Multi-omics delineate growth factor network underlying exercise effects in an Alzheimer's mouse model

Xin Li[1,*], Chaozhong Liu[2,*], Wenbo Li[1,*], Yanwan Dai[2], Chaohao Gu[2], Wenjun Zhou[1], Veronica C. Ciliberto[1], Jing Liang[1,3], Udhaya Kumar. S[1], Dongyin Guan[1], Zhaoyong Hu[4], Hui Zheng[5], Hu Chen[2], Zhandong Liu[2], Ying-Wooi Wan[2,#], Zheng Sun[1,6,#]

[1] Department of Medicine – Endocrinology, Diabetes, and Metabolism, Baylor College of Medicine, Houston, Texas 77030, USA
[2] Department of Pediatrics, Jan and Dan Duncan Neurological Research Institute, Baylor College of Medicine, Houston, Texas 77030, USA
[3] Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Peking University Health Science Center, Beijing 100191, China
[4] Department of Medicine – Nephrology, Baylor College of Medicine, Houston, Texas 77030, USA
[5] Huffington Center on Aging, Baylor College of Medicine, Houston, Texas 77030, USA
[6] Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas77030, USA

* Equal contributions
[#] Correspondence yingwooi@gmail.com and zheng.sun@bcm.edu

ABSTRACT

Physical exercise represents a primary defense against age-related cognitive decline and neurodegenerative disorders like Alzheimer's disease (AD). To impartially investigate the underlying mechanisms, we conducted single-nucleus transcriptomic and chromatin accessibility analyses (snRNA-seq and ATAC-seq) on the hippocampus of mice carrying AD-linked NL-G-F mutations in the amyloid precursor protein gene (APP^{NL-G-F}) following prolonged voluntary wheel-running exercise. Our study reveals that exercise mitigates amyloid-induced changes in both transcriptomic expression and chromatin accessibility through cell type-specific transcriptional regulatory networks. These networks converge on the activation of growth factor signaling pathways, particularly the epidermal growth factor receptor (EGFR) and insulin signaling, correlating with an increased proportion of immature dentate granule cells and oligodendrocytes. Notably, the beneficial effects of exercise on neurocognitive functions can be blocked by pharmacological inhibition of EGFR and the downstream phosphoinositide 3-kinases (PI3K). Furthermore, exercise leads to elevated levels of heparin-binding EGF (HB-EGF) in the blood, and intranasal administration of HB-EGF enhances memory function in sedentary APP^{NL-G-F} mice. These findings offer a panoramic delineation of cell type-specific hippocampal transcriptional networks activated by exercise and suggest EGF-related growth factor signaling as a druggable contributor to exercise-induced memory enhancement, thereby suggesting therapeutic avenues for combatting AD-related cognitive decline.

INTRODUCTION

The beneficial effects of physical exercise on neurocognition are widely observed in human patients with Alzheimer's disease (AD) and animal models[1]. The underlying mechanisms are not completely understood and likely multifaceted through a combination of metabolic, endocrine, immunological, and neuronal changes[2,3]. Transcriptional regulation is particularly interesting because it is a general component of signaling pathways elicited by metabolic, hormonal, or neuronal cues, which is implicated in neurodegenerative diseases such as AD[4]. In addition, transcriptional regulation has a relatively long-lasting effect compared to allosteric regulation or posttranslational modifications of proteins, which is temporally in keeping with the long-lasting effects of physical exercise on neurocognition[5]. Considering that the hippocampus controls memory and that memory dysfunction is a hallmark feature of cognitive decline in AD, we performed single-nucleus multi-omics analyses

55 of the hippocampus to provide a panoramic view of the transcriptomic and chromatin accessibility
56 responses to long-term physical exercise.
57
58 **METHODS**
59
60 **Mice**
61 C57BL/6J wild-type (WT) and APP^{NL-G-F} mice were housed in standard 12 h light/ 12 h dark
62 conditions. For wheel-running, mice were put into cages with a running wheel. The sedentary
63 control group was put into the same type of cages but with a locked wheel. Two mice were housed
64 per cage to avoid social isolation except when mice must be separated and singly housed
65 occasionally due to bullying or fighting. The wheel-running activity was monitored in real time by the
66 Actimetrics ClockLab data collection system. Gefitinib (AdooQ Bioscience A10422) and Wortmannin
67 (AdooQ Bioscience A11161) were dissolved in DMSO and diluted in saline, followed by oral gavage
68 at 50 mg/kg and 0.5 mg/kg, respectively, once every other day. Recombinant mouse HB-EGF
69 protein (Novus NBP2-35069) was dissolved in saline and administered intranasally in awake mice
70 at 3 ug/mouse (around 100 ug/kg) with pipettor at about 3 ul per nostril with alternating rest periods
71 and a total administration volume of 10 ul per day[6]. The sex, age, and duration of the running were
72 indicated in figure legends for each experiment. Animal protocols were approved by the Institutional
73 Animal Care & Use Committee (IACUC) at Baylor College of Medicine.
74
75 **Behavioral tests**
76 All the behavior tests were carried out between 12 PM and 7 PM in a dim light environment (300
77 lumens) except the light-dark test. For the object-in-place test, the task comprised an acquisition
78 and a test phase separated by a 24 h delay. Each mouse was habituated in an arena
79 (40 cm × 40 cm × 30 cm) without objects for 5 min to minimize confounding anxiety and
80 novelty factors. In the acquisition phase, each subject was placed in the center of the arena, which
81 contained 4 different objects in the corners 10 cm away from the walls. Mice were allowed to
82 investigate the objects for 5 min and returned to the home cage. All objects were cleaned with
83 alcohol between each mouse. In the test phase, each subject was replaced in the arena with
84 swapped positions for 2 objects, and the subjects were allowed to investigate the objects for 5
85 minutes. Swapped objects were randomly chosen for each mouse. The discrimination index was
86 calculated as (total exploration time on the novel objects – total exploration time on the familiar
87 objects) / (total exploration time on the novel objects + total exploration time on the familiar
88 objects). Considering that normal aging is probably associated with only a mild cognitive decline, we
89 used a modified object in-place test to increase the difficulty of the test and, therefore, to expose the
90 cognitive difference between the exercise group and the control group in the wild-type mice[7]. Briefly,
91 the mouse was allowed to freely explore for 10 min in an arena with 5 Lego-built objects of different
92 shapes and colors (sample phase). After 24 h, two of the objects were relocated. The mouse was
93 re-introduced to the arena and was allowed to explore for 5 min (choice phase). Mouse behaviors
94 were videotaped, and the discrimination index in the choice phase was calculated as the ratio of the
95 time spent exploring the objects with the new location versus the total time spent exploring any
96 object.
97
98 Novel object recognition (NOR) test was performed as previously described [8]. Briefly, mice were
99 habituated in an arena (22 cm x 44 cm) for 5 min. During the training session, two identical objects
100 built from Lego were placed on the right and left sides of the arena. During the first day, mice were
101 placed in the center of the arena and allowed to explore freely for 5 min. Mice were returned to
102 their home cages. After 24 h, on the test day, one of the objects was replaced by a novel object
103 (built from Lego) with a different color and shape, and mice were allowed to explore the arena for
104 5 min. Animal behavior during the training and test session was tracked by a top camera and
105 analyzed by ANY-maze software (Stoelting). The discrimination index was calculated as
106 (exploration time on the novel object - exploration time on the familiar object) / (exploration time on
107 the novel object + exploration time on the familiar object). Exploration behaviors were defined as
108 sniffing or touching (>1 s) the objects while looking at the objects.

