



OPEN

# Single-cell analysis of the aging female mouse hypothalamus

Kaitlyn H. Hajdarovic<sup>1,8</sup>, Doudou Yu<sup>2,8</sup>, Lexi-Amber Hassell<sup>1,3</sup>, Shane A. Evans<sup>4</sup>, Sarah Packer<sup>3</sup>, Nicola Neretti<sup>3,5</sup> and Ashley E. Webb<sup>1,3,5,6,7</sup>

**Alterations in metabolism, sleep patterns, body composition and hormone status are all key features of aging. While the hypothalamus is a well-conserved brain region that controls these homeostatic and survival-related behaviors, little is known about the intrinsic features of hypothalamic aging. Here, we perform single-nuclei RNA sequencing of 40,064 hypothalamic nuclei from young and aged female mice. We identify cell type-specific signatures of aging in neuronal subtypes as well as astrocytes and microglia. We uncover changes in cell types critical for metabolic regulation and body composition and in an area of the hypothalamus linked to cognition. Our analysis also reveals an unexpected female-specific feature of hypothalamic aging: the master regulator of X inactivation, *Xist*, is elevated with age, particularly in hypothalamic neurons. Moreover, using machine learning, we show that levels of X chromosome genes and *Xist* itself, can accurately predict cellular age. This study identifies critical cell-specific changes of the aging hypothalamus in mammals and uncovers a potential marker of neuronal aging in females.**

While human lifespan has increased dramatically in recent years, improvements in healthspan, the period of life in which a person is disease-free, have been more modest<sup>1</sup>. Susceptibility to a host of diseases increases with aging, including diabetes, stroke<sup>2</sup>, cancer<sup>3</sup> and neurodegenerative diseases<sup>4</sup>. Aging is accompanied by changes in body composition, including decreased lean muscle mass, loss of bone density and increased abdominal fat<sup>1</sup>. Concomitant with these changes are alterations in endocrine states, such as decreased sex hormone production and reduced growth hormone and insulin-like growth factor-1<sup>5</sup>. Endocrine function and homeostatic processes, such as energy homeostasis<sup>6</sup> and release of sex hormones<sup>5</sup>, are controlled by neuropeptidergic neurons in the hypothalamus.

Nutrient sensing is one of several functions of the hypothalamus that implicates this brain region in healthy aging. Specific neuronal subtypes in the hypothalamus respond to circulating cues to organize the response to dietary changes through regulation of energy balance, glucose homeostasis, and growth factor secretion<sup>6</sup>. Caloric restriction (CR) is one of the most well-established behavioral interventions that improves lifespan and healthspan in many model organisms<sup>7</sup>. Genetic models that mimic the effects of CR via modulation of energy sensing pathways have revealed the mechanistic underpinnings of lifespan extension. For example, in *Caenorhabditis elegans*, the effects of dietary restriction are dependent on the function of neuropeptidergic energy sensing neurons; genetic manipulation of energy sensing genes in those neurons is sufficient to increase longevity<sup>8</sup>. Similarly, lifespan extension in the fruit fly *Drosophila* is dependent on specialized neurons called median neurosecretory cells<sup>9</sup>. In rodents, manipulations to the hypothalamus can also alter lifespan. Specifically, brain-specific overexpression of *Sirt1* leads to alterations in the dorsomedial and lateral hypothalamus and increases lifespan<sup>1</sup>. In addition, alteration

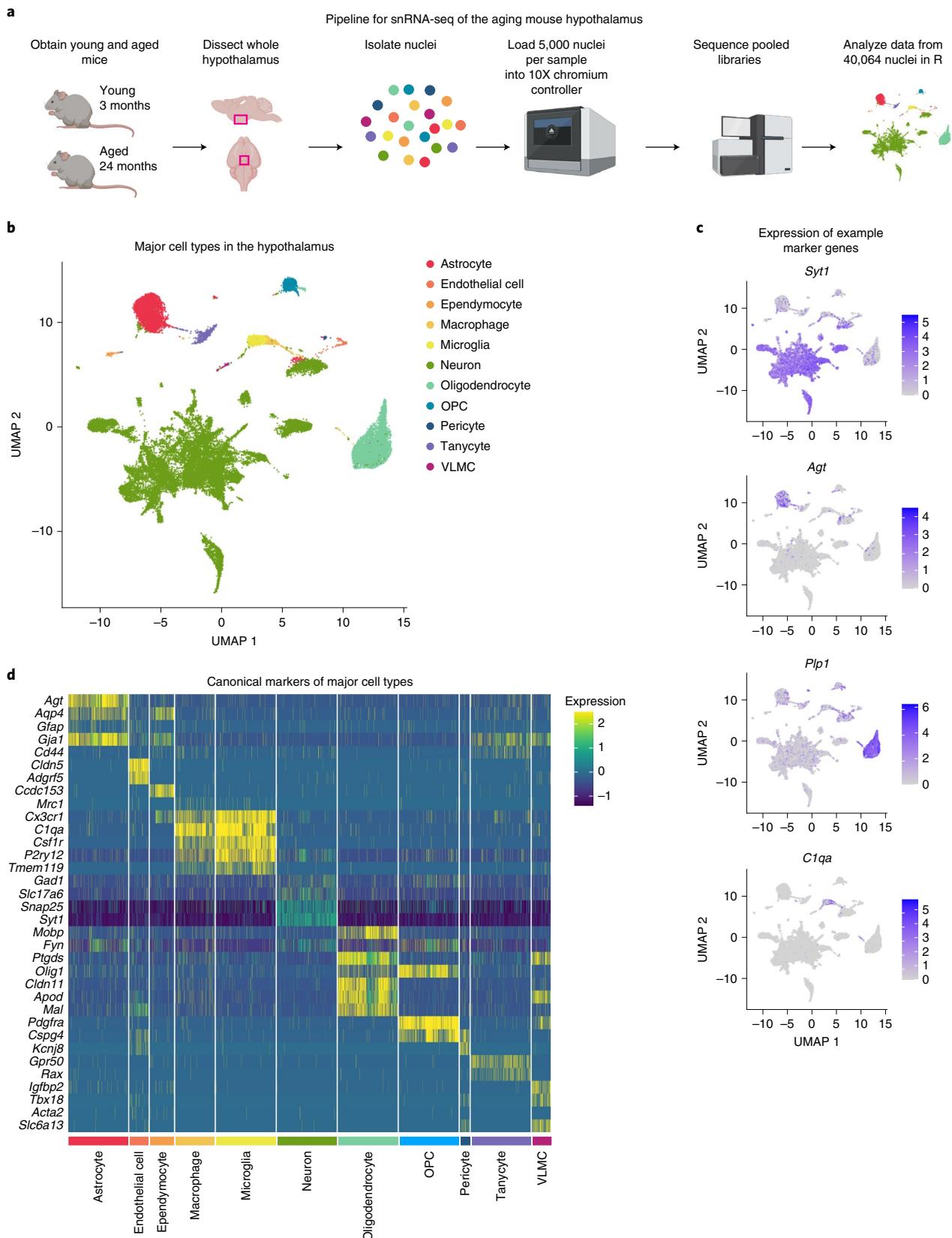
of immune signaling in the mediobasal hypothalamus affects longevity, with a reduction in immune signaling promoting longevity<sup>11</sup>.

Sex differences in lifespan have been documented in many species, including mice<sup>12</sup>. In humans, there is a robust difference in female and male lifespan across countries, with females living an average of 4–10 years longer than males<sup>13</sup>. Intriguingly, interventions that extend lifespan in model organisms do so in a sex-specific manner. For example, CR is one of the most robustly studied interventions and its effects have been observed from yeast to non-human primates<sup>7</sup>. Like many interventions, CR has sex-specific effects, with restricted females generally living longer than males on the same diet<sup>14</sup>. Similarly, the brain-specific *Sirt1* overexpression model results in a larger lifespan increase for females when compared to males<sup>10</sup>.

However, the aging female brain remains critically understudied and we know little about how areas involved in healthy aging, such as the hypothalamus, change with age in females. Epigenetic and transcriptional changes are widespread across tissues during aging, including in the brain<sup>15,16</sup>. Key transcriptional factors such as FOXO/DAF-16, NF-κB and MYC function as conserved regulators of these networks and have been implicated in aging<sup>11,15,17</sup>; however, despite a great interest in how changes in transcriptional programs affect aging, our understanding of how distinct cellular subtypes change transcriptionally with age remains limited. Investigation of how transcriptional programs change in a cell-type-specific manner in the hypothalamus will provide important insight into the aging process across tissues. Recent advances in single-cell RNA sequencing (RNA-seq) have expanded our understanding of the diverse cell types that comprise the hypothalamus<sup>18–24</sup>. This approach allows the investigation of previously unappreciated transcriptional and functional diversity of this brain region. Here, we use a single-nuclei RNA-seq (snRNA-seq) approach to identify aging-associated

<sup>1</sup>Neuroscience Graduate Program, Brown University, Providence, RI, USA. <sup>2</sup>Molecular Biology, Cell Biology, and Biochemistry Graduate Program, Brown University, Providence, RI, USA. <sup>3</sup>Department of Molecular Biology, Cell Biology, and Biochemistry, Brown University, Providence, RI, USA. <sup>4</sup>Graduate program in Computational Biology, Brown University, Providence, RI, USA. <sup>5</sup>Center on the Biology of Aging, Brown University, Providence, RI, USA. <sup>6</sup>Carney Institute for Brain Science, Brown University, Providence, RI, USA. <sup>7</sup>Center for Translational Neuroscience, Brown University, Providence, RI, USA.

<sup>8</sup>These authors contributed equally: Kaitlyn H. Hajdarovic, Doudou Yu.



**Fig. 1 | Single-nuclei analysis of the hypothalamus.** **a**, Schematic detailing the experimental workflow from dissection through analysis (created with BioRender). **b**, Uniform Manifold Approximation and Projection (UMAP) plot of all 40,064 nuclei used for analysis. Clustering analysis revealed 11 broad categories of cell type identity. **c**, UMAP plots of all nuclei labeled for expression of cell type-specific markers. *Syt1*, neurons; *Agt*, astrocytes; *Plp1*, oligodendrocytes; *C1qa*, microglia/macrophages. Color scale indicates level of gene expression. **d**, Heat map highlighting expression of cell type markers in each cluster, a maximum of 500 nuclei per cluster are displayed.

transcriptional changes across the mouse hypothalamus, thereby capturing the diversity of cell types in this brain region.

## Results

**Single-nuclei sequencing of the aging mouse hypothalamus.** We employed snRNA-seq, which is currently the optimal method for single-cell transcriptomic profiling of the diversity of cell types in the adult mammalian brain<sup>25,26</sup>. We isolated nuclei from the hypothalamus of young (3 months) and aged (19–24 months) female mice, with four replicate libraries for each age (Fig. 1a). After quality filtering, we obtained 40,064 high-quality nuclei for analysis: 16,256 and 23,808 nuclei from young and aged animals, respectively (Supplementary Fig. 1a). Cellular composition and data quality were similar across replicates (Supplementary Fig. 1b–d).

Clustering analysis with the Louvain algorithm revealed distinct clusters representing the major cell types of the hypothalamus, which we identified based on expression of canonical markers (Fig. 1b–d). For each individual cluster, we identified the top ten genes that were differentially expressed using the Wilcoxon rank-sum test (Supplementary Table 1). For example, neurons were defined by expression of *Syt1*, astrocytes defined by *Agt* and *Gja1*, oligodendrocytes by *Olig1* and *Plp1* expression and oligodendrocyte precursor cells (OPCs) were identified by expression of *Pdgfra*. The microglia and macrophage clusters were defined by expression of *C1qa* and distinguished by higher expression of *Tmem119* and *P2ry12* in the microglia cluster (Fig. 1c). Less abundant cell types were also observed, including ependymocytes (ependymal cells; *Ccdc153*), pericytes (*Flt1*), endothelial cells (*Clnd5*) and vascular and leptomeningeal cells (VLMCs; *Slc6a13*). We also observed a distinct cluster of tanycytes, which are specific to the hypothalamus and defined by *Rax* expression. Nuclei in these broad categories expressed additional canonical markers associated with their cell type, for example, the astrocyte cluster expressed *Gfap*, further validating the identity of each cluster.

Cell type diversity is achieved through expression of transcriptional regulators that orchestrate cell type-specific gene expression networks. To identify the regulators responsible for distinct expression networks across cell types in the hypothalamus, we used SCENIC, a regulatory network inference tool<sup>27</sup>. SCENIC identifies regulons, defined as a transcription factor and the genes it regulates and scores the activity of the regulons in individual cells. Further, it provides a regulon specificity score (RSS), which indicates whether a given regulon is specific to an individual cell type or shared among clusters. In our analysis, we observed strong shared and cell-type-specific signatures for each cluster. For example, the *Dbx2* regulon is strongly enriched in the astrocyte cluster (RSS = 0.348), with almost all astrocytes expressing the regulon, whereas the *Atf2* and *Creb3l1* regulons are enriched in neurons (Supplementary Fig. 2a,b and Supplementary Table 2) (RSS = 0.639 and 0.626, respectively). Together, this analysis identifies the distinct gene expression signatures in the major cell types in the hypothalamus that are orchestrated by specific combinations of transcriptional regulators associated with cell type identity.

