariates the individual's age at death and postmortem interval. For the analysis of measures of Lewy body and TDP-43 pathology, vascular pathology (macro- and microscopic infarcts, athero- and arteriolar-sclerosis, amyloid angiopathy), medical conditions (diabetes, hypertension, stroke), motor function and gait, and lifestyle variables, we were interested in identifying gene expression alterations that are specifically linked to the measures mentioned and not associated with AD pathology. Therefore, for the analysis of these measures we included global AD pathology, age at death, and the postmortem interval as covariates in the differential expression analysis. We report adjusted p values for multiple testing in all cases by using the p.adjust function with the BH method as implemented in muscat. The multiple testing correction was performed locally, i.e., on each of the cell types separately with the number of tests equal to the number of genes considered. We provide differential expression results for each of the 54 high resolution cell types, as well as for astrocytes and microglia. Differential expression analyses were performed on the entire cohort unless stated otherwise (e.g., when analyzing gene expression changes across multiple stages of Alzheimer's disease progression). We also tested whether the observed associations between gene expression and global AD pathology remain significant after accounting for additional potential confounding factors such as technical variability and sex. For this analysis we included as covariates the individual's age at death, postmortem interval, sex, and factors of unwanted variation as determined by the R package RUVseq.[76] The factors of unwanted variation were determined using the RUVr function with the argument k set to 10 factors of unwanted variation. Moreover, we employed DecontX[25] to estimate and remove ambient RNA from our snRNA-seq data. Subsequently, the differential expression analysis was carried out using muscat, utilizing the decontaminated counts as input. In this analysis, we included the individual's age at death, postmortem interval, sex, and factors of unwanted variation determined by RUVseq as covariates, as described above.

Group-based differential expression analysis

For the group-based differential expression analysis, we categorized individuals into four distinct groups based on categorical variables assessing Alzheimer's disease pathology and cognitive impairment. Initially, we divided the individuals into two groups based on their pathologic diagnosis of Alzheimer's disease (variable niareagansc). Within each of these two groups, we further selected individuals without cognitive impairment (final consensus cognitive diagnosis (cogdx) value of 1) and individuals with a cognitive diagnosis of AD dementia (cogdx value of 4). The following group comparisons were performed:

1) Group 3 (Pathologic diagnosis of AD and No cognitive impairment) versus Group 1 (No pathologic diagnosis of AD and No cognitive impairment)
2) Group 4 (Pathologic diagnosis of AD and Cognitive diagnosis of Alzheimer's dementia) versus Group 2 (No pathologic diagnosis of AD/Cognitive diagnosis of Alzheimer's dementia)
3) Group 4 (Pathologic diagnosis of AD and Cognitive diagnosis of Alzheimer's dementia) versus Group 1 (No pathologic diagnosis of AD and No cognitive impairment)

For the differential expression analysis, we utilized muscat after removing ambient RNA from our snRNA-seq data using DecontX. As covariates, we included the individual's age at death, postmortem interval, sex, and factors of unwanted variation determined by RUVseq, as described above.

Simulation-based power analysis

To estimate the power of our dataset in detecting changes in gene expression, we performed a simulation-based power analysis using the R package RnaSeqSampleSize.[102] This package employs the average read count and dispersion distributions derived from real data to estimate the likelihood of detecting statistical significance in the dataset (power estimation). For this analysis, we utilized individual-level aggregated cell type-specific count matrices, which were used for the subsequent differential expression analysis in this study. We employed the "est_power_distribution" function of the RnaSeqSampleSize package, specifying the following parameters: the number of samples in each group (n = 213), the minimal fold change between the two groups (rho = 1.25), the false discovery rate (f = 0.01), and the number of genes used in estimation of read counts and dispersion distribution (repNumber = 100).

Consideration and evaluation of covariates

To determine the set of covariates to be included in the differential expression analysis, we performed a variance partitioning analysis. This analysis enabled us to identify the covariates that substantially contribute to the expression variance and, consequently, are expected to have an impact on the results of the differential expression analysis. The variance partitioning analysis was used to estimate the fraction of expression variance attributable to known (technical) sources of variation and factors of unwanted variation estimated by RUVseq. The known sources of variation evaluated were: average total transcript (UMI) count per cell (nUMI), average

total gene count per cell (nGenes), average fraction of mitochondrial RNA (percent mito), average fraction of ribosomal RNA (percent ribo), number of cells (nCells), number of reads (nReads), and fraction of reads mapped to the genome (mapped Reads).

The variance partitioning analysis was performed using the function ''fitVarPart'' from the R package dreamlet (DiseaseNeurogenomics.github.io/dreamlet).[103] The median fraction of variance across genes attributed to each covariate was determined and the total fraction of variance attributed to a set of covariates was determined as the sum of the median variance attributed to each individual covariate of the set. The analysis included twenty sets of covariates, each progressively incorporating one additional covariate, with the purpose of evaluating whether the addition of a covariate increases the total fraction of variance explained. The 20 sets of covariates evaluated in the variance partitioning analysis were as follows.

1) pmi,
2) pmi+age,
3) pmi+age+sex,
4) pmi+age+sex+w_1,
5) pmi+age+sex+w_1 + w_2,
6) pmi+age+sex+w_1 + w_2 + w_3,
7) pmi+age+sex+w_1 + w_2 + w_3 + w_4,
8) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5,
9) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6,
10) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7,
11) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8,
12) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 + w_9,
13) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 + w_9 + w_10,
14) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 + w_9 + w_10+nUMI,
15) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 + w_9 + w_10+nUMI+nGenes,
16) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 + w_9 + w_10+nUMI+nGenes+percent_mito,
17) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 + w_9 + w_10+nUMI+nGenes+percent_mito+ percent_ribo,
18) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 + w_9 + w_10+nUMI+nGenes+percent_mito+ percent_ribo+nCells,
19) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 + w_9 + w_10+nUMI+nGenes+percent_mito+ percent_ribo+nCells+nReads,
20) pmi+age+sex+w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 + w_9 + w_10+nUMI+nGenes+percent_mito+ percent_ribo+nCells+nReads+mapped_Reads

The variables w_1 to w_10 represent the factors of unwanted variation estimated by RUVseq. The analysis was performed separately for each major cell type.

As an alternative approach to assess the extent to which the factors of unwanted variation estimated by RUVseq capture the known sources of variation, we conducted a correlation analysis between the two sets of variables. Specifically, we evaluated the correlation between the known sources of variation and the factors of unwanted variation estimated by RUVseq using the cor.test function in R, with the argument ''alternative'' set to ''two.sided'' and the argument ''method'' set to ''spearman''.

Overlap of DEGs

Statistical significance of the overlap of DEGs between cell types and variables was computed using Fisher's exact test using the R package GeneOverlap (v1.28.0) (http://shenlab-sinai.github.io/shenlab-sinai/).[77] The background (argument genome.size) for the gene overlap analysis comparing different variables was defined as the total number of genes evaluated in the differential expression analysis in the cell type under examination. Conversely, when comparing between cell types, the background (argument genome.size) was defined as the union of genes subjected to the differential expression analysis in the two cell types of interest. The p values were adjusted for multiple hypothesis testing using the Bonferroni method. Circos plots illustrating the overlap of DEGs between cell types were generated using the circlize package in R. Venn diagrams illustrating the overlap of DEGs between variables were created using the R package ggVennDiagram.

Overlap of DEGs across the expression level spectrum

Genes were categorized into 10 subsets based on their average expression level in excitatory neurons. Then the overlap of global AD pathology-associated DEGs between excitatory neuron cell types was evaluated within each subset. The average expression level of genes was determined using the ''AverageExpression'' function in the Seurat R package. Statistical significance of the overlap of DEGs between cell types was computed using Fisher's exact test using the R package GeneOverlap (v1.28.0) (http://shenlab-sinai.github.io/shenlab-sinai/).[77] The p values were adjusted for multiple hypothesis testing using the Bonferroni method.

