A spatiotemporal map of the aging mouse brain reveals white matter tracts as vulnerable foci

Oliver Hahn[1,2], Aulden G Foltz[1,2], Micaiah Atkins[1,2*], Blen Kedir[1,2*], Patricia Moran-Losada[1,2], Ian H Guldner[1,2], Christy Munson[1,2,3], Fabian Kern[4,5], Róbert Pálovics[1,2], Nannan Lu[1,2], Achint Kaur[1,2], Jacob Hull[1,2], John R Huguenard[1,2], Andreas Keller[4,5], Benoit Lehallier[6], Tony Wyss-Coray[1,2,7,8,9]

[1]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, California, USA.

[2]Wu Tsai Neurosciences Institute, Stanford University School of Medicine, Stanford, California, USA

[3]Vilcek Institute of Graduate Biomedical Sciences, NYU Langone Health, New York City, New York, USA

[4]Clinical Bioinformatics, Saarland University, Saarbrücken, Germany

[5]Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz-Centre for Infection Research (HZI), Saarbrücken, Germany

[6]Alkahest Inc., San Carlos, CA

[7]Paul F. Glenn Center for the Biology of Aging, Stanford University, Stanford, California, USA

[8]Stanford University, The Knight Initiative for Brain Resilience, Stanford, California, USA

[9]Correspondence: twc@stanford.edu

* These authors contributed equally: Micaiah Atkins, Blen Kedir.

Summary: Aging is the key risk factor for loss of cognitive function and neurodegeneration but our knowledge of molecular dynamics across the aging brain is very limited. Here we perform spatiotemporal RNA-seq of mouse brain aging, encompassing 847 samples from 15 regions spanning 7 ages. We identify a brain-wide gene signature representing aging in glia with spatially-defined magnitudes. By integrating spatial and single-nuclei transcriptomics, we reveal that glia aging is profoundly accelerated in white matter compared to cortical areas. We further discover region-specific expression changes in specialized neuronal populations. Finally, we discover distinct clusters of brain regions that differentially express genes associated with 3 human neurodegenerative diseases, highlighting regional aging as a potential modulator of disease. Our findings identify molecular foci of brain aging, providing a foundation to target age-related cognitive decline.

Introduction

Aging is the predominant risk factor for cognitive dysfunction (Kalache and Gatti, 2003; Partridge et al., 2018) and several neurodegenerative disorders, including Alzheimer's disease (AD) and Parkinson's disease (PD) (Hou et al., 2019; López-Otín et al., 2013; Niccoli et al., 2017). It remains unclear though, how aging contributes to the development of these distinct diseases of the brain, given their differences in pathological hallmarks, time of onset, and, notably, the regions affected (Hou et al., 2019). A quantitative understanding of the dynamics of aging across the brain may provide new insight into the relationship between aging and neurodegeneration. Interestingly, neuroimaging studies using structural and functional magnetic resonance imaging (MRI) data indicate that aging impacts the brain in a region-specific manner (Feng et al., 2020; Pomponio et al., 2020). However, these structural manifestations provide limited insight into the underlying molecular alterations occurring during brain aging. In contrast, changes in gene expression can be a readout of cellular deterioration and molecular processes accompanying aging, permitting quantitative comparisons of aging rates between tissues (Schaum et al., 2020) and cell types (Buckley et al., 2022). Previous studies have profiled age-related gene expression changes in human brain tissue, yet these microarray-based experiments capture a limited set of transcripts and cover usually one to four regions (Colantuoni et al., 2011; Ham and Lee, 2020) or quantify the transcriptome at low temporal resolution (Kang et al., 2011; Soreq et al., 2017). Expression profiling during human brain aging is particularly challenging since it can take hours to days before postmortem tissue is stabilized (Melé et al., 2015; Soreq et al., 2017; Trabzuni et al., 2011). Alternatively, expression profiling in model organisms like *M. musculus* enables quantitative data with minimal confounding factors, but comprehensive studies covering more than a few regions and at high temporal resolution (Hargis and Blalock, 2017; Lee et al., 2000; Tabula Muris Consortium, 2020; Zahn et al., 2007) do - to our knowledge - not exist.

Results

Spatiotemporal quantification of age-related gene expression across the mouse brain

To obtain a deep molecular understanding of the spatiotemporal changes of the aging mammalian brain, we selected 15 brain regions in the mouse with critical functions in cognition, whole-body metabolic control, or human disease. We punched out brain regions from coronal brain sections from the left and right hemisphere, including three cortical regions (motor area, visual area and entorhinal cortex; Mot.cor., Vis.cor and Ent.cor, respectively), anterior (dorsal) and posterior (ventral) hippocampus (Hipp.ant and Hipp.post., respectively), hypothalamus (Hypoth.), thalamus, caudate putamen (part of the striatum; Caud.put.), pons, medulla, cerebellum (Cereb.) and the olfactory bulb (Olf.bulb). We further isolated three regions that were enriched with the corpus callosum (Corp.cal.), choroid plexus (Chor.plx.) and the neurogenic subventricular zone (SVZ), (Figure S1A). We then applied our method to 59 mice (Figure 1A; n = 3-6 males per age; aged 3, 12, 15, 18, 21, 26 and 28 months; n = 5 females per age; aged 3, 12, 15, 18 and 21 months; all C57BL/6JN strain), resulting in a total of 1,770 samples (885 samples from each hemisphere). Isolated regions from the left hemisphere were stored, while all right hemisphere regions were processed through a custom-built bulk RNA-seq (bulk-seq) pipeline (Figure 1B, STAR Methods). We achieved robust tissue sampling with high RNA quality while minimizing perfusion artifacts, as indicated by consistent RNA yields across samples from the same region (Figure S1B), median RNA integrity numbers of 9.45 out of 10 (Figure S1C) (Mueller), and a neglectable fraction of reads mapping to hemoglobin genes (Figure S1D). In comparison, the UK Brain Expression Consortium, one of the largest collections of human brain tissue, reports average RNA integrity numbers of 3.85 for their tissue samples (Soreq et al., 2017; Trabzuni et al., 2011).

After quality control, we obtained 847 single-region transcriptomes. Visualization in uniform manifold approximation and projection (UMAP) space separated samples by region (Figure 1C), but not sex or age, which concurred with deterministic shared-nearest-neighbors graph clustering and hierarchical clustering (Figure S2A-E). However, within individual regions, samples segregated transcriptionally by age. The comparatively subtle effect of aging on gene expression highlights the necessity for precise isolation of brain tissue to avoid confounding cross-region contamination (Figure 1D and S2A).

To assess if the isolated regions accurately captured a given brain structure's transcriptome, we analyzed the expression of region-enriched genes ('marker genes'; STAR Methods and Table S1) in a publicly-available spatial transcriptomics dataset of an adult male mouse brain (https://www.10xgenomics.com/resources). To this end, we combined marker genes of a given region into 'signatures' (DeTomaso et al., 2019) that represent its transcriptional fingerprint. For each signature, we calculated a score per spatial transcriptome spot, summarizing the expression of marker genes into a single value. Signature scores were distinctly elevated in areas corresponding to the anatomical structures annotated in the Allen brain reference atlas (Figure S3A) (Lein et al., 2007). Notably, the corpus callosum-derived signature demarcated fiber tracts throughout the brain, indicating that the sampled transcriptome of this region could be a proxy for white matter tracts in general (Figure S3A). To assess the isolation of the SVZ, we built a signature of activated neural stem cells (aNSC) based on marker genes from single-cell data (Dulken et al., 2019) and calculated the score per each region. These aNSCs are enriched in neurogenic areas at young age and their number declines in old individuals (Dulken et al., 2019). Not only was the score elevated in the aNSC-rich olfactory bulb and SVZ region but we also found a significant decline with age, thus capturing the loss of neurogenesis, a known phenomenon of mouse brain aging (Figure S3B) (Dulken et al., 2019). In summary, our tissue isolation and bulk-seq workflow yielded high-quality transcriptomes that robustly captured a region's gene expression profile across a cohort of mice. The data can be interactively explored at https://twc-stanford.shinyapps.io/spatiotemporal_brain_map/.

Region identity is linked to expression dynamics during aging

RNA-seq permits quantitative comparisons of aging rates between organs and cell types (Schaum et al., 2020; Tabula Muris Consortium, 2020) based on timing and effect size of gene expression shifts. For instance, we found substantial region-dependence in the magnitude and timing of *C4b* expression (Figure 1E), a complement component and major schizophrenia risk factor (Sekar et al., 2016) that is robustly up-regulated in aged mice (Hahn et al., 2020) and models of neurodegeneration (Zhou et al., 2020). Notably, recent single-cell studies revealed that the composition of major cell types remains almost constant throughout the aging mouse brain (Ximerakis et al., 2019), thus the expression dynamics observed in bulk are unlikely to be driven primarily by shifts in cell type abundance. Thus, we were able to use our temporally resolved data to probe the per-region impact of aging on gene expression over time, as this could help to identify structures with selective vulnerability to old age.

We performed pairwise differential expression between 3 months and every following age group to determine when differentially expressed genes arise (DEGs; used from hereon to refer to genes that change with age). To focus on high-confidence expression changes that persist with advancing age, a gene had to pass the statistical cutoff in at least two comparisons to be classified as a DEG (Figure 1E-G). The general trend across regions indicated an increase of DEGs over time that plateaued around 21 months (Figure 1F,G), yet individual regions varied profoundly with respect to their total number of DEGs and the trajectory of DEG accumulation (Figure 1F and Table S2). For instance, the visual cortex showed a steady increase of DEGs until late age, while the transcriptome of the motor cortex already exhibited significant perturbation at 12 months, but there was little increase until a sudden jump at 21 months (Figure 1F,G). In contrast, the transcriptome of the entorhinal cortex appeared largely refractory to the effects of age altogether, with only 13 detectable DEGs in total (Figure 1F,G). This is in line with human MRI (Feng et al., 2020) and microarray (Berchtold et al., 2008) studies demonstrating that the entorhinal cortex displays only mild alterations during cognitively normal human aging, whereas it frequently exhibits the first amyloid deposition in AD patients (Van Hoesen et al., 1991). Together, these results reveal that the effect size of expression shifts during brain aging are strikingly region- and time-dependent, highlighting the necessity for region-resolved quantification and analysis. Notably, the regions with the most profound and earliest shifts in gene expression were the white matter-rich caudate putamen, cerebellum and corpus callosum, the latter showing a tenfold increase in the number of DEGs between 12 and 18 months.

Since pairwise comparisons treat every gene and age group independently, we next validated these results with two independent analyses. First, we probed all genes in the genome for positive or negative correlation with age (Spearman's rho ≥ 0.5 or ≤ -0.5, respectively; padj ≤ 0.05; Table S3), thus taking all age groups into account (Figure 1H). Not only did the regions differ in the number of age-correlated genes, confirming that the effect size of age depends on the region, but also the corpus callosum and cerebellum were the most impacted, while the entorhinal cortex remained largely unaffected (Figure 1H). As a second validation, we performed weighted gene co-expression network analysis (WGCNA) (Langfelder and Horvath, 2008) for each region (STAR Methods; Table S4), clustering genes into modules which might be driven by similar regulation during aging. We filtered for modules exhibiting significant association with age and found the number of modules differed between regions. In line with the above results, we found seven or more modules in the corpus callosum, cerebellum and motor cortex, whereas we detected no age-related modules in the entorhinal cortex (Table S4). To gain biological insight, we performed cell type- and pathway-enrichment for each age-related module and compiled summarized reports for each region as a quick reference resource for the scientific community (https://twc-stanford.shinyapps.io/spatiotemporal_brain_map/). Interestingly, we discovered in 10 regions at least one module with increased expression over time that was enriched for microglia- and inflammation-related genes (Fig 1I,J). Consistent with these findings, we found a small, common set of differentially regulated genes, including neuroinflammatory markers *Fcgr2b*, *Ctss*, *Cst7* (Chen et al., 2020) in modules across regions, suggesting the presence of a minimal group of coregulated genes changing throughout the brain. In summary, we found the results of three independent analyses (pairwise tests, age-correlation and WGCNA) congruent, demonstrating that the observed effects of aging on the transcriptome are region specific.

A minimal gene set forms a common fingerprint of brain aging

WGCNA analysis results indicated the possibility of a shared gene set that changes during aging throughout the brain. Such a minimal age-related gene signature would permit quantitative comparisons of the rates of change in a region's transcriptional age. While the vast majority of DEGs appeared to change only in three or less regions, indicative of region-selective expression patterns, we found 82 genes that were differentially regulated in 10 or more regions (Figure 2A,B; Table S2). These were strongly enriched for up-regulated genes with immune-modulatory functions (Table S5), including MHC I-mediated antigen presentation, interferon-response, cell adhesion and complement cascade, as well as regulators of mouse microglia activity (Figure 2C) including *Cd22* (Pluvinage et al., 2019), *Trem2* and *Tyrobp* (Yeh et al., 2017). Interestingly, of the only 7 down-regulated genes in this set, we found protein homeostasis genes *Dnajb1*, *Hsph1* and *Ahsa1*, as well as collagen synthesis gene *P4ha1*, which is in line with perturbed protein quality control mechanisms as a hallmark of aging (López-Otín et al., 2013) (Figure 2B). We combined these 82 genes into a common RNA aging signature to calculate their expression as a single 'common aging score' (CAS; STAR Methods) for each mouse and region. While the CAS expectedly showed significant increases in every region (Figure 2D, Figure S4A), the shape and amplitude of the trajectories varied profoundly. The pace and direction with which the CAS changed with age is defined from here on as a region's 'CAS velocity'. We employed linear models to approximate these trajectories, using the slope of the linear fit as a metric to comparatively assess the CAS velocity across regions (Figure 2D,E). Of note, the CAS at baseline (i.e. the offset of the linear fit) did not predict a region's CAS velocity (Figure 2F). Our analysis revealed a gradient of velocities across regions, with the three cortical areas and the olfactory bulb ranking last, at approximately one-third of the velocity of the corpus callosum, the 'fastest' region (Figure 2G). Other white matter-rich areas such as the caudate putamen also exhibited high velocities, while the hippocampus, thalamus and hypothalamus - some of the most investigated regions in mouse and human brain aging research (Hargis and Blalock, 2017) - ranked slightly below average. The median CAS across all regions associated with the animals' chronological age (Figure S4B). Yet, the regions' differing velocities resulted in increased per-animal variance, indicating that the transcriptional state of this gene set becomes profoundly desynchronized across the brain. This appeared to be independent of the regions' anatomical location, as the fast-aging corpus callosum stretches between the slow-aging cortical areas and hippocampus (Figure 2H). Interestingly, when we examined the CAS trajectories for the interval between 3 and 21 months, we observed moderate but significant acceleration of the CAS in females compared to males (Figure 2I, S4C,D). In particular, the hypothalamus exhibited the most pronounced acceleration in females (Figure S4D), with the thalamus and corpus callosum indicating a similar trend. These findings are in line with human studies reporting more pronounced expression of immune-related genes in the hippocampus and cortex of aged women (Berchtold et al., 2008; Yuan et al., 2012). Our data could advance the understanding of several sexual-dimorphisms observed in the brain, including the higher age-specific risk of dementia among women (Corrada et al., 2008) in general, and the dynamics of reproductive aging in particular, given the hypothalamus' critical role in regulating reproduction, development and metabolism (Mozhui et al., 2012).

Fiber tracts are foci of accelerated brain aging

Bulk-seq data - even with the regional dissection conducted here - could mask transcriptional changes with age that occur in sub-structures of regions, such as specific layers of the cortex. We thus aimed to validate our CAS analysis with a fine-resolution method that would still capture multiple regions in the same assay. To this end, we performed spatial transcriptomics (Spatial-seq) of the brain across aging, isolating coronal sections from an independent cohort of male mice aged 6, 18, and 21 months (Figure 2J). Using a clustering-based approach to annotate the regional identity of Spatial-seq spots (Figure S5A,B; STAR Methods; Table S6) we identified them as belonging to the hippocampus, cortex (motor and somatosensory area), thalamus, hypothalamus, striatum (including the caudate putamen), choroid plexus and white matter fiber tracts (including the corpus callosum) (Figure 2K and S5C-F). Our data demonstrated robust capture of the same regions across age groups and individuals (Figure S5G-L), thus enabling the comparison of DEGs found in bulk-seq with spatial-seq data (Table S7). We confirmed a more pronounced regulation of DEGs in the white matter cluster (equivalent to the dissected corpus callosum region) compared to the cortex cluster (equivalent to the motor cortex region), including several of the 82 CAS genes (Figure 2L, Table S7) such as *Trem2* (Figure 2M). Calculating CAS for each Spatial-seq spot identified a distinct, spatially-confined increase of the score along the white matter tracts, including not only the corpus callosum but also other fiber tract sub-structures such as the fimbria and internal capsule (Figure 2N,O). In the cortex, however, we observed only a mild increase of CAS with age. In general, CAS velocities calculated via bulk-seq and those calculated via spatial-seq data were well-correlated (Figure 2P), confirming vastly differing aging velocities between proximal regions *in-situ*.

Heterogeneous velocity of CAS is encoded by glial transcripts

We next sought to quantify the activity of the CAS genes at the single-cell level to identify the cell type(s) that shape the heterogeneous expression dynamics across brain regions. We chose the anterior hippocampus as a representative region given its intermediate CAS velocity (Figure 2G), utilizing frozen punches from the left hemispheres of the bulk-seq cohort (Figure 3A). Single-nuclei sequencing (nuc-seq) yielded all major cell types, with no evidence for a shift in cell type composition with age or sex. We observed the highest baseline CAS in microglia, in line with several of the CAS genes being known immune-response genes (Figure 2B,C). While the CAS showed a statistically significant increase in all cell types (Figure 3B), including a mild elevation in several neuronal populations, the most accentuated increase was observed in microglia (Figure 3C), followed by mature oligodendrocytes, brain endothelial cells (BECs), astrocytes and oligodendrocyte progenitor cells (OPCs). Upon closer examination of the 82 genes, it became clear how the CAS could reflect aging dynamics for several cell types beyond microglia by cell type-specific or cell type-selective gene expression shifts (Figure 3D), including *Gfap* (Astrocytes; Figure S6A), *C4b* (Astrocytes and mature oligodendrocytes; Figure S6B-E), *Gpr17* (OPCs; Figure S7A-E) and *H2-Q7* (BECs; Figure S7F). Of note, this analysis also demonstrated that aging can induce expression of genes that are not detected at young age. For instance, *C4b* was mostly expressed in astrocytes at young age, however its expression became detectable and increased foremost with age in mature oligodendrocytes (Figure S6E). Similarly, expression of *H2-Q7* only became detectable in BECs with old age (Figure S7F). We validated our findings in an independent dataset, using publicly available scRNA-seq data from dissected SVZ of young and old male mice (Dulken et al., 2019) (Figure S8A). Though generated using a different cohort, region and method, the CAS increase was most pronounced in microglia which is consistent with our data (Figure S8B-G). There was also a profound increase of CAS in aNSCs, though the very low number of cells at 28 months (less than 50 per animal) complicates robust calculations of CAS at this age. Thus, the region-to-region differences in CAS velocity are predominantly reflecting age effects in non-neuronal cell types, with microglia having the strongest contribution.

Transcriptional aging of microglia is region-dependent

We finally examined if there were varying CAS dynamics between microglia from regions with fast or slow CAS velocity. To this end, we analyzed scRNA-seq data from the *Tabula Muris* consortium (Figure 4A), where comparable numbers of microglia were collected from the freshly-isolated cerebellum, striatum, hippocampus and cortex, which we considered sufficient equivalents to the cerebellum, caudate putamen (both areas with high velocity), anterior hippocampus (medium velocity) and motor/visual cortex (low velocity) regions (compare Figure 2G). Indeed, as predicted by our bulk-seq results, the CAS in aged microglia increased in all four regions significantly, though with greater magnitude in the cerebellum and striatum, followed by the hippocampus and cortex, respectively (Figure 4B). The same trend was detectable on the level of individual CAS genes, like *Trem2* (Figure 4C). Notably, there was no detectable CAS difference among microglia at young age across the striatum, hippocampus and cortex, while the cerebellum-derived microglia exhibited a slightly higher CAS at baseline. We conclude that the CAS velocities observed in bulk- and spatial-seq data are, in part, representing microglia that exhibit region-specific magnitudes of aging. Future studies should examine if other non-neuronal cells, like mature oligodendrocytes or BECs, would exhibit a similar CAS heterogeneity.

Neuronal transcripts encode region-specific expression patterns

Given that the CAS genes represent only 1.5% of all DEGs (Figure 2A; Table S2), we hypothesized that the remainder could represent region-specific expression shifts. We first compared age-related DEGs across mouse organs to construct organ-specific signatures of aging (Figure S9, STAR Methods). The presence of gene sets with specific regulation found in functionally distinct organs led us to investigate whether individual brain regions exhibit a similar degree of specificity during aging. We found that the number of region-specific DEGs varies greatly (Figure 5A), which we utilized to build aging signatures for each region before calculating the respective score across all other regions (Figure 5B,C, Figure S10A,B). As exemplified by the specific signature of the caudate putamen - a region marked by intertwined gray and white matter structures - we found that most region-specific scores increased with age predominantly in the region on which they were based (Figure 5B-D and Figure S10A,B). Except for the thalamus, pons and SVZ, a given region’s signature velocity outperformed those of all other regions. Thus, dozens to hundreds of genes in the brain are regulated in a region-specific or -selective manner, revealing highly compartmentalized effects of aging within a single organ.