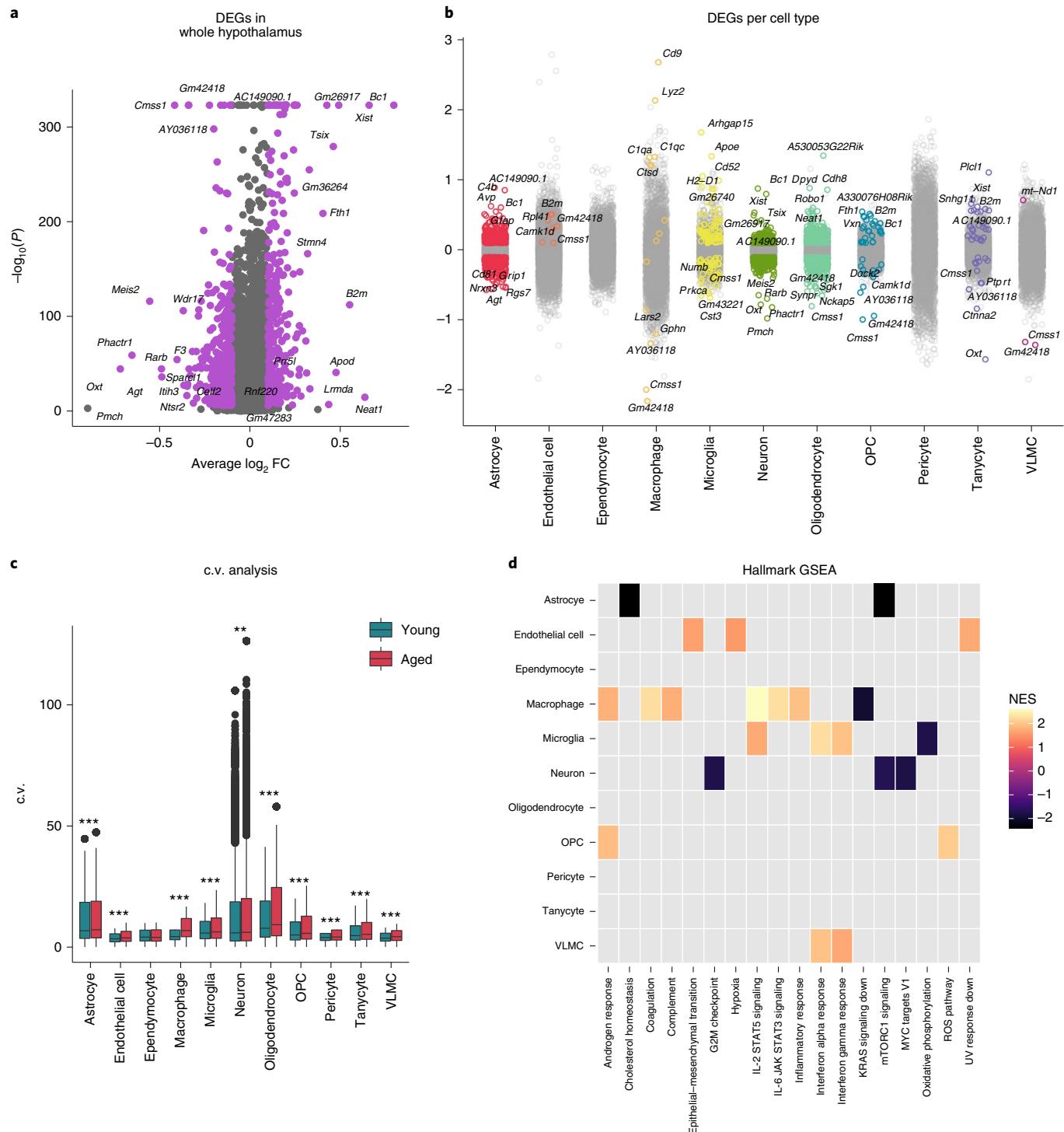
**Gene expression changes in major cell types with age.** We next investigated the changes in gene expression that occur with age in the major cell types of the hypothalamus. As expected, aging was not associated with changes in composition of this brain region and each major cell type was similarly represented in young and aged mice (Supplementary Fig. 1c,d). To gain a global understanding of how gene expression is altered with age, we first performed differential expression analysis on all cells using the model-based analysis of single-cell transcriptomics (MAST)<sup>28,29</sup>, with a random effect for sample of origin. Using this approach, we identified 275 and 342 genes that were upregulated and downregulated with age, respectively (adjusted *P* value (*P*<sub>adj</sub>) < 0.05, fold change (FC) > 0.1) (Fig. 2a

and Supplementary Table 3). As an initial validation, we cross-checked our results with publicly available bulk microarray data on the aging hypothalamus<sup>30</sup>. Although these data differ from ours in regard to strain and sex, this analysis confirmed several changes in our dataset, including downregulation of *Gria1*, *Apoe*, *Camk2a* and *Atp1b2* (Supplementary Table 3). Moreover, we confirmed a key finding from work on male rats showing that oxytocin binding is decreased in several brain regions with age, including the hypothalamus<sup>31</sup>. Here, we show a reduction in *Oxt* expression in aged female mice, suggesting a reduction of oxytocin signaling is a feature of both male and female aging. Notably, the most upregulated genes in the MAST analysis included *Xist* and *Tsix*, which both encode long non-coding RNAs involved in X chromosome inactivation<sup>32,33</sup> and are expressed only in females.

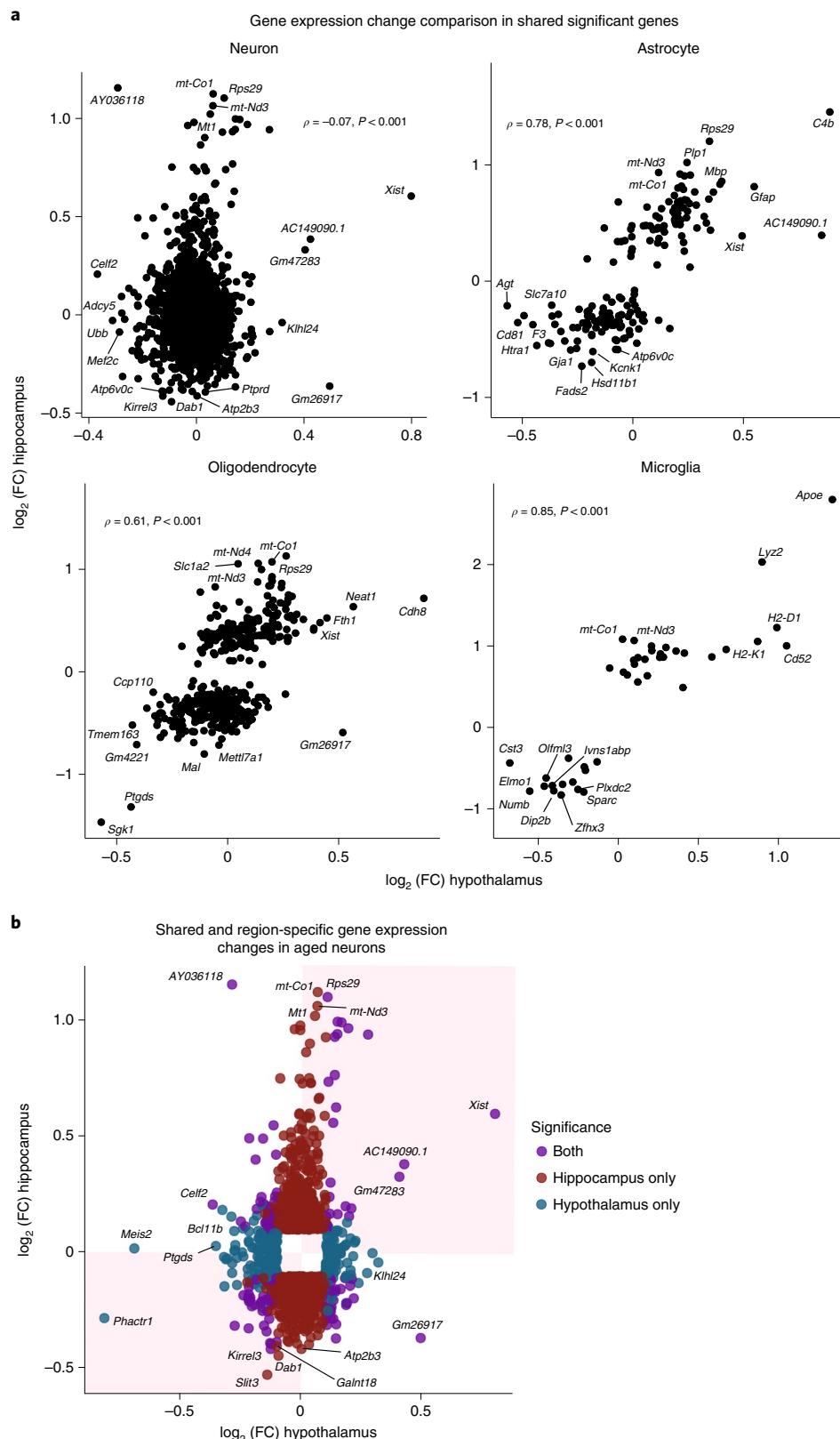
Next, we investigated the impact of age on gene expression in each major cell type. Neurons, astrocytes, oligodendrocytes and microglia showed the greatest numbers of differentially expressed genes (DEGs) with age (Fig. 2b and Supplementary Table 4). Additionally, we performed coefficient of variation analysis on the major cell types and observed a significant difference between ages, with nearly all types showing an increase with age (Fig. 2c). This finding suggests that variability in gene expression increases with age in most cell types, which likely contributes to cellular dysfunction within the aged hypothalamus.

To validate our findings and determine the extent to which the changes we observe are specific to the hypothalamus, we compared our dataset to a publicly available snRNA-seq dataset analyzing the female mouse hippocampus<sup>34</sup>. Astrocytes, oligodendrocytes, and microglia all showed high agreement between the hypothalamic and hippocampal datasets; for example, both datasets show significant upregulation of *C4b* in astrocytes, *Apoe* and *Lyz2* in microglia and *Cdh8* and *Neat1* in oligodendrocytes (Fig. 3a). When we compared the relationship in log<sub>2</sub>(FC) of significant (*P*<sub>adj</sub> < 0.05) genes between the hypothalamus and hippocampus, a statistically significant positive correlation emerged for astrocytes ( $\rho = 0.78$ , *P* < 0.001), oligodendrocytes ( $\rho = 0.61$ , *P* < 0.001) and microglia ( $\rho = 0.85$ , *P* < 0.001). Of note, in contrast to glia, there was less overlap in DEGs between the hypothalamic neurons and hippocampal neurons and there was little correlation in gene expression changes between hypothalamic and hippocampal neurons ( $\rho = -0.07$ , *P* < 0.001). While some genes, such as *Xist*, are upregulated with age in both neuronal subsets, other genes are either unchanged or regulated in opposite directions in the two sets (Fig. 3b). For example, *Phactr1* and *Meis2* are among the most significantly downregulated neuronal genes in the hypothalamic dataset but are not significantly changed with age in the hippocampal dataset (*P*<sub>adj</sub> > 0.05). The genes *Rsrc1*, *Gm26917* and *Kcnip4* are upregulated with age in the hypothalamus, but are downregulated in the hippocampus with age. Notably, the gene *Rps29* is upregulated in both datasets, in agreement with previous single-cell RNA-seq studies of the aged mouse brain<sup>35</sup>. Together these data suggest that glia share general signatures of aging in distinct brain regions, whereas region-specific signatures are predominant in neuronal aging.

To investigate the cellular processes that are altered with age in the different cell types in the hypothalamus, we performed Gene Set Enrichment Analysis (GSEA) using the hallmark gene set<sup>36</sup> (Fig. 2d). We observed unique signatures of aging in each cell type, as well as some shared hallmarks of aging. For example, neurons and astrocytes share an under-enrichment (negative normalized enrichment score) for the mTORC1 signaling gene set. Astrocytes were also under-enriched in genes related to cholesterol homeostasis, which aligns with previous research showing a decrease in expression of cholesterol synthesis genes in aged hypothalamic astrocytes<sup>37</sup>. In contrast, neurons displayed alterations in G<sub>2</sub>M checkpoint genes, which is notable in light of evidence linking Alzheimer's disease to aberrant cell cycle entry driven through activation of mTORC1



**Fig. 2 |** The aging hypothalamus harbors cell-type-specific transcriptional changes. **a**, Volcano plot showing overall differential expression of genes between all young and aged nuclei. Significant genes ( $P_{\text{adj}} < 0.05$ ,  $\text{FC} > 0.1$ ). Genes with a  $P$  value of zero were given an arbitrarily small  $P$  value for plotting purposes. MAST analysis with random effect for sample of origin and sequencing depth, with Bonferroni adjustment of  $P$  values. **b**, Strip plot showing DEGs in each cell type. Significant genes ( $P_{\text{adj}} < 0.05$ ,  $\text{FC} > 0.1$ , MAST analysis with random effect for sequencing depth and sample of origin) in color, nonsignificant genes are in gray. Top five upregulated and top five downregulated genes per cluster labeled. **c**, Coefficient of variation (c.v.) analysis for each cellular subtype. In almost all subtypes the c.v. is significantly higher in the aged condition (two-sided Wilcoxon test with Bonferroni correction, \*\*\* $P_{\text{adj}} < 0.001$ , \*\* $P_{\text{adj}} = 0.001639$ ,  $n = 25,135$  genes per group). Box indicates range from 25th to 75th percentile, with whiskers extending to 1.5 times the interquartile range. Outliers are plotted separately, center indicates median value. **d**, Heat map showing GSEA for hallmark terms. Color indicates normalized enrichment score (NES). Significant gene sets are calculated as  $P_{\text{adj}} < 0.1$ . ROS, reactive oxygen species.