109

110 The Y-Maze short-term spatial memory was measured according to the established procedure [9].
111 The Y-maze was made of white opaque acrylic and had three arms (height: 20 cm; length: 30 cm;
112 width: 8 cm) 120 degrees apart. Spatial cues were placed on the walls in line with the axes of the
113 familiar and novel arms. The paradigm consisted of a 5-minute encoding trial during which one arm
114 was blocked off, followed by a 1-h intertrial interval, then a 5-minute retrieval session. The start arm
115 remained the same in both the encoding and retrieval trials, while the exposed arm during the
116 encoding trial was considered the familiar arm, and the blocked arm was considered the novel arm.
117 The apparatus was wiped with 70% ethanol and dried between each mouse to remove odor cues.
118 The discrimination index was calculated as the comparison between the times spent in the novel
119 and familiar arms during the retrieval paradigm.

120

121 The social interaction and social memory test was performed in a three-chamber apparatus (40.5cm
122 x 60 cm x 22 cm) that had three chambers (left, center, and right) of equal size with 10 × 5 cm
123 openings between the chambers. Mice were given 5 min habituation in the chamber and two
124 consecutive 10 min tests: the first test measured sociability by subjecting the mouse to an intruder
125 under one mesh pencil cup and an empty pencil cup, and the second test measured social novelty
126 by subjecting mice to a novel intruder under the empty pencil cup. A camera and the ANY-maze
127 software program were used to track the mouse in the three-chambered box while the experimenter
128 scored the approaches to the object or partner mouse using a wireless keyboard. Intruders (sex-,
129 age- and weight-matched) habituated to the mesh pencil cups in the apparatus for 1 h per day for 2
130 days before testing. Intruder mice were used up to three times, with one test per day.

131

132 The Morris water maze (MWM) test was performed as described previously [8]. The MWM was
133 virtually divided into four quadrants. During the training session, a transparent rescue platform was
134 submerged under the painted water (0.5 cm −1 cm) and placed in a fixed position between the
135 south and east quadrants of the pool. On the first day of training, mice were first allowed to stand on
136 the platform for 10 s. After that, mice were gently placed into the water facing the wall and allowed
137 to explore for 1 min. Mice were then guided to the rescue platform if they did not find it. Mice were
138 allowed to take a rest on the platform for 10 s, and then re-trained from a different start position
139 with the same procedure. After four training trials, they were dried using a paper towel and returned
140 to home cages. Twenty-four hours later, mice were trained again following the same procedure
141 without the initial habituation session. Mice were trained for five consecutive days. At the end of the
142 fourth trial on day 5, mice were returned to home cages for a rest. One hour later, mice were put
143 into the water maze from the west quadrant and let the mice explore the water maze for 1 min,
144 where the platform had been removed. Mouse behaviors were videotaped and analyzed by the
145 Noldus EthoVision XT. The rescue platform was located in the target quadrant. Mouse memory is
146 evaluated by the escape latency and percentage of time spent in the target quadrant. Escape
147 latency was defined as the time spent before finding the platform. Escape latency during the 5-day
148 training sessions served as an independent measurement of spatial learning and memory.

149

150 The open-field arena (OPA) test was performed using the Versamax animal activity monitor
151 equipped with infrared photo beams as horizontal *X-Y* sensors and/or *Z* sensors. Mice were placed
152 in the center of the open-field arena (40 cm × 40 cm × 30 cm) and allowed to explore for
153 40 min. The locomotor activity and location of the mice were scored automatically by VersaMax
154 software. The percentage of time spent in the center area measures anxiety levels. For the elevated
155 plus maze (EPM) test, we used a plus-shaped platform that was elevated to 40 cm above the floor.
156 Two opposite arms of the maze were walled (15 cm high), whereas the other two arms were open
157 with a 5 mm high ridge to prevent falling. Each arm was 8 cm wide and 25 cm long. The test
158 lasted for 10 min and was started by placing a mouse in the center part of the maze, facing one of
159 the two open arms under a dim environment (300 lumens). An overhead camera and the ANY-maze
160 software program were used to track the mouse. The time spent in the open arms was used as a
161 measure for anxiety. The light-dark (LD) test was performed in a box developed from the open field
162 chamber by placing a dark chamber occupying one-third of the open field box. The light area was