Notably, signature genes appeared to be functionally connected, as exemplified by the caudate putamen-specific signature which was enriched for down-regulated mitochondrial processes and up-regulated cell adhesion and lipid binding functions (Figure 5E and Table S8). To map out the cell types driving this region-specific signature, we analyzed nuc-seq data from the left hemisphere punches of the anterior hippocampus (Figure 3A) and caudate putamen (Figure 4F), where we captured non-neuronal cell types as well as striatum-specific D1- and D2-type medium spiny neurons (D1 and D2 MSNs, respectively). We were able to map several signature genes like *Fgf16*, *S100a10* and *Fabp4* (Figure S10C-E) to distinct cell populations (Figure 5G, STAR Methods and Table S9) suggesting that bulk tissue can indeed capture the expression dynamics of specific cell subsets. Similar to the cell-resolved analysis of the CAS (Figure 3B), we calculated several region-specific signature scores for each cell type in young and old individuals. We found a distinct increase of the caudate putamen-specific signature in D1 and D2 MSNs which was not seen with signatures from other regions (Figure 5H). For instance, expression of muscarinic acetylcholine receptor gene *Chrm3* dropped significantly in the caudate putamen, reflecting down-regulation of this gene specifically in D2 MSNs (Figure 5I,J). In comparison, dentate gyrus granule cells of the hippocampus exhibited a distinct increase of the hippocampus-specific signature (Figure 5K), and we found granule cell-specific regulation of several signature genes such as axon-guidance receptor *Unc5d* (Yamagishi et al., 2011) as well as transcription factor *Onecut1* (Figure 5L,M, Figure S11A,B). Interestingly, *Onecut2*, a member of the same transcription factor family, exhibited selective regulation in the corpus callosum and caudate putamen, namely D1 and D2 MSNs (Figure S11C,D). Onecut transcription factors have only been recently investigated in the CNS as regulators of neuronal differentiation and their motif is a hotspot of DNA single-strand break repair in neurons (Wu et al., 2021). This may indicate that the transcriptome of neuronal populations could be regulated selectively through age-related activity of specific transcription factors.

Finally, we explored whether the biological processes associated with signature genes could indicate differential transcriptional activity across whole pathways or organelles. We observed a significant down-regulation of several mitochondria-related genes in the caudate putamen, including several members of the electron transport chain, which could be indicative of impaired mitochondrial function (Figure 5E). We identified in this region a global, gradual down-regulation of all genes coding for mitochondria-related proteins (Figure S11E), as well as a significant drop in scores for a corresponding mitochondrial signature in aged D2 MSNs, mature oligodendrocytes, and astrocytes (Figure S11F). This was not detected in cell types from the hippocampus or the SVZ (Figure S11F). This specific down-regulation of mitochondrial processes in aged striatum could help to explain previous observations demonstrating selective vulnerability to mitochondrial toxins and stresses in the striatum of old animals (Brouillet et al., 1993; Patki et al., 2009).

In conclusion, we discovered extensive region-specific transcriptional signatures of aging that are largely encoded by expression shifts in distinct neuronal subpopulations reflective of a region's specialization.

Aging results in region-specific expression changes of genes associated with human diseases

Genome-wide association studies (GWAS) can identify candidate genes whose genetic makeup impacts the risk of developing a given disease. We thus wondered if mouse homologues of human GWAS genes implicated in neurodegenerative disorders would be differentially regulated in specific regions during aging. We assembled lists of GWAS genes for AD, PD and multiple sclerosis (MS) (Yang et al., 2021, 2022), and asked if they were significantly enriched among age-related DEGs of a given region (Table S10). We further clustered regions that share differentially expressed disease-associated genes, allowing us to find anatomical 'hubs'. Interestingly, each of the three disease-associated gene sets exhibited a different enrichment pattern (Figure 6A-C) and a varying number of associated genes (Figure 6D-F). MS genes, for instance, showed significant associations with DEGs from nine different regions that fell into two clusters, indicating two disparate subsets. One cluster consisted of several regions, including the corpus callosum and cerebellum, that up-regulated a shared set of inflammation-related genes such as *Stat3*, *Ly86* and *Irf8*, all of which were also part of the CAS (Figure 6A). This raises the intriguing possibility of similarities between the pathophysiology of inflammation and demyelination associated with MS and the accelerated aging observed in white matter-rich areas. The visual and motor cortex regions formed the second cluster, exhibiting even numbers of up- and down-regulated MS genes. This supports recent evidence indicating transcriptional shifts (e.g. of *Cbln2*, Figure 6G) in cortical areas that can occur far away from the actual lesions (Kihara et al., 2022) and highlights the need to broadly study regional patterns of gene expression to understand the role of MS-associated genes.

GWAS hits for AD included genes that are currently under intense mechanistic investigation, including *Apoe*, *Ms4a6d*, *Plcg2* and *Gab2* (Reiman et al., 2007) (Figure 6B,H). They were part of DEGs that are upregulated in a small cluster of three regions: the choroid plexus, corpus callosum, and pons, suggesting a different region-specific regulation of these genes as a potential modulator of disease presentation. In contrast, PD-related genes, like the neuroprotective gene *Ip6k2 (Nagpal et al., 2022)* (Figure 6I), were not concentrated within a cluster of regions but rather distributed across the choroid plexus, cerebellum, SVZ, and visual cortex with limited overlap (Figure 6C). This pattern indicates a disperse regulation of disease-associated genes among the regions studied, which can add to our understanding of how and where PD may progress. Of note, the substantia nigra, a major region where PD typically manifests, was not quantified in our study.

Taken together, our data demonstrate that genetic risk factors linked to three major neurological and neurodegenerative diseases are affected by age in a region-selective manner. While we cannot predict whether the directionality of the regulation itself has a biological consequence, we consider that the region-specific differential regulation of such genes could be an additional factor modulating disease risk.

Discussion

We report here a comprehensive spatiotemporal map of gene expression across the mouse brain throughout the adult lifespan, consisting of 847 bulk, 16,277 spatial spot and 81,616 single-cell transcriptomes - supported by several publicly available spatial- and scRNA-seq datasets. We find that aging affects regional transcriptomes with widely varying magnitude and timing, including the expression of risk genes for neurodegenerative disorders. Gene expression shifts during aging fall into two general categories: a single set of 82 genes that reflects shared aging dynamics of glia cells with spatially-defined magnitude, and a dozen gene sets with region-specific activity encoded in specialized neuronal subpopulations. We establish a CAS for each mouse and region, analyzing its trajectory and rate of change over time as a quantitative measure of a region's transcriptional age. Across analyses, white matter and white matter-rich regions emerge as the most transcriptionally impacted areas during aging. It will be particularly interesting to explore how these changes relate to developmental patterns of myelination and brain function, and whether susceptibility to brain aging and dysfunction are related to developmental processes.

The advent of single-cell technologies and cell dissociation methods have enabled the exploration of an ever increasing number of cell populations in the brain (BRAIN Initiative Cell Census Network (BICCN), 2021), which in turn allows for cell type-specific characterization of gene expression during aging (Buckley et al., 2022). The interplay between cell type and regional niche during aging is, however, yet to be more deeply understood. Our results emphasize the importance of region identity as a profound modulator of gene expression dynamics in the context of aging and neurodegeneration. It will be important for future studies to examine if these heterogeneous expression patterns result in corresponding shifts of the proteome or downstream functional changes in neuronal activity and plasticity. While we established that a common aging signature increased in microglia with region-dependent effect size, we propose further exploration of the CAS in other non-neuronal cell types. This may help clarify if microglia are active drivers of the regional expression dynamics described here, or rather respond to cues provided by other cell type(s) in the region.

Our data reveal that certain brain regions are selectively vulnerable to aging, with the white matter fiber tracts exhibiting a particular sensitivity. These areas are dense with myelinated axons and myelinating cells, forming the basis of neurotransmission across brain regions (Liu et al., 2017). The strong activation of immune- and inflammation-related genes, as well as differential expression of remyelination regulators like *Gpr17* (Dziedzic et al., 2020; Rivera et al., 2021), suggest that the homeostasis of this region is compromised at old age. This could perturb myelin sheath integrity and potentially impair axonal signal transmission between regions as an early event in brain aging. In line with this hypothesis, rejuvenation of hippocampal oligodendrogenesis in aged mice via injection of young cerebrospinal fluid (CSF) improves long-term memory consolidation, thus demonstrating a causal role of compromised myelin on cognition (Iram et al., 2022). The region-specific transcriptome atlas generated here can form a basis for testing rejuvenation strategies such as dietary interventions, plasma and CSF-based therapeutics as well as epigenetic reprogramming to quantify their spatiotemporal impact on the brain at the molecular level.

Our findings strongly support the notion that the impacts of aging on brain function are region specific, potentially relating to the regional vulnerability across different diseases as well as the varied manifestations of neurodegeneration at the level of an individual. We demonstrate that key genetic risk genes are differentially expressed in a region-specific manner, thus locally amplifying, or attenuating their impact on disease pathways. Importantly, our findings also suggest that aging may drive dysfunction in brain regions that are not predominantly affected and studied by classical pathological hallmarks, highlighted by AD risk genes including *Apoe*, *Ms4a6d*, *Plcg2* and *Gab2* which are dysregulated with aging in the mouse choroid plexus, corpus callosum and pons with aging. The translation of these findings to humans may serve as a new brain cartography leading to novel treatment strategies and interventions.

References

Berchtold, N.C., Cribbs, D.H., Coleman, P.D., Rogers, J., Head, E., Kim, R., Beach, T., Miller, C., Troncoso, J., Trojanowski, J.Q., et al. (2008). Gene expression changes in the course of normal brain aging are sexually dimorphic. Proc. Natl. Acad. Sci. U. S. A. *105*, 15605–15610. .

BRAIN Initiative Cell Census Network (BICCN) (2021). A multimodal cell census and atlas of the mammalian primary motor cortex. Nature *598*, 86–102. .

Brouillet, E., Jenkins, B.G., Hyman, B.T., Ferrante, R.J., Kowall, N.W., Srivastava, R., Roy, D.S., Rosen, B.R., and Beal, M.F. (1993). Age-dependent vulnerability of the striatum to the mitochondrial toxin 3-nitropropionic acid. J. Neurochem. *60*, 356–359. .

Buckley, M.T., Sun, E., George, B.M., Liu, L., Schaum, N., Xu, L., Reyes, J.M., Goodell, M.A., Weissman, I.L., Wyss-Coray, T., et al. (2022). Cell type-specific aging clocks to quantify aging and rejuvenation in regenerative regions of the brain.

Chen, W.-T., Lu, A., Craessaerts, K., Pavie, B., Sala Frigerio, C., Corthout, N., Qian, X., Laláková, J., Kühnemund, M., Voytyuk, I., et al. (2020). Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. Cell *182*, 976–991.e19. .

Colantuoni, C., Lipska, B.K., Ye, T., Hyde, T.M., Tao, R., Leek, J.T., Colantuoni, E.A., Elkahloun, A.G., Herman, M.M., Weinberger, D.R., et al. (2011). Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature *478*, 519–523. .

Corrada, M.M., Brookmeyer, R., Berlau, D., Paganini-Hill, A., and Kawas, C.H. (2008). Prevalence of dementia after age 90: results from the 90+ study. Neurology *71*, 337–343. .

De Miguel, Z., Khoury, N., Betley, M.J., Lehallier, B., Willoughby, D., Olsson, N., Yang, A.C., Hahn, O., Lu, N., Vest, R.T., et al. (2021). Exercise plasma boosts memory and dampens brain inflammation via clusterin. Nature *600*, 494–499. .

DeTomaso, D., Jones, M.G., Subramaniam, M., Ashuach, T., Ye, C.J., and Yosef, N. (2019). Functional interpretation of single cell similarity maps. Nat. Commun. *10*, 4376. .

Dulken, B.W., Buckley, M.T., Navarro Negredo, P., Saligrama, N., Cayrol, R., Leeman, D.S., George, B.M., Boutet, S.C., Hebestreit, K., Pluvinage, J.V., et al. (2019). Single-cell analysis reveals T cell infiltration in old neurogenic niches. Nature *571*, 205–210. .

Dziedzic, A., Miller, E., Saluk-Bijak, J., and Bijak, M. (2020). The GPR17 Receptor-A Promising Goal for Therapy and a Potential Marker of the Neurodegenerative Process in Multiple Sclerosis. Int. J. Mol. Sci. *21*. https://doi.org/10.3390/ijms21051852.

Feng, X., Guo, J., Sigmon, H.C., Sloan, R.P., Brickman, A.M., Provenzano, F.A., Small, S.A., and Alzheimer's Disease Neuroimaging Initiative (2020). Brain regions vulnerable and resistant to aging without Alzheimer's disease. PLoS One *15*, e0234255. .

Ferreira, M., Francisco, S., Soares, A.R., Nobre, A., Pinheiro, M., Reis, A., Neto, S., Rodrigues, A.J., Sousa, N., Moura, G., et al. (2021). Integration of segmented regression analysis with weighted gene correlation network analysis identifies genes whose expression is remodeled throughout physiological aging in mouse tissues. Aging *13*, 18150–18190. .

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

Goldmann, T., Wieghofer, P., Jordão, M.J.C., Prutek, F., Hagemeyer, N., Frenzel, K., Amann, L., Staszewski, O., Kierdorf, K., Krueger, M., et al. (2016). Origin, fate and dynamics of macrophages at central nervous system interfaces. Nat. Immunol. *17*, 797–805. .

Hahn, O., Fehlmann, T., Zhang, H., Munson, C.N., Vest, R.T., Borcherding, A., Liu, S., Villarosa, C., Drmanac, S., Drmanac, R., et al. (2020). CoolMPS for robust sequencing of single-nuclear RNAs captured by droplet-based method. Nucleic Acids Res. https://doi.org/10.1093/nar/gkaa1127.

Ham, S., and Lee, S.-J.V. (2020). Advances in transcriptome analysis of human brain aging. Exp. Mol. Med. *52*, 1787–1797. .

Hargis, K.E., and Blalock, E.M. (2017). Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us? Behav. Brain Res. *322*, 311–328. .

Hou, Y., Dan, X., Babbar, M., Wei, Y., Hasselbalch, S.G., Croteau, D.L., and Bohr, V.A. (2019). Ageing as a risk factor for neurodegenerative disease. Nat. Rev. Neurol. *15*, 565–581. .

Iram, T., Kern, F., Kaur, A., Myneni, S., Morningstar, A.R., Shin, H., Garcia, M.A., Yerra, L., Palovics, R., Yang, A.C., et al. (2022). Young CSF restores oligodendrogenesis and memory in aged mice via Fgf17. Nature *605*, 509–515. .

Kalache, A., and Gatti, A. (2003). Active ageing: a policy framework. Adv. Gerontol. *11*, 7–18. .

Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M.M., Pletikos, M., Meyer, K.A., Sedmak, G., et al. (2011). Spatio-temporal transcriptome of the human brain. Nature *478*, 483–489. .

Kihara, Y., Zhu, Y., Jonnalagadda, D., Romanow, W., Palmer, C., Siddoway, B., Rivera, R., Dutta, R., Trapp, B.D., and Chun, J. (2022). Single-Nucleus RNA-seq of Normal-Appearing Brain Regions in Relapsing-Remitting vs. Secondary Progressive Multiple Sclerosis: Implications for the Efficacy of Fingolimod. Front. Cell. Neurosci. *16*, 918041. .

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

Lee, C.K., Weindruch, R., and Prolla, T.A. (2000). Gene-expression profile of the ageing brain in mice. Nat. Genet. *25*, 294–297. .

Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. Nature *445*, 168–176.

Lenth, R.V. (2016). Least-Squares Means: The R Package lsmeans. J. Stat. Softw. *69*, 1–33. .

Liu, H., Yang, Y., Xia, Y., Zhu, W., Leak, R.K., Wei, Z., Wang, J., and Hu, X. (2017). Aging of cerebral white matter. Ageing Res. Rev. *34*, 64–76.

López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M., and Kroemer, G. (2013). The hallmarks of aging. Cell *153*, 1194–1217. .

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.

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. Nature *570*, 332–337. .

Melé, M., Ferreira, P.G., Reverter, F., DeLuca, D.S., Monlong, J., Sammeth, M., Young, T.R., Goldmann, J.M., Pervouchine, D.D., Sullivan, T.J., et al. (2015). Human genomics. The human transcriptome across tissues and individuals. Science *348*, 660–665. .

Mozhui, K., Lu, L., Armstrong, W.E., and Williams, R.W. (2012). Sex-specific modulation of gene expression networks in murine hypothalamus. Front. Neurosci. *6*, 63. .

Mueller, O. RNA Integrity Number (RIN) – Standardization of RNA Quality Control.

Nagpal, L., Kornberg, M.D., and Snyder, S.H. (2022). Inositol hexakisphosphate kinase-2 non-catalytically regulates mitophagy by attenuating PINK1 signaling. Proc. Natl. Acad. Sci. U. S. A. *119*, e2121946119. .

Niccoli, T., Partridge, L., and Isaacs, A.M. (2017). Ageing as a risk factor for ALS/FTD. Hum. Mol. Genet. *26*, R105–R113. .

Pálovics, R., Keller, A., Schaum, N., Tan, W., Fehlmann, T., Borja, M., Kern, F., Bonanno, L., Calcuttawala, K., Webber, J., et al. (2022). Molecular hallmarks of heterochronic parabiosis at single-cell resolution. Nature *603*, 309–314. .

Partridge, L., Deelen, J., and Slagboom, P.E. (2018). Facing up to the global challenges of ageing. Nature *561*, 45–56. .

Patki, G., Che, Y., and Lau, Y.-S. (2009). Mitochondrial dysfunction in the striatum of aged chronic mouse model of Parkinson’s disease. Front. Aging Neurosci. *1*, 3. .

Picelli, S., Björklund, Å.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods *10*, 1096–1098. .

Pluvinage, J.V., Haney, M.S., Smith, B.A.H., Sun, J., Iram, T., Bonanno, L., Li, L., Lee, D.P., Morgens, D.W., Yang, A.C., et al. (2019). CD22 blockade restores homeostatic microglial phagocytosis in ageing brains. Nature *568*, 187–192. .

Pomponio, R., Erus, G., Habes, M., Doshi, J., Srinivasan, D., Mamourian, E., Bashyam, V., Nasrallah, I.M., Satterthwaite, T.D., Fan, Y., et al. (2020). Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. Neuroimage *208*, 116450. .

Reiman, E.M., Webster, J.A., Myers, A.J., Hardy, J., Dunckley, T., Zismann, V.L., Joshipura, K.D., Pearson, J.V., Hu-Lince, D., Huentelman, M.J., et al. (2007). GAB2 alleles modify Alzheimer’s risk in APOE epsilon4 carriers. Neuron *54*, 713–720. .

Rivera, A.D., Pieropan, F., Chacon-De-La-Rocha, I., Lecca, D., Abbracchio, M.P., Azim, K., and Butt, A.M. (2021). Functional genomic analyses highlight a shift in Gpr17-regulated cellular processes in oligodendrocyte progenitor cells and underlying myelin dysregulation in the aged mouse cerebrum. Aging Cell *20*, e13335. .

Schaum, N., Lehallier, B., Hahn, O., Pálovics, R., Hosseinzadeh, S., Lee, S.E., Sit, R., Lee, D.P., Losada, P.M., Zardeneta, M.E., et al. (2020). Ageing hallmarks exhibit organ-specific temporal signatures. Nature *583*, 596–602. .

Sekar, A., Bialas, A.R., de Rivera, H., Davis, A., Hammond, T.R., Kamitaki, N., Tooley, K., Presumey, J., Baum, M., Van Doren, V., et al. (2016). Schizophrenia risk from complex variation of complement component 4. Nature *530*, 177–183. .

Soreq, L., UK Brain Expression Consortium, North American Brain Expression Consortium, Rose, J., Soreq, E., Hardy, J., Trabzuni, D., Cookson, M.R., Smith, C., Ryten, M., et al. (2017). Major Shifts in Glial Regional Identity Are a Transcriptional Hallmark of Human Brain Aging. Cell Rep. *18*, 557–570. .

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888–1902.e21. .

Tabula Muris Consortium (2020). A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. Nature *583*, 590–595. .

Trabzuni, D., Ryten, M., Walker, R., Smith, C., Imran, S., Ramasamy, A., Weale, M.E., and Hardy, J. (2011). Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. J. Neurochem. *119*, 275–282. .

Van Hoesen, G.W., Hyman, B.T., and Damasio, A.R. (1991). Entorhinal cortex pathology in Alzheimer’s disease. Hippocampus *1*, 1–8. .

Wager-Miller, J., Murphy Green, M., Shafique, H., and Mackie, K. (2020). Collection of Frozen Rodent Brain Regions for Downstream Analyses. J. Vis. Exp. https://doi.org/10.3791/60474.

Wu, W., Hill, S.E., Nathan, W.J., Paiano, J., Callen, E., Wang, D., Shinoda, K., van Wietmarschen, N., Colón-Mercado, J.M., Zong, D., et al. (2021). Neuronal enhancers are hotspots for DNA single-strand break repair. Nature *593*, 440–444. .

Ximerakis, M., Lipnick, S.L., Innes, B.T., Simmons, S.K., Adiconis, X., Dionne, D., Mayweather, B.A., Nguyen, L., Niziolek, Z., Ozek, C., et al. (2019). Single-cell transcriptomic profiling of the aging mouse brain. Nat. Neurosci. *22*, 1696–1708. .