**Fig. 3 | Shared and region-specific aging signatures between hypothalamus and hippocampus.** **a**, Two-sided Spearman's rank correlation of  $\log_2(\text{FC})$  in significant ( $P_{\text{adj}} < 0.05$ , MAST analysis with random effect for sample of origin and sequencing depth, with Bonferroni adjustment of  $P$  values) genes shared between the hypothalamus and hippocampus. Correlation of expression change in neurons is not correlated ( $\rho = -0.069, P < 0.001$ ), whereas gene expression changes in other cell types are strongly positively correlated: astrocytes ( $\rho = 0.78, P < 0.001$ ), oligodendrocytes ( $\rho = 0.61, P < 0.001$ ) and microglia ( $\rho = 0.85, P < 0.001$ ). **b**, Scatter-plot of genes significant in at least one dataset ( $P_{\text{adj}} < 0.05, \log_2(\text{FC}) > 0.1$ ). Color indicates whether a gene is significant in both datasets (purple), hippocampus only (red) or hypothalamus only (blue). Shading indicates genes that are changing in the same direction for both datasets.

(ref. <sup>38</sup>). Of note, mTORC1 activation in the arcuate nucleus of the hypothalamus is reduced in ovariectomized mice<sup>39</sup>. Thus, mTORC1 under-enrichment in the aged female hypothalamus represents a key a target for further study into the intersection of sex, aging and neurodegenerative disease.

Microglia and macrophages function as immune cells in the brain and both cell types show cell type-specific gene set enrichments with age (Fig. 2d). Macrophages are enriched in interleukin IL-2 STAT5 signaling, IL-6 JAK STAT3 signaling and inflammatory response gene sets. Microglia are enriched in genes related to the interferon- $\alpha$  response and interferon- $\gamma$  response and are under-enriched in the oxidative phosphorylation gene set with age. These data suggest distinct changes in different subsets of immune cells in the aging hypothalamus.

Glial cells, including microglia, are critical regulators of neuronal function. To understand how the relationships between these cells and neurons are changed with age, we utilized the ligand-receptor repository CellPhoneDB<sup>40</sup> to infer cell-cell interactions (Supplementary Fig. 3a). Notably, we found that loss of ligand-receptor interactions involving growth factors was a theme across the cell types studied. For example, in the young astrocyte-neuronal and tanycyte-neuronal cell pairs, TGFB2\_TGFB3 was enriched. Both cell pairs lose this enrichment with age. Similarly, there is a loss of enrichment in pathways involving FGF9. Specifically, while FGF9-FGFR3 and FGF9-FGFR1 are enriched in young astrocyte-neuronal and tanycyte-neuronal cell pairs, this enrichment is lost with age. The astrocyte-neuronal and oligodendrocyte-neuronal pairs also lose enrichment of the FGF9-FGFR2 ligand-receptor interaction with age. These factors are involved in diverse processes such as repair, learning and memory and neurogenesis<sup>41</sup>. Additionally, the aged astrocyte-neuron cell pair loses SEMA3A\_NRP1 signaling with age, whereas the microglia-neuronal cell pair loses both SEMA3A-NRP1 enrichment and SEMA3A-plexinA4 enrichment. Semaphorin3a is a known player in synaptogenesis<sup>42</sup>. Taken together, the loss of these cell signaling pathways may represent a mechanism for alterations in synaptogenesis and neuronal homeostasis in the aged hypothalamus.

**Aged microglia represent a progressive aging trajectory.** Microglia are macrophage-like cells found throughout the brain and are critical for the immune response, including release of cytokines and chemokines, antigen presentation and phagocytosis of debris<sup>43</sup>. Recent studies have revealed gene expression changes and microglial activation in the aged brain, which likely contribute to neurodegeneration<sup>43</sup>. Based on our findings that microglial-neuronal interactions involving the Alzheimer's-associated gene APP, as well as MIF were enriched with age (Supplementary Fig. 3a), we sought additional strategies to uncover changes in these cells over time. Using Monocle3 (ref. <sup>44</sup>), we performed pseudotemporal ordering of nuclei from the microglia and macrophage clusters. The trajectory accurately captures the transition from young to aged nuclei, suggesting a gradual progression toward aging in this cell type and a significant increase in the proportion of aged nuclei

across pseudotime (Fig. 4a,b). To confirm this analysis, we freshly isolated CD11b<sup>+</sup>hypothalamic microglia from mice at three time points (3, 12 and 24 months) and performed qPCR for candidate genes discovered in the pseudotime analysis. This experiment recapitulated specific genes trajectories (Supplementary Fig. 4a), validating this approach.

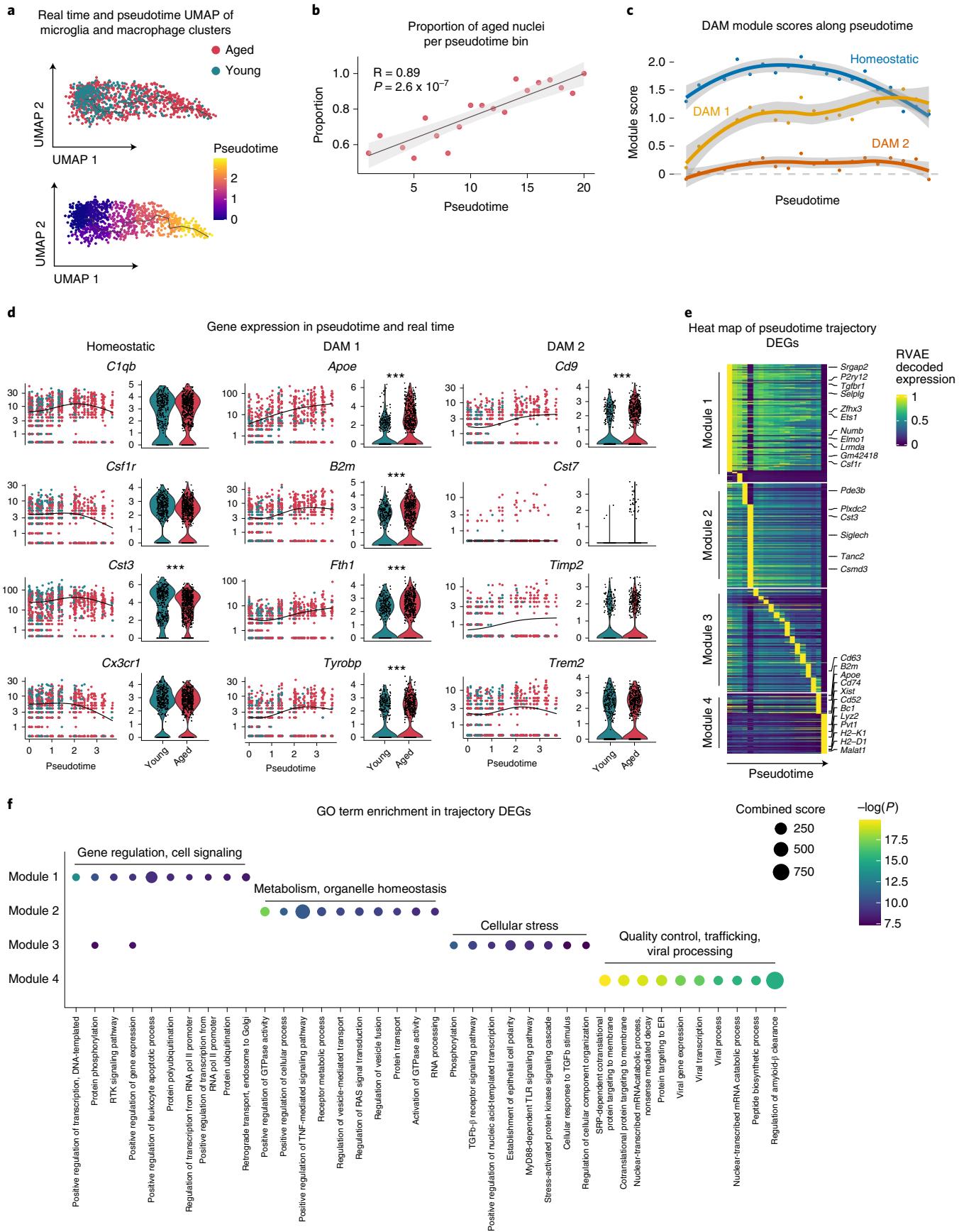
Since changes in microglia have been implicated in both physiological aging and neurodegeneration, we examined how disease-associated microglia (DAM) genes change as a function of pseudotime. We aggregated the expression of key genes from three microglia gene sets identified in the literature<sup>45,46</sup>, homeostatic microglia (homeostatic), TREM2 independent stage 1 DAM (DAM 1) and TREM2 dependent stage 2 DAM (DAM 2), and plotted the aggregated expression as a module score along the pseudotime trajectory. Overall, there was a decrease in the module score for the homeostatic module over pseudotime suggesting a loss of maintenance of healthy microglia over time. In contrast, we observed an increase in the DAM 1 disease module score near the end of pseudotime (Fig. 4c). The DAM 2 module does not seem to play a role in steady-state hypothalamic aging, as the module score remains low throughout pseudotime.

To further understand the role of these gene modules in aging hypothalamic microglia, we visualized gene expression across pseudotime and through real time (Fig. 4d). While young microglia generally cluster earlier in pseudotime (pseudotime 0.0 through 1.5), aged microglia expressing these genes are distributed throughout pseudotime. Thus, hypothalamic microglia from aged animals have increased heterogeneity representing a progressive aging trajectory. While a subset of aged microglia retaining a youthful gene expression signature, many aged microglia highly express disease-associated genes.

To fully capture gene expression changes along the trajectory, we performed Moran's *I* test on microglia and macrophage genes and found 2,112 statistically significant trajectory-dependent genes (Supplementary Table 5). To characterize their expression dynamics along pseudotime, we applied RVAgene<sup>47</sup>, an autoencoder neural network framework to reconstruct and smooth the pseudotime-dependent genes expression. We then visualized the recurrent variational autoencoder (RVAE) decoded expression along pseudotime in a heat map and manually grouped the genes into four modules according to their pseudotemporal expression patterns (Fig. 4e). For example, genes in the module 1 are highly expressed in early pseudotime while genes in the module 4 are expressed in late pseudotime. To understand the biological processes enriched in each module, we performed Gene Ontology (GO) analysis (Fig. 4f). Notably, gene modules are transitioning through pseudotime from positive regulation of biological processes to immune responses and finally to the amyloid- $\beta$  clearance and viral infection corresponding to known phenotypes of normal aging.

**Expression of X inactivation genes is altered with age.** Our initial differential expression analysis revealed the unexpected finding that the long non-coding RNA *Xist* is one of the most highly upregulated

**Fig. 4 | Trajectory analysis of aging hypothalamic microglia.** **a**, Monocle3 pseudotemporal ordering of microglia and macrophage clusters ( $n=1,121$  nuclei) defining a single trajectory from young to aged nuclei. Nuclei are colored by age (top) and pseudotime (bottom). **b**, Scatter-plot showing the proportion of aged nuclei along the pseudotime timeline in 20 time bins (sized 0.15 per time bin). Pearson correlation of the proportion of aged nuclei and pseudotime timeline,  $R=0.89$  and  $P=2.6 \times 10^{-7}$  (two-sided), 95% CI (in gray) shown. **c**, Plot showing the module expression score of three microglia states (homeostatic, DAM 1 and DAM 2). The darker lines are the local regression result for individual time bins (20 total), with the gray shadow depicting the 95% CIs. **d**, Kinetics plot showing the relative expression of representative genes for microglia states (left). The lines approximate expression along the trajectory using polynomial regressions. Violin plots of gene expression and the results of MAST with random effect for sample of origin and sequencing depth, with Bonferroni adjustment of  $P$  values ( $^{***}P_{adj} < 0.001$ , exact  $P_{adj}$  values reported in Supplementary Table 4) (right). **e**, Heat map showing modules of trajectory DEGs (t-DEGs) in the microglia cluster ( $n=2,112$  genes). The expression value is RVAE decoded expression. The genes were grouped into four modules after ranking by RVAE decoded expression. Module 1 (645 genes), module 2 (570 genes), module 3 (566 genes) and module 4 (331 genes). **f**, Dot plot showing the top ten GO biological process terms for genes in individual modules.



genes in the female hypothalamus with age (Fig. 3a). Differential expression analysis of each major cell type indicated upregulation of *Xist* with age in astrocytes, macrophages, microglia, neurons, oligodendrocytes, as well as tanycytes (Fig. 5a), and we also observed upregulation of *Xist* in the aging hippocampus<sup>34</sup> (Fig. 3a,b). *Xist* is a key player in X chromosome inactivation in females and is encoded on the X inactivation center (XIC), which harbors additional non-coding RNA genes involved in the same process<sup>32,33,48</sup>. Of note, we observed age-related upregulation of related RNAs in some cell types: *Ftx*, *Jpx*, and *Tsix* (Fig. 5a). We validated the upregulation of *Xist* using RNA extracted from independent tissue samples of different brain regions (hypothalamus, cerebellum, cortex, and olfactory bulb). Notably, although *Xist* trended up in all brain regions we tested, the upregulation only reached significance in the female hypothalamus, revealing a feature of female hypothalamic aging (Fig. 5b). As expected, we did not detect *Xist* expression in adult male mice and there was no upregulation of this gene with age in males (Fig. 5b). We further confirmed this finding using RNAScope to detect the *Xist* transcript in situ in coronal sections through the mouse hypothalamus. The average intensity of *Xist* expression in aged female hypothalamus (25 months) was significantly higher than in young female hypothalamus (3 months) using this method (Fig. 5c).