163 connected by a small opening to allow mice to move from one area to the other. The test lasted for
164 10 min in a bright environment and started by placing a mouse in the bright area. The activity and
165 location of the mouse were scored automatically by VersaMax software. The number of transitions
166 between dark and light zones and the time spent in the light and dark areas were the index for
167 anxiety.
168
169 **Histology, RNAscope, and ELISA**
170 Mice were anesthetized with isoflurane (3-4% for induction, 1.5-2.5% for maintenance) for
171 transcardiac perfusion with cold PBS and 4% paraformaldehyde. Overnight-fixed brains were
172 immersed in 30% sucrose, embedded in the optimal cutting temperature (OCT) compound, and
173 frozen in isopentane in dry ice. Coronal brain sections (30 μm) were prepared on the Leica
174 CM1850 cryostat slicer. The coronal sections were collected in cryoprotectant solution (25%
175 glycerol, 25% ethylene glycol, 50% PBS pH 7.4). Anti-β-Amyloid (6E10) antibodies (Biolegend,
176 803001;1:500), Fluor 488 goat anti-mouse IgG(H+L) (Life Technologies, A11029; 1:1,000) were
177 diluted in TBS blocking buffer separately before use. Antigen retrieval was performed for 6E10
178 antibody by formic acid treatment (90%formic acid for 5 min for 5 min at room temperature (25 °C)).
179 Brain sections were incubated with the primary antibodies at 4 °C overnight. Sections were then
180 washed three times in TBS at room temperature and incubated further with fluorescence-
181 conjugated secondary antibodies for 1 h at room temperature. These sections were then washed
182 three times in TBS at room temperature, mounted with DAPI Fluoromount-G (SouthernBiotech,
183 0100-20), and sealed with the coverslip. Immunofluorescence of brain sections was viewed and
184 captured with the Zeiss Axio imager.M2m microscope (Axiovision 4.8) and processed by ImageJ
185 software (v 1.53e). The immunoreactive areas were quantified using ImageJ as previously
186 described[10]. The average data of at least three sections per mouse was used to reduce the variance
187 among tissue sections. A two-way ANOVA followed by a post-hoc Fisher's LSD test was used to
188 analyze the quantification data.
189
190 For RNAscope analysis, mice were anesthetized for transcardiac perfusion with cold PBS and 4%
191 paraformaldehyde. Overnight post-fixed brains were immersed in 30% sucrose, embedded in OCT,
192 and frozen in precooled isopentane. Coronal brain sections (around 12 μm) were prepared on the
193 Leica CM1850 cryostat slicer. The coronal sections were collected. Stxbp1 was assayed using
194 RNAscope following the standard protocol from ACD with minor modifications. In brief, brain
195 sections were rinsed with PBS to remove OCT. The brain sections were incubated at 60 °C for 30
196 min. Then, the brain sections were post-fixed in 4% PFA at 4 °C for 15 min. After the post-fixation,
197 the brain sections were dried in ethanol. The brain sections were then incubated with hydrogen
198 peroxide at room temperature for 10 min. The sections were rinsed for 2 min three times in distilled
199 water, and then the brain sections were retrieved in RNAscope 1× target retrieval reagent at
200 100 °C for 5 min. The slides were then rinsed in distilled water for 2 min three times and re-dried in
201 100% alcohol for storage. The pretreated brain sections were incubated with protease III for 30 min
202 at 40 °C. The protease III was removed, and the brain sections were rinsed in distilled water for 2
203 min three times. The brain sections were hybridized with the probe of Stxbp1 (ACD) for 2 h at
204 40 °C. After that, the brain sections were rinsed for 2 min three times in the wash buffer to remove
205 the excessive probes. The RNAscope Multiplex FL v2 Amp1 was added to the brain sections and
206 incubated at 40 °C to amplify the signal for one probe. The brain sections were rinsed with wash
207 buffer after 30 min. The probe signals were detected using the RNAscope Multiplex Fluorescent
208 Detection Reagents V2 (ACD 323110). Brain sections were treated by the Multiplex FLV2 HRP
209 blocker and washed in PBS for 2 times. Brain sections are viewed and captured with the Zeiss Axio
210 imager M2m microscope and processed by ImageJ software. To quantify the expression of Stxbp1,
211 we measured the counts of the signal dots in each cell within the specific regions using ImageJ
212 software. At least 10 neurons were counted for each mouse, and the averaged number from those
213 cells represents the expression intensity for each mouse. The results were calculated as counts per
214 cell. A student's two-tailed t-test was used to analyze the quantification data.
215
216 Mouse serum HB-EGF levels were determined by the Mouse HB-EGF ELISA kit (ABclonal,

4

217 RK02882). Blood was collected through the tail vein using a microvette capillary blood collection
218 tube. Samples were incubated at room temperature for 40 min, and centrifuged at 3,000 g for 15
219 min to collect serum. ELISA is performed according to the manufacturer's instructions.
220

221 **snRNA-seq and snATAC-seq**
222 Mice were euthanized with CO_2 after wheel-running for 3 months at the age of 6 months old. The
223 hippocampus was isolated immediately, washed with cold D-PBS, and snap-frozen with liquid
224 nitrogen. Hippocampi from 3 mice were pooled together before nuclei isolation and capture for each
225 of the analyses (snRNA-seq and snATAC-seq). The nuclei isolation for snRNA-seq and snATAC-
226 seq were performed in different batches from different mouse cohorts. The Frankenstein protocol
227 was adopted to isolate nuclei from frozen tissues (dx.doi.org/10.17504/protocols.io.3fkgjkw). Briefly,
228 the hippocampus was sliced into 2-5 mm^3 pieces on ice, and the samples were transferred to a 1 ml
229 Dounce homogenizer with prechilled nuclei isolation buffer (10 mM Tris-HCl, 10 mM NaCl, 3 mM
230 $MgCl_2$, and 0.01% NP40). The tissue was homogenized with five strokes of the loose pestle and 10
231 strokes of the tight pestle. The homogenate was filtered through a BD Falcon 40 µm cell strainer
232 and centrifuged at 300g for 5 minutes at 4 °C. The supernatant was discarded. The nuclei pellet
233 was resuspended with D-PBS with 20% percoll (GE) and centrifuged at 500 g for 15 minutes at 4 °C.
234 The supernatant was discarded, and the nuclei pellet was resuspended D-PBS with 1% BSA for
235 snRNA-seq and with the Nuclei Buffer for snATAC-seq. Nuclei concentration was measured with a
236 hemocytometer, and the samples were further processed using the NovaSeq system at the Single
237 Cell Genomics Core at Baylor College of Medicine for 10x single cell 3' v3 RNAseq and 10x Single
238 Cell ATAC sequencing. An average of 10k cells were targeted for capture per sample per assay,
239 with over 50,000 reads per cell. snRNA-seq and snATAC-seq were performed separately on
240 different cohorts of mice, generating 246,765,198 reads from 34,950 nuclei for snRNA-seq and
241 211,976,010 reads from 31,793 nuclei for snATAC-seq after quality control.
242

243 **snRNA-seq data initial processing**
244 Sequenced reads were first processed and quality controlled using the Cell Ranger Single Cell
245 Software Suite provided by 10x Genomics (version 3.1.0). Reads were aligned to the mouse
246 genome (mm10), and read counts per gene for every cell in a sample were obtained. We used
247 CellBender (version 0.1.0, https://github.com/broadinstitute/ CellBender.git) to identify nuclei with
248 ambient RNA further using the remove-background function with the 'total-droplets-included' set at
249 1,000,000. Nuclei kept by both Cell Ranger and CellBender were retained for subsequent analyses.
250 A total of 37303 nuclei were included (8804 APP_EX, 10411 APP_RT, 6891 WT_EX, 11197
251 WT_RT). Seurat pipeline was used to analyze the kept nuclei by first removing cells with more than
252 5% of reads mapped to mitochondrial genes and those with less than 500 genes or more than 5500
253 genes. On the remaining 32285 cells, the top 2000 variable genes were identified and used to
254 perform dimension reduction using PCA with a maximum of 200 PCs. The first 50 PCs were then
255 used for further clustering and UMAP. To decide the cell populations, we used clustering with a
256 resolution of 0.5. The top marker genes were identified empirically using the FindMarkers function.
257 To identify differentially expressed genes (DEG) between the genotypes in each cell population, we
258 first subset the cells of the particular cell population and then used the FindMarkers function from
259 Seurat.
260