Permutation test for evaluating the significance of the overlap of DEGs between excitatory neuron subtypes

The average expression of each gene in each excitatory neuron subtype was determined using the 'AverageExpression' function from the Seurat R package. The genes considered in the differential expression analysis for each excitatory neuron subtype were categorized into 10 subsets based on their average expression level within the corresponding subtype. Next, the genes in each of

the 10 subsets were intersected with the genes identified to be significantly associated with global AD pathology. This intersection allowed us to determine the number of significant DEGs (Differentially Expressed Genes) in each subset. The process was performed separately for genes positively and negatively associated with global AD pathology. Subsequently, we randomly sampled the determined number of significant DEGs from each of the 10 subsets, ensuring that the expression level distribution of the DEGs was preserved in the random samples. This random sampling step was repeated for a total of 1000 iterations. For each of the 1000 random samples, we determined the overlap of genes between excitatory neuron subtypes and compared it to the observed overlap between the two cell types. To assess the significance of the observed overlap, we computed z scores, which represent the difference between the observed value of overlap and the mean value of overlap based on the permutation results, divided by the standard deviation of the permutation results.

Permutation test for evaluating the significance of the overlap of DEGs between our dataset and the De Jager dataset

The permutation test to assess the significance of the overlap of differentially expressed genes between our dataset and the De Jager dataset was conducted similarly as described above. Briefly, we computed the average expression level of each gene within each cell type and categorized the genes considered in the differential expression analysis into 10 subsets based on their average expression level. Subsequently, we determined the number of significant differentially expressed genes (DEGs) within each subset and randomly sampled the same number of DEGs, preserving the expression level distribution of DEGs, for a total of 1000 iterations. These steps were performed separately for both our dataset and the De Jager dataset.

Empirical evaluation of the effect of cell number on the number of DEGs detected and the overlap of DEGs between cell types

To evaluate the impact of cell number on the detection of DEGs and the overlap of DEGs between cell types, we employed the following approach. For each excitatory neuron subtype, we downscaled the number of cells per individual using the ''subset'' function of the R package Seurat, specifying the desired number of cells with the ''downsample'' argument. Next, we performed differential expression analysis using the R package muscat, following the previously described methodology. To determine the statistical significance of the DEG overlap between cell types, we employed Fisher's exact test, utilizing the R package GeneOverlap (v1.28.0) (http://shenlabsinai.github.io/shenlab-sinai/).[77] The background, represented by the ''genome.size'' argument, was defined as the union of genes subjected to the differential expression analysis in the two cell types of interest. To account for multiple hypothesis testing, we adjusted the p values using the Bonferroni method.

Gene ontology term enrichment analysis

Gene Ontology enrichment analyses were performed using metascape[72] (http://metascape.org). For all cases we used the complete proteome as the default background in Metascape and report multi-test corrected q-values.

Consensus signatures and SynGO analyses

Consensus signatures were defined as sets of genes significantly associated with global AD pathology in at least three different excitatory neuron subtypes. Consensus signatures were defined for genes positively and negatively associated with global AD pathology, respectively. The synaptic location of consensus signature genes was examined using the curated ontology of the SynGO consortium.[29] SynGO analyses were performed via the SynGO portal (https://www.syngoportal.org/) using default settings.

Analysis of consistency with random subsets of the cohort

To test the robustness and generalizability of our results, we divided the full cohort into two non-overlapping random subsets and performed the differential expression analysis separately on both sub-cohorts. The differential expression analyses was performed using the R package muscat.[100] Lowly expressed genes were excluded and only genes with more than one count in at least 10 cells were considered. For the analysis of the variable measuring global AD pathology, we included as covariates the individual's age at death, postmortem interval, sex, and factors of unwanted variation as determined by the R package RUVseq.[76] The factors of unwanted variation were determined using the RUVr function with the argument k set to 10 factors of unwanted variation. The multiple testing correction was performed locally using the p.adjust function with the BH method as implemented in muscat. Statistical significance of the overlap of DEGs between the results of the original analysis across the entire cohort and the two random sub-cohorts was computed using Fisher's exact test using the R package GeneOverlap.[77] The p values were adjusted for multiple hypothesis testing using the Bonferroni method. The correlation of association scores and effect sizes (log fold change) was computed using the cor.test function in R with the argument ''method'' set to "spearman". For the correlation analysis significant DEGs as determined by the original differential expression analysis across the entire cohort were considered.

Comparison of effect sizes between AD-associated variables

The correlation of the effect size of up- and downregulated genes between multiple AD pathology-related variables including global AD pathology, neuritic plaque burden, neurofibrillary tangle burden, tangle density, overall amyloid level, and diffuse plaque burden was computed using the cor.test function in R with the argument ''method'' set to "pearson". The union of genes significantly associated with the two variables of interest were considered, and the correlation was computed based on the log fold change values. The analysis was performed separately for positively and negatively associated genes.

Comparison with previously published snRNA-seq studies of AD

To validate our differential expression results, we conducted a comparison with multiple previously published single-nucleus RNA sequencing studies of Alzheimer's disease.[14,15,23,24] The results of differential expression analyses from the studies conducted by Lau et al., Morabito et al., Sadick et al., and Su et al. were obtained by retrieving the supplemental information associated with

each respective study. We evaluated the correlation between the effect sizes of genes associated with global Alzheimer's disease pathology in our study and the effect sizes of overlapping DEGs in the previously published studies. The correlation was computed using the cor.test function in R with the argument ''alternative'' set to ''two.sided" and the argument ''method'' set to "pearson". p values were adjusted for multiple testing using the BH method as implemented in the R function p.adjust. Association scores were determined by multiplying the correlation coefficient by the negative log10 of the BH-adjusted p value and dividing the product by the absolute value of the correlation coefficient. Thus, the association score was defined as the signed negative log10 BH-adjusted p value, where the sign was derived from the sign of the correlation coefficient.

Correlation analysis of gene expression and clinical and pathological traits using self-organizing maps

First, we computed the average expression per study participant for each gene in each cell type. Then, for each cell type, a gene-wise correlation coefficient (Spearman's rank correlation coefficient) was computed using gene expression and selected traits across all individuals as variables. The traits included in this analysis were apoe_genotype (ApoeE genotype), cogdx (final cognitive consensus diagnosis), cogdx_stroke (Clinical diagnosis of stroke), dxpark (Clinical diagnosis of Parkinson's disease), cogn_global_lv (Global cognitive function), cogn_ep_lv (Episodic memory), cogn_po_lv (Perceptual orientation), cogn_ps_lv (Perceptual speed), cogn_se_lv (Semantic memory), cogn_wo_lv (Working memory), cognep_random_slope (rate of change of episodic memory), cogng_random_slope (rate of change of global cognitive function), cognpo_random_slope (rate of change of perceptual orientation), cognps_random_slope (rate of change of perceptual speed), cognse_random_slope (rate of change of semantic memory), cognwo_random_slope (rate of change of working memory), age_death (age at death), msex (Sex), hypertension_bl (Hypertension), cancer_bl (Cancer at baseline), diabetes_sr_rx_bl (Diabetes), headinjrloc_bl (Head injury), heart_bl (Heart), stroke_bl (Stroke), bradysc_lv (Bradykinesia score), gaitsc_lv (Gait score), parksc_lv (Parkinsonian summary score), braaksc (Braak stage), ceradsc (CERAD score), gpath (Global AD pathology), gpath_3neocort (global measure of neocortical AD pathology), niareagansc (NIA-Reagan diagnosis of AD), plaq_d_mf (Diffuse plaque burden in the midfrontal cortex), pmi (postmortem interval), amyloid (overall amyloid level), plaq_d (Diffuse plaque burden), plaq_n (Neuritic plaque burden), plaq_n_mf (Neuritic plaque burden in the midfrontal cortex), dlbdx (Lewy Body disease), nft (Neurofibrillary tangle burden), nft_mf (Neurofibrillary tangle burden in the midfrontal cortex), tangles (Tangle density), tdp_st4 (TDP-43 stage), arteriol_scler (Arteriolosclerosis), caa_4gp (Cerebral amyloid angiopathy), cvda_4gp2 (Cerebral atherosclerosis), ci_num2_gct (Gross chronic infarcts), ci_num2_mct (Chronic microinfarcts), ci_num2_gtt (Gross infarcts), ci_num2_mtt (Microinfarcts), gpath_CR_score (Global AD pathology cognitive resilience score), age_death_CR_score (Age at death cognitive resilience score). The resulting correlation matrices for each cell type were concatenated and analyzed using a computational algorithm (SOM).[104] All SOMs were created using the kohonen R package.[78]

Protein complexes analysis

Identification of protein complexes associated with AD

Using the CORUM database, a comprehensive resource of mammalian protein complexes,[33] 2417 identified core complexes in humans were subsetted out of 3512 complexes. A module score was computed for each complex in each cell by the AddModuleScore function from Seurat.[70,105–107] An averaged module score was assigned to each individual for each cell subtype in each brain region. This was followed by fitting a generalized linear model to the data to identify directionality of the module scores in each cell subtype and in each brain region relative to pathology and cognition variables. Age at death and postmortem interval (pmi) were included as confounding variables with each predictor variable of interest in the model. The Benjamini & Hochberg (BH) method was used for adjusting the p values for multiple comparisons as implemented by the p.adjust function from the R stats package (each heatmap comprises adjusted p values for the shown comparisons or more comparisons that were not shown). Association scores were computed by multiplying the sign of the regression coefficient by the negative log10 p value (signed negative log10 p value) for each comparison. Following the same approach, we also computed the association between the expression level of protein complexes and Lewy body pathology, TDP-43 pathology, vascular pathology, and other medical conditions.