Although most genes on the inactive X chromosome are not expressed in females, a small number of genes are well known to 'escape' inactivation and are expressed from both X chromosomes. These X escape genes are species and tissue specific<sup>49,50</sup>. In the mouse, 14 genes escape X inactivation in brain tissue<sup>49</sup>. This list includes both *Xist* and *Ftx*, which have increased expression with age in our dataset. To determine whether increased XIC gene expression with age might be affecting escape genes, we interrogated expression of genes known to escape X inactivation in mice. We compiled a list of genes that are both known to escape X inactivation in any tissue context in mice and are expressed in our dataset. We found that although changes to XIC genes seems to be uniform across cell types in our data, age-related changes to expression of X escape genes are cell-type-specific (Fig. 5d). For example, in astrocytes, *Idh3g* is downregulated with age, whereas *Firre*, *Plp1* and *Tmsb4x* are upregulated. In neurons, *Gprasp1* and *Huwel* are downregulated with age, whereas *5530601H04Rik* and *Kdm5c* are upregulated. Of the 39 X escape genes expressed in the dataset, 15 were differentially regulated with age in at least one cell type. These data suggest that global changes to X chromosome regulation may be a feature of female hypothalamic aging.

Finally, to understand whether the changes in *Xist* that we observed in mouse aging might be related to age-associated pathologies in humans, we assessed changes in *XIST* expression between control and Alzheimer's disease human brain samples across two brain regions using publicly available snRNA-seq datasets<sup>51,52</sup>. Using MAST with a random effect for sample of origin, we compared *XIST* expression across all cells from females in two independent datasets (Fig. 5e). Notably, *XIST* is upregulated in human entorhinal cortex in women with Alzheimer's disease, which is one of the earliest and most affected regions in this disease ( $\log_2(\text{FC})=0.574$ ,  $P_{\text{adj}}<0.001$ ,

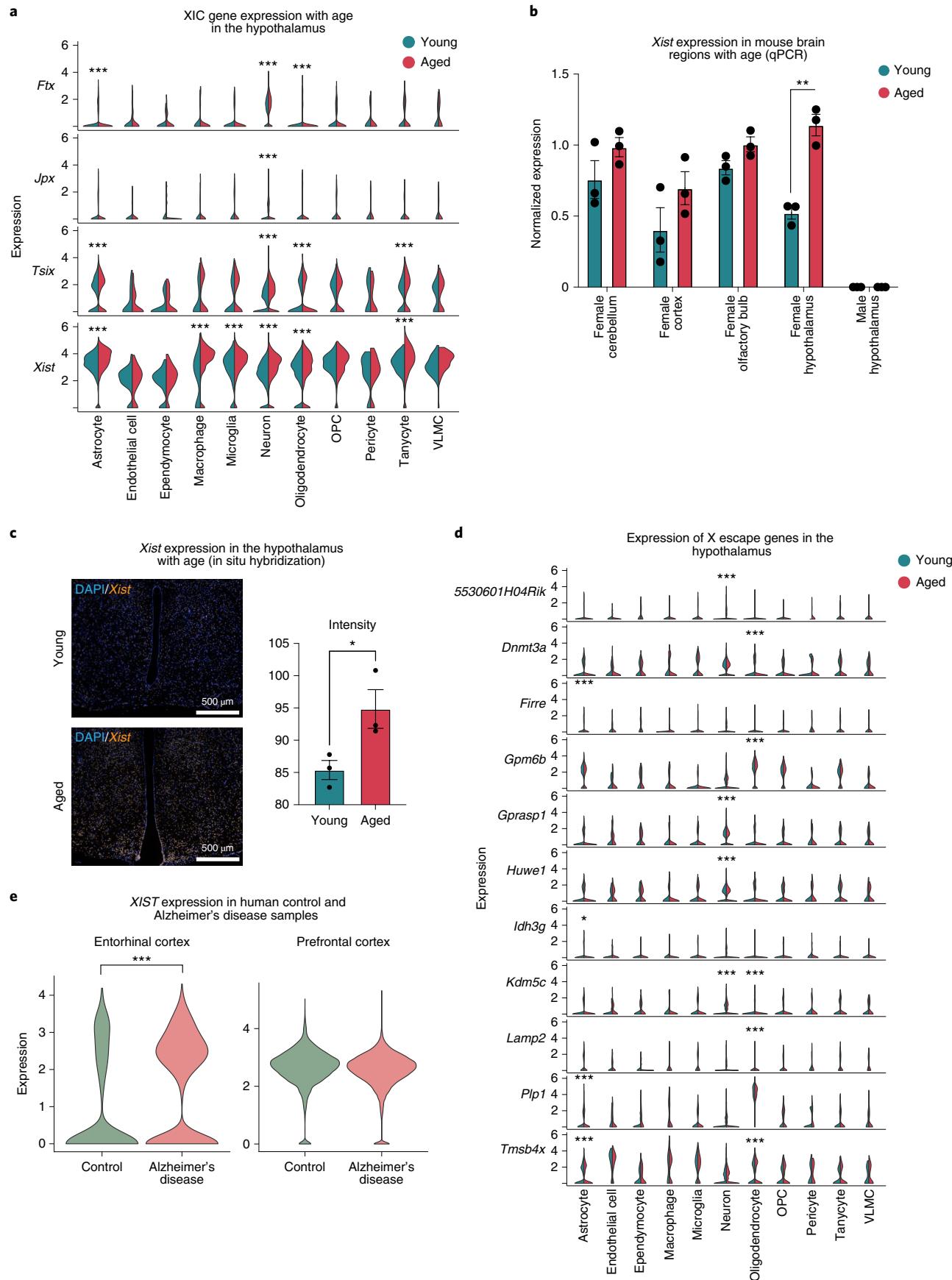
$n=3,942$  nuclei). In contrast, nuclei derived from human prefrontal cortex shows no changes in *XIST* expression between control and Alzheimer's disease samples ( $P_{\text{adj}}>0.05$ ,  $n=26,212$  nuclei). Thus, changes in *XIST* expression may be a brain-region-specific feature of Alzheimer's disease in female patients.

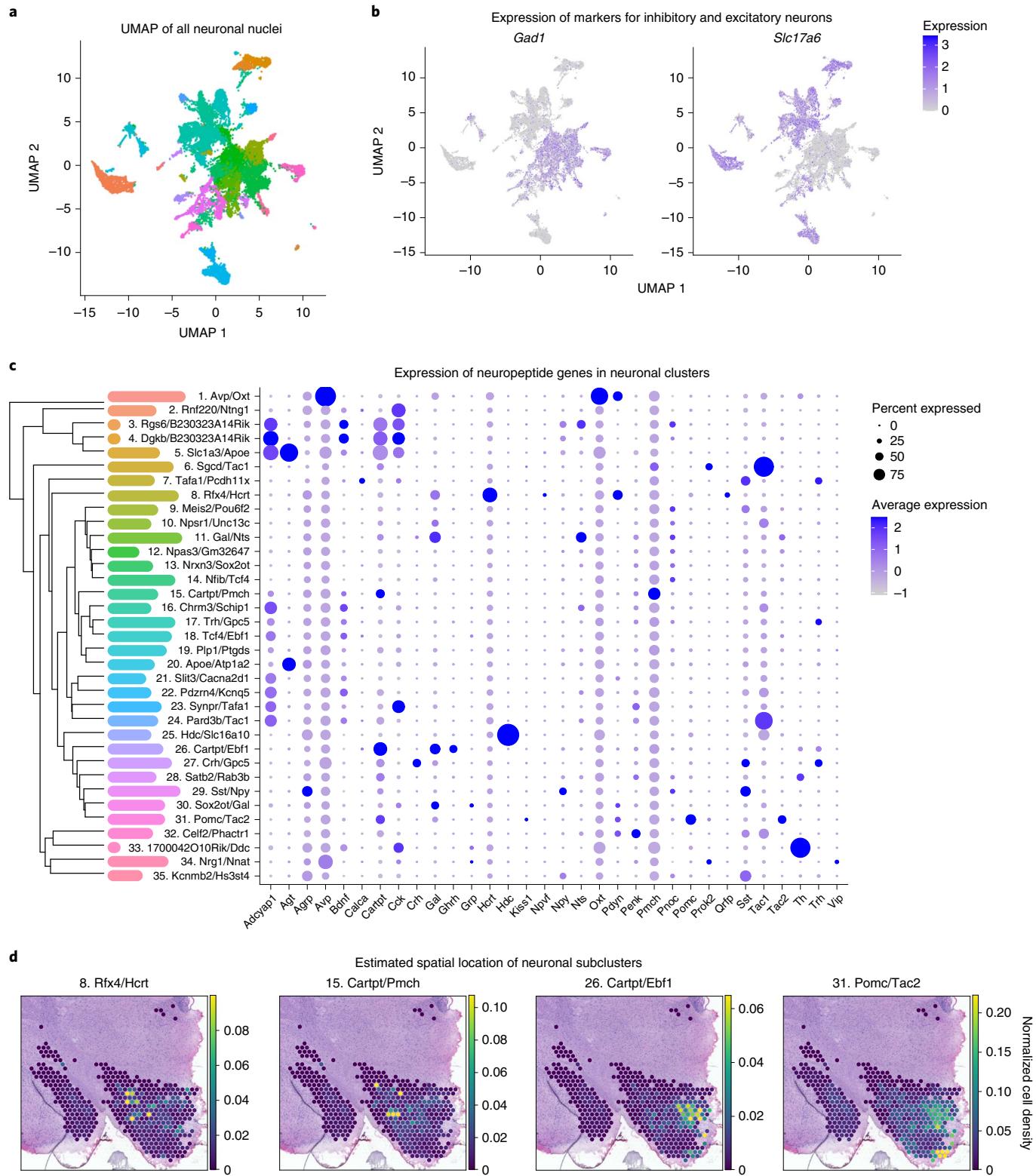
**Neuronal subtype-specific changes during aging.** Hypothalamic neurons are highly diverse and function to orchestrate a wide range of processes and behaviors necessary for organismal survival<sup>53</sup>. This diversity of function is accomplished by cell-type-specific gene expression programs, with each area of the hypothalamus containing a range of transcriptionally dissimilar neuronal subtypes<sup>18–23</sup>. Indeed, even neurons expressing the same neuropeptide gene may comprise functionally distinct subpopulations<sup>54</sup>. To address this complexity, we sub-clustered the neuronal nuclei to identify transcriptionally distinct populations. This analysis revealed 35 transcriptionally distinct clusters (Fig. 6a) and broadly separated the nuclei into inhibitory (*Gad1* expressing GABAergic) or excitatory (*Slc17a6/vGLUT2* expressing glutamatergic) identities (Fig. 6b). The 35 clusters represent both known and undefined neuronal subtypes (Supplementary Table 6 shows markers of cluster identity). To discern the relationship between the clusters, we organized them according to transcriptional similarity using a cluster tree analysis (Fig. 6c; left). Neurons with similar functions clustered closely to one another. For example, some AgRP/NPY neurons and POMC neurons may arise from common progenitors<sup>6</sup> and the Sst/Npy (29, expressing *Agrp*) and *Pomc/Tac2* (31) clusters are near to one another on the cluster tree.

We next investigated expression of specific neuropeptide genes across the clusters to functionally define the distinct neuronal subpopulations (Fig. 6c; right). These clusters generally correspond to known cell types expressing one or two hallmark neuropeptides. We were able to identify neuronal clusters expressing genes encoding neuropeptides controlling processes that are altered with age (Supplementary Table 7). For example, we observed notable changes in clusters associated with feeding and energy homeostasis<sup>55</sup>, including those expressing the peptides agouti-related peptide (*Agrp*), cocaine and amphetamine related transcript (*Cartpt*), cholecystokinin (*Cck*), neuropeptide Y (*Npy*), proopiomelanocortin (*Pomc*), galanin (*Gal*) and hypocretin/orexin (*Hcrt*). Based on neuropeptide gene expression, these clusters most likely represent known neuronal populations with defined functions. For example, cluster Sst/Npy (29) most likely consists of AgRP/NPY neurons from the arcuate nucleus of the hypothalamus.