Yamagishi, S., Hampel, F., Hata, K., Del Toro, D., Schwark, M., Kvachnina, E., Bastmeyer, M., Yamashita, T., Tarabykin, V., Klein, R., et al. (2011). FLRT2 and FLRT3 act as repulsive guidance cues for Unc5-positive neurons. EMBO J. *30*, 2920–2933. .

Yang, A.C., Kern, F., Losada, P.M., Agam, M.R., Maat, C.A., Schmartz, G.P., Fehlmann, T., Stein, J.A., Schaum, N., Lee, D.P., et al. (2021). Dysregulation of brain and choroid plexus cell types in severe COVID-19. Nature *595*, 565–571. .

Yang, A.C., Vest, R.T., Kern, F., Lee, D.P., Agam, M., Maat, C.A., Losada, P.M., Chen, M.B., Schaum, N., Khoury, N., et al. (2022). A human brain vascular atlas reveals diverse mediators of Alzheimer’s risk. Nature *603*, 885–892. .

Yeh, F.L., Hansen, D.V., and Sheng, M. (2017). TREM2, Microglia, and Neurodegenerative Diseases. Trends Mol. Med. *23*, 512–533. .

Yuan, Y., Chen, Y.-P.P., Boyd-Kirkup, J., Khaitovich, P., and Somel, M. (2012). Accelerated aging-related transcriptome changes in the female prefrontal cortex. Aging Cell *11*, 894–901. .

Zahn, J.M., Poosala, S., Owen, A.B., Ingram, D.K., Lustig, A., Carter, A., Weeraratna, A.T., Taub, D.D., Gorospe, M., Mazan-Mamczarz, K., et al. (2007). AGEMAP: a gene expression database for aging in mice. PLoS Genet. *3*, e201. .

Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science *347*, 1138–1142. .

Zhang, Y., Sloan, S.A., Clarke, L.E., Caneda, C., Plaza, C.A., Blumenthal, P.D., Vogel, H., Steinberg, G.K., Edwards, M.S.B., Li, G., et al. (2016). Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. Neuron *89*, 37–53. .

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. .

Babraham Bioinformatics - SeqMonk Mapped Sequence Analysis Tool.

RStudio.

CellPlot: R package for the integrated visualisation of functional enrichment and expression data (Github).

Acknowledgments: We thank the members of the Wyss-Coray laboratory for feedback and support, and H. Zhang, D. Berdnik, K. Dickey and D. Channappa for laboratory management. We further thank the Neuroscience Microscopy Service center and director G. Wang for microscopy training and assistance in acquiring images for Spatial-seq. We thank A. Bastian for advice in establishing the Spatial-seq experiments. This work was supported by the Schaller-Nikolich Foundation (A.Ke.), the Wu Tsai Neurosciences Institute and Bertarelli Foundation (T.W.-C.), the Cure Alzheimer's Fund (T.W.-C.), a National Institute of Aging grant R01-AG072255 (T.W.-C.), the Milky Way Research Foundation (T.W.-C.), the American Heart Association-Allen Initiative in Brain Health and Cognitive Impairment (T.W.-C.), the Simons Foundation - Simons Collaboration on the Plasticity and the Aging Brain (T.W.-C.), The Phil and Penny Knight Initiative for Brain Resilience (T.W.-C.), Michael J. Fox Foundation for Parkinson's Research grants 125491594 (A.Ke and F.K.) and MJFF-021418 (T.W.-C., A.Ke. and F.K.).

Author contributions: O.H. and T.W.-C. conceptualized the study. O.H. designed and led experiments, conducted bioinformatic analyses and prepared the manuscript. A.G.F. and O.H. selected regions and devised the brain dissociation method. J.H. and J.R.H aided in region selection and method design. O.H., I.H.G., C.M., A.G.F., M.A., B.K. and A.Ka. conducted the tissue collection. A.G.F. dissociated and organized brain regions, and extracted RNA. N.L. and O.H. established the bulk-seq workflow. A.G.F., M.A. and O.H. processed libraries for bulk-seq. B.K. and O.H. conducted the Spatial-seq experiments. I.H.G., C.M. and A.Ka. aided with the H&E workflow. M.A. and A.G.F. performed the nuc-seq experiments. P.M.L., B.L., R.P., F.K. and A.Ke. assisted with analysis and bioinformatics procedures. B.K. developed the searchable web interface (Shiny app). O.H. and T.W.-C. edited the manuscript with input from all authors. All authors read and approved the final manuscript.

Declaration of interests: The authors declare no competing interests.

Main Figures

Fig 1

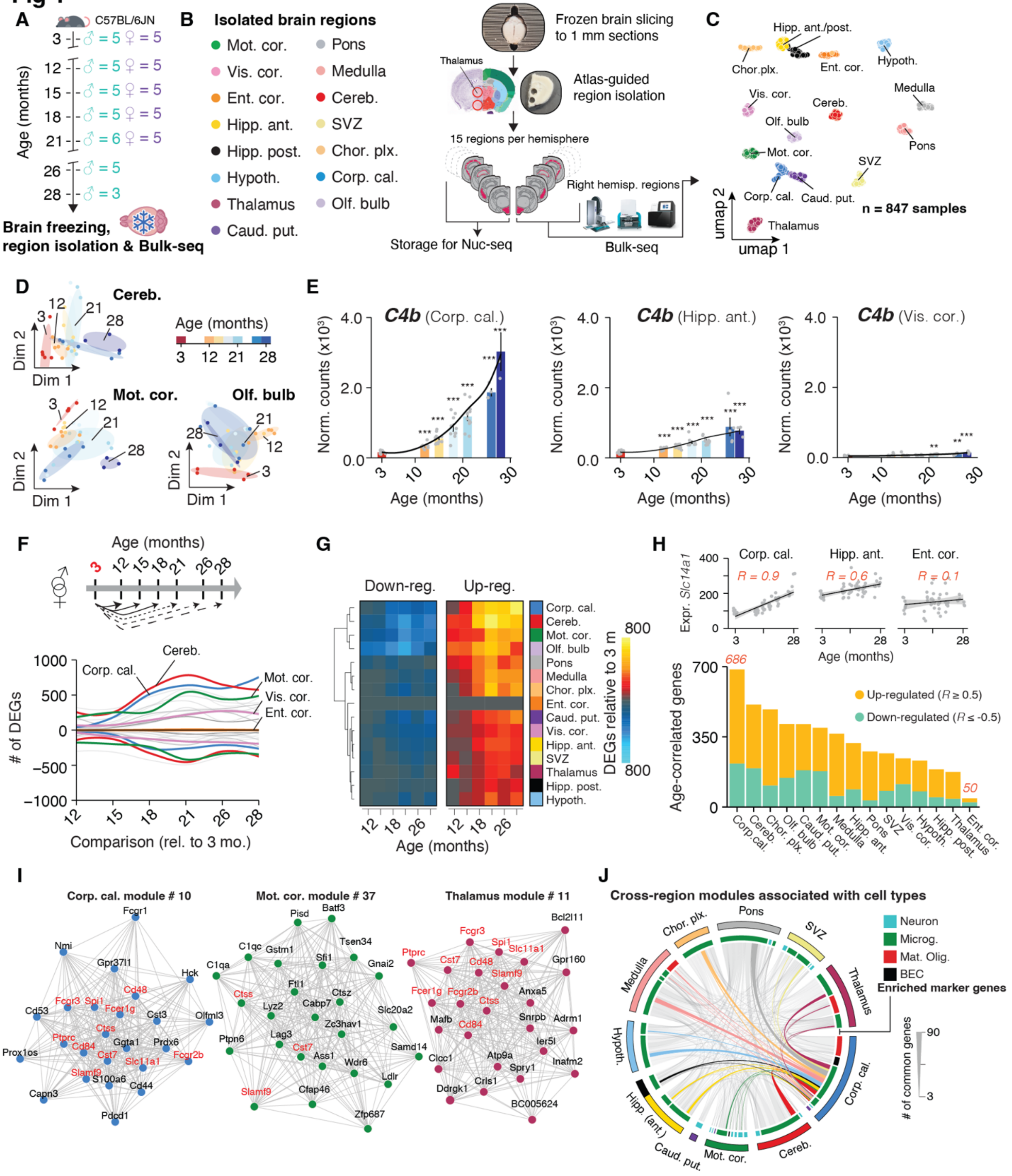

Figure 1 Brain regions exhibit distinct transcriptional patterns of aging independent of anatomical proximity

(A) Cohort overview. Whole brains were collected from male (n = 3-5, 3–28 months) and female (n = 5, 3–21 months) C57BL/6JN mice. (B) Study outline. 15 brain regions were isolated from each hemisphere of the brains collected in (A). Regions from the right hemisphere were analyzed using Bulk-seq. (C) UMAP representation of brain region transcriptomes (n = 847 total samples), based on the first 40 principal components. (D) Diffusion maps of region transcriptomes from male cerebellum, motor cortex, and olfactory bulb. Dim., dimension. (E) *C4b* expression in corpus callosum, anterior hippocampus and visual cortex. Black lines indicate averaged-smoothed gene expression. Differential expression relative to the 3 months group is indicated. Data are mean ± s.e.m. Two-sided Wald test, adjusted for multiple testing. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. (F) Smoothed line plot displaying the number of DEGs for pairwise comparisons, referenced to data at 3 months. Positive (negative) values represent upregulated (downregulated) genes, gray lines represent non-labelled regions. DEGs that reached significance in ≥ 2 pairwise comparisons were included. (G) Heat map of the data in (F). (H) Number of genes that significantly correlate with age (Spearman's rho ≥ 0.5), colored by up- and down-regulation. (I) Networks of the most highly connected genes ('eigengenes') of three exemplary modules with significant age-association identified in corpus callosum, motor cortex and thalamus. Networks display connections of the corresponding topological overlap above a threshold of 0.08. (J) Chord diagram representation of genes shared in age-associated modules across regions. Modules with significant enrichment of cell type markers are displayed. Modules and associated genes are listed in Table S4.

Fig 2

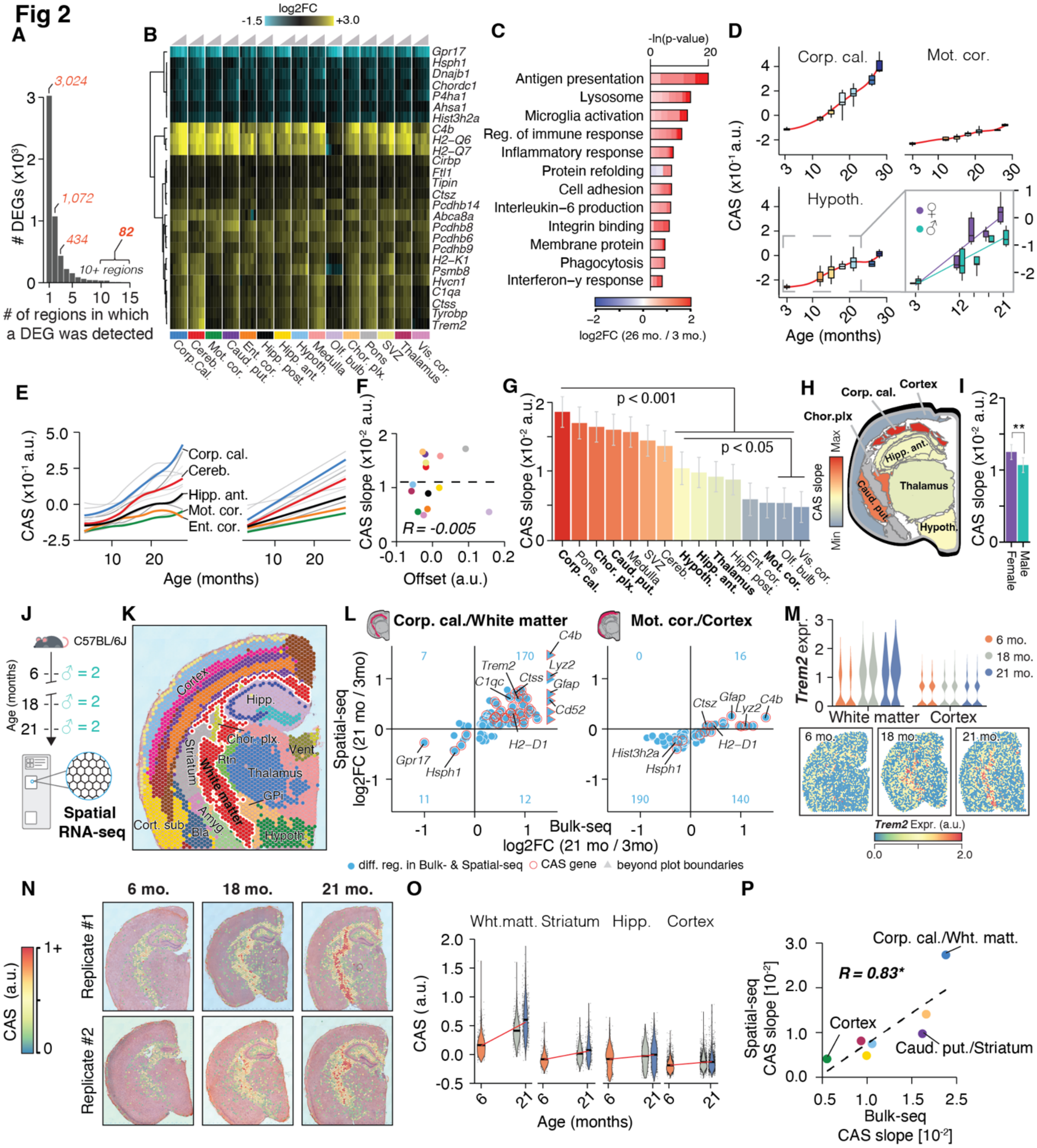

Figure 2 A common gene signature detects accelerated aging in white matter tracts

(A) Bar graph indicating the number of regions in which a given DEG was detected (Table S2). (B) Region-wise expression changes with age (column-wise from left to right) for genes with shifts in at least 10 of the 15 collected regions. (C) Representative GO analysis of 82 genes with shifts in at least 10 of the 15 collected regions that make up the CAS. Lengths of bars represent negative ln-transformed Padj using two-sided Fisher's exact test. Colors indicate gene-wise log2 fold-changes (log2(FC)) between 26 and 3 months old mice as measured in the corpus callosum. Numbers beside bars indicate differentially expressed genes in that GO category. The complete list of enriched GO terms can be found in Table S5. (D) CAS trajectories in the corpus callosum, motor cortex, and hypothalamus. Insert indicates trajectories for male and females in the hypothalamus from 3 to 21 months. (E) CAS trajectories of all regions approximated via local estimate (LOESS) and linear regression, colored by region; gray lines represent non-labelled regions. (F) Offset and slope comparison for linear models in (E), colored by region. Linear regression (dashed line) and Spearman correlation coefficient are indicated. (G) Slope of linear regressions in (D), colored by slope. Data are mean ± 95% confidence intervals. Bolded regions are highlighted in the following panel. (H) Coronal cross-section sketch of the mouse brain, with regions colored according to CAS linear slopes. Corpus callosum was chosen to represent white matter tracts. (I) Slope of linear regression across all brain regions from 3 to 21 months, colored by sex. Data are mean ± 95% confidence intervals. Two-sided Tukey's HSD test, adjusted for multiple testing, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. The highest (least significant) Pval is indicated. (J) Spatial-seq experiment overview. Brain tissue was collected from an independent male C57BL/6J mouse cohort (n = 2 mice; 6, 18 and 21 months). (K) Representative spatial transcriptome data (6 months replicate #2), colored by cluster-based annotation, according to Fig S5. Labels represent region-level annotation according to Fig S5. Complete data description and abbreviations are in Fig S5. (L) Comparison of Bulk- and Spatial-seq differential expression results in white matter cluster/corpus callosum punch; cortex cluster/motor cortex punch. DEGs (Padj < 0.05) found in both datasets are shown, with their log2-transformed expression ratios (21 rel. to 3 months) in Bulk- and Spatial-seq data. CAS genes are highlighted. The number of overlapping DEGs in each quadrant is indicated in blue. (M) Spatially-resolved expression of *Trem2* across age. Violin plots represent expression in white matter- and cortex-associated spots, split by replicates. (N) Spatial representation of CAS. Spots with values ≥ 0 are shown. (O) Violin plot representing CAS across spatial clusters of white matter, striatum, hippocampus and cortex. Red line indicates linear regression fit. (P) Comparison of CAS slopes for linear models in Bulk- and Spatial-seq, colored by region. Linear regression (dashed line) and Spearman correlation coefficient are indicated. Corpus callosum, caudate putamen and motor cortex regions were chosen to represent white matter, striatum and cortex, respectively.

Fig 3

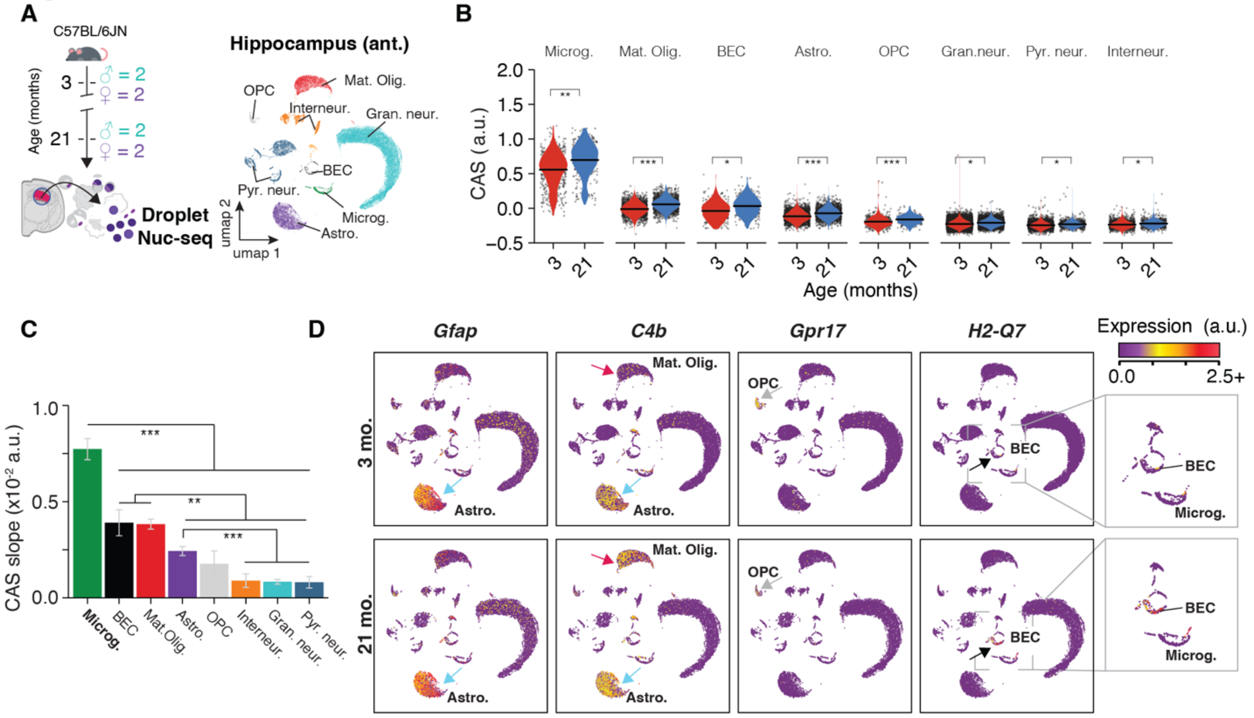

Figure 3 Aging in glia and endothelial cells is the major contributor to CAS increase

(A) Nuc-seq experiment overview. Nuc-seq of left-hemisphere regions of the anterior hippocampus from the same mice used for bulk RNA-seq (n = 2 males, n = 2 females; 3, and 21 months). UMAP representation of all nuclei populations (n = 36,339 cells). (B) Violin plot representing CAS across hippocampal cell types. Points indicate nuclei-wise expression levels, and the violin indicates average distribution of expression split by age. *P* values calculated with two-tailed t-test on per-replicate median of CAS, adjusted for multiple testing. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$ (C) CAS slope of linear regressions in (B), colored by cell type. Data are mean ± 95% confidence intervals. Two-sided Tukey's HSD test, adjusted for multiple testing, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. The highest (least significant) Pval is indicated. (D) Expression of CAS genes *Gfap*, *C4b*, *Gpr17* and *H2-Q7*. Quantification and statistical analysis can be found in Figures S6 and S7.

Fig 4

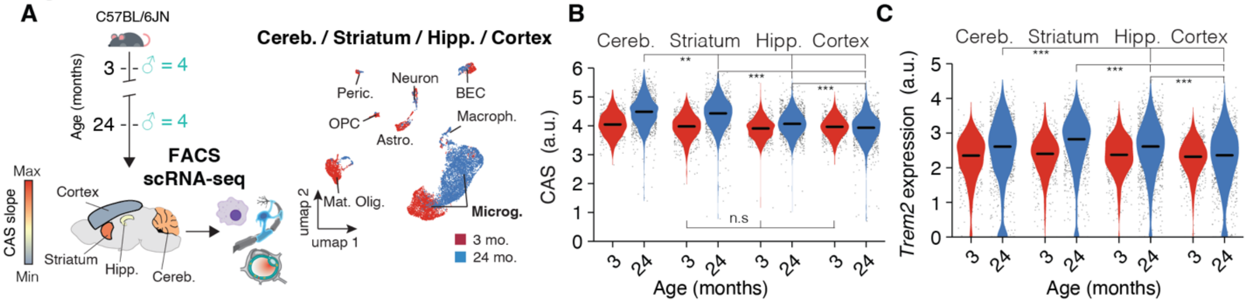

Figure 4 CAS analysis reveals that transcriptional aging of microglia depends on their region of origin.