To increase the stringency of our selection, we tested whether genes encoding protein complex components of complexes significantly associated with AD pathology, were also detected as significantly differentially expressed genes. For each protein complex and each cell type, we determined the fraction of complex components that were identified as significantly differentially expressed. Based on this analysis we selected 14 protein complexes that were prominently associated with pathology at the complex level and at the single gene level.

Aggregation of p Values across excitatory neuron subtypes

To determine which protein complexes exhibit a conserved alteration in response to global pathology in excitatory neurons, averaged regression coefficient and aggregated p value were computed for each complex across 12 cell subtypes. Aggregated p value was derived using sumlog function from metap package,[79] which combine p values by the sum of logs method (Fisher's method). Averaged regression coefficients and negative log10 of aggregated p values for all complexes were plotted using a volcano plot (n = 2338). In addition to Fisher's exact test, we employed the Cauchy combination test to address potential violations of the independence assumption and correlated p values.[108] P-values were combined using the CCT function of the R package STAAR, which aggregates p values using the Cauchy method.[109,110] Our analysis using the Cauchy combination test yielded consistent findings with those obtained from Fisher's exact test, confirming the upregulation of the Cohesin-SA1 complex as the most significant among the analyzed complexes.

Identification of genes co-regulated with the cohesin complex in excitatory neurons

To identify genes co-regulated with the cohesin complex, we determined the correlation between the expression level of each gene and the module score of the protein complex. The correlation between genes and complexes was determined across all the individual cells of a cell type. The Spearman correlation was computed using the rcorr function from the Harrell Miscellaneous (Hmisc) R package. p values were adjusted for multiple testing (the number of genes tested per cell type) using the BH method as implemented in the R function p.adjust. Association scores for each gene in each cell type were determined by multiplying the correlation coefficient by the negative log10 of the BH-adjusted p value and dividing the product by the absolute value of the correlation coefficient. Thus, the association score was defined as the signed negative log10 BH-adjusted p value, where the sign was derived from the sign of the correlation coefficient.

To identify genes co-regulated with the cohesin complex in excitatory neurons, we focused our analysis on genes significantly positively associated with global AD pathology in at least three excitatory neuron cell types (consensus signature genes). For each of these genes the vector of association scores across the 12 excitatory neuron subtypes was used to determine pairwise gene-gene Pearson correlation coefficients using the cor function in R. Finally, k-means clustering of the resulting gene-gene correlation matrix as implemented in the Heatmap function of the R package ComplexHeatmap[80] was used to identify sets of genes with a similar association with cohesin complex expression across excitatory neuron subtypes. Gene ontology term enrichment analysis of identified gene sets co-regulated with the cohesin complex was performed using metascape.[72]

Identification of genes co-regulated with the cohesin complex in oligodendrocytes

To identify genes co-regulated with the cohesin complex in oligodendrocytes, gene-complex association scores for the cohesin-SA1, cohesin-SA2, and the IKBA-p50-p65 Nf(kappa)B complex in oligodendrocytes were determined as described above. For each gene, the vector of association scores across the three protein complexes was used for the downstream analysis. For the downstream analysis, only genes significantly positively associated with global AD pathology in oligodendrocytes were considered. Genes were grouped by k-means clustering of the vectors of association scores described above using the Heatmap function of the R package ComplexHeatmap.[80]

Identification of genes co-regulated with the cohesin complex across brain regions

In addition to the analysis performed based on tissue from the prefrontal cortex, we analyzed the co-regulation between genes and the cohesin complex in the mid-temporal cortex, the angular gyrus, the entorhinal cortex, and the hippocampus. A module score was computed for each complex in each cell by the AddModuleScore function from Seurat.[70,105–107] To identify genes co-regulated with the cohesin complex, we determined the correlation between the expression level of each gene and the module score of the cohesin complex. The correlation between genes and complexes was determined across all the individual cells of a major cell type. The major cell types considered in this analysis were excitatory neurons (Exc), inhibitory neurons (Inh), astrocytes (Ast), oligodendrocytes (Oli), oligodendrocyte precursor cells (OPCs), and microglia (Mic). The Spearman correlation was computed using the cor.test function in R with the argument ''alternative'' set to "two.sided". p values were adjusted for multiple testing (the number of genes tested per cell type) using the BH method as implemented in the R function p.adjust. Association scores for each gene in each cell type were determined by multiplying the correlation coefficient by the negative log10 of the BH-adjusted p value and dividing the product by the absolute value of the correlation coefficient.

We also clustered consensus signature genes positively associated with global AD pathology in excitatory neurons in the prefrontal cortex based on their correlation with the cohesin-SA1 or the cohesin-SA2 complex in the excitatory neurons of the four cortical brain regions mentioned above (mid-temporal cortex, angular gyrus, entorhinal cortex, hippocampus). For each gene the vector of association scores across excitatory neurons of the mid-temporal cortex, angular gyrus, entorhinal cortex, and hippocampus was used to compute pairwise gene-gene Pearson correlation coefficients using the cor function in R. Finally, k-means clustering of the resulting gene-gene correlation matrix as implemented in the Heatmap function of the R package ComplexHeatmap[80] was used to identify sets of genes with a similar association with cohesin complex expression across excitatory neurons of four distinct brain regions.

Bulk RNA-seq differential expression analysis

Differential expression analysis of bulk RNA sequencing (RNA-seq) data from the ROSMAP cohorts was performed with DESeq2[73] (plotted) and edgeR.[101,111,112] Age at death and postmortem interval (pmi) were converted into z-scores and included as confounding variables in the regression equation in the differential expression analysis with each variable of interest. Both approaches provided similar results.

Cognitive resilience analysis

Quantification of cognitive resilience

To quantify cognitive resilience, linear regression analysis was used to model the relation between cognition and pathology (Prism 9, Version 9.3.1). For each individual, a score was calculated as the difference between the observed cognition (y) and the predicted cognition ($\hat{y}$) by the model based on their level of pathology (x) (Methods S1, page 5A and 5B). Cognitive Resilience (CR) was measured based on global cognitive function. On the other hand, to quantify the Cognitive Decline Resilience (CDR), the random slope of global cognition (the estimated person-specific rate of change in the global cognition variable over time) was used as the dependent variable in the model.

To study the resilience to AD pathology along disease progression, CR and CDR scores were derived based on four measures of AD pathology, neuritic plaque burden (plaq_n), neurofibrillary tangle burden (nft), tangle density (tangles) and global AD pathology

(gpath). Each individual in the cohort was assigned 8 new measures of cognitive resilience, based on the available clinical-pathological measures. Differential gene expression analysis was performed using muscat,[100] correcting for the confounding variables age at death and postmortem interval (pmi) to retrieve cell type-specific gene expression alterations associated with cognitive resilience. To confirm the differential gene expression results based on the CR and CDR scores, we also evaluated the association between gene expression and global cognitive function or the rate of change of global cognitive function adjusting for AD pathology as a covariate. The AD pathology measures considered as a covariate were global AD pathology (gpath), neuritic plaque burden (plaq_n), neurofibrillary tangle burden (nft), or tangle density (tangles). Thus, together with the DGE analysis based on CR and CDR scores, we performed a total of 16 different tests assessing the association between gene expression and cognitive resilience.