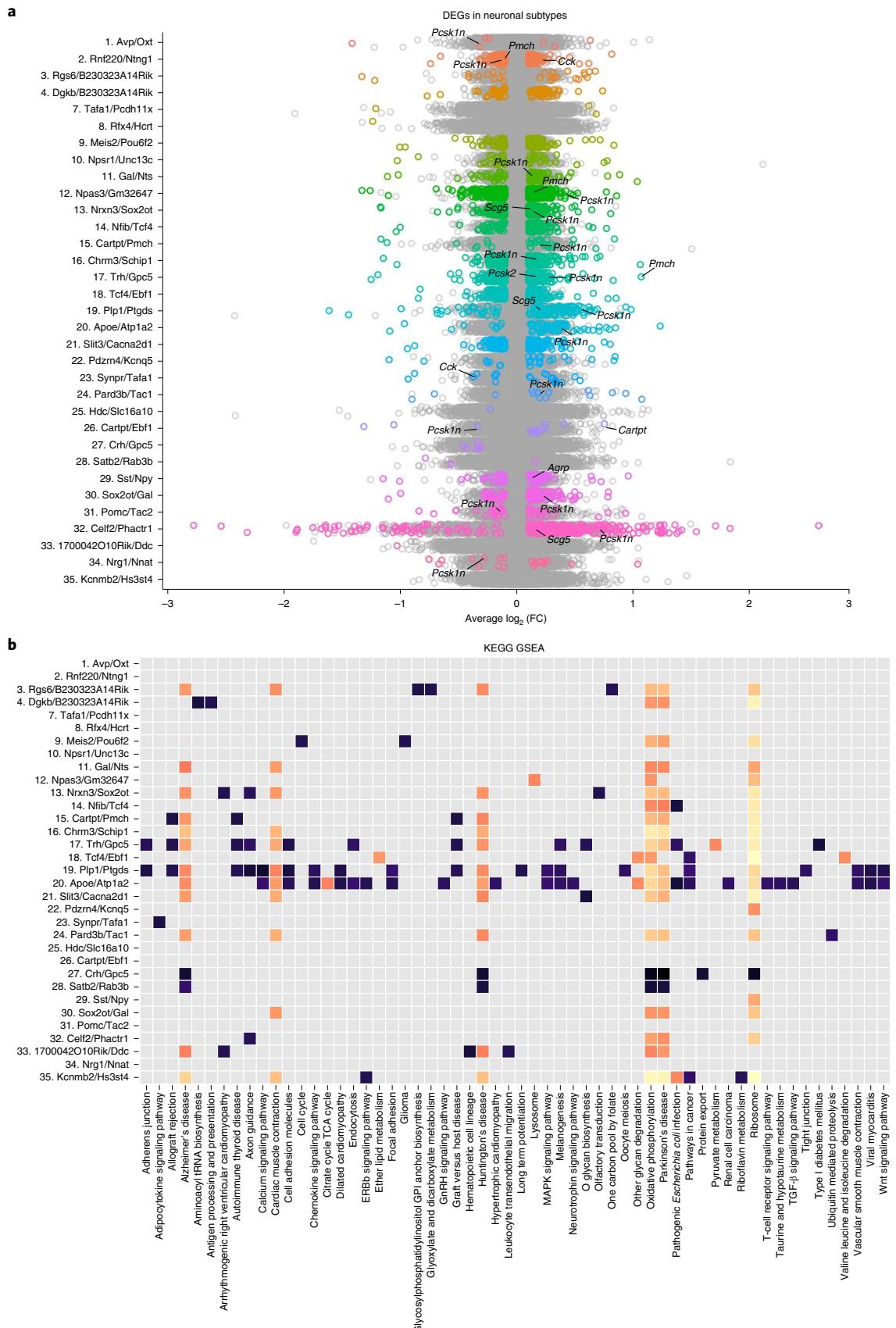
To further confirm neuronal subtype identity, we compared our dataset with publicly available spatial transcriptomic data from Cell2Location<sup>56</sup>. While messenger RNA signatures from broad cluster categories such as astrocytes (Supplementary Fig. 5) do not show restriction to one or more hypothalamic subnuclei, the mRNA signatures of specific neuronal subclusters are localized in discrete locations. (Fig. 6d and Supplementary Fig. 4a,b). For example, the *Pomc/Tac2* (31) cluster localizes to the most ventral portion of the coronal section (Fig. 6d). Notably, two clusters expressing *Cartpt* (15, *Cartpt/Pmch* and 26, *Cartpt/Ebf1*) show little spatial overlap

**Fig. 5 | Alterations to X chromosome inactivation center are a feature of the aged female hypothalamus.** **a**, Expression of genes involved in X chromosome inactivation by age and cell type. Differential expression between young and aged samples was calculated using MAST with random effect for sample of origin and sequencing depth, with Bonferroni adjustment of  $P$  values ( $***P_{\text{adj}}<0.001$ , exact  $P$  values are reported in Supplementary Table 4). **b**, RT-qPCR of *Xist* expression in specific brain regions. *Xist* expression is significantly higher in the hypothalamus ( $n=3$  animals per age group,  $**P=0.008$ , two-sided unpaired Student's  $t$ -test with Bonferroni-Dunn correction, bars indicate mean  $\pm$  s.e.m.). **c**, Representative images from RNAScope for *Xist* transcript in young and aged female hypothalamus (left). Quantification of signal intensity ( $n=3$  animals per group,  $*P=0.0473$ , two-sided unpaired Student's  $t$ -test, bars indicate mean  $\pm$  s.e.m.) (right). **d**, Expression of genes known to escape X chromosome inactivation by age and cell type. Differential expression between young and aged samples was calculated using MAST with random effect for sample of origin and sequencing depth, with Bonferroni adjustment of  $P$  values ( $^*P_{\text{adj}}=.043$ ,  $***P_{\text{adj}}<0.001$ , exact  $P_{\text{adj}}$  values are reported in Supplementary Table 4). **e**, *XIST* expression in human entorhinal cortex (left) and prefrontal cortex (right). ( $***P_{\text{adj}}<0.001$ , MAST with random effect for sample of origin and Bonferroni adjustment of  $P$  values).

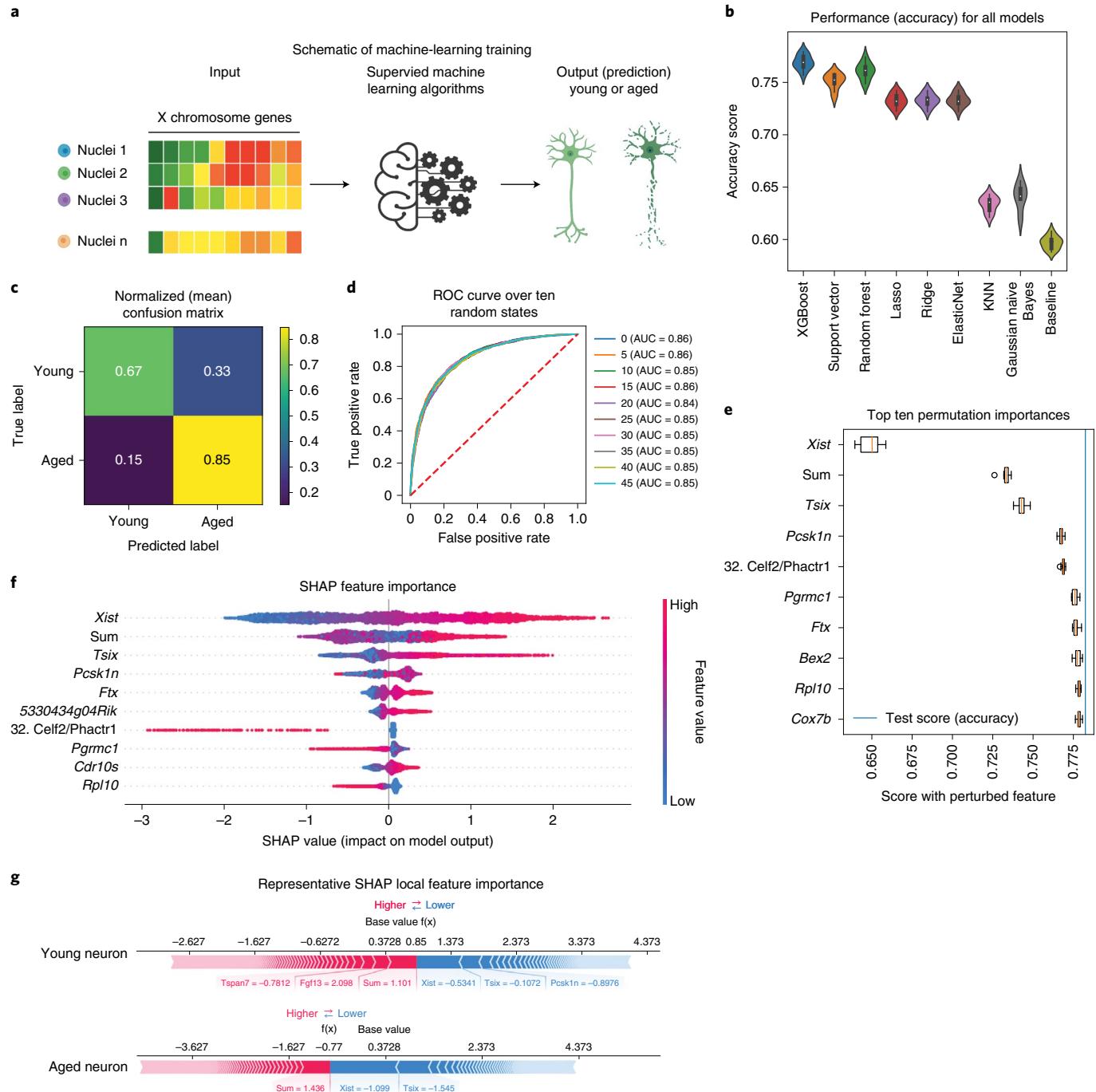




**Fig. 6 | Identification of transcriptionally distinct neuronal subtypes.** **a**, UMAP of all neuronal nuclei. Distinct clusters are identified by color, with identities listed in **c**. **b**, UMAPs highlighting GABAergic (*Gad1*) and glutamatergic (*Slc17a6*) nuclei neuronal clusters. Color scale indicates expression level. **c**, Neuronal clusters are labeled according to the top two marker genes and ordered based on overall transcriptional similarity (left). Expression of neuropeptide genes in each cluster (right). Dot size indicates percent of nuclei expressing the gene and color indicates intensity of expression. **d**, Estimated spatial locations of neuronal subclusters.



**Fig. 7 | Neuronal subtypes exhibit distinct transcriptional changes with age.** **a**, Strip plot showing DEGs per cluster. Significant genes ( $FC > 0.1$ ,  $P_{adj} < 0.05$ ) are colored, nonsignificant genes in gray. Genes discussed in text are labeled. **b**, Heat map of GSEA results for each neuronal cluster. Significantly enriched terms ( $P_{adj} < 0.1$ ) are colored according to the NES.



**Fig. 8 | Xist expression predicts neuronal age in the mouse hypothalamus.** **a**, Schematic of the machine-learning approach. **b**, Violin plots showing model test set performance. Model accuracy across ten repeats with different random states is shown, with Xgbc outperforming the other models. Xgbc was then retrained on new data splits across 50 repeats with different random states. Whisker box plots (inset) show median, lower and upper quartiles and whiskers extending to 1.5 times the interquartile range. Data are presented as mean  $\pm$  s.d. **c,d**, Confusion matrix and ROC curve depicting Xgbc model accuracy across 50 and 10 random states, respectively. **e**, Top ten most important features of the Xgbc model. Permutation was repeated ten times. Whisker box plots show median, lower and upper quartiles and whiskers extending to 1.5 times the interquartile range. Data are presented as mean values  $\pm$  s.d. Note the strong influence of Xist on model accuracy score. **f**, SHAP summary plot showing feature importance for the top ten features that predict cellular age in the model. **g**, SHAP force plot showing the most impactful features on the model prediction for example observations in young and aged neurons.

despite their shared neuropeptide profile, highlighting the strength of this method to define cell types both spatially and transcriptionally. Thus, this spatial analysis further validates the identity and function of the identified neuronal subclusters.

We next performed differential expression on clusters in which there were at least 20 nuclei per condition (Fig. 7a and Supplementary Table 8). For each cluster, we also performed GSEA using the KEGG gene set. Most clusters tested exhibited transcriptional changes with

age, although the number of DEGs varied by subtype. We observed that clusters expressing peptides involved in feeding and energy homeostasis were particularly altered with age in this analysis (such as 15. *Cartpt/Pmch*, 34 DEGs; 29. *Sst/Npy*, 65 DEGs; and 31. *Pomc/Tac2*, 42 DEGs). Among the many transcriptional changes found in the *Pomc/Tac2* cluster (31) there was an intriguing downregulation of *Pcsk1n*. *Pcsk1n* encodes proprotein convertase subtilisin/kexin type 1 inhibitor, also called proSAAS, a propeptide that inhibits processing of other neuropeptides such as POMC<sup>57</sup>. This gene was also downregulated in a cluster of neurons expressing *Cartpt* (*Cartpt/Ebf1* (26)). Notably, in a different *Cartpt*-expressing cluster (*Cartpt/Pmch* (15)), the gene is upregulated with age, suggesting that changes to neuropeptide processing pathways with age are cell-type specific.