(A) Meta-analysis of scRNA-seq data from (Tabula Muris Consortium, 2020) of microglia from cerebellum, striatum, hippocampus and cortex. UMAP representation of all cell populations (n = 6,373 cells), colored by age. Regions are colored according to CAS slopes in Figure 2G. (B,C) Violin plot representing (B) CAS and (C) *Trem2* expression across microglia from four different brain regions. Points indicate nuclei-wise expression levels, and the violin indicates average distribution of expression split by age. (MAST, Benjamini–Hochberg correction; false discovery rate (FDR) < 0.05 and logFC > 0.2 to be significant). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Fig 5

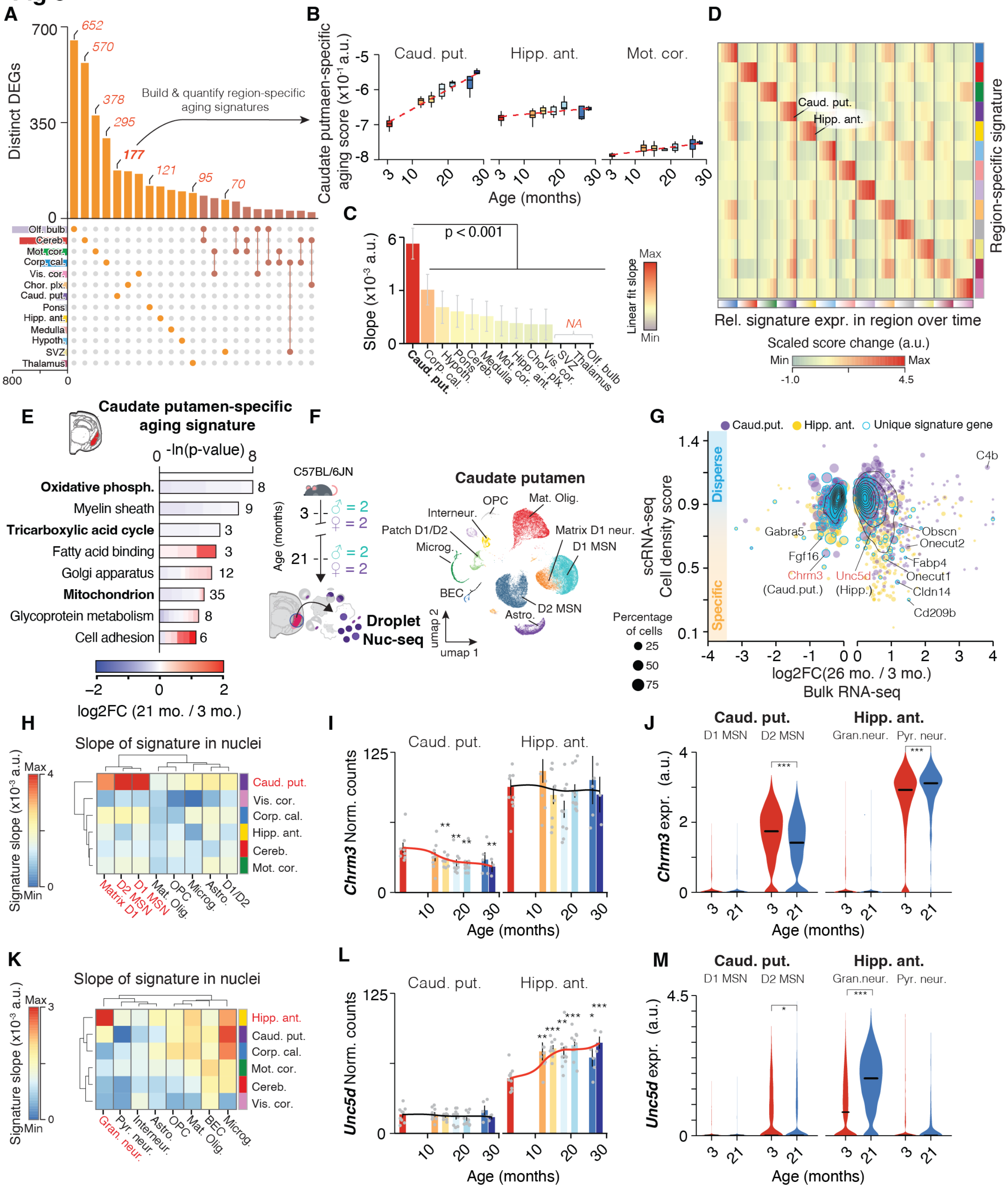

Figure 5 Region-specific expression shifts are encoded by neuronal transcripts

(A) Regional specificity of DEGs. UpSet plot showing a matrix layout of DEGs shared across and specific to each region. Each matrix column represents either DEGs specific to a region (single circle with no vertical lines) or DEGs shared between regions, with the vertical line indicating the regions that share that given DEG. Top, bar graph displays the number of DEGs in each combination of regions. Left bottom, bar graph displays the total number of DEGs for a given region. Gene sets with ≥ 25 genes are shown. Unique gene sets were used to construct region-specific aging signatures. (B) Trajectories of caudate putamen-specific aging score in the caudate putamen, anterior hippocampus and motor cortex. Linear fit is indicated as dashed lines. (C) Slope of linear regressions in (B), colored by slope. Data are mean ± 95% confidence intervals. Two-sided Tukey's HSD test, adjusted for multiple testing, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. The highest (least significant) Pval is indicated. (D) Region-wise score changes with age relative to 3 months (column-wise from left to right) for region-specific signatures. Score changes are z-scaled within a row. Quantification and statistical analysis can be found in Fig S10. (E) Representative GO enrichment as in (Figure 2C) for 177 DEGs unique to the caudate putamen that make up its specific signature. The complete list of enriched GO terms can be found in Table S8. (F) Nuc-seq experiment overview. Nuc-seq of left-hemisphere regions of the caudate putamen from the same mice used for bulk RNA-seq ($n = 2$ males, $n = 2$ females; 3, and 21 months). UMAP representation of all nuclei populations ($n = 45{,}277$ cells). (G) Single-nuclei dispersion scores plotted against log2-transformed expression ratios between 21 and 3 months (bulk RNA-seq) for the caudate putamen and anterior hippocampus. The colors represent different organ types and size of the dots corresponds to the percentage of cells that express a given gene. Genes that make up the region-specific score are highlighted. (H) Slope of cell type-wise changes with age for representative region-specific signatures from (D). D1 and D2 Medium spiny neuron populations (MSN) are highlighted, as they exhibit a distinct increase with age exclusively for the caudate putamen-specific signature. (I) Bulk expression across caudate putamen and anterior hippocampus for *Chrm3*. Black lines indicate averaged-smoothed gene expression. The trajectory with significant age effect is highlighted. Data are mean ± s.e.m. (J) Violin plot representing *Chrm3* expression across neuronal cell types in caudate putamen and anterior hippocampus. Points indicate nuclei-wise expression levels, and the violin indicates average distribution of expression split by age. (MAST, Benjamini–Hochberg correction; false discovery rate (FDR) < 0.05 and logFC > 0.2 to be significant). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. (K) Slope of cell type-wise changes with age for representative region-specific signatures from (D). Granule cells are highlighted, as they exhibit a distinct increase with age exclusively for the anterior hippocampus-specific signature. (L) Same as (I) for *Unc5d*. (M) Same as (J) for *Unc5d*.

Fig 6

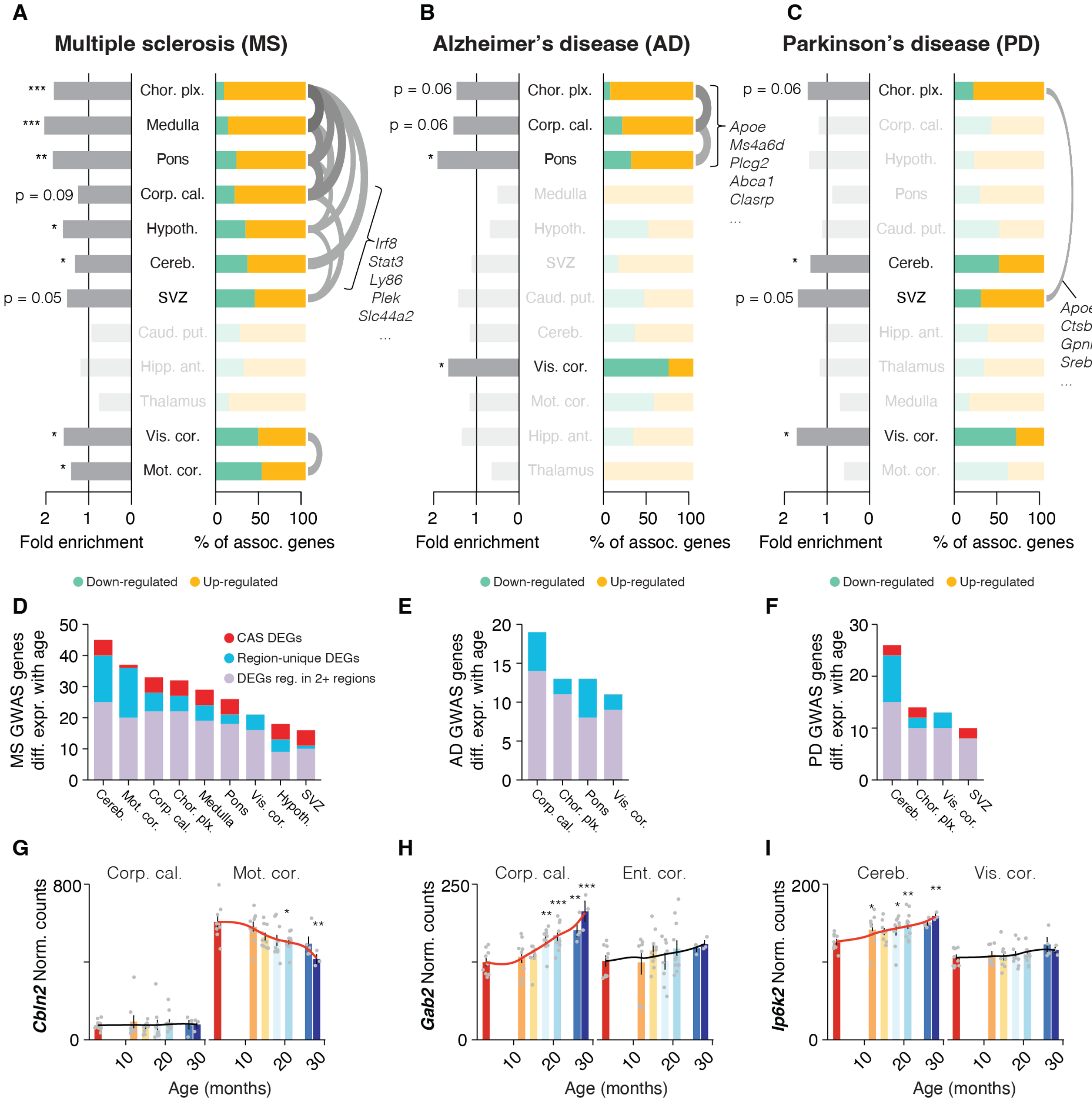

Figure 6 Interplay of region and age shapes expression of disease variant homologues

(A-C) Enrichment analysis of region-resolved DEGs for human GWAS variants for (A) MS, (B) AD and (C) PD. Associated genes are listed in Table S10. Fold enrichment (left bars) and the relative composition of disease-associated DEGs with respect to their regulation (right bars) is indicated. Regions with no significant enrichment are transparent. The vertical order of regions results from hierarchical clustering on a pairwise Jacquard Distance matrix, so regions with overlapping DEGs will cluster together. Gene overlaps with a Jaccard index ≥ 0.25 are indicated with an arc. One-sided Fisher's exact test with hypergeometric distribution, Benjamini–Hochberg correction. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. (D-F) Number of DEGs per region that are homologues of human GWAS variant for (D) MS, (E) AD and (F) PD. Colors group the genes into CAS DEGs, region-specific DEGs or other (DEG in 2 or more but less than 10 regions). (G-I) Bulk expression in selected regions for (G) *Cbln2*, (H) *Gab2* and (I) *Ip6k2*. Black lines indicate averaged-smoothed gene expression. The trajectory with significant age effect is highlighted. Data are mean ± s.e.m.

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact

Requests for resources and reagents should be directed to the lead contact, Tony Wyss-Coray (twc@stanford.edu).

Data and Code Availability

The sequencing datasets analyzed during the current study are available in the Gene Expression Omnibus repository under accession numbers GSE212336, GSE212576, and GSE212903.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animal husbandry and organ collection

For Bulk-seq and nuc-seq, male and virgin female C57BL/6JN mice were shipped from the National Institute on Aging colony at Charles River. 5-6 male and 5 female mice were used for each 3, 12, 15, 18, and 21 months group, while only 5 and 3 male mice were used for the 26 and 28 months groups, respectively. For Spatial-seq, aged C57BL/6J mice (000664, Jackson Laboratory) were shipped from Jackson Laboratory. 2 male mice/age were used for the 6, 18, and 21 months groups. All mice were housed in cages of 2-3 mice at the Stanford ChEM-H animal facility under a 12 h/12 h light/dark cycle at 67–73 °F and provided with food and water ad libitum. Mice were housed in the ChEM-H animal facility for one month before euthanasia, except for mice older than 18 months, which were housed at the ChEM-H animal facility beginning at 18 months. Takedown of the bulk- and nuc-seq cohort was conducted between 10:00am-12:00pm over four days. Takedown of mice for Spatial-seq was conducted between 10:00am-10:15am on a single day. Age groups and sexes were rotated through over the duration of the takedowns to average out the impact of takedown time. After anaesthetization with 2.5% v/v Avertin, ~700ul blood was drawn via cardiac puncture before transcardial perfusion with 20 ml cold PBS. The brain was immediately removed and snap-frozen by submerging for 60 seconds in liquid nitrogen-cooled isopentane. Brains were stored at -80°C until further processing. All animal care and procedures complied with the Animal Welfare Act and were in accordance with institutional guidelines and approved by the institutional administrative panel of laboratory animal care at Stanford University.

METHOD DETAILS

Brain region dissection

Dissociating the mouse brain at scale poses several challenges, as the tissue consists of a multitude of biologically distinct structures that require careful, time-consuming separation to avoid cross-region contamination - all while avoiding tissue degradation and loss of RNA quality. We systematically assessed several isolation, dissection, and freezing strategies, most of which yielded low-quality RNA or were not scalable to the intended set of samples and regions. We found success in perfusing the animal before isolating and freezing the whole brain in under 5 minutes, thus rapidly stabilizing the tissue and RNA. Region isolation via slicing and atlas-guided tissue punching was subsequently conducted at sub-0C temperatures. In detail, brain regions were dissected from frozen mouse brains through a modification of a previously developed protocol (Wager-Miller et al., 2020). Frozen brains were sliced into 1mm thick coronal slices at -20°C using a metal brain matrix and .22mm razor blades (Ted Pella, 15045; VWR, 55411-050) and were then placed on dry ice and covered to prevent condensation. One slice at a time was placed on a metal block cooled on wet ice and 1.5mm and 2mm diameter regions of interest were dissected quickly via disposable biopsy punches (Alimed, 98PUN6-2, 98PUN6-3) from the left and right hemispheres guided by visual landmarks and the Allen Mouse Brain Atlas. The same biopsy punch was used for identical regions between left and right hemispheres, but replaced between regions and mice. 15 regions were collected: three cortical regions (motor cortex, visual cortex and entorhinal cortex), anterior (dorsal) and posterior (ventral) hippocampus, hypothalamus, thalamus, caudate putamen (part of the striatum), pons, medulla, cerebellum and the olfactory bulb, corpus callosum, choroid plexus and the subventricular zone. The following regions required overlapping punches and were thus sequentially collected: (1) motor cortex, (2) caudate putamen, (3) subventricular zone, (4) corpus callosum.

Regions were stored at -80°C until further processing.

Bulk-seq preparation and sequencing

We isolated RNA from the right hemisphere brain regions described above using the RNeasy 96 kit (Qiagen, 74181) and a TissueLyser II (Qiagen, 85300), according to RNeasy 96 Handbook protocol “Purification of Total RNA from Animal Tissues using Spin Technology” without the optional on-plate DNase digestion. Quality control of RNA was conducted using a Bioanalyzer

(Agilent) at the Stanford Protein and Nucleic Acid Facility for three randomly selected samples per brain region.

cDNA and library syntheses were performed in house using the Smart-seq2 protocol as previously described (De Miguel et al., 2021; Schaum et al., 2020) with the following modifications: Extracted RNA (2 ul at a concentration of 25 ng/ul) was reverse-transcribed and the resulting cDNA amplified using 10 cycles. After bead clean-up using 0.7x ratio with AMPure beads (Thermo Fisher, A63881), cDNA concentration was measured using the Quant-iT dsDNA HS kit (Thermo Fisher, Q33120) and normalized to 0.4 ng/ul as input for library prep. 0.8 ul of each normalized sample was mixed with 2.4 ul of tagmentation mix containing Tn5 Tagmentation enzyme (20034198, Illumina) and then incubated at 55°C for 12 minutes. The reaction was stopped by burying the plate in ice for 2 minutes followed by quenching with 0.8 ul 0.1% sodium dodecyl sulfate (Teknova, S0180). 1.6 ul indexing primer (IDT) was added and amplified using 12 cycles. Libraries were pooled and purified using two purification rounds with a ratio of 0.8x and 0.7x AMPure beads. Library quantity and quality was assessed using a Bioanalyzer (Agilent) and Qubit dsDNA HS kit. Pipetting steps were performed using the liquid-handling robots Dragonfly or Mosquito HV (SPT Labtech) using 384 well-plates and PCR reactions were carried out on a 384-plate Thermal Cycler (BioRad). Illumina sequencing of the resulting libraries was performed by Novogene (https://en.novogene.com/) on an Illumina NovaSeq S4 (Illumina). Base calling, demultiplexing, and generation of FastQ files were conducted by Novogene.

Spatial-seq preparation and sequencing

Frozen brains (n = 2 males per age; aged 6, 18 and 21 months; C57BL/6J strain) were embedded in OCT for cryosectioning at 16 micron thickness (app. Bregma -1.655mm; Allen brain reference atlas coronal section 71). Reactions were carried out with the Visum Spatial Gene Expression (GEX) and Tissue Optimization (TO) Slide & Reagent Kits according to the manufacturer's protocol with recommended reagents (10X Genomics, 1000193 and 1000184). Sections were placed on designated capture areas of slides for TO and GEX and stored at -80°C until further processing. TO and GEX slides were fixed with methanol and stained with hematoxylin and eosin (H&E) for visualization of tissue morphology on a AxioImager Widefield Fluorescence Microscope (Zeiss) at 10-fold magnification. To determine the optimal permeabilization time, TO slides were incubated with permeabilization enzyme for various timeframes followed by incubation with reverse transcriptase (RT) and fluorescently labeled nucleotides (FLNs) and enzymatic tissue removal. After visualizing cDNA signal via fluorescence microscopy, we selected 20 minutes as the optimal permeabilization time. GEX slides were incubated with permeabilization enzyme for 20 minutes followed by incubation with RT. cDNA was then transferred into tubes and amplified for 15 cycles using a Thermal Cycler (BioRad). Library construction steps were performed according to the manufacturer's protocol and included cDNA fragmentation, end repair and A-tailing, adaptor ligation, and sample indexing and amplification. Quality control of the constructed library was conducted via Bioanalyzer (Agilent). Illumina sequencing of the resulting libraries was performed by Novogene (https://en.novogene.com/) on an Illumina NovaSeq S4 (Illumina). Base calling, demultiplexing, and generation of FastQ files were conducted by Novogene.

Nuc-seq preparation and sequencing

Single-nuclei preparation (n = 2 males and females per age and region; aged 3 and 21 months; all C57BL/6JN strain) and sequencing was performed as previously described(Hahn et al., 2020) with the following modifications: Nuclei from left hemisphere brain region punches were isolated with EZ Prep lysis buffer (Sigma, NUC-101) on ice. Samples were placed into 2 ml cold EZ lysis buffer in a 2 ml glass dounce tissue grinder (Sigma, D8938) and homogenized by hand 25 times with pestle A followed by 25 times with pestle B while incorporating a 180-degree twist. Tissue homogenate was transferred to a fresh 15 ml tube on ice. The tissue grinder was rinsed with 2 ml fresh lysis buffer and transferred to the tube holding the homogenate for a total volume of 4 ml. Samples were incubated on ice for 5 minutes. Nuclei were centrifuged at 500 x g for 5 minutes at 4°C, supernatant removed and pellet resuspended with 4 ml EZ lysis buffer, and incubated on ice for 5 minutes. Centrifugation at 500 x g for 5 minutes at 4°C was repeated. After removing supernatant, the pellet was resuspended with 4 ml chilled PBS and filtered through a 35-um cell strainer into a 5 ml round bottom FACS tube (Corning, 352235). Following centrifugation at 300 x g for 10 minutes at 4°C with break 3, supernatant was gently poured out leaving behind the nuclei pellet. Pellet was resuspended in 400 ul PBS containing 1% BSA (Thermo Fisher, BP9700100), 0.2 ul Hoechst dye (Thermo Fisher, H3570), and 2 ul recombinant RNase inhibitor (Takara, 2313B). Isolated nuclei were sorted on a MA900 Multi-Application Cell Sorter (Sony Biotechnology). 25,000 single nuclei per sample were collected into 1.5 ml DNA lo-bind tubes (Eppendorf, 022431021) containing 1 ml buffer mix with PBS, UltraPure BSA (Thermo Fisher, AM2618), and RNase inhibitor (Takara, 2313B). One male and one female sample from the same time point and region were pooled at this stage by FACS collecting into the same sample tube (thus yielding 50,000 nuclei per tube). Collected nuclei were centrifuged at 400 x g for 5 minutes at 4°C with break 2. Supernatant was removed leaving 40 ul suspended nuclei. Nuclei were counted using a hemocytometer (Sigma, Z359629-1EA) and assessed for concentration and quality.