261 Gene Set Enrichment Analysis (GSEA) v4.2.2 Mac App was used. Two libraries specific to mouse,
262 mouse pathway (893 gene sets) and GO libraries (12526 gene sets) were obtained from http://ge-
263 lab.org/gskb/; along with C2 (6290 gene sets) and C5 (14998 gene sets) from MSigDB
264 (https://www.gsea-msigdb.org/ gsea/msigdb/ index.jsp) were used in the analysis. GSEPreparked
265 analysis was adopted with the ranked files generated from logFC of all genes except for the ones
266 with no expression in each cell type obtained from the FindMarkers function from the Seurat
267 package. The logFC was calculated with SCTransformed data in the Seurat object. GSEA was run
268 for each cell type in three contrasts: APP_EX vs. APP_RT, APP_RT vs. WT_RT, and WT_EX vs.
269 WT_RT. Parameters used in the GSEA were: max size = 5,000; min size = 5; collapsing mode for
270 probe sets => 1 gene = max_probe; normalization mode = meandiv. The GSEA analysis results

271   were examined at different thresholds: 1. Soft threshold with $p < 0.001$; 2. Median threshold with
272   FDR < 0.25 and $p < 0.001$; 3. Harsh threshold with FDR < 0.1 and $p < 0.001$; 4. Extreme threshold
273   top 10 in normalized enrichment score (NES) (positive/negative) and FDR < 0.25 and $p < 0.001$.
274   Three gene sets were shortlisted for further analysis. They were picked out of all gene sets that
275   were significantly over-represented at the median threshold in more than 4 cell types and were
276   reversed in NES in APP_EX vs. APP_RT and APP_RT vs. WT_RT contrasts.
277
278   Sub-clustering analysis was done in excitatory neuron (EN) nuclei using Seurat version 3.2.2. No
279   further filters were applied for nuclei in the sub-clustering analysis. The SCT assay matrix of all
280   excitatory neuron nuclei was used. Variable features were identified within the chosen cell type's
281   nuclei with the function FindVariableFeatures with the vst method, and the top 2,000 features were
282   used in the analysis. Then, dimensionality reduction was run using the RunPCA function. Data was
283   visualized by UMAP with dim 1:30. Then EN nuclei were clustered with the function FindNeighbors
284   with dims 1:30 and FindClusters at resolution 0.2. Sub-cluster markers curated from literature[11].
285
286 **snATAC-seq data initial processing**
287   The data from sorted bam files created by CellRanger (version 3.1.0) was processed with package
288   snaptools to generate the bin-by-nuclei matrix for ADEX, ADRT, WTEX, and WTRT samples
289   separately, following the online tutorial: https://github.com/r3fang/SnapTools. Bam files were
290   processed by the snap-pre function with parameters settings as min-mapq = 30, min-flen = 50, max-
291   flen = 1000, keep-chrm = TRUE, keep-single = FALSE, keep-secondary = False, overwrite = True,
292   num = 20000, min-cov = 500. Then snap-add-bmat function was used to call the bin matrix with bin-
293   size-list set as 5,000. All samples were then combined for analysis with the package snapATAC
294   following the pipeline tutorial https://github.com/r3fang/SnapATAC/
295   blob/master/examples/10X_brain_5k/README.md. Original filters of nuclei were set as UMI >= 3 &
296   <= 5; promoter_ratio >= 0.15 & <= 0.6; peak_region_fragments > 1,000 & < 20,000;
297   frag_in_peak_ratio > 0.15. Bins in the blacklist or mitochondrial chromatin or over top 95% coverage
298   among all bins were filtered out. Dimensionality reduction was done with the function
299   runDiffusionMaps with num.eig set as 50. Clustering was done with the function runKNN (eigs.dims
300   set as 1:30 and k as 15) and the function runCluster (leiden version 0.3.5 and resolution as 1.5).
301   Bins were binarized for further analysis. Peaks were then called on bins in each cluster with
302   buffer.size set as 500. Peaks were then merged by the function reduce. Lastly, peaks were scaled
303   by the function scaleCountMatrix with the RPM method. Snaptools and SnapATAC are installed
304   from GitHub: https://github.com/r3fang/SnapATAC (Snaptools version v1.2.3 and SnapATAC
305   version 1.0.0).
306
307   The peak-by-nuclei matrix was then imported to the Seurat pipeline for further analysis. Peaks from
308   the blacklist, unwanted chromosomes, and small clusters (Cluster 18, 19, and 20) were filtered out.
309   Furthermore, nuclei with less than 1,000 peaks were removed. UMAP was run again in Seurat to
310   generate the embedding for visualization. ACTIVITY matrix for gene-by-nuclei matrix was generated
311   by summing up peaks within 2kb upstream of the TSS (transcription starting site). Based on the
312   ACTIVITY matrix, marker plots and enrichment analysis were carried out to determine the final cell
313   type annotations. Sub-clustering analysis was done in excitatory neuron nuclei using Seurat version
314   3.2.2. No further filters were applied for nuclei in the sub-clustering analysis. The ACTIVITY matrix
315   of all excitatory neuron nuclei was used. Variable features were identified within the chosen cell
316   type's nuclei with the function FindVariableFeatures with the vst method, and the top 30,000
317   features were used in the analysis. Then, dimensionality reduction was run using the RunLSI
318   function. Data was visualized by UMAP with dim 1:30. Then EN nuclei were clustered with the
319   function FindNeighbors with dims 1:30 and FindClusters at resolution 0.2. Sub-cluster markers[11].
320
321 **Integrated multi-omics data processing**
322   The snRNA-seq and snATAC-seq data were integrated following the Seurat integration pipeline.
323   The assay used for snRNA-seq was RNA, and for snATAC-seq, it was ACTIVITY. Anchors were
324   found with the function FindTransferAnchors with variable genes from RNA data as features,

6

325   reference.assay as "RNA", query.assay as "ACTIVITY", reduction method as cca, and dims with
326   1:40. ATAC data was then imputed with the function TransferData using anchors found, weight
327   reduction with LSI method, and variable genes as variable features from RNA object. Then snATAC
328   and snRNA data were merged with the function merge. RunPCA was performed on integrated data
329   with all features in the combined data and RunUMAP with dim 1:30.
330