Changes to expression of neuropeptide genes was also evident, with upregulation of *Agrp* in cluster *Sst/Npy* (29), upregulation of *Cartpt* in *Cartpt/Ebf1* (26) and downregulation of *Cck* in two *Cck*<sup>+</sup> subclusters (*Synpr/Tafal* (23) and *Rnf220/Ntng1* (2)). Thus, our dataset links neuron-specific gene expression changes in the hypothalamus with key features of organismal aging, such as weight and metabolic changes.

On the basis of expression of specific peptide genes (*Adcyap1*, *Cartpt* and *Cck*) and other established marker genes (*Foxb1* and *Cpne9*)<sup>21</sup>, we identified three clusters representing the medial mammillary nucleus of the hypothalamus: *Rgs6/B230323A14Rik* (3), *Dgkb/B230323A14Rik* (4) and *Slc1a3/Apoe* (5). This region is notable because unlike most areas of the hypothalamus, it is involved in memory via connections with the hippocampus<sup>58</sup>. While cluster *Slc1a3/Apoe* (5) had too few cells to meet our criteria for performing differential expression, both cluster *Rgs6/B230323A14Rik* (3) and *Dgkb/B230323A14Rik* (4) were significantly altered with age (Fig. 7a). GSEA using the KEGG gene set revealed enrichment for genes related to Alzheimer's disease, cardiac muscle contraction, Huntington's disease, oxidative phosphorylation, Parkinson's disease and the ribosome. There was an additional de-enrichment in genes related to glycosylphosphatidylinositol (GPI)-anchor biosynthesis and glyoxylate and dicarboxylate metabolism (Fig. 7b). The identification of changes in this brain region is significant, as they may contribute to cognitive impairments with age.

Through GSEA, a shared aging signature emerged among many hypothalamic neuronal subtypes. This included enrichment in pathways related to Alzheimer's disease (14 clusters), Huntington's disease (11 clusters), oxidative phosphorylation (19 clusters), Parkinson's disease (17 clusters) and the ribosome (21 clusters) (Fig. 7b). A notable exception to this signature is the cluster most likely representing corticotropin-releasing hormone (CRH) neurons of the paraventricular nucleus of the hypothalamus (*Crh/Gpc5* (27)). CRH neurons are an integral component of the hypothalamic–pituitary–adrenal axis in the stress response<sup>59</sup>. Decreased CRH has been studied for several decades as a potential hallmark of Alzheimer's disease<sup>60</sup> and CRH itself has been shown to be neuroprotective against Aβ toxicity<sup>61</sup>. In cluster *Crh/Gpc5* (27) of our dataset, gene sets related to Alzheimer's disease, Huntington's disease, oxidative phosphorylation, Parkinson's disease, protein export and the ribosome were all strongly under-enriched (Fig. 7b), suggesting that this neuronal subtype has a distinct disease-associated expression signature compared to other neurons. Together, these data highlight the hypothalamic transcriptional changes unique to individual neuronal subtypes or common across neurons, which may contribute to age-related neurodegenerative disease.

Finally, we sought to understand the role of *Xist* in defining the aged neuronal state. To do so, we tested whether expression of X chromosome genes was sufficient to predict neuronal age in our dataset (Fig. 8a). We trained eight different supervised machine-learning models to classify neurons as either young or aged. Based on the accuracy score (Fig. 8b), the XGBoost classifier (Xgbc)<sup>62</sup>

outperformed others with  $77.8 \pm 0.6224\%$  accuracy. We then fine-tuned the model to optimize hyperparameters and retrained it on new data splits across 50 random states to measure the uncertainties due to splitting and the non-deterministic model. The confusion matrix (Fig. 8c) and the area under the receiver operating characteristic (ROC) curve (AUC) (Fig. 8d) confirmed model performance. Notably, when we randomly shuffled the feature '*Xist* expression', the model performance dropped dramatically down to the near-baseline level (Fig. 8e). We then applied Shapley additive explanations (SHAP)<sup>63</sup> to further interpret the predictions and rank the features by importance. Consistent with our findings, *Xist* was the most important feature in the prediction, followed by the sum of all genes detected in the dataset (Fig. 8f). Local feature importance of two randomly selected individual neurons (young and aged) also showed that *Xist* had the most significant impact on driving the model prediction (Fig. 8g). These data suggest that *Xist* upregulation is a key feature of hypothalamic neuronal aging and may predict female neuronal aging in the hypothalamus.

While comparing our dataset to hippocampal data, we noted that *Xist* is highly upregulated with age in both brain regions. To understand the role of *Xist* and the X chromosome in aging across brain regions, we tested whether X chromosome genes were sufficient to predict cellular age in the hippocampus. We reran the Xgbc model using snRNA-seq data from female mouse hippocampal neurons (Fig. 3a and Supplementary Fig. 6). Although overall gene expression changes with age in the hypothalamus and hippocampus do not correlate (Fig. 2a), X chromosome gene expression is still sufficient to predict cellular age with  $82.5 \pm 0.7080\%$  accuracy in hippocampal neurons. Interestingly, *Xist* was the second most important predictor of age, based on permutation importance and SHAP (Supplementary Fig. 6c,d), confirming *Xist* upregulation as a shared feature of neuronal aging across two brain regions.

## Discussion

In this work, we used snRNA-seq to identify the age-associated transcriptional changes in the mouse hypothalamus. This brain region is critical for the regulation of physiological homeostasis, including sleep, circadian rhythms, feeding, and metabolism. These functions are well known to be disrupted during aging and our findings implicate widespread transcriptional changes concomitant with physiological changes.

Our approach successfully captured the major cell types in the brain, as well as hypothalamus-specific cell types such as tanycytes. We found that cellular subtypes in this region acquire distinct aging signatures and discovered that increased transcriptional heterogeneity is a common feature across all cell types with age. Consistent with our findings, age-related transcriptional alterations have been observed in aging human brains and increased transcriptional noise is thought to be a hallmark of aging. Our finding that different neuronal subtypes have distinct aging signatures is consistent with recent reports identifying differential susceptibility to neurodegeneration<sup>64</sup>. Identification of the transcriptional signatures involved may pave the way for therapeutics targeted at subpopulations most susceptible to dysregulation with age.

We observed notable changes in the microglial population with age. Microglia are resident immune cells in the brain and previous research has shown that microglia-mediated inflammation in the hypothalamus can affect lifespan<sup>11</sup>. By utilizing trajectory inference analysis, we uncovered that while some aging microglia retain features of young cells, the population shows a progression toward an aged phenotype based on distinct gene expression modules. Notably, DAM genes such as *Apoe* change throughout both age and pseudotime.

Sex differences in aging have been observed across taxa, including in mice<sup>12,13</sup>. In mammals, females generally live longer than males<sup>12</sup> and many aging interventions, such as CR, are

more effective in females<sup>13,14</sup>. In addition, the sexually dimorphic response to aging interventions seems to be a conserved phenomenon, with female *Drosophila* responding more strongly to dietary restriction paradigms than males<sup>65</sup> and hermaphroditic *C. elegans* responding more strongly to dietary restriction than males<sup>66</sup>. In mice, males and females differ in regard to sex chromosome content (males are XY and females are XX) and the presence of gonadal hormones such as higher androgens in males and estrogens in females. Of note, X chromosome content has been linked to longevity and the presence of two X chromosomes contributes to increased longevity regardless of hormonal status<sup>67</sup>. This study from the Dubal laboratory was performed using the four core genomes mouse line, in which the *Sry* gene (which induces testes development) exists on an autosome rather than the Y chromosome, allowing for chromosomal sex to be disambiguated from gonadal sex/hormone status. In our study, we uncover a potential mechanism by which the X chromosome affects aging. We observed widespread upregulation of *Xist* in aged female animals, as well as upregulation of other XIC genes, including *Tsix*, *Jpx*, and *Ftx*. Notably, this increased expression was highly prominent in neurons, although upregulation of *Xist* was observed in oligodendrocytes, astrocytes, and tanycytes as well. In a machine-learning algorithm, *Xist* expression was the most important variable in a model to classify whether a hypothalamic neuron was young or aged. Together, our findings reveal an unexplored feature of aging in the female brain.

Future work will be required to address the limitations of this study. We analyzed only female mice and an additional comparison between male and female mice will be needed to fully elucidate sex-specific mechanisms of aging in the hypothalamus. Further, only two time points were investigated using snRNA-seq. Additional time points throughout the lifespan will be useful in understanding when and how age-related hypothalamic transcriptional changes occur. A main strength of this study was the use of single nuclei rather than cells to analyze a larger population of neurons that may not have otherwise survive dissociation and processing. Despite the large number of neuronal nuclei we captured, some populations of interest, such as kisspeptin-expressing neurons, were not highly represented in the sample populations. Due to the diversity of neuronal subtypes in the hypothalamus, future sequencing of larger amounts or specific populations of nuclei is needed to adequately isolate and analyze rare populations of interest.

In summary, our study reveals the major transcriptional features of hypothalamic aging. We observed transcriptional variation across cell types, cell-type specific aging signatures and features of aging in females. Understanding how individual populations of cells in this region contribute to overall loss of homeostasis with age will be vital to identifying treatments for aging and age-related disease.

## Methods

**Animals.** Single-nuclei isolation. Young (3 month) and aged (19–24 month) C57/Bl6 female mice were obtained from the National Institute on Aging. Mice were housed and used according to protocols approved by Brown University IACUC and in accordance with institutional and national guidelines. Animals were housed at  $70 \pm 2^\circ\text{F}$  with humidity from 50–70%. Animals were fed ad libitum LabDiet 5010 chow and water. The light cycle was 12 h on/12 h off; lights were on between 07:00 and 19:00.

Animals were exposed to male bedding 3 d before being euthanization to synchronize the estrous cycle. Animals were sacrificed at Zeitgeber time ZT2–ZT3.

**Statistics and reproducibility.** No statistical methods were used to predetermine sample sizes but our sample sizes are similar to those reported in previous publications using single-cell RNA-seq on brain tissue<sup>25,36</sup>. No animals were excluded from the study. Individual nuclei were excluded from analysis based on quality control metrics (feature count and mitochondrial read count). Data collection and analysis were not performed blind to the conditions of the experiments, except in the case of the RNAScope image analysis, in which the experiment was performed and analyzed randomized and blinded.

**snRNA-seq.** To reduce noise stemming from differences in estrous state, two whole hypothalamuses were pooled into each biological replicate, for a total of two replicates for the young and aged conditions. Nuclei extraction was performing using the Nuclei PURE Prep kit (Millipore Sigma) according to the manufacturer's instructions with the modifications that for each sample, two hypothalamuses were dissected out of the animals and rinsed in cold PBS. Tissue was transferred using a transfer pipette into a refrigerated Dounce homogenizer with 5 ml lysis solution following kit instructions. Tissue was homogenized with the Dounce B and the lysate was transferred into a 15-ml Falcon tube through a 40-μm filter. The sucrose purification step was performed with the modifications that half the volume of all reagents was used to account for the small tissue sample sizes, an SW34 rotor was used and samples were spun for 45 min at 30,000 (13,000 r.p.m.) at 4°C. Nuclei were counted using a hemocytometer and 5,000 cells per sample were loaded onto the Chromium Single Cell 3' Chip (10X Genomics) and processed with the Chromium Controller (10X Genomics). Samples Young\_1, Young\_2, Aged\_1 and Aged\_2 were prepared using the Chromium Single Cell 3' Library and Gel Bead kit v2 according to manufacturer's instructions. Samples were sequenced at GENEWIZ on an Illumina HiSeq, with a target of 50,000 reads per sample. The Aged\_1 and Young\_2 samples underwent an additional round of sequencing to obtain sufficient read depth. Samples Young\_3, Young\_4, Aged\_3 and Aged\_4 were prepared with the Next GEM Single Cell 3' Reagent kit (10X Genomics) and sequenced at GENEWIZ on an Illumina NovaSeq.

**Quality control, data processing and analysis.** We performed sequence alignment to the mm10 genome (2020) using the CellRanger (CellRanger/6.0.0) software from 10X Genomics with the –include introns flag. The resulting feature–barcode matrices were read into R v.4.1.0, excluding any nuclei expressing fewer than 200 genes and any gene expressed in fewer than three nuclei.

Filtering and visualization were performed using Seurat (v.4.0.3)<sup>68</sup>. For samples sequenced on an Illumina HiSeq, nuclei with fewer than 200 or more than 3,000 features were filtered out. For samples sequenced on the NovaSeq, nuclei with fewer than 200 or more than 7,500 features were filtered out. Similarly, nuclei with greater than 10% mitochondrial mapping were removed, resulting in 23,808 nuclei in the aged condition and 16,256 nuclei in the young condition. Integration of the datasets was performed using the IntegrateData function on 5,000 variable features. The number of nuclei, unique molecular identifiers and unique genes per sample are reported in Supplementary Fig. 1. To assign identities to clusters, the FindAllMarkers() command with default parameters was used. This finds the top genes that define a cluster identity. We named each cluster using the top two genes to come out of the FindAllMarkers() analysis.