Reagents of the Chromium Single Cell 3' GEM & Gel Bead Kit v3.1 (10X Genomics, 1000121) were thawed and prepared according to the manufacturer's protocol. Nuclei and master mix solution was adjusted to target 10,000 nuclei per sample and loaded on a standard Chromium Controller (10X Genomics, 1000204) according to manufacturer protocols. We applied 11 PCR cycles to generate cDNA. Library construction was conducted using Chromium Single Cell 3' Library Construction Kit v3 (10X Genomics, 1000121). All reaction and quality control steps were carried out according to the manufacturer's protocol and with recommended reagents, consumables, and instruments. We chose 11 PCR cycles for library generation. Quality control of cDNA and libraries was conducted using a Bioanalyzer (Agilent) at the Stanford Protein and Nucleic Acid Facility. Illumina sequencing of the resulting libraries was performed by Novogene (https://en.novogene.com/) on an Illumina NovaSeq S4 (Illumina). Base calling, demultiplexing, and generation of FastQ files were conducted by Novogene.

QUANTIFICATION AND STATISTICAL ANALYSIS

Bulk-seq quantification, quality control

Raw sequence reads were trimmed to remove adaptor contamination and poor-quality reads using Trim Galore! (v.0.4.4, parameters: --paired --length 20 --phred33 --q 30). Trimmed sequences were aligned using STAR (v.2.5.3, default parameters). Multi-mapped reads were filtered. Read quality control and counting were performed using SeqMonk v.1.48.0 and RStudio v.3.6. Data visualization and analysis were performed using custom Rstudio scripts and the following Bioconductor packages: Deseq2 (Love et al., 2014), topGO, destiny and org.Mm.eg.db. Finally, we excluded pseudogenes and predicted genes from the count matrix to focus predominantly on well-annotated, protein-coding genes. In total, all of the following analyses were performed on the same set of 21,076 genes.

To assess the quality of our dataset, the count matrix was analyzed using Seurat's built-in, default dimensionality reduction workflow (Stuart et al., 2019) (Normalization: 'LogNormalize'; Variable feature discovery: selection.method='vst', features=2000). Umaps were calculated using Seurat's built-in functions, based on the first 40 principle components (PC) dimensions (Figure 1C; Figure S2B-D). A shared-nearest-neighbors graph was constructed using the first 40 PC dimensions before clustering samples using Seurat's built-in FindClusters function with a resolution of 0.8 to identify samples that would not cluster with their region of origin.

We corroborated the Seurat-based quality assessment by loading and normalizing the count matrix using DEseq2 before conducting the built-in variance stabilizing transformation (Love et al., 2014). We then performed hierarchical sample-to-sample clustering using Ward's clustering algorithm across all 21,076 genes (Figure S2E). To detect whether samples within a given tissue would show profound clustering by age, we finally calculated diffusion maps using the R package destiny with default parameters (Figure 1D).

For bar graph visualization of gene expression (e.g. Figure 1E), we used DEseq2-normalized counts after calculating factors and dispersion estimates across all regions using the factor design ~age + region. Trajectories were smoothed via triangular moving average across the interval between 3 and 28 months. This quantification and smoothing was solely used for visualization and was not the basis for any statistical testing in this study.

Bulk-seq differential expression

To identify significant differential expression changes with age, we used the raw count matrix as recommended for the DEseq2 standard analysis pipeline. Factors and dispersion estimates were calculated for each region separately. We conducted differential expression analysis comparing samples from 3 months to each consecutive time point, using sex as covariate. This is consistent with previously published differential expression analyses performed across whole organs in mice(Schaum et al., 2020). *P* values were adjusted for multiple testing, and genes with an adjusted *P* value of less than 0.05 were determined to be statistically significant. Finally, we required a gene to reach statistical significance (after multiple testing correction) in at least 2 pairwise comparisons (e.g. 3 months vs 18 months and 3 months vs 21 months) to be called a differentially expressed gene (DEG). We chose this criterion to retain only genes with robust differential expression patterns across age groups. We recognize that this tends to select against genes that are differentially expressed very late in life (i.e. 3 months vs 28 months).

DEG Gene Ontology functional enrichment

Unless stated otherwise, we performed functional enrichment analysis for DEGs using the Biocunductor package topGO. Unless stated otherwise, the set of expressed genes (defined as passing the independent filtering criterion of DEseq2 (Love et al., 2014)) was used as background for all functional enrichment analyses involving expression data. Top-ranked, representative Gene Ontology (GO) terms were selected and visualized using the CellPlot package. The full-length GO terms were shortened to fit into the figure format.

Bulk-seq GWAS gene enrichment

We analyzed if DEGs of a given region would be enriched for disease-associated genes using a previously assembled list of GWAS hits for several neurodegenerative diseases (Yang et al., 2022). The analysis was focused on Alzheimer's disease and Parkinson's disease (both age-related forms of dementia) and multiple sclerosis as we had observed several white matter-related effects in our dataset. We refer to these as 'disease-associated genes'. Disease-associated genes that were expressed in a given region (defined as passing the independent filtering criterion of DESeq2 (Love et al., 2014)) were analyzed. To determine if disease-associated genes were enriched among the DEGs of a given region, we used a one-sided hypergeometric test with expressed genes as background. Resulting *P* values were corrected for multiple testing. We chose the anterior hippocampus region as representative for the hippocampus, and further excluded the entorhinal cortex (too few DEGs) and olfactory bulb. For each disease, we plotted the enrichment and the relative composition of disease-associated DEGs with respect to their regulation (i.e., up- or down-regulated) using the CellPlot package. We clustered regions for using a pairwise Jacquard Distance matrix, so that regions with overlapping diseases-associated DEGs will cluster together. Gene overlaps with a Jaccard index ≥ 0.25 were indicated with an arc.

Bulk-seq correlation of gene expression with age

For each region separately, we probed the expression of each gene (using DEseq2-normalized counts) for positive or negative correlation with age using Spearman's method and tested for significant association. *P* values were adjusted for multiple testing using the Benjamini-Hochberg method. Genes with Spearman's rho ≥ 0.5 or ≤ -0.5, respectively, and padj ≤ 0.05 were called as significantly age-correlated in a given region. The total number of age-correlated genes was used to evaluate the impact of aging on a given region.

Weighted gene co-expression network analysis (WGCNA)

Network analysis was performed with the Weighted Gene Correlation Network Analysis (WGCNA) (Ferreira et al., 2021) package to identify significant modules that were associated with a specific aging group and brain region. Modules were independently detected in each brain region. For each brain region the soft-thresholding (ß value) was set based on scale-free topology (R2>0.8) to construct a correlation adjacency matrix. ß values 18, 10, 9, 8, 12, 4, 4, 5, 7, 9, 24, 14, 4 and 13 were used for the corpus callosum, cerebellum, motor cortex, entorhinal cortex, anterior hippocampus, posterior hippocampus, hypothalamus, medulla, olfactory bulb, choroid plexus, pons, SVZ, thalamus and visual cortex respectively. The 'blockwiseModules' function was used to construct the network. Biweight midcorrelation ('bicor') was used to compute the correlation between each pair of genes. Network analysis was performed with the "signed" network. The "deepSplit" argument value was 2 and a minimum cluster size was 25. (blockwiseModules parameters: datExpr=(datExpr), maxBlockSize=22000, networkType="signed", corType="bicor", power=ß, saveTOMFileBase=(file='TOM_signed'), minModuleSize=25, deepSplit=2, saveTOMs=TRUE).

The average linkage hierarchical clustering of the topological overlap dissimilarity matrix (1-TOM) was used to generate the network dendrogram. The hybrid dynamic tree-cutting method was used to define modules. Modules were summarized by their first principal component (ME, module eigengene) and modules with eigengene correlations >0.9 were merged.

Module-aging group associations were evaluated using a linear model within each brain region. Significance values were corrected for multiple testing using Benjamini-Hochberg method. Results from module-eigengene association tests are shown in Table S4. Genes within each module were prioritized based on their module membership (kME), defined as correlation to the module eigengene. The top 'hub' genes for several of the modules are shown in supplementary Table S4. Cell type enrichment analyses were performed using several mouse derived cell type specific expression datasets (Goldmann et al., 2016; Zeisel et al., 2015; Zhang et al., 2016). Enrichment was performed for cell type specific marker genes using Fisher's exact test, followed by Benjamini-Hochberg-correction for multiple testing. The WGCNA results were assembled in summarizing figures that can browsed through our interactive shiny app website (https://twc-stanford.shinyapps.io/spatiotemporal_brain_map/).

Estimating the variance of the data depending on metadata

To estimate the variance in the data depending on age, tissue or gender we made use of principal variance component analysis (PVCA) as implemented in the Bioconductor Package pvca. PVCA combines the strength of principal component analysis and variance components analysis (VCA). Originally it was applied to quantify batch effects in microarray data. In our case, however, we do not provide experimental batches but rather groups of meta data as input

Gene signature generation and score calculation

Gene *signatures* are used in this study to quantify the expression of a gene set, thus representing the aggregated expression of multiple genes in a given transcriptome (e.g. a regional bulk-seq transcriptome or a single-nuclei transcriptome). The resulting value is defined as a *score*. Throughout the manuscript we generated signatures and quantified scores using the VISION (v.3.0) package as detailed in the original study (DeTomaso et al., 2019). Notably, VISION z-normalizes signature scores with random gene signatures to account for global sample-level metrics (such as total number of counts/UMIs, which can be affected by age (Pálovics et al., 2022)). While VISION was originally intended for the analysis of signatures in single-cell data we found its analysis workflow applicable for bulk, spatial and single-cell/-nuclei datasets. We note that due to differences in baseline expression across regions or cell types as well as the z-normalization mentioned above, VISION scores can be negative. However, our analyses are focused - unless stated otherwise - on the relative score changes (i.e. increase or decrease relative to 3 months) occurring with age in a given region or cell type.

Bulk-seq marker genes and score calculation

Seurat's FindAllMarkers function was run using the 'DESeq2' test with parameters and Bonferroni correction for multiple testing to identify region-specific marker genes (*P* value of less than 0.05; Figure S3). For each region, we constructed unsigned signatures (DeTomaso et al., 2019) based on a given region's significant marker genes. For each signature, we calculated scores across a publicly available spatial transcriptome dataset from 10X Genomics (https://www.10xgenomics.com/resources) and compared the patterns to structural annotations in Allen Mouse Brain Atlas.

Bulk-seq Common Aging Score (CAS) calculation and CAS velocity comparison

We ranked genes on the basis of their regulation across regions, to summarize in how many regions a given gene would be called as a DEG (i.e. reach statistical significance in at least two comparisons between samples from 3 months and any following age group). We included only the anterior hippocampus region in the selection of cross-region DEGs to prevent a potential bias towards aging effects in the hippocampus. This led to the identification of 82 genes that were marked as DEG in at least 10 out of 14 regions (15 regions minus the posterior hippocampus region). We constructed a signed gene signature(DeTomaso et al., 2019) based on 75 up- and 7 down-regulated DEGs. We used the signature to calculate CAS for each single-region transcriptome. To quantify a region's score increase over time (aging velocity), we constructed a linear model with the design: score ~ age + region + age:region (score explained by a two factor model including interaction term) using the linear model function in R. We used the lstrends function of the lsmeans package(Lenth, 2016) that utilizes least-square means to estimate and compare the slopes of fitted lines for each region. We subsequently used Tukey's range test across all possible region-to-region comparisons to assess which regions exhibited statistically significant (*P* value < 0.05) slope differences. In addition, we repeated the analysis resolved for sex-specific effects across the 3, 12, 15, 18 and 21 months groups (for which we had both male and female samples). We assessed if there was a differential aging velocity between sexes across all regions (Figure 2I), for which analyzed a linear model with the design: score ~ age + sex + age:sex + region. We further performed the same analysis iteratively for each region individually (Figure S4D) using a model with the design: score ~ age + sex + age:sex. We corrected the resulting *P* values for each region-wise analysis using the Benjamini-Hochberg method.

Organ-specific aging signature identification and velocity comparison

To explore the feasibility of detecting gene expression patterns with organ-specific regulation during aging, we re-analyzed a previously published bulk RNA-seq dataset of 17 mouse tissues profiled across ten age groups (n = 4 males; aged 1, 3, 6, 9, 12, 15, 18, 21, 24 and 27 months; n = 2 females; aged 1, 3, 6, 9, 12, 15, 18 and 21 months)(Schaum et al., 2020). The dataset comprised the following organs: bone, brain, brown adipose tissue (BAT), gonadal adipose tissue (GAT), heart, kidney, limb muscle (muscle), liver, lung, bone marrow (marrow), mesenteric adipose tissue (MAT), pancreas, skin, small intestine (intestine), spleen, subcutaneous adipose tissue (SCAT), and white blood cells (WBC). We obtained pre-processed data as described in the original study and performed differential expression analysis accordingly(Schaum et al., 2020). We identified age-related DEGs in the same manner as described for the bulk-seq data: we used the raw count matrix as recommended for the DEseq2 standard analysis pipeline. Factors and dispersion estimates were calculated for each tissue separately. We conducted differential expression analysis comparing samples from 3-months-old mice to each consecutive time point, using age and sex as covariates. *P* values were adjusted for multiple testing, and genes with an adjusted *P* value of less than 0.05 were determined to be statistically significant. Finally, we required a gene to reach statistical significance (after multiple testing correction) in at least 2 pairwise comparisons (e.g. 3-months-old vs 12 months-old and 3-months-old vs 21 months-old) to be called a differentially expressed gene (DEG). We analyzed age groups that would be comparable to the age groups profiled in our study (3, 12, 15, 18, 21, 24 and 27 months). We ranked genes on the basis of their regulation across organs, to summarize in how many organs a given gene would be called as a DEG (i.e. reach statistical significance in at least two comparisons between samples from 3-months-old mice and any following age group). DEGs that were only detected in a single organ were assembled into signed, organ-specific aging signatures using VISION (DeTomaso et al., 2019), comparable to the CAS. For organs that exhibited fewer than 25 unique DEGs we did not construct a signature. For each organ-specific signature, we performed the following analysis: We first tested for each organ separately, if the respective signature would show a significant correlation with age using linear models with the design: score ~ age. Organs that showed no significant (*P* val < 0.05, *t*-test) association with the age were excluded. Next, we constructed a linear model with the design: score ~ age + organ + age:organ (organ-specific score explained by a two factor model including interaction term) using the linear model function in R. We used the lstrends function of the lsmeans package(Lenth, 2016) that utilizes least-square means to estimate and compare the slopes of fitted lines for each organ. We subsequently used Tukey's range test across all possible organ-to-organ comparisons to assess which organs exhibited statistically significant (*P* value < 0.05) slope differences. Notably, we asked if the organ where the signature was identified (the 'reference' organ) would show a significantly higher slope compared to all other organs. The summarized results are displayed in the heatmap in Figure S9.

Bulk-seq region-specific aging signature identification and velocity comparison

We ranked genes on the basis of their regulation across regions, to summarize in how many regions a given gene would be called as a DEG (i.e. reach statistical significance in at least two comparisons between samples from 3-months-old mice and any following age group). DEGs that were only detected in a single region were assembled into signed, region-specific aging signatures using VISION(DeTomaso et al., 2019), comparable to the CAS. We excluded the posterior hippocampus region in the selection of region-specific DEGs. Further, there were less than 20 unique DEGs found for the entorhinal cortex, which we considered too small to construct a signature. For each region-specific signature, we performed the following analysis: We first tested for each region separately, if the respective signature would show a significant correlation with age using linear models with the design: score ~ age. Regions that showed no significant (P val < 0.05, *t*-test) association with age were excluded. Next, we constructed a linear model with the design: score ~ age + region + age:region (region-specific score explained by a two factor model including interaction term) using the linear model function in R. We used the lstrends function of the lsmeans package(Lenth, 2016) that utilizes least-square means to estimate and compare the slopes of fitted lines for each region. We subsequently used Tukey's range test across all possible region-to-region comparisons to assess which regions exhibited statistically significant (P value < 0.05) slope differences. Notably, we asked if the region where the signature was identified (the 'reference region') would show a significantly higher slope compared to all other regions. The summarized results are displayed in the heatmap in Figure S10.

Spatial-seq mapping, embedding, clustering and region identification

Space Ranger analysis pipelines were utilized to align image and FASTQ files, detect tissue and fiducial frames, count barcodes/UMIs. Throughout the manuscript we refer to the barcoded areas from a given dataset as 'spots'. Spots with less than 5 UMIs were removed as well as all spots at the outline of the tissue as these can be affected by RNA diffusion. We integrated all six sample-wise datasets (two from 6 months, two from 18 months and two from 21 months), using Seurat's built-in SCTransform and integration workflow(Stuart et al., 2019), with 2000 genes set as integration features. Integrated datasets were then used as input for spot embedding and clustering. A shared-nearest-neighbors graph was constructed using the first 30 PC dimensions before clustering spots using Seurat's built-in FindClusters function with a resolution of 0.8 and default parameters. Umaps and tSNEs were calculated using Seurat's built-in functions, based on the first 30 PC dimensions. Count data was subsequently normalized and scaled using SCTransform across all spots to allow for visualization of expression values and differential gene expression analysis.

We chose a data-driven approach to group spots and map them to anatomical structures of the brain (Figure S5A): Transcriptional clustering yielded 29 clusters and we used Seurat's FindAllMarkers function (parameters: min.pct=0.1, thresh.use=0.1 assay='SCT', only.pos=TRUE) to identify cluster markers. We compared the expression of marker genes to in-situ hybridization (ISH) data from the Allen Mouse Brain Atlas(Lein et al., 2007) and visual landmarks from the H&E microscopy images (e.g. *cornu Ammonis* and dentate gyrus of the hippocampus; cell-sparse structure of the white matter fiber tracts). To enable comparisons with the regions isolated for Bulk-seq, we additionally grouped annotated clusters into meaningful region-level sets, guided by their anatomical location and the hierarchical ordering of structures in the Allen Mouse Brain Atlas. Ontology and nomenclature of clusters is indicated in Fig S5B.

Spatial-seq differential expression analysis and comparison with Bulk-seq data

Given comparable representation of clusters across all samples and age groups, we considered differential expression analysis across age groups feasible. We analyzed differential expression in the white matter and cortex cluster as we considered them comparable to the corpus callosum (high CAS velocity) and motor cortex (low CAS velocity) region from the bulk-seq dataset. Differential gene expression of genes comparing Spatial-seq data from 6 months to 21 months was done using the 'DESeq2' algorithm implemented in Seurat on Spatial count data. Seurat natural log(fold change) > 0.2 (absolute value), adjusted *P* value (Benjamini-Hochberg correction) < 0.05, and expression in greater than 10% of spots in both comparison groups were required to consider a gene differentially expressed. To test for a potential association between gene-expression changes measured in Spatial-seq and bulk-seq data, we considered only genes that changed significantly in both datasets. For both regions, we confirmed significant overlap between the DEGs found in Bulk- and Spatial-seq dataset (Fisher's exact test, *P* Value < 0.05). Next, we plotted $\log_2$ fold expression changes during aging as measured via Spatial-seq versus $\log_2$ fold expression changes on the bulk-seq level. The distribution of genes among the four resulting quadrants was tested for directionality using Fisher's exact test.

Spatial-seq CAS calculation

Calculation of CAS and CAS velocities for Spatial-seq data was carried out in a similar manner as described above for bulk-seq data: CAS for each Spatial-seq spot were calculated and score increase over time was calculated for region clusters with equivalent Bulk-seq regions: White matter (compared to corpus callosum), cortex (compared to motor cortex), striatum (compared to caudate putamen), hippocampus, hypothalamus, choroid plexus and thalamus. To quantify a Spatial-seq region's score increase over time (aging velocity), we constructed a linear model with the design: score ~ age + region + age:region and carried out slope estimation and differential analysis as described above. We acknowledge that this analysis does not account for biological replicates but treats each spot belonging to the same region as replicate. We therefore visualized CAS in Spatial-seq for each replicate, to demonstrate that age-related changes in CAS supersede the intra-replicate CAS differences.