331   Trajectory analysis was done in oligodendrocytes following the monocle3 trajectory analysis pipeline
332   (monocle3 version 1.0.0). Oligodendrocyte nuclei were sub-clustered into two sub-clusters with
333   FindNeighbors (dims 1:30) and FindClusters (resolution 0.05). The bigger sub-cluster with 10,480
334   nuclei out of all 11,252 nuclei was used in trajectory analysis. The learn_graph and order_cells were
335   applied with default monocle3 settings. Calculated pseudotime was then binned in equal 50 frames
336   for analysis. In plotting, each bin on the x-axis was represented by the middle value of each binned
337   frame. Proportions for nuclei within each bin were calculated by the formula shown below:

$$(Within\ genotype) Proportion\ for\ genotype_A$$
$$= cell\ number\ for\ genotype_A in\ bin_n\ /\ total\ cell\ number\ for\ genotype_A$$
$$(Across\ genotype) Proportion\ for\ genotype_A$$
$$= (cell\ number\ for\ genotype_A\ in\ bin_n\ /\ total\ cell\ number\ for\ genotype_A)$$
$$/ \sum_{x}^{4\ genotypes} (cell\ number\ for\ genotype_x\ in\ bin_n\ /\ total\ cell\ number\ for\ genotyp_x)$$

338   Module scores were calculated by the Seurat function AddModuleScore for nuclei in each of the 50
339   binned pseudotimes. Oligodendrocyte marker lists are from literature [12]. Newly formed
340   oligodendrocyte markers: Tcf7l2, Casr, Cemip2, Itpr2; Myelin forming oligodendrocyte markers: Mal,
341   Mog, Plp1, Opalin, Serinc5, Ctps1; Mature oligodendrocyte: Klk6, Apod, Slc5a11, Pde1a.
342

343   Differential Accessible Regions detection by a Siamese neural network model
344   Due to the lack of power in using statistical tests to discover DARs, we instead trained a simple
345   neural network to distinguish different conditions, followed by interpreting the trained model to
346   extract the differentially accessible regions. We detected DARs for each cell type separately.
347   Preprocessed and cell-type annotated snRNA-seq and snATAC-seq were loaded as Seurat v3
348   objects. Then, cell-type-specific snRNA-seq and snATAC-seq Seurat objects were created to
349   perform omics data integration by Seurat v3. After integration, anchor cell pairs inferred by Seurat
350   were saved, together with a highly-variable gene matrix (vst method, 5000 genes) and highly-
351   accessible peaks matrix (detected in more than 5% of the cells). Gene matrix, peak matrix, anchor
352   pairs, and genotype label were provided to train a customized Siamese neural network (explained in
353   the next section). After training, the network was able to project all cells from different omics data
354   into a common low-dimensional space, where snRNA-seq and snATAC-seq data are mixed, and
355   different genotypes are separated from each other. With the goal of separating genotypes achieved,
356   we believe the neural network has identified peaks that vary among genotypes. To extract these
357   peaks from the black box model learned, we first applied the Activation Maximization algorithm
358   (explained later) to construct 12 pseudo-cell snATAC-seq data which are identified by the model
359   with a high possibility of belonging to each genotype. Then t-test was applied to find the peaks that
360   are important exclusively in only one genotype. Directions of these peaks in APP_RT vs. WT_RT
361   and APP_EX vs. APP_RT were determined by checking our real snATAC-seq data. Finally, peaks
362   with log2 foldchange smaller than log2(1.1) or not falling in gene regions were removed. With all the
363   above being done, we got a set of differential accessible regions among different genotypes in each
364   cell type. In detecting DARs in this way, we believe it can bear the noises of single-cell data and
365   utilize the information from snRNA-seq data. The regions found by this method were validated in the
366   original data.
367

368   Customized Siamese neural network to separate multi-omics data by genotypes
369   The goal of this neural network is to separate genotypes while mixing snRNA-seq and snATAC-seq
370   data. After training, we will be able to find important peaks that can separate genotypes not only
371   based on chromatin accessibility but also transcriptome. To do this, we created a model based on

7

372   scDGN[13] neural network framework, including several modules.

373

374   The encoder module is used to project datasets into a common lower dimensional space and
375   contains two fully connected layers that produce the hidden features $x' = f_{e2}(f_{e1x}(x; \theta_{e1x}); \theta_{e2})$ for
376   snRNA-seq or $y' = f_{e2}(f_{e1y}(y; \theta_{e1y}); \theta_{e2})$ for snATAC-seq, where θ represents the parameters in
377   these layers. The label classifier, $f_{lx}(x'; \theta_{lx})$ and $f_{ly}(y'; \theta_{ly})$, ensures genotypes are separated in
378   the common space. The goal of the domain discriminator $f_d(x'; \theta_d)$ and $f_d(y'; \theta_d)$ is to determine
379   whether a pair of inputs ((xi,xj), (xi,yj), (yi,xj), (yi,yj)) are from the same domain or not. The overall
380   objective function to be minimized is:

$$E = L_l(f_{lz}(z_1'; \theta_{lz}), lz) + \lambda L_d(f_d(z_1'; \theta_d), f_d(z_2'; \theta_d)),$$

381   where $z_1, z_2 \in \{x, y\}$, λ can control the trade-off between the goals of domain invariance and higher
382   classification accuracy. Inspired by Siamese networks, the domain loss adopts a contrastive loss for
383   a pair z_1 and z_2, where $z \in \{x, y\}$:

$$L_d(f_d(x'; \theta_d), f_d(y'; \theta_d)) = U(1 - \cos(f_d(z_1'), f_d(z_2'))) + (1 - U)\max\{0, \cos(f_d(z_1'), f_d(z_2')) - m\},$$

384   where U=0 indicates the two cells are from the same modality, but different genotypes, and U=1
385   indicates that they are identified as anchors by Seurat. $\cos(\cdot)$ is the cosine embedding loss, and m
386   is the margin that indicates the prediction boundary. Overall, the aim of the objective function is to
387   minimize the label classification loss and the domain loss. In this way, genotypes are separated,
388   and important peaks are learned by the model based on not only snATAC-seq but also snRNA-seq
389   data.

390

391   **Activation Maximization to find important peaks**
392   With genotypes separated, we used activation maximization to extract important peaks for each
393   genotype. Given a particular genotype i and a trained neural network f, activation maximization
394   looks for important input genes x_m and peaks y_m by solving the following optimization problem:

$$x_m = \max_x \log(f(x)) - \lambda\|x\|^2, \quad y_m = \max_y \log(f(y)) - \lambda\|y\|^2.$$

395   Twelve such pseudo-cells were constructed for each of the 3 genotypes, including WT_RT,
396   APP_RT, and APP_EX. A t-test was performed to identify exclusively important peaks in genotypes,
397   and their log-fold-change is determined by our snATAC-seq data.