Differential expression was performed using MAST (v.1.18.0)<sup>28,29</sup>, with random effect for sequencing depth and sample of origin<sup>69</sup>. Genes were considered significant if the adjusted *P* value was  $<0.05$  and the  $\log_2(\text{FC})$  was  $>0.1$  or  $<-0.1$ . For re-analysis of publicly available data, raw cell/count matrices were downloaded and data were reprocessed according to the above-mentioned workflow. MAST was performed with random effect for sample of origin.

**Gene set enrichment analysis.** GSEA was performed using the fgsea package (v.1.18.0)<sup>36</sup> using the hallmark gene set list and KEGG gene set list from MSigDB (v.7.2.)<sup>69</sup>. For each cluster, genes were ranked by  $\log_2(\text{FC})$  after MAST analysis and the analysis was performed using the fgseaMultilevel command with default settings and seed set at 1,000. Gene sets were considered to be enriched if the adjusted *P* value was  $<0.1$ . Conversions between mouse and human annotation was performed using biomaRt (v.2.48.2).

**Trajectory inference and analysis using Monocle3.** To infer the aging process for the microglia/macrophage clusters ( $n = 1,121$  nuclei) generated in Seurat, we applied Monocle3 (refs. <sup>44,70</sup>). Monocle3 uses dimensionality reduction to place single cells in a two-dimensional space, removes batch effects by mutual nearest neighbor alignment and connects single cells to construct a trajectory in a semi-supervised way. For the microglia/macrophage cluster, we use the integrated Seurat object with no further batch correction or dimensionality reduction in Monocle3. We subsetted the microglia and macrophage cluster and programmatically specified the root of the trajectory by selecting the node most enriched for young cells. The trajectory and its direction calculated by Monocle3 are in agreement with the distribution of young and aged cells. Spatial differential expression analysis along the trajectory was performed with Moran's *I* test in Monocle3 and selected genes with  $q < 0.05$  as trajectory-dependent genes (2,112 genes). The set of genes were grouped into four modules according to its RVAE decoded expression<sup>49</sup> along the trajectory.

**Functional enrichment analysis.** EnrichR<sup>71</sup> v.3.0 was applied to perform the functional enrichment analysis of 2,112 genes in individual modules, resulting in lists ('4\_modules\_q\_moranI') of statistically significant enriched terms ( $P_{\text{adj}} < 0.05$  with Benjamini–Hochberg correction) for individual modules. We checked the gene set databases GO\_Biological\_Process\_2018, GO\_Cellular\_Component\_2018 and GO\_Molecular\_Function\_2018. We kept GO terms with  $P < 0.05$  and visualized the ten most significant terms for each module and visualized in the dot plot. The Python package RVAGene (v.1.0, in Python v.3.9.6 with PyTorch v.1.9.0)

RVAE implementation was used to decode the t-DEGs ( $n=2,112$ ) along the pseudotime trajectory. Expression was averaged in individual time bins and then rescaled to the value in  $[-1,1]$  and input to RVAgene. For the neural network, the parameters used were symmetrical architecture with two hidden layers (48 nodes per layer) and two latent variable dimensions. The output reconstructed trajectory for the t-DEGs was used to plot the heat map.

**Single-cell regulatory network inference and clustering (SCENIC).** RNA counts from samples Young\_3, Young\_4, Aged\_3 and Aged\_4 were exported into a loom file using SCoPeLoomR\_0.11.0. The standard pySCENIC (v.0.11.0) workflow was run using Brown University's cloud computing resource. The workflow was completed 50 times and the resulting loom files were loaded back into R. Only regulons and genes within the regulons appearing 10 of 50 times or more were retained for further analysis. AUCell analysis was performed in R using SCENIC\_1.2.4 and AUCell\_1.14.0. Using the previously defined regulons, AUCell analysis was performed on all cells from the dataset following the default pipeline. For binarization of regulons, default thresholds were used. Regulon specificity scores were generated using the calcRSS() command.

**CellPhoneDB.** RNA counts for young and aged datasets were analyzed separately to allow for comparison. Conversions between mouse and human annotation was performed using biomaRt (v.2.48.2). CellPhoneDB (v.2.1.7) was run in a conda environment (anaconda/2020.02) using the statistical\_analysis method with 1,000 iterations and a 0.1 threshold. For visualization in R, only ligand-receptor pairs in which direction could be inferred were retained for analysis.

**Microglia isolation and RT-qPCR analysis.** Young (2–3 months), middle aged (8–13 months) and aged (20–24 months) C57/BL6 wild-type and POMC-EGFP reporter mice (Jax Stock no. 009593) were housed and used according to protocols approved by Brown University and in accordance with institutional and national guidelines. Animals were sacrificed at ZT4. For each biological replicate, four animals were pooled, with genotypes and estrous state balanced across conditions. Tissue was dissociated with the Adult Brain Dissociation kit (Miltenyi Biotec, 130-107-677) according to manufacturer's instructions. Dissociated tissue was incubated with CD11b MicroBeads (Miltenyi Biotec, 130-049-601) for 15 min at 4°C. Labeled cells were isolated using Miltenyi Biotec MS columns (130-042-201) on the OctoMACS Separator. RNA was purified using the RNeasy micro kit (74004) and complementary DNA was generated with the High-Capacity Reverse Transcription kit (Applied Biosystems, 4374966). A negative control (−RT) for each sample was also generated by excluding the Multiscribe Reverse Transcriptase component of the reaction.

**Cell2location.** Cell2location is a Bayesian model that uses snRNA-seq cell-type signatures to infer cell types in Visium spatial transcriptomics by decomposing mRNA counts in each Visium voxel into cell types. We performed the three main steps in the cell2location workflow: estimate reference expression signatures of cell types using our dataset, map the learned cell type signatures onto the slides and perform downstream analysis. The code, model parameters and training evaluations can be found in the jupyter notebooks in our GitHub repository. In brief, we used the default parameters to train the cell2location model.

**Neuronal age prediction using machine learning.** Neuronal nuclei (25,002) were selected for young or aged classification. All genes were annotated with their chromosomal location. For each neuron, one categorical feature (neuronal subtype) and 281 numerical features were used for machine learning: 278 X chromosome genes (mean expression >0.1 read per cell and detected in >3,000 nuclei), aggregated X chromosome gene expression 'x\_sum', aggregated all gene expression 'sum' and their ratio 'x\_prop'. The pipeline and functions were implemented in Scikit-learn<sup>72</sup>. For data splitting, 20% of nuclei were first split into the testing set and the remaining 80% were further split into training and validation sets using fivefold cross-validation, resulting in train-validation-testing of 64–16–20. For preprocessing, OneHotEncoder was applied for the categorical feature and StandardScaler was applied for the numerical features.

Eight machine-learning models were tested over ten different random states. The best hyperparameters were selected using GridSearchCV and the model performance was evaluated using accuracy score of the test sets. XGBoost classifier<sup>62</sup> was selected, fine-tuned (max\_depth of 5 with early stop) and then retrained on new splits across 50 different random states. The baseline accuracy was  $0.596 \pm 0.00765$  and the model accuracy was  $0.778 \pm 0.006224$ . Model interpretation was performed using permutation feature importance and SHAP<sup>63</sup>.

For the hippocampus dataset, neuronal nuclei (11,204) were selected. For each neuron, 253 X chromosome genes (mean expression >0.1 read per cell and detected in >1,000 nuclei), aggregated X chromosome gene expression 'x\_sum', aggregated all gene expression 'sum' and their ratio 'x\_prop' were used as features for model training. The rest of the processes were the same as above except that the max\_depth of 4 for the final 50 different random states. The baseline accuracy was:  $0.594 \pm 0.009167$  and the model accuracy was:  $0.825 \pm 0.007080$ .

**Whole-brain RNA isolation, cDNA generation and qRT-PCR.** Hypothalamus, olfactory bulb, cerebellum and cortex were dissected in cold PBS from the brains

of 3-month-old and 24-month-old C57BL/6 mice ( $n=6$ ; 3 male and 3 female for each age) and snap-frozen in liquid nitrogen. RNA was purified using the QIAGEN RNeasy Lipid Tissue Mini kit (QIAGEN, 74804). cDNA was generated using 500 ng of RNA and the High-Capacity Reverse Transcription kit (Applied Biosystems, 4374966). A negative control (−RT) for each sample was also generated by excluding the Multiscribe Reverse Transcriptase component of the reaction. qPCR reactions were completed using the PowerUp SYBR Green Master Mix (Invitrogen, A25918). Stock primers were diluted to 10 mM in sterile water and cDNA was diluted 1:5 in sterile water (whole brain) or 1:3 in sterile water (microglia). Expression levels of the genes of interest (see table below) were quantified using a ViiA 7 Real Time PCR System with QuantStudio software. For whole brain, *Actin* was used as a housekeeping gene. For microglia, *Itgam* (CD11b) was used. Each sample, water control and −RT control sample was run in triplicate for each primer set. CT values were normalized to the housekeeping gene and  $\Delta CT$  values were plotted as  $2^{-\Delta CT}$ . Technical replicates were averaged per biological replicate.

Gene	F sequence	R sequence
<i>Actin</i>	TGTTACCAACTGGGACGACA	CTCTCAGCTGTGGTGGTGA
<i>Apoe</i>	GATCAGCTCGAGTGGCAA	CTCTGCAGCTTCCCTGGAC
<i>Arhgap15</i>	AAAGCCAAAATTGCAGATGG	GAGCTTGCTGCTTGGAGTCT
<i>Cst3</i>	CTGACTGCTCTTCATGACC	TGCAGCTGAATTTCAGTCAGG
<i>Cx3cr1</i>	AAGCTCACGACTGCCTCTT	CCGGTTGTTCATGGAGTTGG
<i>Itgam</i> ( <i>CD11b</i> )	CTTCTGGTCACAGCCTAGC	TGGACCAACTGTCACAA
<i>Lyz2</i>	ACTGCTCAGGCCAAGGTCTA	TGCTCTCGTGCTGAGCTAAA
<i>Xist</i>	GGTTCTCTCCAGAACG	TGGTAGATGGCATTGTG TAGGAAG

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Fastq files for raw snRNA-seq and Seurat objects were deposited at the Gene Expression Omnibus under accession code GSE188646. Publicly available datasets are available for hippocampal single-nuclei data, GSE161340; human entorhinal cortex data, GSE138852 (samples AD3-AD4 and Ct1-Ct2); human prefrontal cortex data, GSE174367 (samples 17, 19, 37, 43, 45, 50, 66 and 90); and Spatial data, ArrayExpress E-MTAB-1114.

## Code availability

All code is available at [https://github.com/Webb-Laboratory/Hajdarovic\\_And\\_Yu\\_et\\_al\\_2022](https://github.com/Webb-Laboratory/Hajdarovic_And_Yu_et_al_2022).

Received: 6 March 2021; Accepted: 2 June 2022;

Published online: 04 July 2022

## References

- Partridge, L., Deelen, J. & Slagboom, P. E. Facing up to the global challenges of ageing. *Nature* **561**, 45–56 (2018).
- Yousufuddin, M. & Young, N. Aging and ischemic stroke. *Aging* **11**, 2542–2544 (2019).
- Baldacci, L. & Aapro, M. in *Biological Basis of Geriatric Oncology* 1–15 (Springer US, 2005).
- Farooqui, T. & Farooqui, A. A. Aging: an important factor for the pathogenesis of neurodegenerative diseases. *Mech. Ageing Dev.* **130**, 203–215 (2009).
- Chahal, H. & Drake, W. The endocrine system and ageing. *J. Pathol.* **211**, 173–180 (2007).
- Toda, C., Santoro, A., Kim, J. D. & Diano, S. POMC neurons: from birth to death. *Annu. Rev. Physiol.* **79**, 209–236 (2017).
- de Cabo, R., Carmona-Gutierrez, D., Bernier, M., Hall, M. N. & Madeo, F. The search for antiaging interventions: from elixirs to fasting regimens. *Cell* **157**, 1515–1526 (2014).
- Bishop, N. A. & Guarente, L. Two neurons mediate diet-restriction-induced longevity in *C. elegans*. *Nature* **447**, 545–549 (2007).
- Broughton, S. J. et al. DILP-producing median neurosecretory cells in the drosophila brain mediate the response of lifespan to dietary restriction. *Aging Cell* **9**, 336–346 (2010).
- Satoh, A. et al. Sirt1 extends life span and delays aging in mice through the regulation of Nk2 homeobox 1 in the DMH and LH. *Cell Metab.* **18**, 416–430 (2013).
- Zhang, G. et al. Hypothalamic programming of systemic ageing involving IKK-β, NF-κB and GnRH. *Nature* **497**, 211–216 (2013).
- Lemaître, J.-F. et al. Sex differences in adult lifespan and aging rates of mortality across wild mammals. *Proc. Natl Acad. Sci. USA* **117**, 8546–8553 (2020).