Bulk RNA-seq differential expression analysis

Differential expression analysis of bulk RNA sequencing (RNA-seq) data from the ROSMAP cohorts was performed with DESeq2[73] (plotted) and edgeR.[101,111,112] Both approaches provided similar results.

Overrepresentation analysis

Overrepresentation analysis of resilience genes expressing cells was performed for each cell type and the following brain regions: prefrontal cortex, mid-temporal cortex, angular gyrus, entorhinal cortex, hippocampus, and thalamus. Individuals were split into high/low subgroups for each CR/CDR variable based on the median value determined using the quantile function with probability of 0.5 from the R stats package. The distribution function for the hypergeometric distribution (phyper in R) was used for computing the significance of the enrichment of cells detectably expressing the gene of interest in each cell subtype in the high CR/CDR subgroup. The phyper arguments were defined as, q: number of gene-expressing cells in the high subgroup, k: total number of cells in the high subgroup, n: number of gene-non-expressing cells in in both subgroups (total population), m: number of gene-expressing cells in both subgroups (total population) and lower.tail = FALSE to compute probabilities, $P[X > x]$. The Benjamini & Hochberg (BH) method was used for adjusting the p values for multiple comparisons as implemented in the p.adjust function from the R stats package (each heatmap comprises adjusted p values for the shown comparisons or more comparisons that were not shown).

Cell type composition

Overrepresentation analysis

To determine cell types enriched for cells isolated from individuals with low AD pathology, high cognition, or high cognitive resilience to AD pathology, we first split the cohort into four-quarters based on the quartiles of the variable of interest. Then, for each high-resolution cell type the overrepresentation of cells isolated from the first, second, third, or fourth quarter of individuals was computed separately using the hypergeometric cumulative distribution function (phyper() function) in R. The arguments for the phyper function were defined as, q: number of cells in the cell type of interest isolated from the quarter of individuals of interest minus 1, m: the total number of cells isolated from the quarter of individuals of interest, n: the total number of cells minus the total number of cells isolated from the quarter of individuals of interest, k: number of cells in the cell type of interest and lower.tail = FALSE to compute probabilities, $P[X > x]$. The p values were adjusted for multiple hypothesis testing using the Bonferroni method with the number of tests equal to the product of all cell types and conditions analyzed (6954 tests).

Comparison of the relative abundance of cell types between groups

For comparing the relative abundance of cell types between groups of individuals, the fraction of cells of a cell type was computed relative to all the cells isolated from an individual. To analyze changes in the relative abundance of cell types with disease progression, we first defined groups of individuals by splitting the cohort into four-quarters based on the quartiles of the variable of interest. The statistical analysis of differences in the relative abundance of cell types between groups was performed using Kruskal-Wallis's test followed by Dunn's multiple comparison test. For comparisons between two groups (e.g., high versus low cognitive resilience), the statistical analysis of differences in the relative abundance of cell types between the groups was performed using multiple Mann-Whitney tests and p values were corrected for multiple hypothesis testing using the Holm-Šídák method. The statistical analyses comparing the relative abundance of cell types was performed using Prism 9 software.

Cell type composition changes in association with continuous variables

We investigated the relationship between the relative abundance of cell types and measures of AD pathology, cognition, and cognitive resilience by modeling the number of cells of a certain cell type in a specific sample (individual study participant) relative to the total number of cells (in that sample) using a quasi-binomial regression model. We fitted the regression model using the glm function in R, including age, sex, and postmortem interval as covariates. p values were corrected for multiple testing using the Benjamini–Hochberg procedure as implemented in the R function p.adjust. The results are presented in the form of association scores (signed negative log10 BH-adjusted p value, where the sign was determined by the direction (positive or negative) of the association). To compare individuals with and without cognitive impairment among subjects with a pathologic diagnosis of Alzheimer's disease, we selected individuals with a pathologic diagnosis of AD, and among those further selected individuals without cognitive impairment (final consensus cognitive diagnosis (cogdx) value of 1) and individuals with a cognitive diagnosis of AD dementia (cogdx value of 4). The cogdx variable was then transformed into a binary variable, where a cogdx score of 4 was assigned the value of 0, and a score of 1 was assigned the value of 1. Finally, the quasi-binomial regression analysis was performed as described above using the glm function in R with the argument "family" set to "quasibinomial", and including age, sex, and postmortem interval as covariates.

External data sources

To validate our findings, we analyzed a separate large single-nucleus RNA-seq dataset derived from the dorsolateral prefrontal cortex (DLPFC) of 465 ROSMAP study participants (referred to as the De Jager dataset).[26,27] Processed snRNAseq data (DLPFC, Experiment 2) were obtained from Synapse (https://www.synapse.org/#!Synapse:syn51123521).

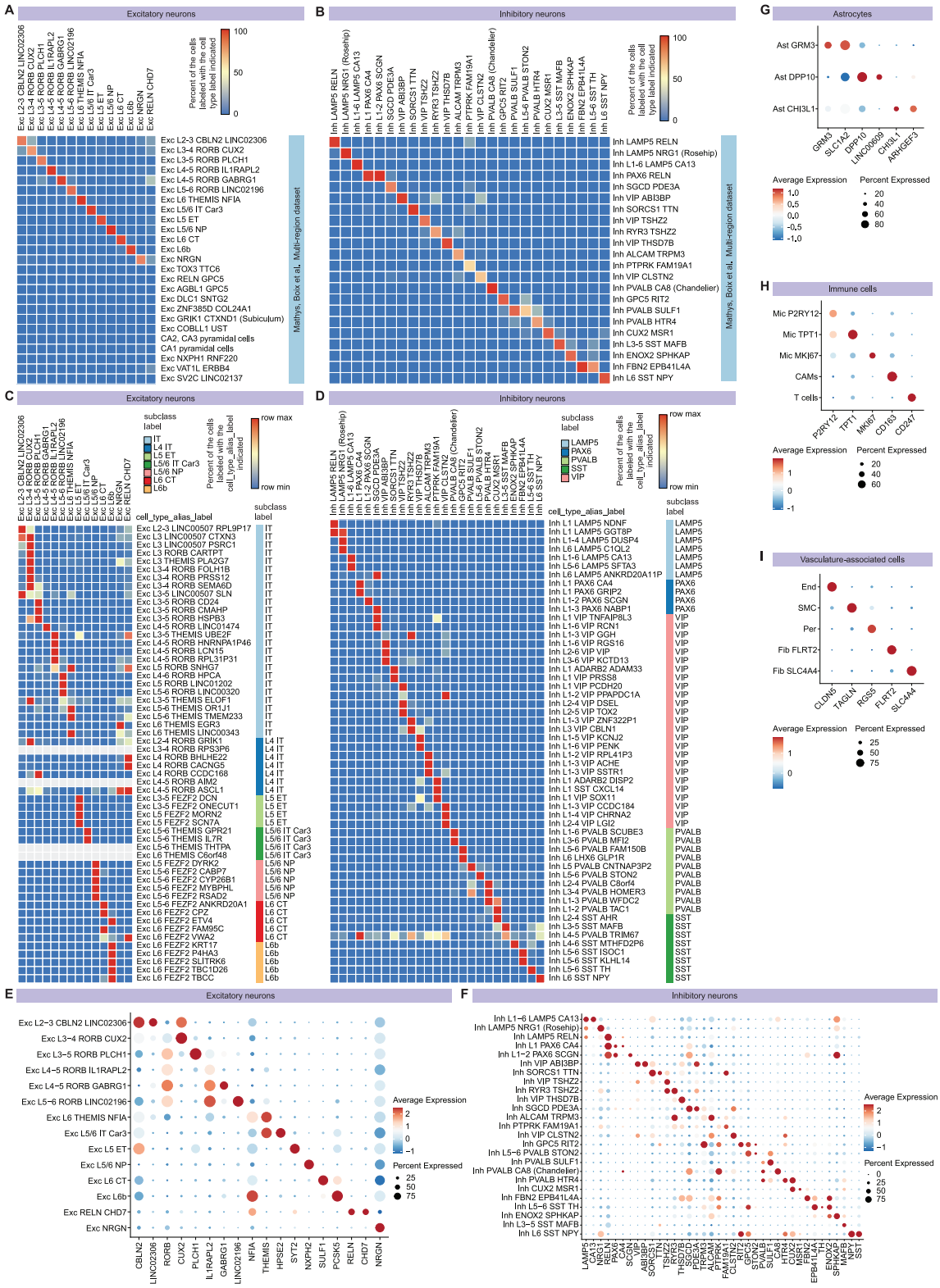
To further validate our findings with data from a different brain region, we downloaded a snRNA-seq dataset derived from the middle temporal gyrus (MTG).[28] Specifically, processed data (SEAAD_MTG_RNAseq_final-nuclei.2023-05-05.h5ad) were obtained from the Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD) (https://sea-ad-single-cell-profiling.s3.amazonaws.com/index.html#MTG/RNAseq/). To project our cell type annotations onto the De Jager dataset and the SEA-AD dataset, we followed the reference mapping workflow of Seurat using the commands FindTransferAnchors and MapQuery. Differential gene expression analyses on these datasets were performed using muscat, as described above, after estimating and removing ambient RNA from the snRNA-seq data using DecontX. In these analyses, we included the individual's age at death, postmortem interval, sex, and factors of unwanted variation determined by RUVseq as covariates.