398

399   **Homer motif analysis and regulatory network construction**
400   With exercise reversed DARs of each cell type, we ran Homer[14] *findMotifsGenome* function to
401   perform known motif enrichment analysis. All parameters are set to default. The summary plot
402   containing the p-value and peak ratio in each cell type, and reverse direction is drawn with ggplot2
403   in R. With discovered motifs in each cell type, we collected their downstream genes from TRASFAC
404   database. Network per cell type was constructed, with edges indicating TF-downstream gene
405   relationship. The networks were saved as .gml files and visualized in Cytoscape with color indicating
406   log fold changes in the two comparisons.

407

408   **Integrated analysis with spatial transcriptomics data**
409   The count matrix of spatial transcriptomics data was filtered by removing spots with tissue coverage
410   of less than 30% in the HE images and then removing genes that were detected in less than 10
411   spots. The edgeR function "cpm" was used for the normalization of the filtered matrix. The output
412   log cpm matrix was used for the following analysis. For assigning cells in the snRNA-seq data back
413   to the locations in the spatial transcriptomics data, we suppose that the reference atlas has n
414   positions with p genes, and the snRNA-seq data set has m cells with the same number of p genes.
415   We aimed to assign the m cells into n positions using a linear regression model with L1 norm and
416   generalized L2 norm via graph Laplacian. We created a random walk normalized graph Laplacian
417   matrix based on the location information and anatomical structures from the ST data. If two spots
418   belong to one anatomical structure and the distance between the spots is smaller than a specific
419   threshold, then the spots will be connected in the Laplacian matrix. Our model uses a linear method
420   to measure the differences in gene expression levels in assigning cells to locations. The optimal
421   solution minimizes the differences between gene expression levels of individual cells and gene

expression levels of locations. For each individual cell, we want to minimize the following objective function,

$$\sum_{i}^{n}(\left\|s_i - \sum_{j=1}^{m}\beta_{ij}T_j\right\|^2 + \lambda_1\|\beta_i\|_1 + \lambda_2\beta_i{}^t L\beta_i)$$

where $s \in R^{n \times g}$ is the single-nucleus expression matrix, $T \in R^{m \times g}$ is marker genes expression matrix of reference atlas (spatial transcriptomics data), L is normalized graph laplacian computed based on the location distance matrix and anatomical structures. The L1 norm penalization encourages sparsity on the coefficients, which guarantees that one cell can only be assigned to a small number of locations. The generalized L2 norm encourages the smoothness of the coefficients, which guarantees that cells with similar gene expression levels are more likely to be assigned to closer locations. To map the snRNA-seq data back to the mouse hippocampus, the reference profile was first established by using the spatial transcriptomics expression atlas of 500 highly variable genes at the 6-month age retrieved from Navarro et al. Reconstructed spatial expression patterns were validated by well-known genetic markers, such as Prox1 and Ociad2. Reconstructed patterns of marker genes were consistent with the FISH images from Allen Brain Atlas.

To reconstruct spatial expression patterns, we used the following steps: (1) Read the gene expression matrix from snRNA-seq and expression matrix from reference atlas (spatial transcriptomics data). (2) Construct the Laplacian matrix based on the location information. (3) Use CVX to solve the convex function with L1 norm and generalized L2 norm. (4) Assign cells to target locations based on the distribution of marker genes in the objective function. (5) Reconstruct the spatial patterns based on the expression profiles in the snRNA-seq data and cell locations from the mapped results. Fill the expression profiles from the snRNA-seq data in the assigned locations.

To calculate cell proportion in different anatomical regions, we first performed glm-SMA algorithm to assign the cells from snRNA-seq data back to the locations from the spatial transcriptomics data. Then, we counted the cell number in different anatomical regions and calculated the cell proportion of each cell type in DG and CA regions. To confirm the result was not generated by chance, we randomly shuffled the genotypes in the APP_EX and APP_WT samples and repeated the shuffling 100 times. Then, we recalculated the cell proportions and did the student t-test. Granule cell proportion increased in the DG region from AD_EX samples with a p-value < 2.2e-16.

Statistical analysis
Statistical analyses were performed using SPSS (V.21.0, IBM) unless described otherwise in the above sections. No statistical methods were used to pre-determine sample sizes. Instead, sample sizes were determined based on previous publications for the relevant assays. Normality was tested by the Shapiro-Wilk test (n < 10) or D'Agostino-Pearson omnibus test (n > 10). For non-normal data or data with nonequivalent variances, the comparisons between two or multiple groups were tested with the Mann-Whitney test or the Kruskal-Wallis test, respectively. All tests were two-sided. All measurements were taken from distinct biological samples (mice or human subjects). Most comparisons between the two groups were analyzed using a two-sided, unpaired t-test. Body weight with multiple time points or Morris water maze tests were analyzed with repeated-measures ANOVA with Tukey's post hoc test. For statistical significance, a two-tailed unpaired t-test, or one-way repeated ANOVA with Fisher's LSD test multicomparisons, was used for experiments with two groups. The behavior test experimenter was blinded to the exercise or pharmacological treatment conditions during the early stage of analysis, such as counting the time duration from video clips. The statistical analysis and data plotting was then done by experimentalists who knew both genotype and treatment information. Animals were excluded and euthanized before behavior tests if they showed distress, infection, bleeding, or anorexia. All data were expressed as mean ± SEM. All data were individually plotted (Prism 9, GraphPad). The exact numbers of animals are reported in the figure legends. P < 0.05 is set as significance.