Nuc-seq mapping, embedding, clustering, sample demultiplexing and cell type identification

Cell Ranger (v.6.1.2) analysis pipelines were utilized to align reads to mm10 reference genome and count barcodes/UMIs. To account for unspliced nuclear transcripts, reads mapping to pre-mRNA were counted. Throughout the manuscript we use nuclei and 'cells' synonymously. Outliers with a high ratio of mitochondrial (more than 5%, fewer than 400 features) relative to endogenous RNAs and homotypic doublets (more than 6,000 features in hippocampus; more than 7,000 features in caudate putamen) were removed in Seurat(Stuart et al., 2019). We integrated all sample-wise datasets, using Seurat's built-in SCTransform and integration workflow(Stuart et al., 2019), with 500 genes set as integration features. Integrated datasets were then used as input for cell embedding and clustering. A shared-nearest-neighbors graph was constructed using the first 12 PC dimensions before clustering spots using Seurat's built-in FindClusters function with a resolution of 0.4 and default parameters. Umaps and tSNEs were calculated using Seurat's built-in functions, based on the first 12 PC dimensions. A given nuc-seq sample represented nuclei from a male and female animal (of the same age) that were pooled in equal numbers during the nuclei isolation steps. To demultiplex a sample by sex, we calculated the ratio of counts belonging to female- (*Xist*, *Tsix*) and male-specific (*Ddx3y*, *Eif2s3y*, *Uty*, *Kdm5d*) genes, and identified nuclei with a $\log_2$ cutoff of 1 and -1 as female- and male-derived nuclei, respectively. Ambiguous nuclei, which had reads from both female and male nuclei, were removed from the analysis. Count data was subsequently normalized and scaled to allow for visualization of expression values and differential gene expression analysis. Seurat's FindAllMarkers function (parameters: min.pct=0.15, thresh.use=0.15 assay='SCT') was run to identify cluster markers. Clusters were annotated based on marker genes. Finally, nuclei were manually inspected using known cell type-specific marker genes and nuclei expressing more than one cell type-specific marker were defined as doublets and removed(Mathys et al., 2019).

Publicly available scRNA-seq data embedding

We re-analyzed two previously published single-cell RNA-seq datasets: (1) Droplet-based scRNA-seq of freshly dissected SVZ at young and old age (n = 3 males per age; aged 3 and 28 months; all C57BL/6JN strain)(Dulken et al., 2019); (2) Smart-seq2-based (Picelli et al., 2013) scRNA-seq of freshly isolated cells from the myeloid and non-myeloid fraction of the striatum, cerebellum, hippocampus and cortex at young and old age (n = 4 males per age; aged 3 and 24 months; all C57BL/6JN strain) (Tabula Muris Consortium, 2020). For the SVZ data, we obtained and analyzed pre-processed count matrices. For visualization purposes, we integrated all sample-wise datasets, using Seurat's built-in SCTransform and integration workflow(Stuart et al., 2019), with 2000 genes set as integration features. Integrated datasets were then used as input for cell embedding. Umaps and tSNEs were calculated using Seurat's built-in functions, based on the first 12 PC dimensions. Cell annotations were transferred from the original study. Count data was subsequently normalized and scaled to allow for visualization of expression values.

For the second dataset, we obtained and analyzed pre-processed count matrices. We followed previous analyses on the same dataset(Iram et al., 2022). For visualization purposes, we integrated all sample-wise datasets, using Seurat's built-in SCTransform workflow(Stuart et al., 2019). Integrated datasets were then used as input for cell embedding. Umaps and tSNEs were calculated using Seurat's built-in functions, based on the first 12 PC dimensions. Cell annotations were transferred from the original study. Count data was subsequently normalized and scaled to allow for visualization of expression values.

Differential expression in scRNA- and Nuc-seq data

Differential gene expression of genes comparing young and old samples was done using the MAST(Finak et al., 2015) algorithm, which implements a two-part hurdle model. Seurat natural log(fold change) > 0.2 (absolute value), adjusted *P* value (Benjamini-Hochberg correction) < 0.05, and expression in greater than 10% of cells in both comparison young and old samples.

Signature calculations in scRNA- and Nuc-seq data

Calculation of CAS and CAS velocities for scRNA- and Nuc-seq data was carried out in a similar manner as described above for Spatial-seq: CAS for each cell were calculated and CAS increase over time was calculated for cell types. To quantify a cell type's score increase over time (aging velocity), we constructed a linear model with the design: score ~ age + cell type + age:cell type and carried out slope estimation and differential analysis as described above. We acknowledge that this analysis does not account for biological replicates but treats each cell belonging to the same cell type as replicate. To account for this, we further calculated the cell type-median CAS for each biological replicate and tested for differential CAS regulation with two-tailed t-test on per-replicate median of CAS. For the comparison of CAS increase with age across microglia from different brain regions (Figure 4A-C), we used two-sided Wilcoxon rank-sum tests to test for CAS differences between microglia from the same age group.

For region-specific signatures, we performed a per-cell type slope quantifications as detailed for the CAS and clustered the resulting slope estimates using hierarchical clustering (Figure 5H,K).

Single-nuclei dispersion score

We employed a previously published strategy to quantify if a given DEG detected in bulk data would be expressed in a specific cell type(Schaum et al., 2020). For each gene in each brain region, we selected cells expressing the gene (log-CPM expression > 0). Next, we assigned to the cells the log-CPM expression values of the gene as weights. Based on these, we calculated the weighted center of the cells in the single-cell landscape defined by the UMAP embeddings. We defined the 'single-cell dispersion' of the gene as the weighted mean distance of the cells from their weighted center. Finally, we introduced region specific factors to account for differences between brain region specific embeddings. Per region, we set pseudo log-CPM count 1 to all cells and calculated the dispersion of them. We normalized the dispersion scores by these region-specific factors.

Supplementary Figures

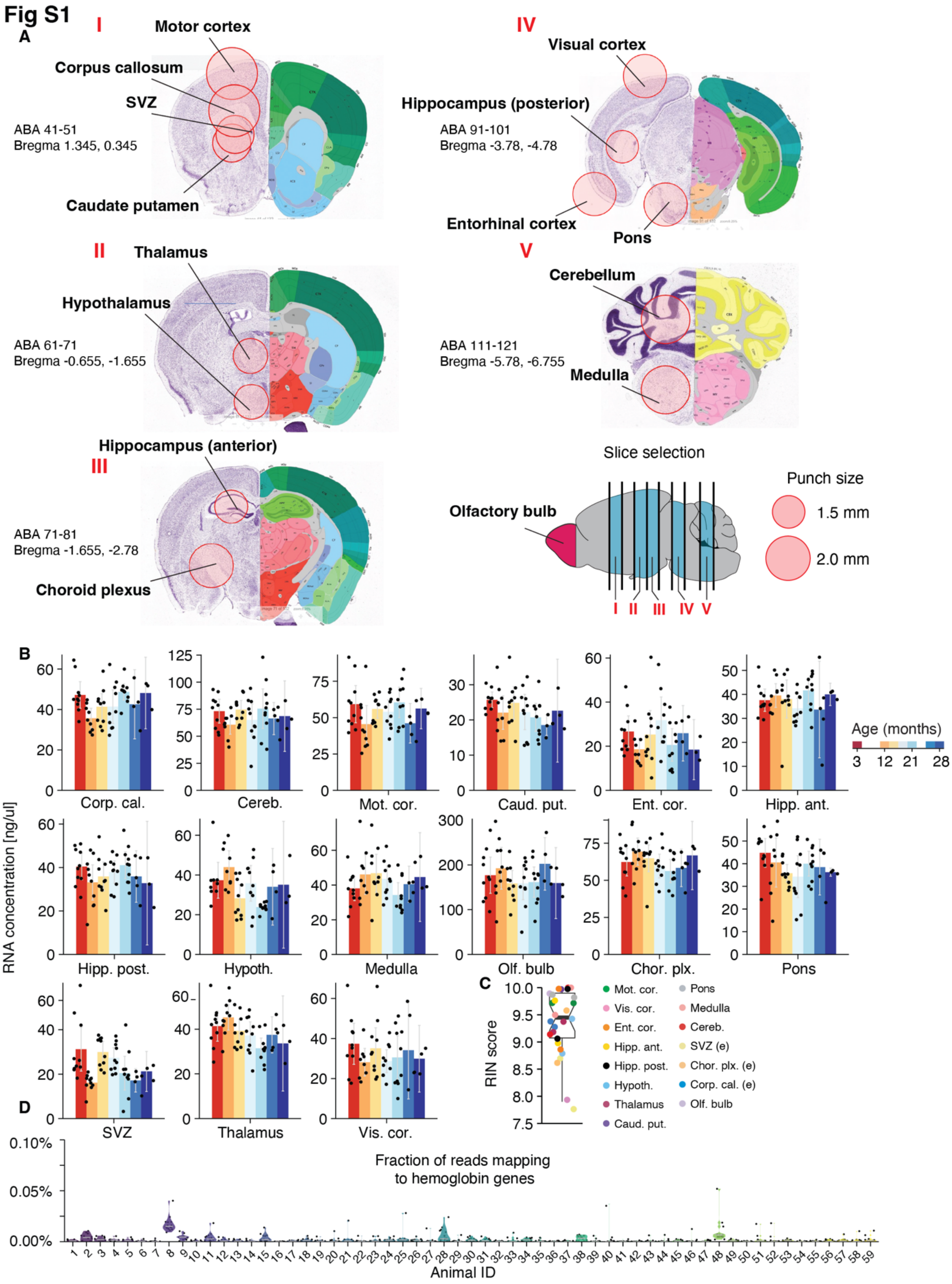

Figure S1 Capturing transcriptional heterogeneity of the brain with high RNA integrity

(A) Location of tissue punches overlaid on the Allen Institute Mouse Coronal Reference Atlas. Equivalent regions were taken from both the left and right hemisphere. Depth of each brain slice is indicated as: ABA, Allen brain atlas coronal layer number range, and Bregma, stereotaxic coordinates of mm. from bregma. The following regions required overlapping punches and were thus sequentially collected: (1) motor cortex, (2) caudate putamen, (3) subventricular zone, (4) corpus callosum. (B) RNA concentration measured with Quant-iT broad range RNA Assay Kit in all regions, colored by age. Data are mean ± s.e.m. (C) RNA integrity number, determined by Bioanalyzer eukaryote total RNA, for randomly-selected samples, colored by region. Each region is represented by at least 3 samples. RINs above 7 indicate good quality RNA. (D) Violin plot of fraction of reads mapping to hemoglobin genes for each sample, grouped by animal.

Fig S2

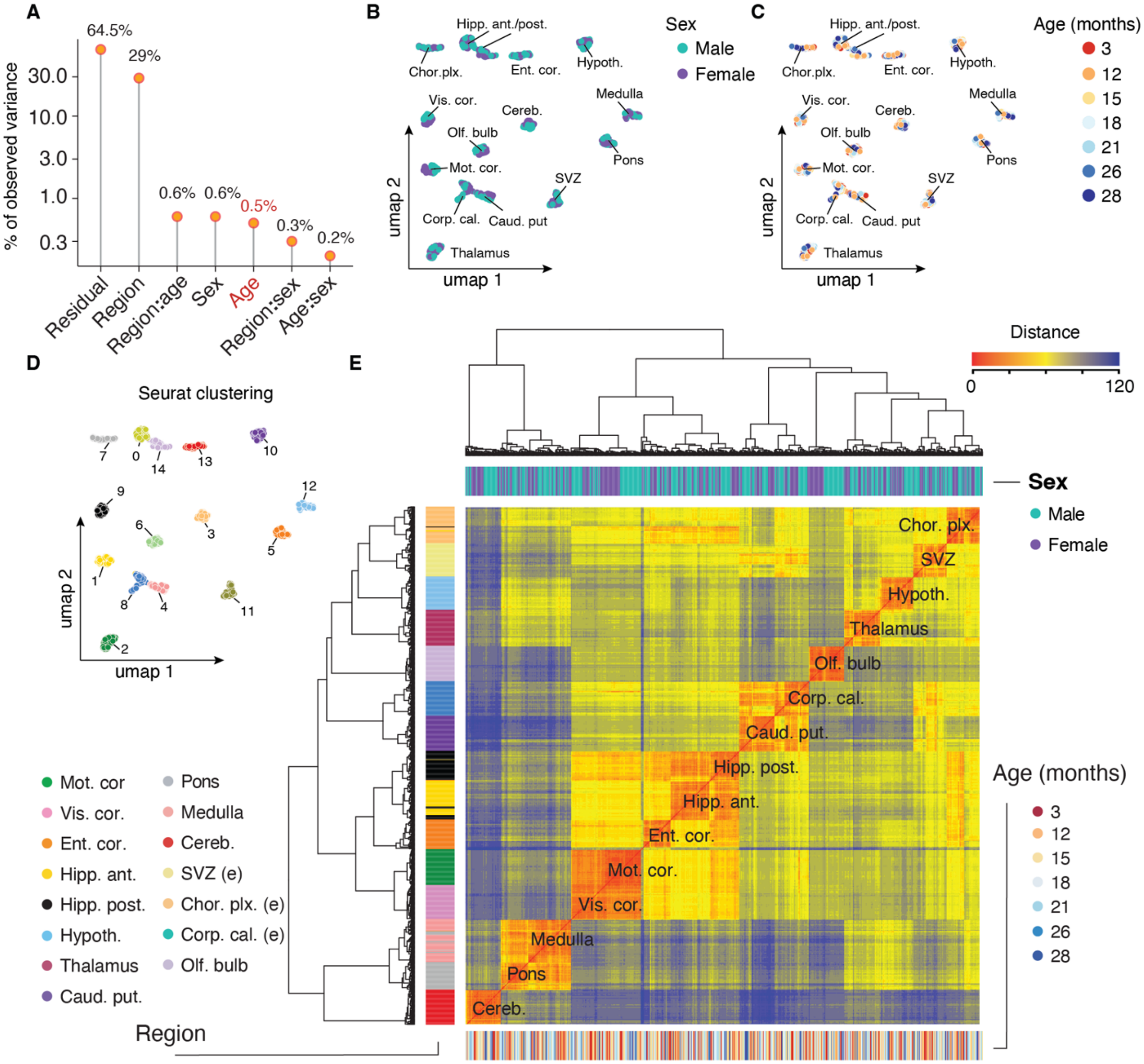

Figure S2 Gene expression variance analysis across regions

(A) Visualization of the principal variance component analysis, displaying the gene expression variance explained by residuals (that is, biological and technical noise) or experimental factors such as region, age, sex and respective interactions. Highlighted is variance explained by age. (B,C,D) UMAP representation of bulk RNA-seq samples (n = 847 total samples) based on the first 40 principal components colored according to (B) sex, (C) age, and (D) Seurat clustering with default settings. (E) Hierarchical clustering of all samples using Ward's algorithm. Samples are annotated by region, sex and age.

Fig S3

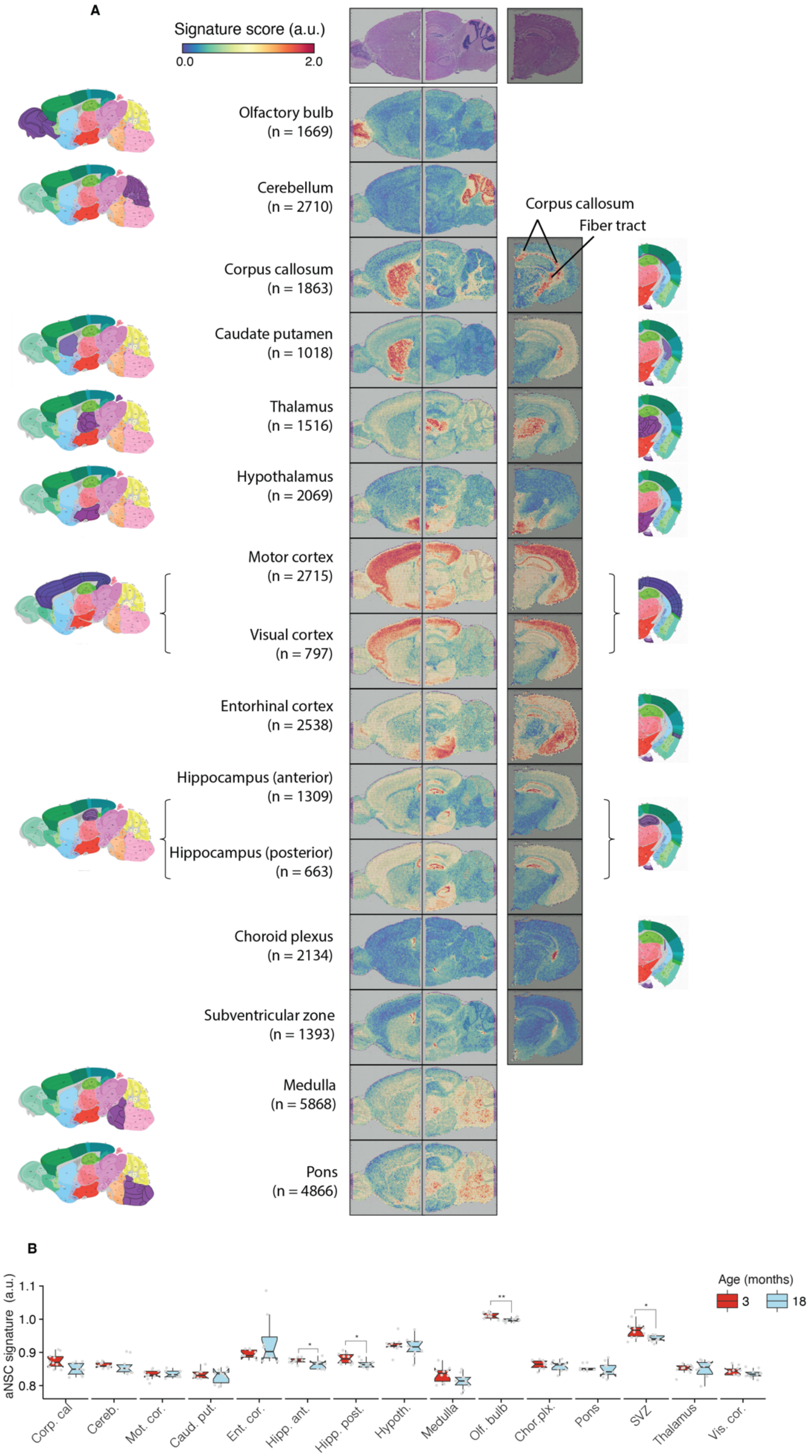
A
Sagittal$_{anterior}$
Sagittal$_{posterior}$
Coronal
Signature score (a.u.)
0.0
2.0
Olfactory bulb (n = 1669)
Cerebellum (n = 2710)
Corpus callosum
Fiber tract
Corpus callosum (n = 1863)
Caudate putamen (n = 1018)
Thalamus (n = 1516)
Hypothalamus (n = 2069)
Motor cortex (n = 2715)
Visual cortex (n = 797)
Entorhinal cortex (n = 2538)
Hippocampus (anterior) (n = 1309)
Hippocampus (posterior) (n = 663)
Choroid plexus (n = 2134)
Subventricular zone (n = 1393)
Medulla (n = 5868)
Pons (n = 4866)
B
aNSC signature (a.u.)
1.1
1.0
0.9
0.8
Age (months)
3
18
*
*
**
*
Corp. cal
Cereb.
Mot. cor.
Caud. put.
Ent. cor.
Hipp. ant.
Hipp. post.
Hypoth.
Medulla
Olf. bulb
Chor.plx.
Pons
SVZ
Thalamus
Vis. cor.