IMARIS colocalization quantification

Imarisx64 8.1.2 (Bitplane, Zurich, Switzerland) was used to analyze BDH1 expression levels within GFAP-positive astrocytes, and CHCHD6 expression levels within NEUN-positive neurons. The IMARIS surface functions was first applied to 3D render GFAP and NEUN signals in respective experiment, allowing to define regions of interest for further analyses. The "IMAIRSColoc" function was then utilized to create a new channel, referred to as the colocalized channel, from the 2 source channels. Following this, to quantify the volume of the colocalized signals, additional surfaces were produced specifically for the colocalized channel.

CellPress
OPEN ACCESS

Supplemental figures



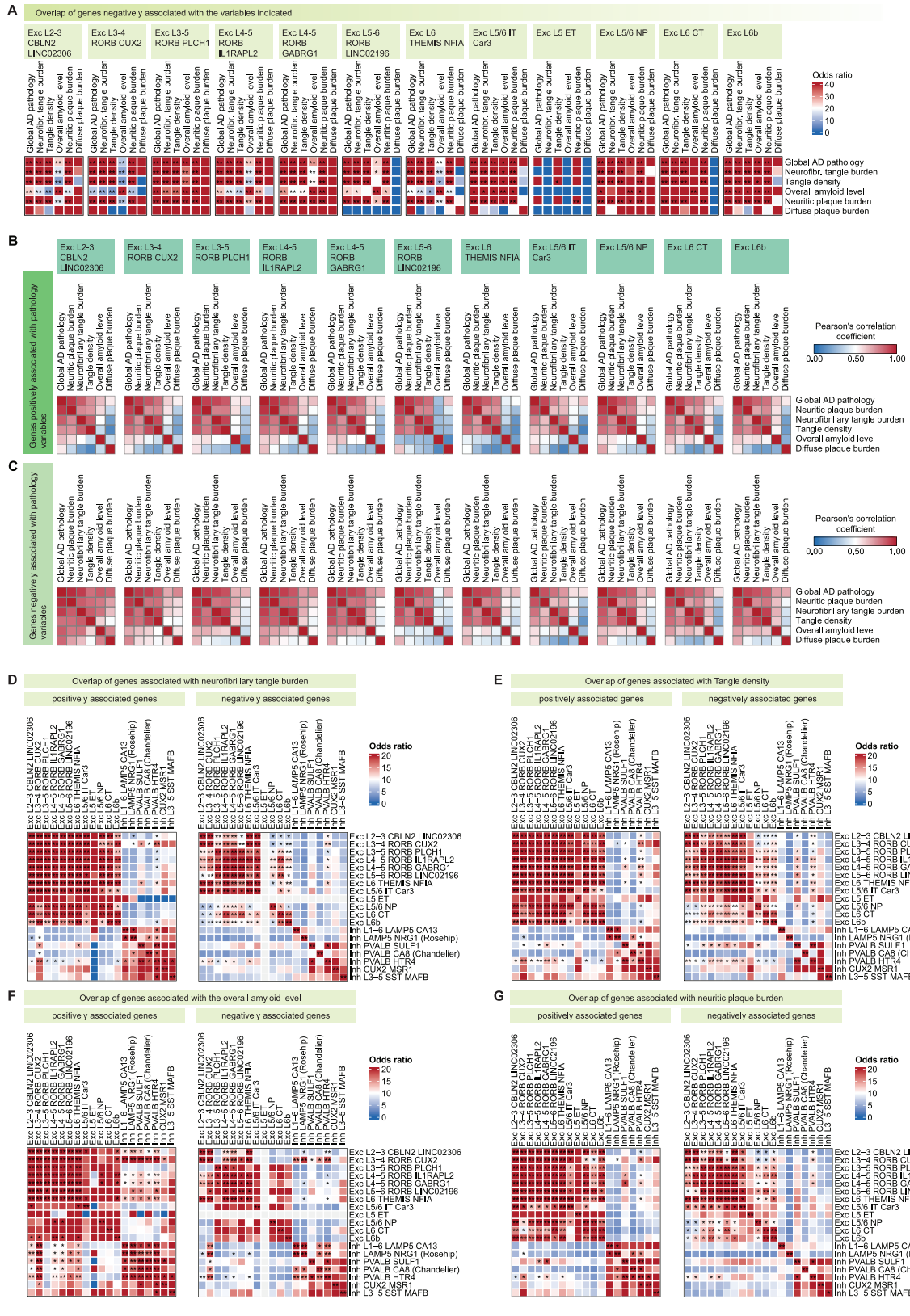
(legend on next page)

Figure S1. Cortical excitatory and inhibitory neuron subtypes align with previously identified neuronal subpopulations, related to Figure 1

(A and B) Projection of cortical excitatory (A) and inhibitory (B) neuron subpopulation annotations previously defined (rows) onto the neuronal subtypes identified in this study (columns). The heatmap shows the proportion of cells (columns) that were labeled with the cell type label (rows) indicated.

(C and D) Projection of cortical excitatory (C) and inhibitory (D) neuron subpopulation annotations previously defined by the Allen Brain Institute (rows) onto the neuronal subtypes identified in this study (columns). The heatmap shows the proportion of cells (columns) that were labeled with the cell type alias label (rows) indicated.

(E–I) Marker genes of excitatory neuron subtypes (E), inhibitory neuron subtypes (F), astrocyte subtypes (G), immune cell subtypes (H), and vasculature-associated cell types (I). The dot plot shows the average scaled expression of the marker genes indicated across subtypes.