Sex as a biological variable

474 We used both male and female mice and clarified the sex in the figure legends and methods
475 sections. Most experiments were done in male mice, while female mice were also used. We did not
476 find sex differences.
477
478 **Data access**
479 The snATAC-seq (GSE237884) and snRNA-seq (GSE237885) data will be available at NCBI's
480 Gene Expression Omnibus (GEO) under the GSE237925 SuperSeries after this manuscript is
481 officially published.
482
483 **Data availability**
484 Data is available upon request to the corresponding authors.
485
486 **Code availability**
487 Code is available upon request to the corresponding authors.
488
489 **RESULTS**
490
491 **Exercise improves memory functions and induces cell-type-specific transcriptomic changes.**
492 We used the C57BL/6J homozygous knock-in mice containing the Swedish (NL), Beyreuther/Iberian
493 (F), and Arctic (G) mutations in the gene for amyloid precursor protein (APP^{NL-G-F})[15,16] as a model for
494 AD because it lacks the artificial hyperactivity phenotype from APP overexpression, which
495 resembles human AD pathophysiology[17,18]. For chronic physical exercise, we used voluntary wheel-
496 running to minimize stress on mice. Wild-type (WT) and APP^{NL-G-F} mice were put in cages with
497 running wheels at 3 months old. After exercise for around 6 months exercise, the exercise groups
498 (WT_EX and APP_EX) shared similar wheel-running exercise volumes and lost similar amounts of
499 body weight compared to their respective rest controls (WT_RT and APP_RT) (**Suppl Fig S1a-b**).
500 The chronic exercise improved memory functions in APP^{NL-G-F} mice but not in WT mice in the
501 classical object-in-place test (**Fig 1a-c**) and Y-maze test (**Fig 1d-f**) at the age of 10 months. Total
502 travel distances in these tests remained similar between WT and APP^{NL-G-F} mice (**Suppl Fig S1c-d**).
503 Although exercise appeared to have no effects in WT mice in these standard tests, a modified
504 object-in-place test with 5 objects[7] demonstrated a clear memory-enhancing effect of exercise in
505 WT mice (**Fig 1g-i**). In summary, chronic wheel-running exercise improves learning and memory in
506 both WT and APP^{NL-G-F} mice, with a more robust effect on the APP^{NL-G-F} mice.
507
508 We reason that the molecular changes in the brain would precede the behavioral changes and
509 exercise for too long might cause secondary changes that are outcome rather than the cause of the
510 behavioral changes. Therefore, we subjected the hippocampus from 6-month-old WT and APP^{NL-G-F}
511 mice to nuclei isolation and capture after exercise training for 3 months. The snRNA-seq and
512 snATAC-seq analyses were performed in different batches with different mice. For each analysis,
513 hippocampi from 3 mice were pooled together before nuclei isolation and capture. After quality
514 control, we obtained 31,793 nuclei for snATAC-seq and 34,950 nuclei for snRNA-seq, which were
515 clustered into 6 major cell types, including excitatory neurons (EN), inhibitory neurons (IN),
516 oligodendrocytes (OLG), oligodendrocyte progenitor cells (OPC), astrocytes (AST), and microglia
517 (MG) (**Fig 1j-k**). Exercise did not cause obvious changes in the overall cellular composition in WT or
518 APP^{NL-G-F} mice (**Suppl Fig S2a**). We focus on the exercise effect (APP_EX vs. APP_RT and
519 WT_EX vs. WT_RT comparisons) and the amyloid effect (APP/RT vs. WT/RT comparison).
520 Differentially expressed genes (DEGs) from these comparisons were identified from each cell type.
521 Exercise led to over 3 - 6 folds more DEGs in APP mice than in WT mice in most cell types (**Suppl
522 Fig S2b, Suppl Table S1-6**), consistent with more robust improvement of the cognitive functions in
523 APP^{NL-G-F} mice compared to WT mice.
524
525 Exercise caused a predominant upregulation of gene expression, while amyloid led to a
526 predominant downregulation of gene expression (**Suppl Fig S2b-c**). DEGs altered by exercise or
527 amyloid were enriched in different functional pathways in a cell type-specific manner (**Suppl Fig**

528   **S2c-d, Suppl Table S7-9**). For example, DEGs in microglia were enriched in the complement and
529   IL5/IL6 signaling pathways; those in oligodendrocytes were enriched in prostaglandin signaling;
530   those in OPCs were enriched in the IL12/STAT4 signaling; those in astrocytes were enriched in lipid
531   metabolism, angiogenesis, and cytoskeletal regulation; and those in neurons were enriched calcium
532   signaling, protein translation, and mRNA processing (**Suppl Fig S2d**). Notably, the receptor tyrosine
533   kinase signaling, especially insulin, c-KIT, epidermal growth factor receptor (EGF), and the
534   downstream phosphoinositide 3-kinases (PI3K) signaling pathways were universally enriched
535   across multiple cell types (**Suppl Fig S2d**).
536
537   **Exercise counteracts amyloid-dependent transcriptomic changes in growth factor signaling.**
538   The exercise-induced genes (APP_EX vs. APP_RT) showed a robust negative correlation with the
539   amyloid-induced genes (APP_RT vs. WT_RT) across all cell types (**Fig 2a**). Over 833 reversed
540   DEGs were found in at least one cell type, with significant changes in both comparisons but in
541   opposite directions. Most of these DEGs were upregulated by exercise and downregulated by
542   amyloid (**Fig 2b**), suggesting that exercise 'reversed' the amyloid-induced transcriptomic changes
543   by activating transcription. Interestingly, less than 9 reversed genes were shared by any 4 clusters
544   (**Fig 2c**), suggesting that the reversal effects are highly cell type-specific at the gene level. However,
545   at the pathway level, the EGF receptor (EGFR) and insulin pathways stood out as the top common
546   functional pathways with the reversal pattern across most cell types (**Fig 2d-f, Suppl Table S10-11**).
547
548   The EGFR/insulin pathway was suppressed by amyloid and upregulated by exercise (**Fig 3a** and
549   **Suppl Fig S3**). Stxbp1 is a top gene in neurons with a reversed expression pattern within the
550   EGFR/insulin pathway and encodes a syntaxin-binding protein involved in synaptic vesicle cycling.
551   RNAscope verified that hippocampal Stxbp1 was suppressed by amyloid and upregulated by
552   exercise (**Fig 3b-c**). EGFR and insulin signaling share many downstream players, including PI3K,
553   AKT, and MAPK (**Fig 3d-e**). The EGFR/insulin signaling pathway functions downstream of several
554   growth factors, such as insulin-like growth factor (IGF), fibroblast growth factors (FGF), hepatocyte
555   growth factor (HGF), transforming growth factor (TGF), vascular endothelial growth factor (VEGF),
556   and platelet-derived growth factor (PDGF) (**Fig 3e**). In summary, exercise counteracts amyloid-
557   induced repression of the growth factor signaling in multiple cell types.
558
559   **Exercise counteracts amyloid-induced transcriptional regulatory networks.**
560   snATAC-seq identified differentially accessible regions (DARs) in response to exercise or amyloid
561   deposition in each cell type (**Suppl Fig S4a**). The top enriched pathways and motifs in these DARs
562   were cell type-specific (**Suppl Fig S4b-c**). Among the enriched transcription factors (TFs) are those
563   related to growth factor signaling, cell proliferation, and neuron differentiation, including EGR1, MYB,
564   ATOH1, and ASCL1 (**Suppl Fig S4c**). Consistent with transcriptomics data, exercise-induced and
565   amyloid-induced genome accessibility changes displayed negative correlations across all cell types
566   (**Fig S4d**). Unlike the transcriptomics data, DARs with the reversal phenotype were more evenly
567   distributed in both directions (**Fig 4a**). However, motif analyses of these reversed DARs revealed
568   direction-specific transcription regulator networks centered on TFs (**Fig 4b**). DARs upregulated by
569   amyloid and downregulated by exercise were referred to as "U>D" (from 'Upregulation' to
570   'Downregulation'), while those DARs downregulated by amyloid and upregulated by exercise were
571   referred to as "D>U" (from 'Downregulation' to 'Upregulation'). TFs involved in growth and
572   differentiation, such as EGR1, MYB, MEF2, and ASCL1, show direction-specific enrichment in
573   neurons (**Fig 4b**). These TFs are downstream of growth factors, PI3K, or MAPK signaling pathways
574   and are involved in synaptic plasticity or cell growth[19–22]. WT1 and NEUROG2, transcription factors
575   in cell growth and neurogenesis[23–25], were enriched in the D>U DARs in both excitatory neurons and
576   inhibitory neurons. ATOH1, the transcription factor essential for cerebellar granule cell formation[26],
577   was enriched in D>U DARs of excitatory neurons, and U>D DARs of inhibitory neurons (**Fig 4b**).
578   DARs of both directions were enriched with the ETS family of TFs in microglia, the SOX family of
579   TFs in oligodendrocytes, and the LHX family of TFs in astrocytes (**Fig 4b**).
580
581   Transcriptional regulatory network analysis of the excitatory neurons suggested that EGR1, WT1,