Figure S3 Collected brain regions reflect transcriptomes of anatomically distinct structures (A) Spatial expression of gene signatures derived from marker genes enriched in regions from Fig S1. For each region, marker gene analysis was performed and enriched genes were grouped into a marker signature. Each signature was quantified in publicly available Visium spatial transcriptomics data from a young mouse brain. Dataset contained anterior and posterior sagittal sections, as well as a coronal section. Number of marker genes that make up the respective signature is indicated in parentheses and marker genes are listed in Table S1. Quantification of gene signatures in spatial transcriptome yields scores for each spatial spot. Scores for each signature were compared to the Allen Institute Mouse Sagittal and Coronal Reference Atlases where possible. Positions of regions that we aimed to capture are highlighted in purple. (B) Expression of activated neural stem cell (aNSC)-derived marker gene signature in young (3 months) and aged (18 months) samples. Signature exhibits the highest baseline expression in the olfactory bulb and subventricular zone (SVZ) derived samples. A significant drop in the signature score is notable in olfactory bulb, SVZ and both anterior and posterior hippocampus regions. *P* values calculated with two-tailed t-test on per-replicate median of signature score, adjusted for multiple testing. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Fig S4

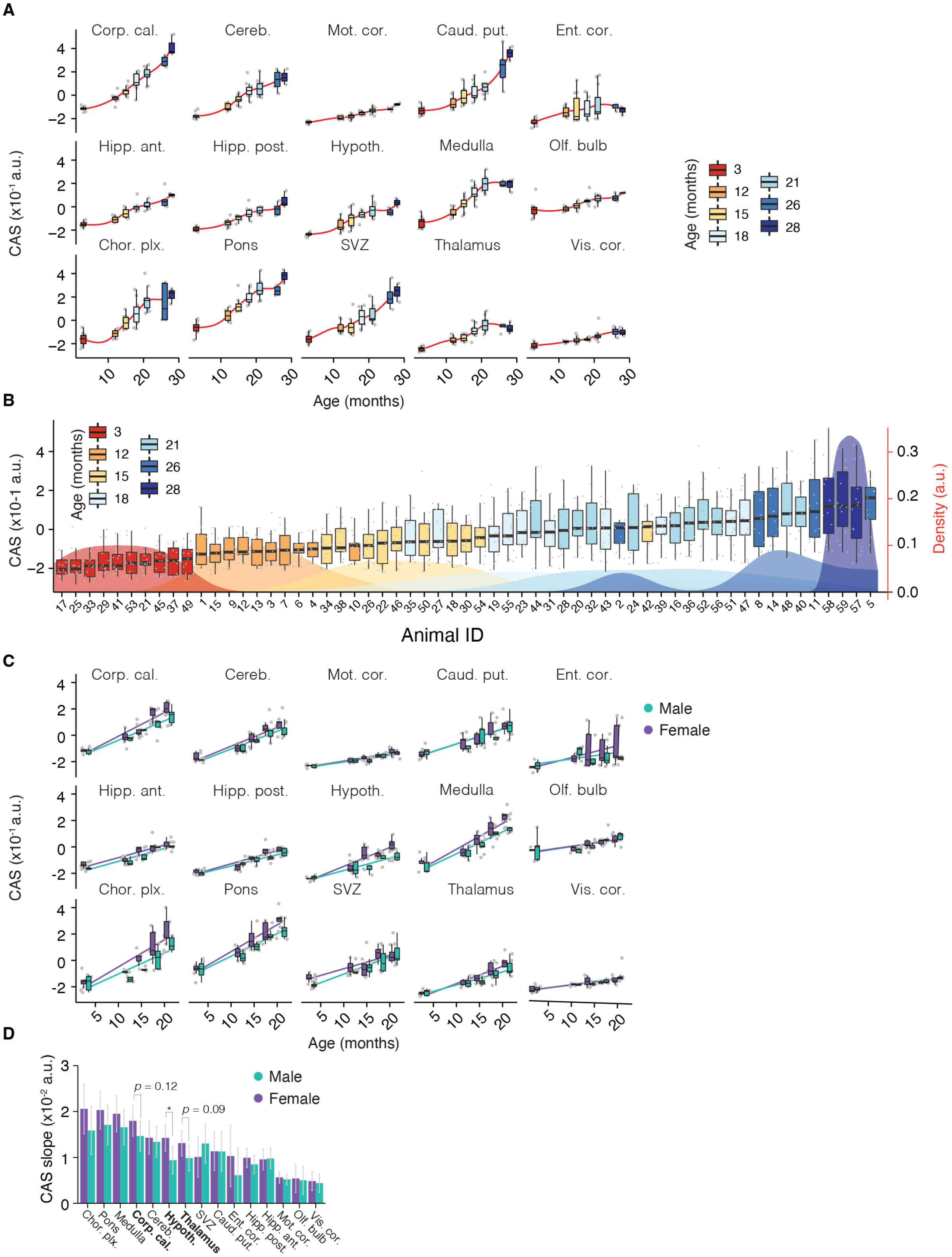

Figure S4 A common RNA aging signature quantifies the region-specific pace and magnitude of transcriptional shifts in the brain

(A) CAS trajectories of all regions, colored by age. Red lines indicate averaged-smoothed values. (B) CAS for each sample grouped by animal. Each boxplot represents distribution of CAS across the 15 regions as measured by bulk RNA-seq. Animals are ranked by median score and colored by chronological age. Distribution of chronological age groups is indicated in the background. (C) CAS trajectories of all regions for the interval of 3 to 21 months, colored by sex. (D) Slope of linear regressions in (C), colored by sex. Data are mean ± 95% confidence intervals. Pair-wise, two-sided Tukey's HSD test, adjusted for multiple testing, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. The highest (least significant) Pval is indicated.

Fig S5

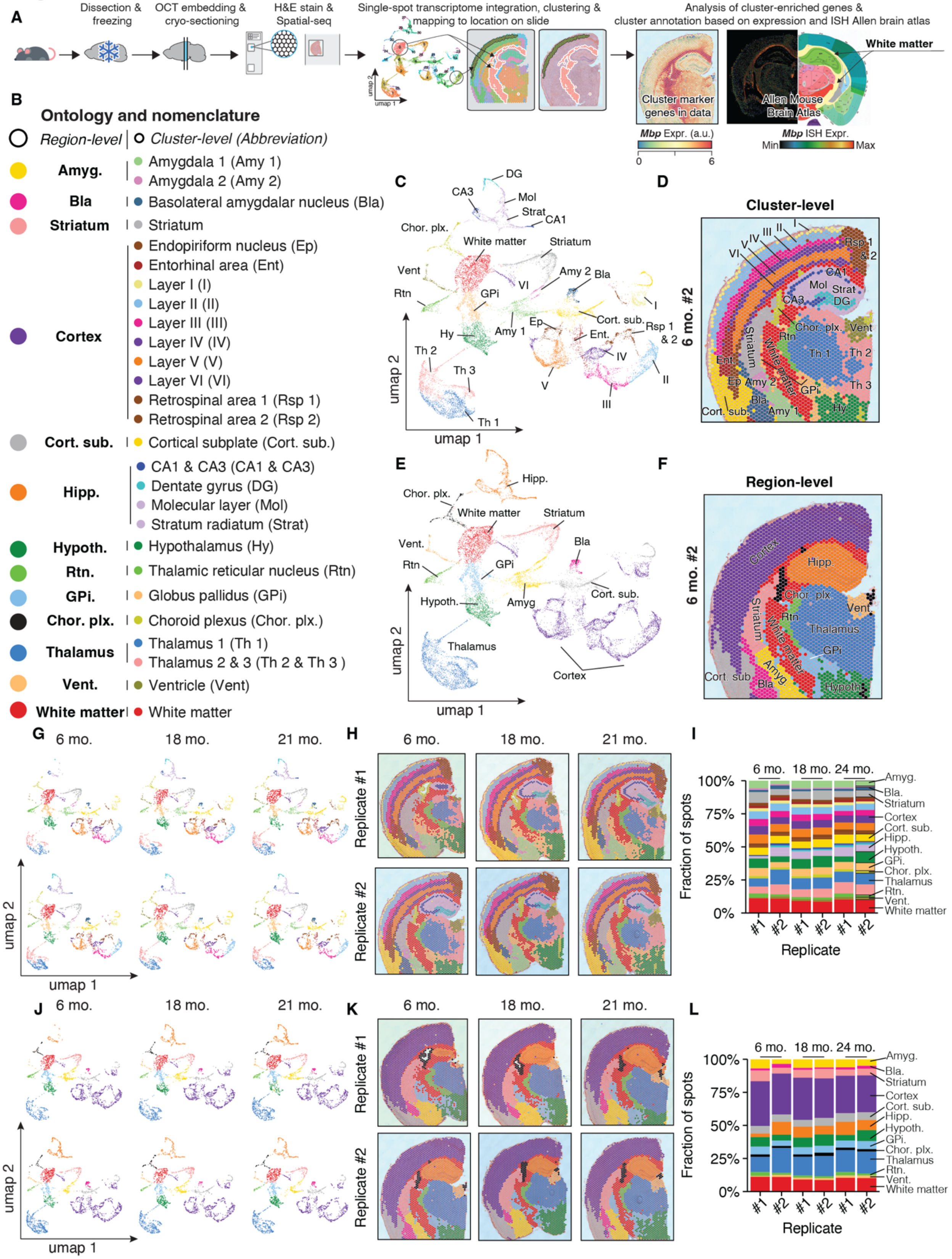
A
Dissection & freezing
OCT embedding & cryo-sectioning
H&E stain & Spatial-seq
Single-spot transcriptome integration, clustering & mapping to location on slide
umap 1
umap 2
Analysis of cluster-enriched genes & cluster annotation based on expression and ISH Allen brain atlas
Cluster marker genes in data
Mbp Expr. (a.u.)
0
6
Allen Mouse Brain Atlas
Mbp ISH Expr.
Min
Max
White matter
B
Ontology and nomenclature
Region-level
Cluster-level (Abbreviation)
Amyg.
Amygdala 1 (Amy 1)
Amygdala 2 (Amy 2)
Bla
Basolateral amygdalar nucleus (Bla)
Striatum
Striatum
Cortex
Endopiriform nucleus (Ep)
Entorhinal area (Ent)
Layer I (I)
Layer II (II)
Layer III (III)
Layer IV (IV)
Layer V (V)
Layer VI (VI)
Retrospinal area 1 (Rsp 1)
Retrospinal area 2 (Rsp 2)
Cort. sub.
Cortical subplate (Cort. sub.)
Hipp.
CA1 & CA3 (CA1 & CA3)
Dentate gyrus (DG)
Molecular layer (Mol)
Stratum radiatum (Strat)
Hypoth.
Hypothalamus (Hy)
Rtn.
Thalamic reticular nucleus (Rtn)
GPi.
Globus pallidus (GPi)
Chor. plx.
Choroid plexus (Chor. plx.)
Thalamus
Thalamus 1 (Th 1)
Thalamus 2 & 3 (Th 2 & Th 3)
Vent.
Ventricle (Vent)
White matter
White matter
C
DG
CA3
Mol
Strat
CA1
Chor. plx.
White matter
Striatum
Vent
Amy 2
Bla
Rtn
GPi
VI
I
Ep
Cort. sub.
Rsp 1 & 2
Hy
Amy 1
Ent.
IV
Th 2
Th 3
V
II
III
Th 1
umap 1
umap 2
D
Cluster-level
6 mo. #2
E
Hipp.
Chor. plx.
White matter
Striatum
Vent.
Rtn
GPi
Bla
Hypoth.
Amyg
Cort. sub.
Thalamus
Cortex
umap 1
umap 2
F
Region-level
6 mo. #2
G
6 mo.
18 mo.
21 mo.
umap 1
umap 2
H
6 mo.
18 mo.
21 mo.
Replicate #1
Replicate #2
I
6 mo.
18 mo.
24 mo.
100%
75%
50%
25%
0%
Fraction of spots
#1
#2
Replicate
Amyg.
Bla.
Striatum
Cortex
Cort. sub.
Hipp.
Hypoth.
GPi.
Chor. plx.
Thalamus
Rtn.
Vent.
White matter
J
6 mo.
18 mo.
21 mo.
umap 1
umap 2
K
6 mo.
18 mo.
21 mo.
Replicate #1
Replicate #2
L
6 mo.
18 mo.
24 mo.
100%
75%
50%
25%
0%
Fraction of spots
#1
#2
Replicate
Amyg.
Bla.
Striatum
Cortex
Cort. sub.
Hipp.
Hypoth.
GPi.
Chor. plx.
Thalamus
Rtn.
Vent.
White matter

Figure S5 Robust capture of spatial transcriptomes across age
(A) Spatial-seq processing and analysis overview. Whole brains were frozen prior to OCT embedding and cryo-sectioning. Coronal sections were placed on a 10X Visium Spatial Gene Expression slide, followed by H&E staining and spatial reverse-transcription reaction. Single-spot transcriptomes were integrated, clustered with default settings and visualized as UMAP. Clustered spatial spot transcriptomes were mapped to their original location. To annotate the clusters, their marker genes (Table S6) were visualized, compared to the Allen Brain Atlas (*23*). (B) Complete data description and abbreviations of ontology and nomenclature for spatial transcriptome data. Regional-level annotated manually, and cluster-level determined by Seurat clustering. (C,D) Representative spatial transcriptome data (6 months replicate #2), colored by cluster-level annotation and represented as (C) UMAP and (D) spatial transcriptome. (E,F) Representative spatial transcriptome data (6 months replicate #2), colored by region-level annotation and represented as (E) UMAP and (F) spatial transcriptome. (G,H,I) Cluster-level annotation across replicates and datasets represented as (G) UMAP and (H) spatial transcriptome. (I) Fraction of spots corresponding to each cluster. (J,K,L) Region-level annotation across replicates and datasets represented as (J) UMAP and (K) spatial transcriptome. (L) Fraction of spots corresponding to each region.

Fig S6

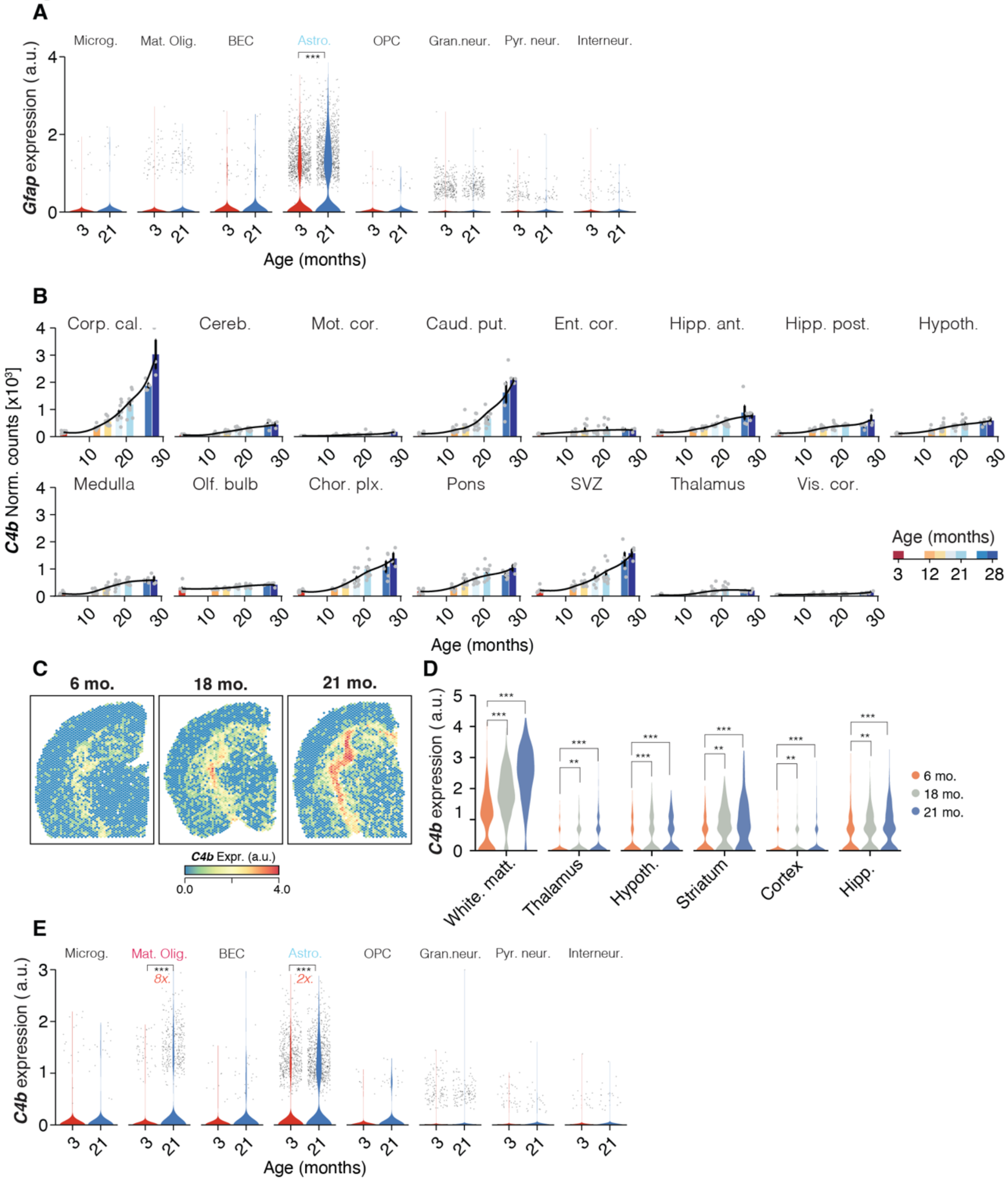

Figure S6 Cell type-specific quantification of *Gfap* and *C4b*

(A) Violin plot of *Gfap* expression across hippocampal cell types. Points indicate nuclei-wise expression levels, and the violin indicates average distribution of expression split by age. (B) *C4b* expression across all bulk regions, colored by age. Black lines indicate averaged-smoothed gene expression. Data are mean ± s.e.m. (C) *C4b* expression in spatial transcriptome across age. (D) Violin plot of *C4b* expression in exemplary region-level clusters of spatial transcriptome data, according to Fig S5. (E) Violin plot of *C4b* expression across hippocampal cell types. Points indicate nuclei-wise expression levels, and the violin indicates average distribution of expression split by age. (MAST, Benjamini–Hochberg correction; false discovery rate (FDR) < 0.05 and logFC > 0.2 to be significant). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Fig S7

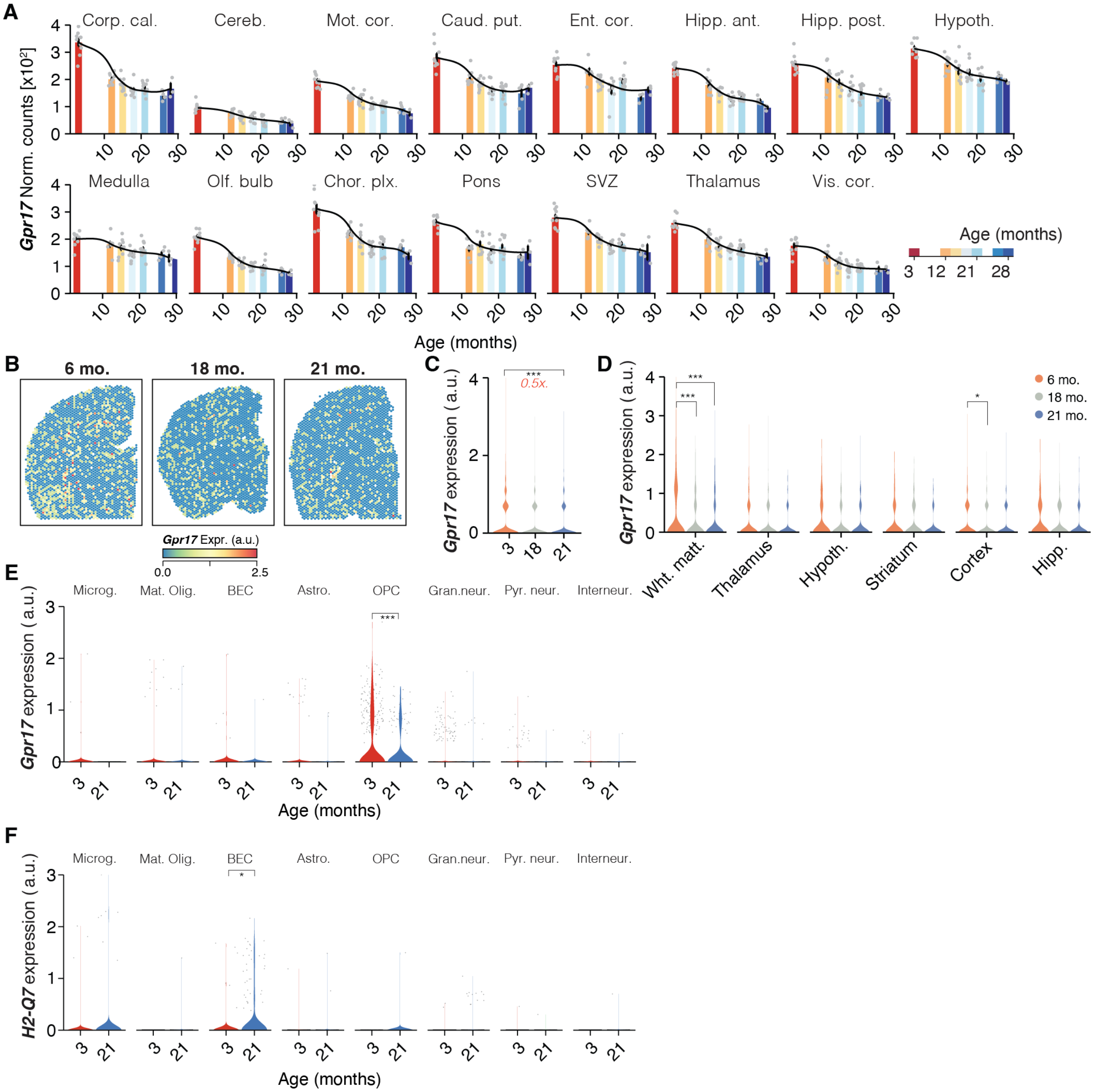

Figure S7 Cell type-specific quantification of *Gpr17* and *H2-Q7*

(A) *Gpr17* expression across all bulk regions, colored by age. Black lines indicate averaged-smoothed gene expression. Data are mean ± s.e.m. (B) *Gpr17* expression in spatial transcriptome across age. (C) Violin plot of *Gpr17* expression full spatial transcriptome data across ages. (D) Violin plot of *Gpr17* expression in exemplary region-level clusters of spatial transcriptome data, according to Fig S5. (E,F) Violin plot of (E) *Gpr17* and (F) *H2-Q7* expression across hippocampal cell types. Points indicate nuclei-wise expression levels, and the violin indicates average distribution of expression split by age. (MAST, Benjamini–Hochberg correction; false discovery rate (FDR) < 0.05 and logFC > 0.2 to be significant). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Fig S8