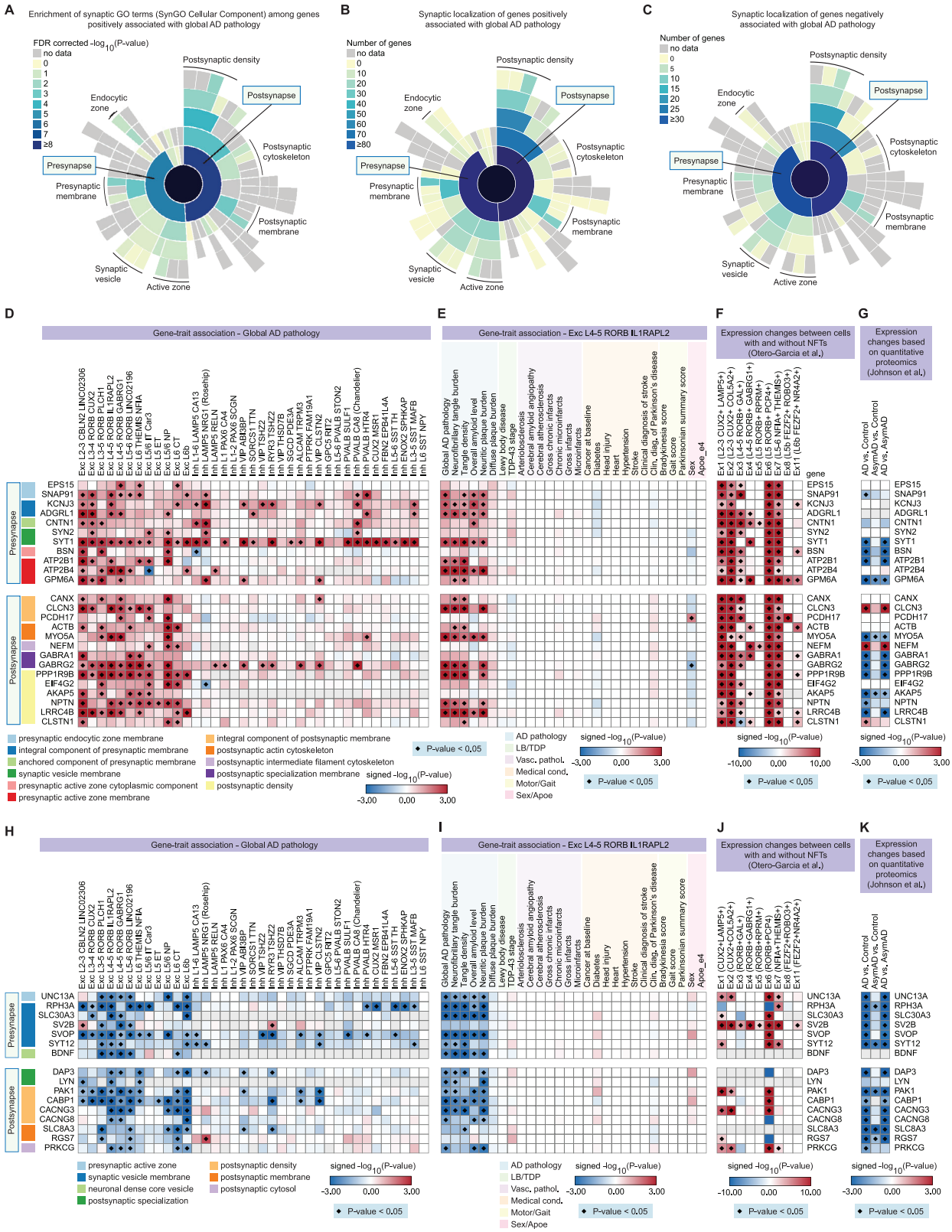
(legend on next page)

Figure S2. Overlap of significantly differentially expressed genes between variables and cell types, related to Figure 2

(A) Overlap (OR, one-sided Fisher's exact test) of genes significantly negatively associated with the variables indicated. The p values have been adjusted for multiple hypothesis testing; **−log10(p value) > 200, *−log10(p value) > 100 (one-sided Fisher's exact test).

(B and C) Correlation of the effect size of (B) up- and (C) downregulated genes between multiple AD pathology-related variables (rows and columns) across excitatory neuron subtypes. The union of genes significantly associated with the two variables of interest were considered in this analysis. The Pearson correlation was computed based on the log fold change values.

(D–G) Overlap (OR, one-sided Fisher's exact test) of genes significantly positively (left panel) or negatively (right panel) associated with the variables indicated between the excitatory neuron subtypes indicated. The p values have been adjusted for multiple hypothesis testing; **−log10(p value) > 200, *−log10(p value) > 100 (one-sided Fisher's exact test). The variables considered are NFT burden (D), tangle density (E), overall amyloid level (F), and neuritic plaque burden (G).



(legend on next page)

Figure S3. Genes involved in synaptic signaling associated with AD pathology, related to Figure 3

(A–C) Sunburst plots depict SynGo ontology terms organized by synaptic locations.

(A) Enrichment of synaptic GO terms (SynGO Cellular Component) among genes positively associated with global AD pathology.

(B and C) Synaptic localization of genes positively (B) or negatively (C) associated with global AD pathology (number of genes).

(D and H) Association between the expression level of genes (rows) and global AD pathology across the cell types indicated (columns).

(D) Genes positively associated with AD pathology.

(H) genes negatively associated with AD pathology.

(E and I) Association between the synaptic genes (rows) and variables indicated (columns) in L4–5 excitatory neurons (Exc L4–5 RORB IL1RAPL2).

(F and J) Expression changes (association scores) of synaptic genes between cells with and without NFTs, as determined by Otero-Garcia et al.[30]

(G and K) Changes in protein expression level comparing DLPFC tissue from control and AD brains as reported by Johnson et al.[31] are shown. The direction and significance of the protein expression level changes are shown in the form of signed −log10 Holm adjusted p values, where the sign was determined by the difference in the expression level (AD minus control).

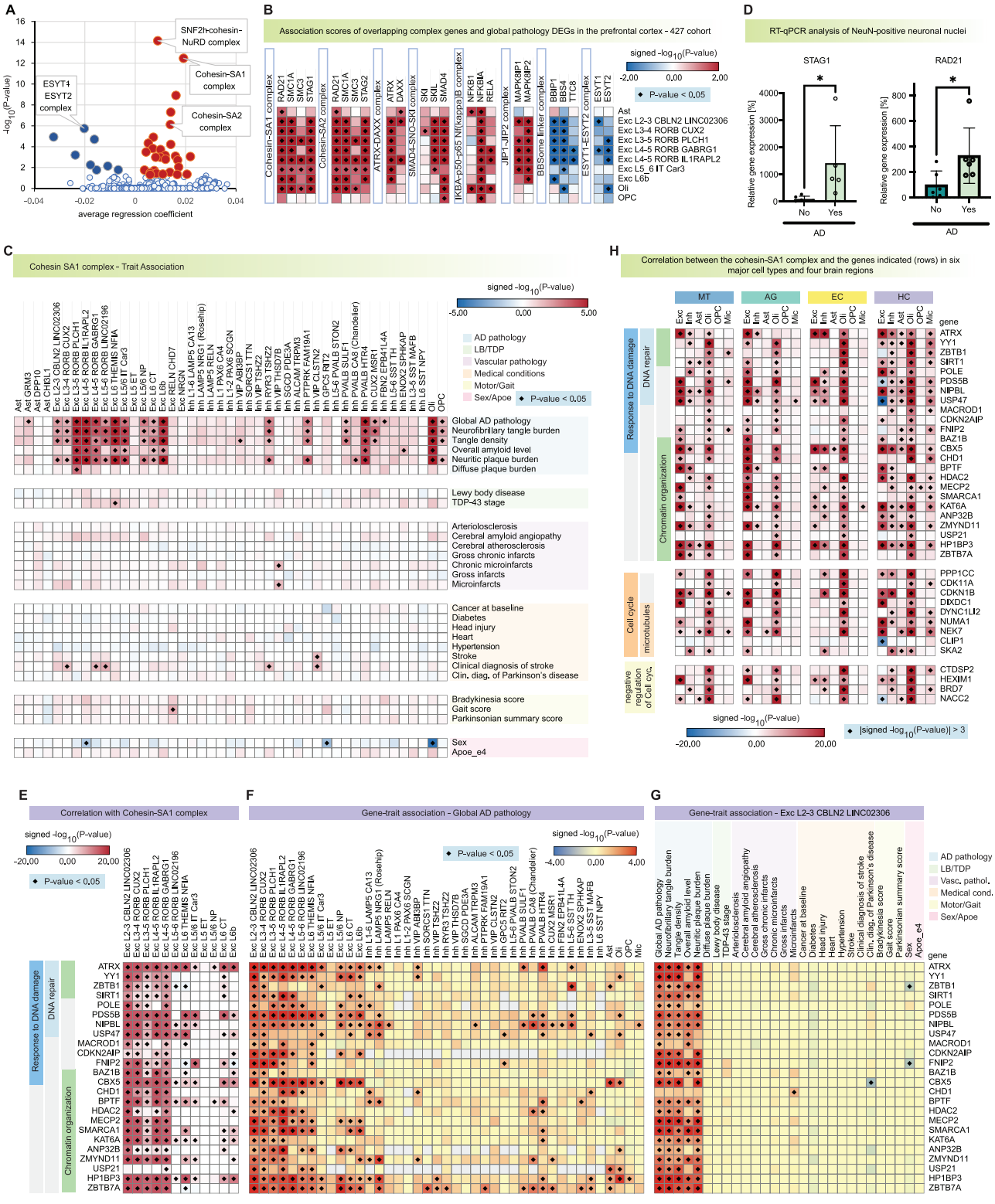


Figure S4. Systematic analysis of protein complexes altered in AD, related to Figure 4