13. Austad, S. N. & Bartke, A. Sex differences in longevity and in responses to anti-aging interventions: a mini-review. *Gerontology* **62**, 40–46 (2016).
14. Mitchell, S. J. et al. Effects of sex, strain, and energy intake on hallmarks of aging in mice. *Cell Metab.* **23**, 1093–1112 (2016).
15. Webb, A. E., Kundaje, A. & Brunet, A. Characterization of the direct targets of FOXO transcription factors throughout evolution. *Aging Cell* **15**, 673–685 (2016).
16. Benayoun, B. A. et al. Remodeling of epigenome and transcriptome landscapes with aging in mice reveals widespread induction of inflammatory responses. *Genome Res.* **29**, 697–709 (2019).
17. Hofmann, J. W. et al. Reduced expression of MYC increases longevity and enhances healthspan. *Cell* **160**, 477–488 (2015).
18. Chen, R., Wu, X., Jiang, L. & Zhang, Y. Single-cell RNA-seq reveals hypothalamic cell diversity. *Cell Rep.* **18**, 3227–3241 (2017).
19. Campbell, J. N. et al. A molecular census of arcuate hypothalamus and median eminence cell types. *Nat. Neurosci.* **20**, 484–496 (2017).
20. Romanov, R. A. et al. Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. *Nat. Neurosci.* **20**, 176–188 (2017).
21. Mickelsen, L. E. et al. Cellular taxonomy and spatial organization of the murine ventral posterior hypothalamus. *eLife* **9**, e58901 (2020).
22. Mickelsen, L. E. et al. Single-cell transcriptomic analysis of the lateral hypothalamic area reveals molecularly distinct populations of inhibitory and excitatory neurons. *Nat. Neurosci.* **22**, 642–656 (2019).
23. Kim, D. W. et al. The cellular and molecular landscape of hypothalamic patterning and differentiation from embryonic to late postnatal development. *Nat. Commun.* **11**, 4360 (2020).
24. Wen, S. et al. Spatiotemporal single-cell analysis of gene expression in the mouse suprachiasmatic nucleus. *Nat. Neurosci.* **23**, 456–467 (2020).
25. Bakken, T. E. et al. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS ONE* **13**, e0209648 (2018).
26. Ding, J. et al. Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nat. Biotechnol.* **38**, 737–746 (2020).
27. Van de Sande, B. et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* **15**, 2247–2276 (2020).
28. Finak, G. et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278 (2015).
29. Zimmerman, K. D., Espeland, M. A. & Langefeld, C. D. A practical solution to pseudoreplication bias in single-cell studies. *Nat. Commun.* **12**, 738 (2021).
30. Jiang, C. H., Tsien, J. Z., Schultz, P. G. & Hu, Y. The effects of aging on gene expression in the hypothalamus and cortex of mice. *Proc. Natl Acad. Sci. USA* **98**, 1930–1934 (2001).
31. Arsenijevic, Y., Dreifuss, J. J., Vallet, P., Marguerat, A. & Tribollet, E. Reduced binding of oxytocin in the rat brain during aging. *Brain Res.* **698**, 275–279 (1995).
32. Brockdorff, N. & Duthie, S. M. X chromosome inactivation and the Xist gene. *Cell. Mol. Life Sci.* **54**, 104–112 (1998).
33. Lee, J. T., Davidow, L. S. & Warshawsky, D. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat. Genet.* **21**, 400–404 (1999).
34. Ogrodnik, M. et al. Whole-body senescent cell clearance alleviates age-related brain inflammation and cognitive impairment in mice. *Aging Cell* **20**, e13296 (2021).
35. Ximerakis, M. et al. Single-cell transcriptomic profiling of the aging mouse brain. *Nat. Neurosci.* **22**, 1696–1708 (2019).
36. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at *bioRxiv* <https://doi.org/10.1101/060012> (2021).
37. Boisvert, M. M., Erikson, G. A., Shokhirev, M. N. & Allen, N. J. The aging astrocyte transcriptome from multiple regions of the mouse brain. *Cell Rep.* **22**, 269–285 (2018).
38. Koseoglu, M. M., Norambuena, A., Sharlow, E. R., Lazo, J. S. & Bloom, G. S. Aberrant neuronal cell cycle re-entry: the pathological confluence of Alzheimer's disease and brain insulin resistance, and its relation to cancer. *J. Alzheimers Dis.* **67**, 1–11 (2019).
39. González-García, I. et al. mTOR signaling in the arcuate nucleus of the hypothalamus mediates the anorectic action of estradiol. *J. Endocrinol.* **238**, 177–186 (2018).
40. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
41. Reuss, B. & von Bohlen und Halbach, O. Fibroblast growth factors and their receptors in the central nervous system. *Cell Tissue Res.* **313**, 139–157 (2003).
42. Baldwin, K. T. & Eroglu, C. Molecular mechanisms of astrocyte-induced synaptogenesis. *Curr. Opin. Neurobiol.* **45**, 113–120 (2017).
43. Luo, X.-G., Ding, J.-Q. & Chen, S.-D. Microglia in the aging brain: relevance to neurodegeneration. *Mol. Neurodegener.* **5**, 12 (2010).
44. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
45. Deczkowska, A. et al. Disease-associated microglia: a universal immune sensor of neurodegeneration. *Cell* **173**, 1073–1081 (2018).
46. Keren-Shaul, H. et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell* **169**, 1276–1290 (2017).
47. Mitra, R. & MacLean, A. L. RVAGene: generative modeling of gene expression time series data. *Bioinformatics* **37**, 3252–3262 (2021).
48. Chureau, C. et al. Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. *Hum. Mol. Genet.* **20**, 705–718 (2011).
49. Berleth, J. B. et al. Escape from X inactivation varies in mouse tissues. *PLoS Genet.* <https://doi.org/10.1371/journal.pgen.1005079> (2015).
50. Tukiainen, T. et al. Landscape of X chromosome inactivation across human tissues. *Nature* **550**, 244–248 (2017).
51. Grubman, A. et al. A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. *Nat. Neurosci.* **22**, 2087–2097 (2019).
52. Morabito, S. et al. Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nat. Genet.* **53**, 1143–1155 (2021).
53. Sternson, S. M. Hypothalamic survival circuits: blueprints for purposive behaviors. *Neuron* **77**, 810–824 (2013).
54. Lam, B. Y. H. et al. Heterogeneity of hypothalamic pro-opiomelanocortin-expressing neurons revealed by single-cell RNA sequencing. *Mol. Metab.* **6**, 383–392 (2017).
55. Leibowitz, S. F. & Wortley, K. E. Hypothalamic control of energy balance: different peptides, different functions. *Peptides* **25**, 473–504 (2004).
56. Kleshchevnikov, V. et al. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-021-01139-4> (2022).
57. Fricker, L. D. et al. Identification and characterization of proSAAS, a granin-like neuroendocrine peptide precursor that inhibits prohormone processing. *J. Neurosci.* **20**, 639–648 (2000).
58. Vann, S. D. & Nelson, A. J. D. in *Progress in Brain Research* Vol. 219 163–185 (Elsevier, 2015).
59. Pedersen, W. A. et al. Corticotropin-releasing hormone protects neurons against insults relevant to the pathogenesis of Alzheimer's disease. *Neurobiol. Dis.* **8**, 492–503 (2001).
60. Mouradian, M. M. et al. Spinal fluid CRF reduction in Alzheimer's disease. *Neuropeptides* **8**, 393–400 (1986).
61. Bayatti, N. & Behl, C. The neuroprotective actions of corticotropin releasing hormone. *Ageing Res. Rev.* **4**, 258–270 (2005).
62. Chen, T., He, T., Benesty, M., Khotilovich, V. & Tang, Y. XGBoost—Introduction to Boosted Trees. XGBoost, <https://xgboost.readthedocs.io/en/latest/tutorials/model.html> (2017).
63. Lundberg, S. & Lee, S.-I. A unified approach to interpreting model predictions. Preprint at *ArXiv* <https://doi.org/10.48550/arXiv.1705.07874> (2017).
64. Leng, K. et al. Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. *Nat. Neurosci.* **24**, 276–287 (2021).
65. Burger, J. M. S. Sex-specific effects of interventions that extend fly life span. *Sci. Aging Knowl. Environ.* **28**, 30 (2004).
66. Honjoh, S., Ihara, A., Kajiwara, Y., Yamamoto, T. & Nishida, E. The sexual dimorphism of dietary restriction responsiveness in *Caenorhabditis elegans*. *Cell Rep.* **21**, 3646–3652 (2017).
67. Davis, E. J., Lobach, I. & Dubal, D. B. Female XX sex chromosomes increase survival and extend lifespan in aging mice. *Aging Cell* **18**, e12871 (2019).
68. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).
69. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
70. Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
71. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
72. Pedregos, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

## Acknowledgements

This work was supported by a Carney Institute for Brain Science Zimmerman Innovation Award and National Institutes of Health/NIA R21AG070527 to A.W. The N.N. lab was supported by grant R01AG050582 from the NIH/NIA. K.H. was funded by a Neustein Graduate Fellowship from the Carney Institute for Brain Science at Brown University. D.Y. was funded by the Open Graduate Education from Brown Graduate School. This research was conducted using computational resources and services at the Center for Computation and Visualization, Brown University.

## Author contributions

K.H. and A.W. designed the study. K.H., D.Y., S.-E. and L.H. performed snRNA-seq experiments. K.H. and D.Y. processed and analyzed the snRNA-seq data. K.H. performed microglia isolation and RT-qPCR. L.H. performed RT-qPCR and RNAscope experiments and analyzed the data. S.P. performed data analysis of the snRNA-seq

dataset. S.E. assisted with data analysis of the hippocampal snRNA-seq dataset. D.Y. performed the machine-learning analysis. K.H., D.Y. and A.W. wrote and edited the manuscript. N.N. supervised hippocampal snRNA-seq data collection and analysis. A.W. supervised the entire study. K.H and D.Y. contributed equally to this work.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s43587-022-00246-4>.

**Correspondence and requests for materials** should be addressed to Ashley E. Webb.

**Peer review information** *Nature Aging* thanks Han Kyoung Choe and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection No software was used to collect the data in this study.

Data analysis Analysis used R version 4.1.0 and Python 3.7.10. We performed sequence alignment to the mm10 genome (2020) using the CellRanger (cellranger/6.0.0) software from 10x Genomics with the –include introns flag. The resulting feature-barcode matrices were read into R version 4.1.0, excluding any nuclei expressing fewer than 200 genes, and any gene expressed in fewer than three nuclei. The following R packages are used in this study: Seurat (4.0.3), MAST(1.18.0), fgsea(1.18.0), biomart (2.48.2), monocle3\_1.0.0, enrichR (3.0) SCopeLoomR\_0.11.0, SCENIC\_1.2.4 , AUCell\_1.14.0, CellPhoneDB (version 2.1.7) Python packages: RVAgene (version 1.0, in python version 3.9.6 with PyTorch version 1.9.0), Cell2location (0.05), Scikit-learn (0.24.2). Statistics for RT-qPCR experiments were performed using Graphpad Prism (Prism 9.3.1).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Fastq files for raw single nuclei RNA sequencing and Seurat object were deposited at GEO accession GSE188646. Publicly available datasets are available on GEO:

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications(25,36).
Data exclusions	Our data were collected at two different points in time, using two different sample preparation kits and sequenced on difference sequencers. Because of differences in sequencing depth, we use different data exclusion parameters for the samples. Samples sequenced on an Illumina HiSeq, nuclei with fewer than 200 or more than 3000 features were filtered out. Nuclei with less than 200 features are likely to be artefacts of sample preparation, and nuclei above 3000 features are likely to be doublets, as these nuclei contain more than double the average number of features of nuclei in these samples. For samples sequenced on the NovaSeq, nuclei with fewer than 200 or more than 7500 features were filtered out, following the same logic as above. Similarly, nuclei with greater than 10% mitochondrial mapping were removed as these are likely to represent cells that were dead or dying during sample preparation. This resulted in 23,808 nuclei in the aged condition, and 16,256 nuclei in the young condition.
Replication	The single cell sequencing experiments were performed in two batches using two different sample preparation kits and sequenced on difference sequencers. While our initial batch was sequenced at a lower depth, all findings (DE genes) remained significant when a second cohort of animals was added to the analysis.
Randomization	It was not possible to allocate mice randomly to the experimental condition (aging) due to the nature of the study, as aged and young mice needed to be sacrificed and processed at the same time to avoid batch effects.
Blinding	For the RNA-scope experiments, the experimenter was blinded to the age of animals from which the sections were obtained. All image analysis was performed blinded. Blinding was not possible for data collection for other experiments. Data were collected in batches of four biological replicates for the snRNA-seq, and processed in an alternating fashion (young sample, aged sample, young sample, aged sample) to ensure that lysis timing, incubation timing, and time from animal sacrifice to sample collection was balanced across the two conditions.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input checked="" type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology and archaeology
<input type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	Human research participants
<input checked="" type="checkbox"/>	Clinical data
<input checked="" type="checkbox"/>	Dual use research of concern

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Mouse (*Mus musculus*); C57/Bl6 females; 3 and 19-24 months

Wild animals

No wild animals were used in this study

Field-collected samples

No field collected samples were used in this study

Note that full information on the approval of the study protocol must also be provided in the manuscript.