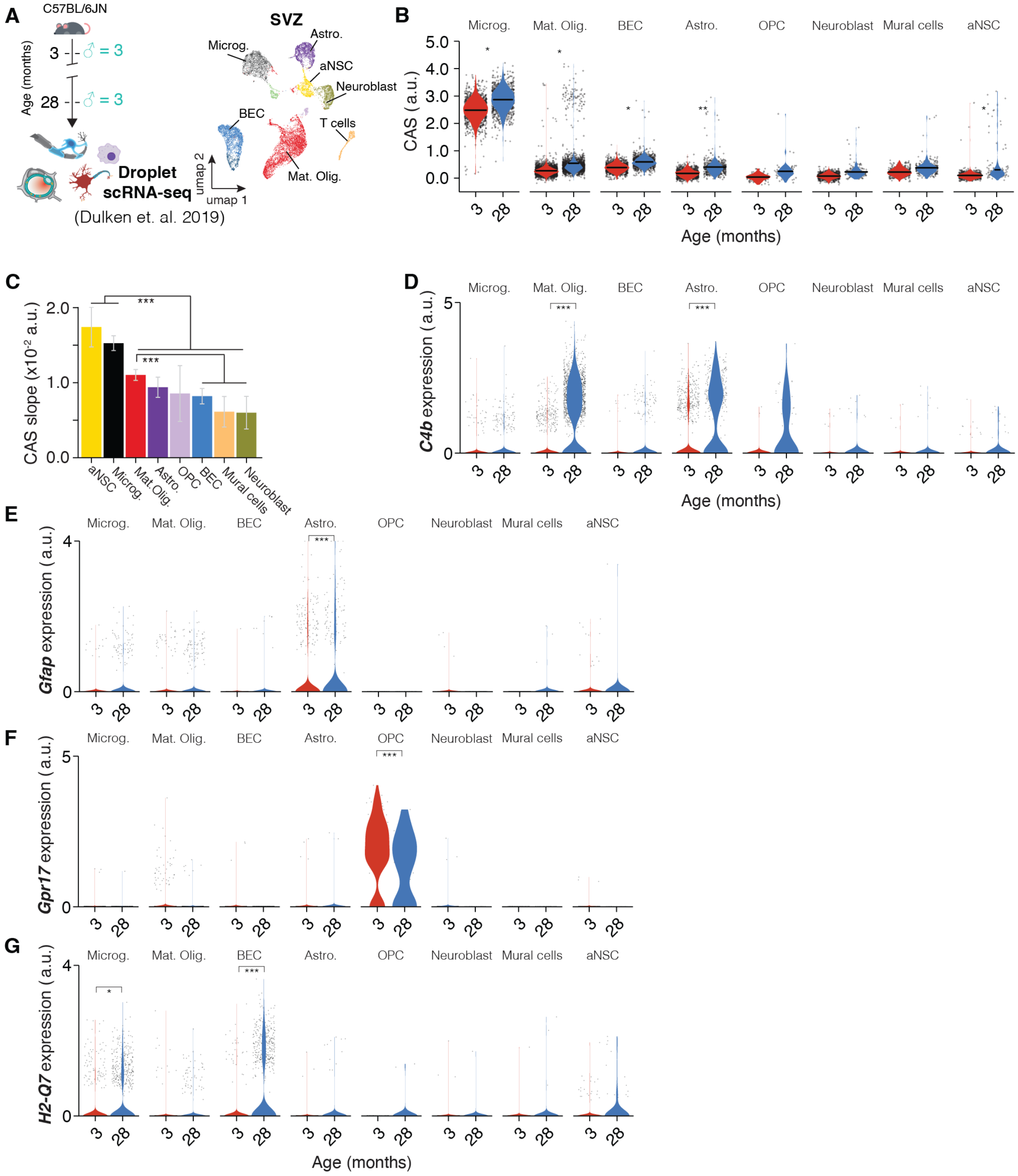

Figure S8 CAS analysis of SVZ scRNA-seq data

(A) Meta-analysis of droplet scRNA-seq data from (*24*) of cells from the SVZ. UMAP representation of all cell populations (n = 15,684 cells). (B) Violin plot representing CAS across cell types in the SVZ. Points indicate nuclei-wise expression levels, and the violin indicates average distribution of expression split by age. P values calculated with two-tailed t-test on per-replicate median of score. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. (C) CAS slope of linear regressions in (B), colored by cell type. Data are mean ± 95% confidence intervals. Two-sided Tukey's HSD test, adjusted for multiple testing, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. The highest (least significant) Pval is indicated. (D-G) Violin plot representing (D) *C4b*, (E) *Gfap*, (F) *Gpr17* and (G) *H2-Q7* expression across cell types in the SVZ. Points indicate nuclei-wise expression levels, and the violin indicates average distribution of expression split by age. (MAST, Benjamini–Hochberg correction; false discovery rate (FDR) < 0.05 and logFC > 0.2 to be significant). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Fig S9

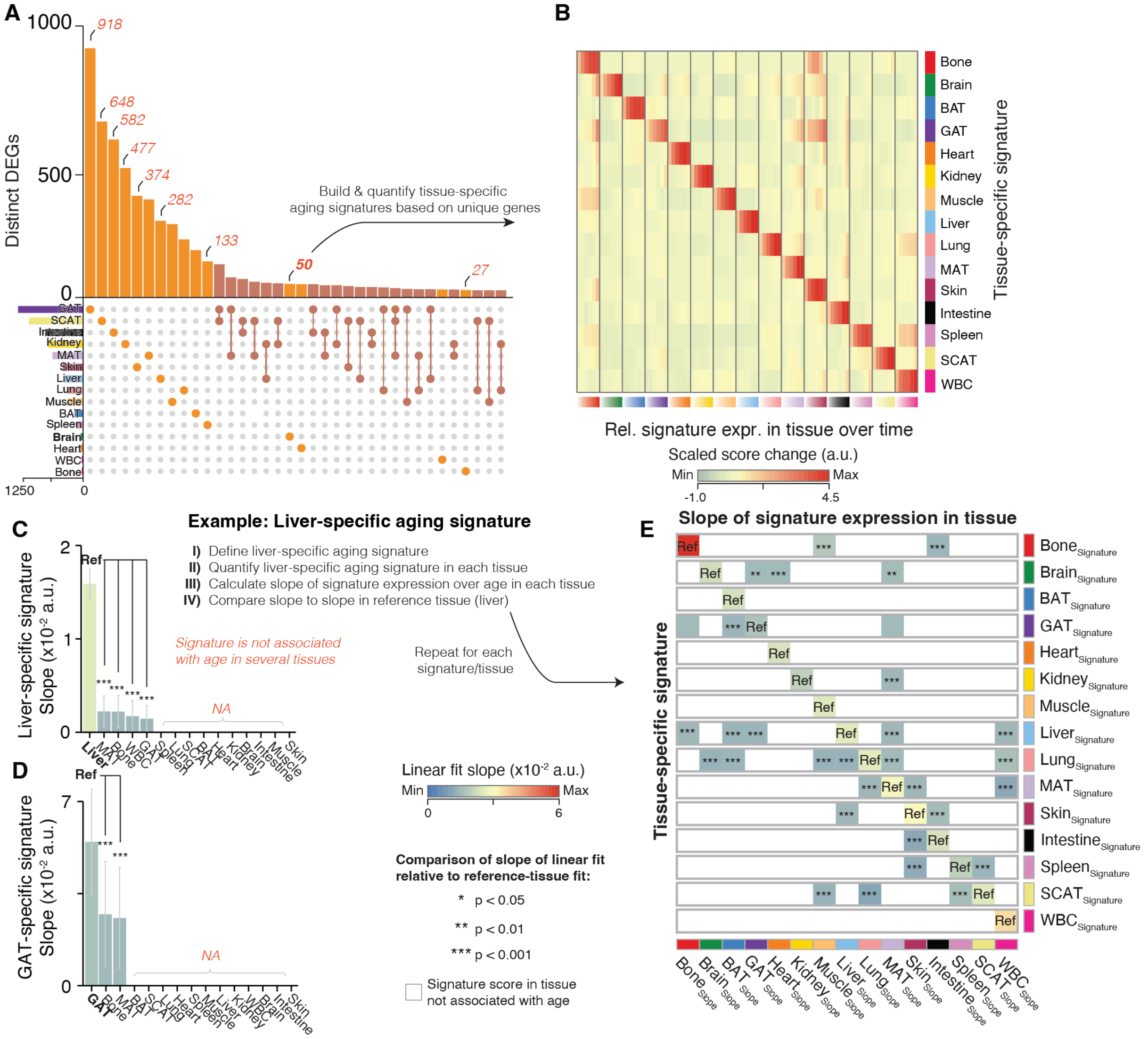

Figure S9 Identifying unique transcriptional signatures of aging across peripheral organs

(A) Meta-analysis from *(8)* for cross-organ specificity of DEGs. UpSet plot showing a matrix layout of DEGs shared across and specific to a tissue. Each matrix column represents either DEGs specific to a tissue (single circle with no vertical lines) or DEGs shared between tissue, with the vertical line indicating the tissues that share that given DEG. Top, bar graph displays the number of DEGs in each combination of tissue. Left bottom, bar graph displays the total number of DEGs for a given tissue. Only gene sets ≥ 25 genes are shown. Unique gene sets were used to construct tissue-specific aging signatures. (B) Tissue-wise score changes with age relative to 3 months (column-wise from left to right) for tissue-specific signatures. Score changes are z-scaled within a row. (C) Slope of linear regressions for the liver-specific signature, colored by slope. The liver signature was quantified in each organ and tested for correlation with age. Tissues indicated with NA exhibited no significant association with age. Subsequently, the slope of the linear fit for each organ that exhibited association with age was compared against the reference tissue, where the signature was first identified in (i.e., the liver). Pairwise comparisons across all tissues were run and Pval corrected for multiple testing using Tukey’s method. Comparisons relative to the reference tissue were inspected to determine if the tissue-specific signature exhibits an age-related increase distinct to the reference tissue. Data are mean ± 95% confidence intervals. Two-sided Tukey’s HSD test, adjusted for multiple testing, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. The highest (least significant) Pval is indicated. (D) Analysis as in (C) for the gonadal adipose tissue (GAT) - specific signature. (E) Summarized results for each tissue-specific signature analysis. Reference tissues are indicated with ‘Ref” and slopes for each tissue with significant age-association in a given signature is indicated by color. Tissues that exhibited no association with age (NA) for a given signature were left blank. Adjusted *P* values for slope comparison of a given tissue with the reference tissue are indicated.

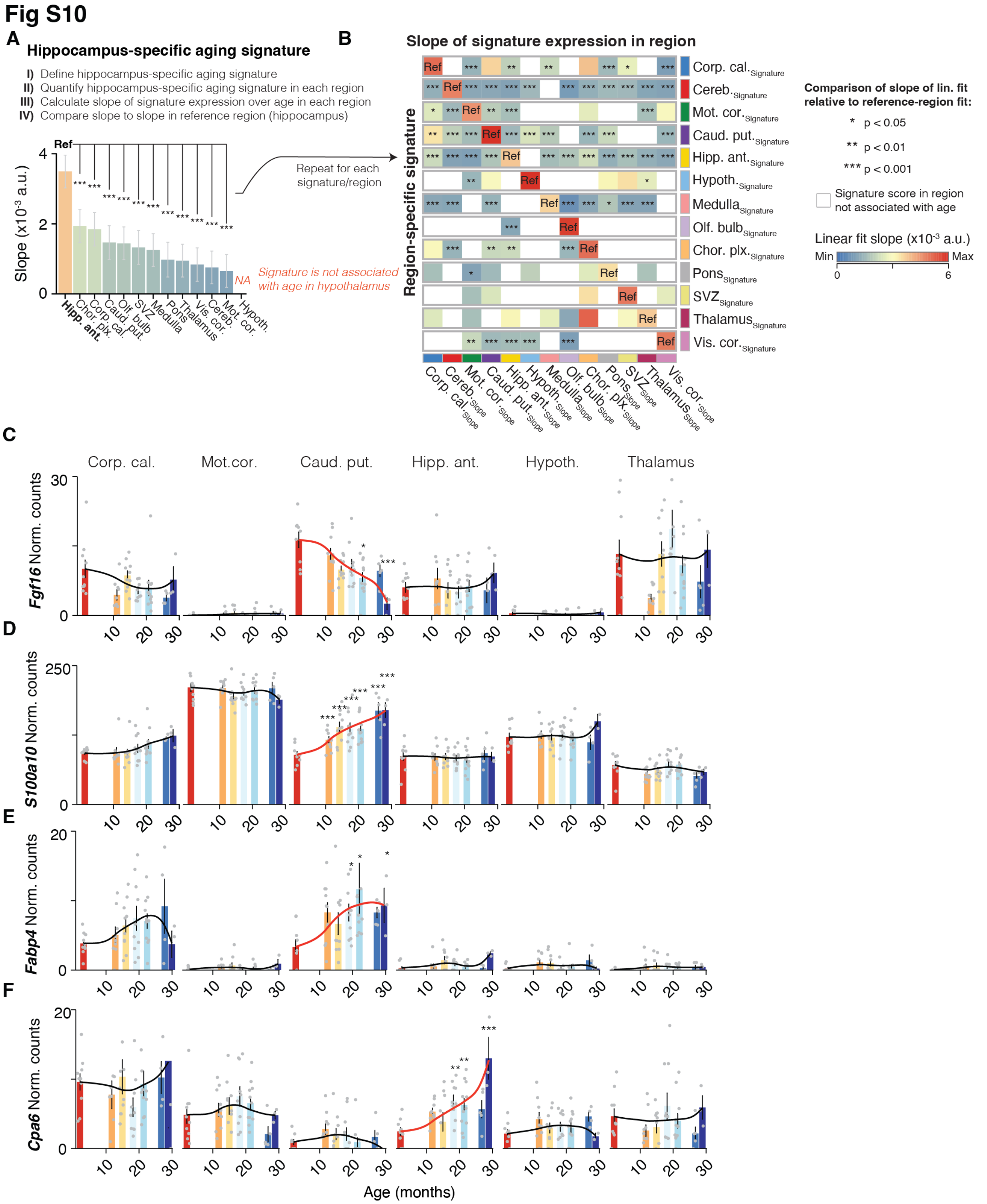

Figure S10 Identifying unique transcriptional signatures of aging across brain regions

(A) Slope of linear regressions for the hippocampus-specific signature, colored by slope. The hippocampus signature was quantified in each region and tested for correlation with age. Regions indicated with NA exhibited no significant association with age. Subsequently, the slope of the linear fit for each region that exhibited association with age was compared against the reference region, where the signature was first identified in (i.e., the hippocampus). Pairwise comparisons across all regions were run and *P* values corrected for multiple testing using Tukey's method. Comparisons relative to the reference region were inspected to determine if the region-specific signature exhibits an age-related increase distinct to the reference region. Data are mean ± 95% confidence intervals. Two-sided Tukey's HSD test, adjusted for multiple testing, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. The highest (least significant) Pval is indicated. (B) Summarized results for each region-specific signature analysis. Reference regions are indicated with 'Ref' and slopes for each region with significant age-association in a given signature is indicated by color. Regions that exhibited no association with age (NA) for a given signature were left blank. Adjusted *P* values for slope comparison of a given region with the reference region are indicated. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. (C-G) Bulk expression across proximal brain regions of genes (C) *Fgf17*, (D) *S100a10*, (E) *Fabp4*, (F) *Cpa6* that exhibit distinct changes in caudate putamen or hippocampus. Black lines indicate averaged-smoothed gene expression. The trajectory with significant age effect is highlighted. Data are mean ± s.e.m.

Fig S11

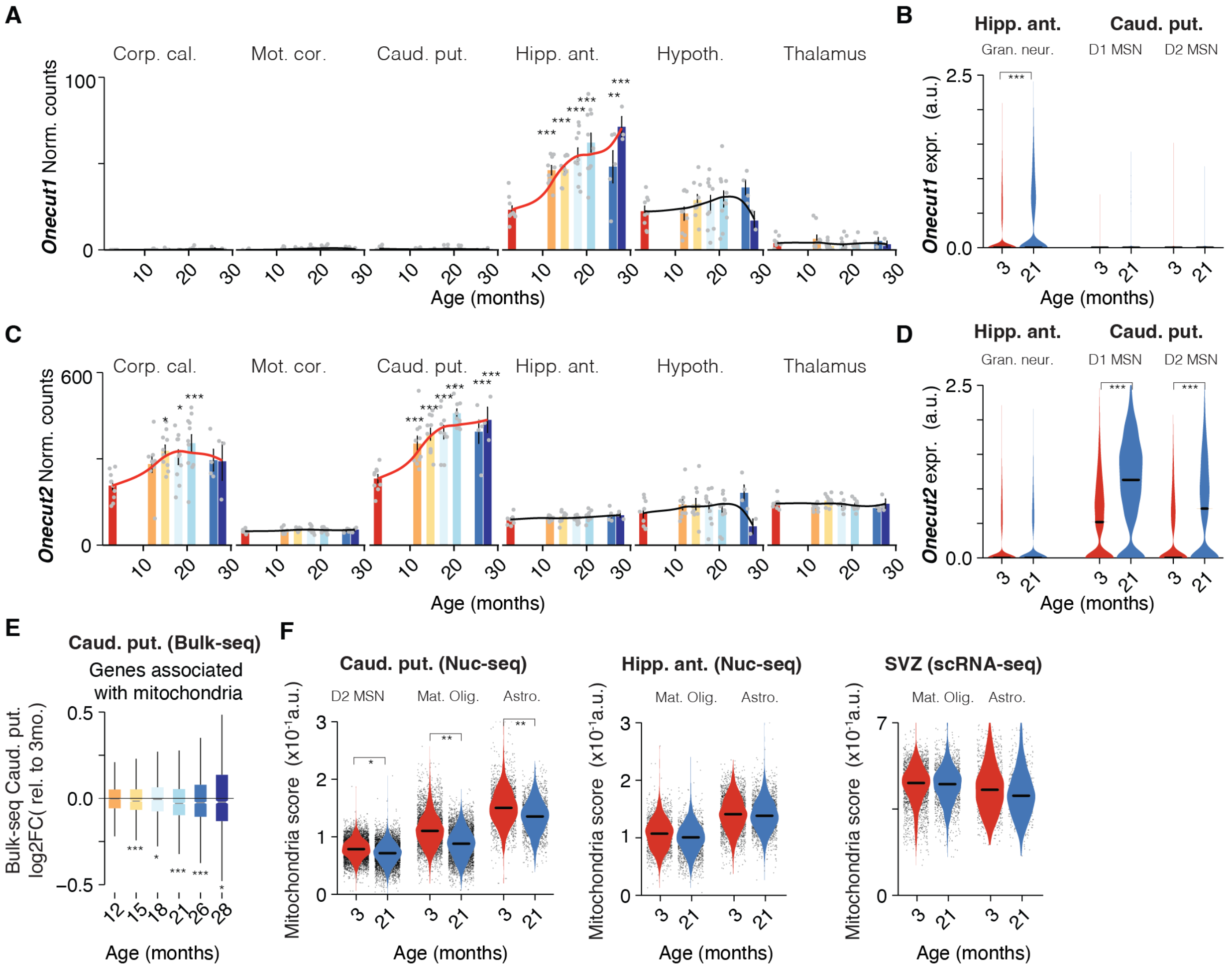

Figure S11 Region-specific expression of Onecut transcription factors during aging

(A) Bulk expression across proximal brain regions of *Onecut1*. Black lines indicate averaged-smoothed gene expression. The trajectory with significant age effect is highlighted. Data are mean ± s.e.m. (B) Violin plot of *Onecut1* expression across neuronal cell types in hippocampus and caudate putamen. Points indicate nuclei-wise expression levels, and the violin indicates average distribution of expression split by age. (MAST, Benjamini–Hochberg correction; false discovery rate (FDR) < 0.05 and logFC > 0.2 (log2FC > 0.7) to be significant). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. (C) Same as (A) for *Onecut2*. Data are mean ± s.e.m. (D) Same as (B) for *Onecut2*. (E) Distribution of gene-wise expression changes in caudate putamen with age relative to 3 months for genes associated with the GO term 'Mitochondrion' (n = 1,733 genes). Two-sided Wilcoxon rank-sum test, adjusted for multiple testing. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. (F) Violin plot representing mitochondria RNA score across D2 MSN, mature oligodendrocytes and astrocytes in caudate putamen (left). Same for astrocytes and oligodendrocytes in the anterior hippocampus (center) and SVZ (right). Points indicate nuclei-/cell-wise expression levels, and the violin indicates average distribution of expression split by age. P values calculated with two-tailed t-test on per-replicate median of score. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Supplementary table legends

Table S1. Bulk-seq region marker genes (separate file)
Marker gene analysis for bulk-seq regions from Figure 1C and Figure S3A.

Table S2. DEG detection across regions (separate file)

DEGs and the number of tissues in which they were detected in. Further indicates CAS and region-specific DEGs from Figure 2A,B and Figure 5A.

Table S3. Age-correlated genes (separate file)

List of genes with significant correlation with age from Figure 1H.

Table S4. WGCNA modules (separate file)

List of modules as detected by WGCNA from Figure 1 I,J. Enriched Gene Ontology terms and cell type markers are indicated. Modules with significant age-related change are noted.

Table S5. Functional enrichment of CAS genes (separate file)

Complete list of enriched Gene Ontology terms for CAS genes from Figure 2C.

Table S6. Spatial-seq cluster marker genes (separate file)

Marker gene analysis for Spatial-seq clusters from Figure S5.

Table S7. DEGs detected in bulk- and Spatial-seq (separate file)

Comparison of age-related DEGs found in bulk- and Spatial-seq data of corpus callosum/white matter cluster and motor cortex/cortex cluster, from Figure 2L.

Table S8. Functional enrichment of caudate putamen-specific DEGs (separate file)

Complete list of enriched Gene Ontology terms for caudate putamen-specific DEGs from Figure 5E.

Table S9. Cell dispersion analysis for hippocampus and caudate putamen (separate file)

Quantification of cell-specificity of DEGs found in hippocampus or caudate putamen from Figure 5G

Table S10. DEGs with human GWAS homologue (separate file)

List of DEGs with human disease GWAS homologue for AD, PD and MS from Figure 6. Colored cells indicate regions where we discovered a significant enrichment of disease-related genes.