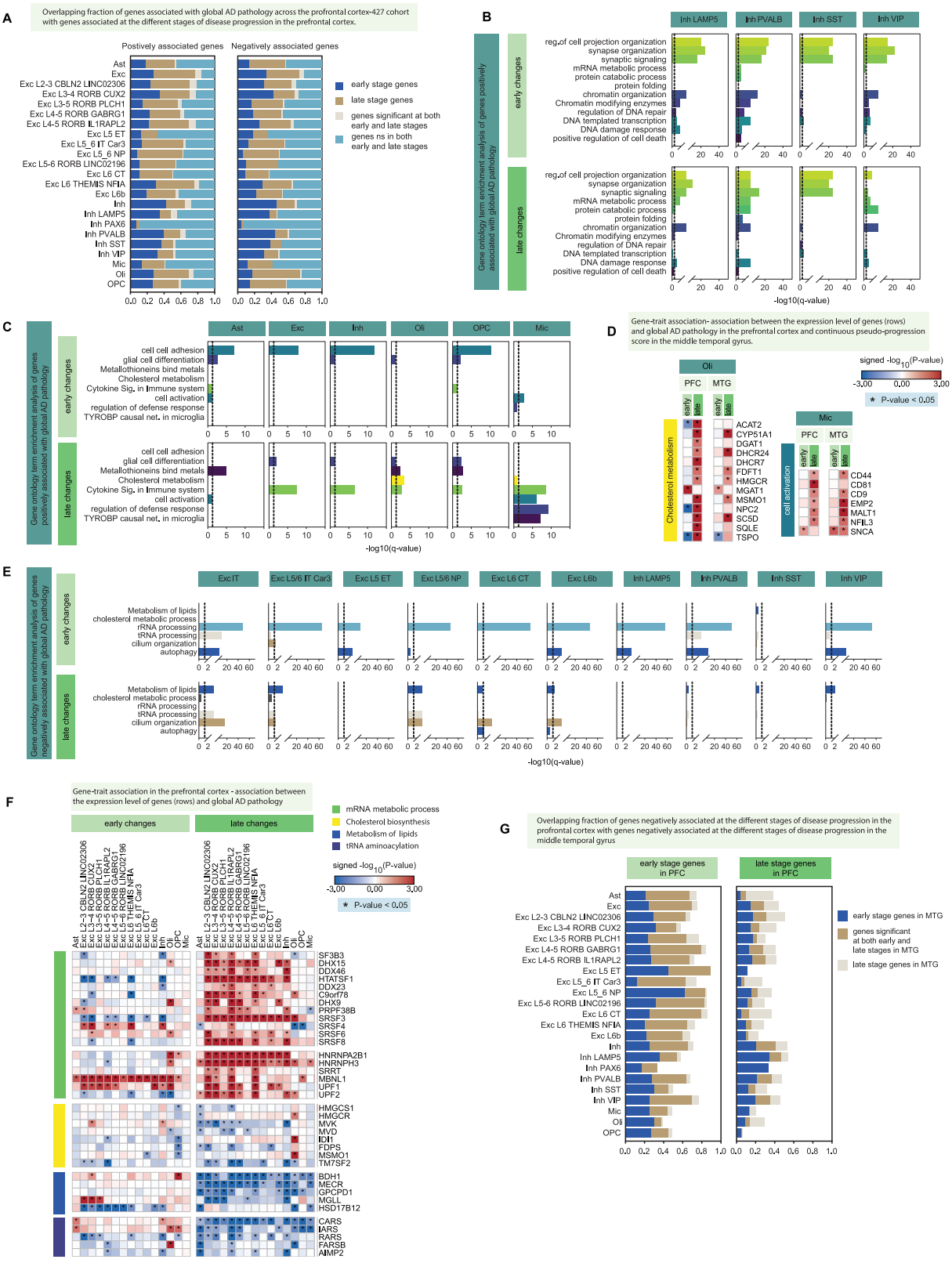
(A) Volcano plot showing averaged regression coefficients and −log10 of aggregated p values for protein complexes analyzed.

(B) Expression changes (association scores) of individual genes encoding components of the protein complexes shown in Figure 4A.

(C) Association between the expression level of the cohesin-SA1 complex and the variables indicated.

(legend continued on next page)

(D) RT-qPCR validation of genes associated with the cohesin complex. NeuN-positive nuclei were isolated by fluorescence-activated cell sorting (FACS) (n = 6 individuals without a pathologic diagnosis of AD and n = 6 individuals with a pathologic diagnosis of AD; for analysis of STAG1, n = 5 individuals with a pathologic diagnosis of AD). Data are mean ± s.e.m.; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; n.s.; $p > 0.05$ (Student's two-tailed t test).
(E) Correlation (association scores) between the cohesin complex and the genes indicated (rows) across excitatory neuron subtypes (columns).
(F) Association between global AD pathology and the expression level of genes co-regulated with the cohesin complex (rows) across the cell types indicated (columns).
(G) Association between the expression level of genes co-regulated with the cohesin complex (rows) and the variables indicated (columns) in L2–3 excitatory neurons (Exc L2–3 CBLN2 LINC02306).
(H) Correlation between the expression level of the cohesin-SA1 complex and the genes indicated (rows) across six major cell types and four distinct regions of the human brain.



(legend on next page)

Figure S5. Temporal changes in gene expression across multiple stages of Alzheimer's disease progression, related to Figure 5

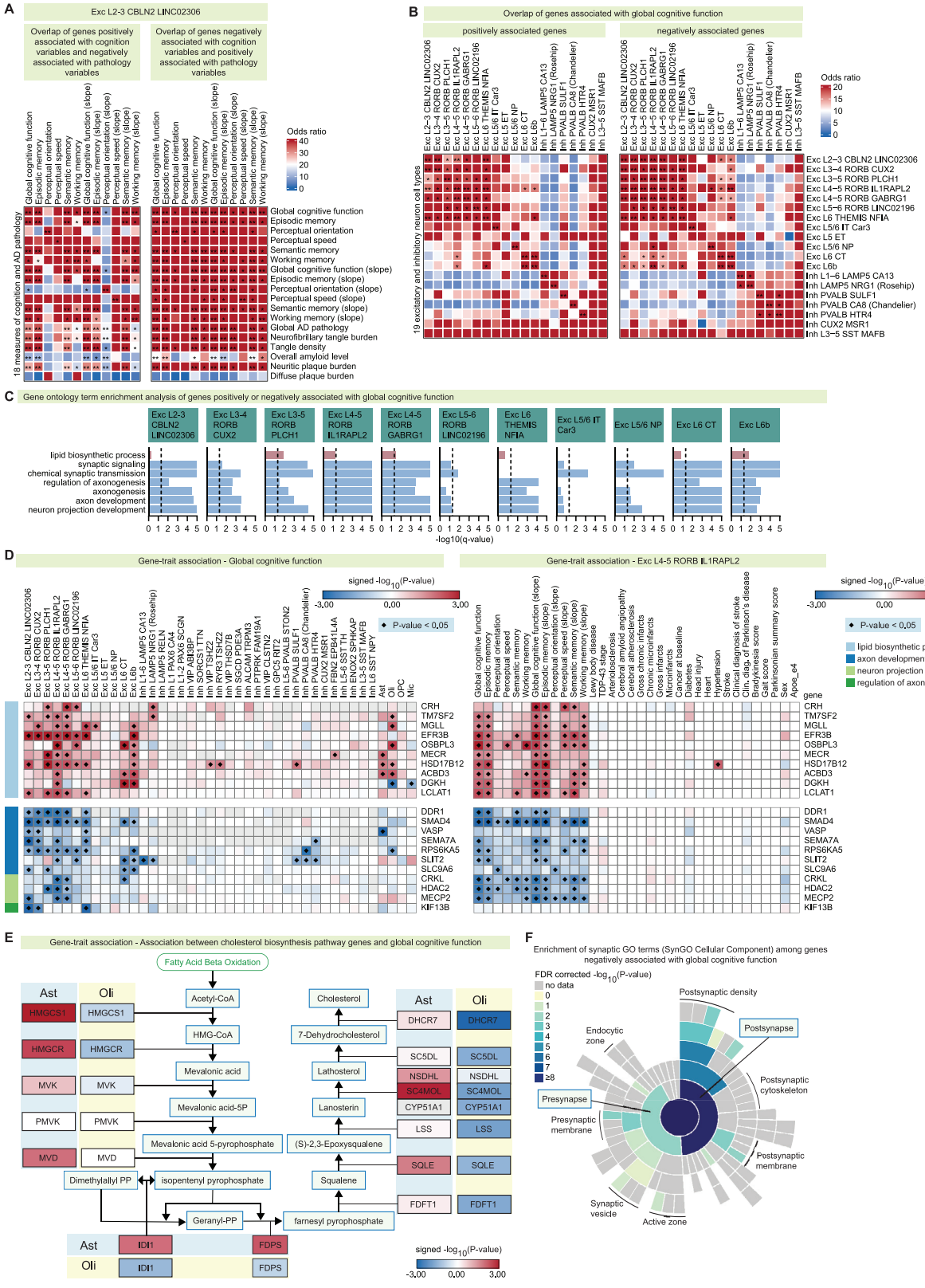
(A) Proportion of genes associated with global AD pathology within the entire cohort, also showing significant associations with AD pathology at distinct stages of disease progression.

(B, C, and E) Selected gene ontology (GO) terms (rows) enriched among genes significantly (B and C) positively or (E) negatively associated with global AD pathology across the cell types indicated (columns). Upper panels: early changes (second versus first group), lower panels: late changes (third versus second group).

(D) Association (association scores) between the expression level of selected genes involved in the biological processes indicated (rows) and global AD pathology and continuous pseudo-progression score in the prefrontal cortex and middle temporal cortex, respectively.

(F) Association (association scores) between the expression level of selected genes involved in the biological processes indicated (rows) and global AD pathology.

(G) Overlapping fraction of genes negatively associated at the different stages of disease progression in the PFC with genes negatively associated at the different stages of disease progression in the MTG. Left column: early-stage genes in PFC (second versus first group); right column: late-stage changes (third vs. second group).



(legend on next page)

Figure S6. Molecular correlates of cognitive impairment in the aged human brain, related to Figure 6

(A) The heatmaps show the overlap (OR, one-sided Fisher's exact test) of genes significantly associated with the variables indicated in L2–3 excitatory neurons (Exc L2–3 CBLN2 LINC02306). **−log10(p value) > 200, *−log10(p value) > 100 (one-sided Fisher's exact test).

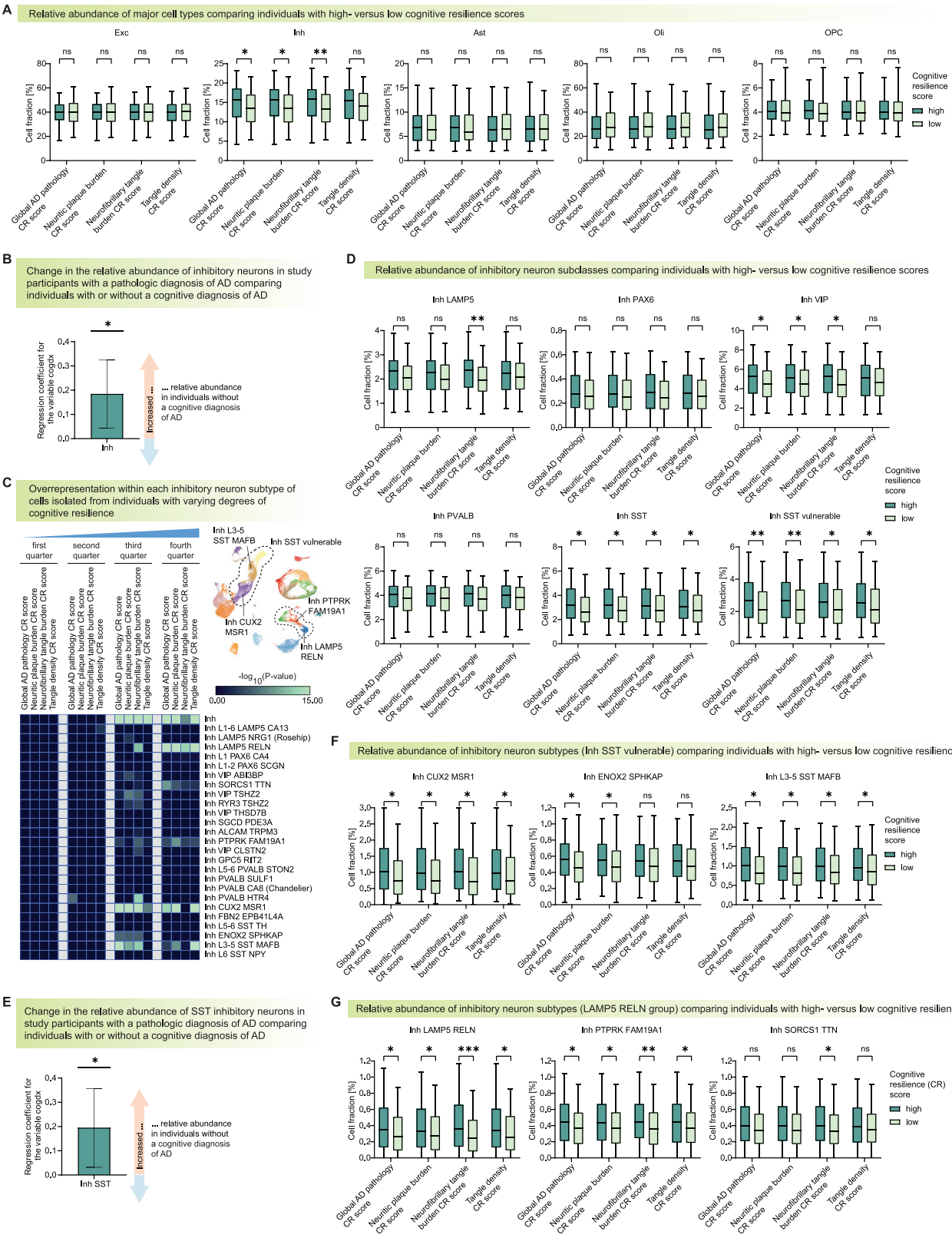
(B) Overlap (OR, one-sided Fisher's exact test) of genes significantly positively (left panel) or negatively (right panel) associated with global cognitive function between the excitatory neuron subtypes indicated. **−log10(p value) > 200, *−log10(p value) > 100 (one-sided Fisher's exact test).

(C) Selected GO terms enriched among genes significantly positively (red) or negatively (blue) associated with global cognitive function.

(D) Association (association scores) between the expression level of genes (rows) and global cognitive function across the cell types indicated (columns).

(E) Association between the expression level of cholesterol biosynthesis genes and global cognitive function in astrocytes and oligodendrocytes. Boxes are colored by the association scores.

(F) Enrichment of synaptic GO terms (SynGO Cellular Component) among genes negatively associated with global cognitive function.



(legend on next page)

Figure S7. Cell types associated with cognitive resilience to AD pathology, related to Figure 7

(A, D, F, and G) Relative abundance of (A) major cell types, (D) inhibitory neuron subclasses, (F) SST inhibitory neuron subtypes, or (G) inhibitory neuron subtypes (LAMP5 RELN group) comparing individuals with high versus low cognitive resilience scores. Within each box, horizontal lines denote median values; boxes extend from the 25th to the 75th percentile of each group's distribution of values; whiskers extend from the 5th to the 95th percentile. $^{**}p < 0.01$, $^{*}p < 0.05$; n.s., $p > 0.05$ (multiple Mann-Whitney tests corrected for multiple hypothesis testing using the Holm-Šídák method).

(B and E) Association (quasi-binomial regression) between the relative abundance of (B) inhibitory neurons or (E) SST inhibitory neurons and the clinical consensus diagnosis of cognitive status at the time of death (cogdx). Bar plot showing the regression coefficient for the predictor variable cogdx. Error bars represent the 95% confidence interval. $^{*}p < 0.05$ (glm function in R for quasi-binomial regression). Only individuals with a pathologic diagnosis of AD were considered in this analysis.

(C) Overrepresentation within each inhibitory neuron subtype of cells isolated from individuals with varying degrees of CR (hypergeometric test; the p values have been adjusted for multiple hypothesis testing; −log10[Bonferroni-corrected p values] are shown).