582 MYB, and MEF2 are central TFs that reversed amyloid-mediated transcriptomic changes by
583 activating transcription in response to exercise (**Fig 4c-d** and **Suppl Fig S5**). The downstream
584 genes of these central TFs in the network overlapped significantly with genes in the EGFR/insulin
585 pathway (**Suppl Table S12-13**). Many transcription factors can serve as transcription activators and
586 repressors in a context-dependent manner. Therefore, it makes sense that the same transcription
587 factor network can drive opposite reversal directions. In inhibitory neurons, ELK1/4 and TCF4
588 replace EGR1/MYB as key TFs working with MEF2A and WT1 for exercise-induced transcriptional
589 remodeling (**Fig 4e-f** and **Suppl Fig S6**). By comparison, the network analysis suggests that the
590 SOX family TFs work with ELK4 to drive the reversal of the APP-induced oligodendrocyte over-
591 maturation by exercise (**Fig 4g-h** and **Suppl Fig S7**), in line with the known role of the SOX family in
592 oligodendrocyte differentiation[27]. Similar patterns were observed in other cell types, with cell type-
593 specific TFs driving distinct downstream genes that converge on the EGFR/insulin pathway (**Suppl
594 Fig S8-S10** and **Suppl Table S12-13**). In summary, exercise stimulates cell type-specific
595 transcriptional regulatory networks, counteracting amyloid-induced transcriptomic changes by
596 activating growth factor signaling.
597 One cellular manifestation of growth factor signaling activation is neurogenesis. Granule cells (GC)
598 in the dentate gyrus (DG) of the hippocampus constitute the primary niche for adult hippocampal
599 neurogenesis[28]. A more focused analysis of the snRNA-seq data within the hippocampal excitatory
600 neurons revealed that exercise causes a more drastic gene expression change in the GC
601 population than the pyramidal cell population (**Fig 5a-c**), leading to a higher proportion of granule
602 cells within the EN cluster (**Fig 5d**). Trajectory analysis of neurons did not recapitulate the
603 neurogenesis process, probably due to the scarcity of nascent neurons in the adult brain. Therefore,
604 we sought to resolve the snRNA-seq data spatially to address whether exercise-induced differences
605 show spatial preference towards DG GCs. We integrated the EN snRNA-seq data with the previous
606 spatial transcriptomic data[29] to group EN nuclei into sub-hippocampal regions, such as DG (8813
607 nuclei) and CA1 (3379 nuclei) (**Suppl Fig S11**). We found that exercise increased the GC
608 proportion in the DG, but not IN proportion in the DG (**Fig 5e**). Production of immature GCs (imGCs)
609 is a hallmark of adult hippocampal neurogenesis. To predict imGCs, we applied the logistic
610 regression model trained on mice prototype imGCs (with a gene signature of Ascl1, Dcx, Tubb3,
611 Neurod1, and Tbr1)[30] to our snRNA-seq data. The imGCs proportion was reduced by amyloid,
612 which was rescued by exercise (**Fig 5f**). These results suggest that multiple growth factor signaling
613 and neurotrophic pathways participate in exercise-stimulated neurogenesis in the DG.
614
615 In addition to neurons, oligodendrocytes stood out as another cell type with a growth and
616 proliferation phenotype amenable to exercise. We integrated snRNA-seq and snATAC-seq and took
617 the co-embedding space for trajectory analysis[31] (**Fig 5g**). The trajectory recapitulated
618 oligodendrocyte maturation because the expression signature of the new, myelin-forming, and
619 mature oligodendrocyte showed a monotonic correlation with the pseudo-time scale (**Fig 5h**).
620 Amyloid increased the proportion of mature oligodendrocytes (**Fig 5i**), which is in line with previous
621 reports that amyloid oligomers promote oligodendrocyte differentiation and lead to thicker myelin[32,33].
622 We speculate that this over-maturation phenotype could be due to senescence or a lack of
623 replenishment. Interestingly, exercise reduced the mature oligodendrocyte proportion while
624 increasing myelin-forming and new oligodendrocyte proportions (**Fig 5i**).
625
626 **Growth factor signaling contributes to the memory-enhancing effects of exercise.**
627 The single-nucleus multi-omics suggest that several TFs-centered networks contribute to exercise-
628 stimulated activation of growth factor signaling pathways in different cell types. To address whether
629 the EGFR-related growth factor signaling is required for exercise-mediated cognitive improvement,
630 we administered EGFR inhibitor Gefitinib at 50 mg/kg and PI3K inhibitor Wortmannin at 0.5 mg/kg
631 through oral gavage once every other day from 4 to 8 months old in APP^{NL-G-F}mice while they were
632 subjected to wheel-running starting at 4 months old. These pharmacologic manipulations did not
633 affect the wheel-running exercise volume (**Fig 6a-b**) or body weight (**Fig 6c**). However, the drugs
634 efficiently blunted the exercise-mediated cognitive improvement in the object-in-place test (**Fig 6d-e**),
635 Y-maze test (**Fig 6f-g**), social memory (**Fig 6h-j**), and Morris water maze test (**Fig 6k-l**) without

636 affecting sociability (**Fig 6i**) or the total locomotor activity during these tests. The inhibitors did not
637 alter anxiety-related behaviors in the elevated plus maze test, open field arena test, or light-dark test
638 (**Fig 6m-o**). Interestingly, the inhibitors also blunted exercise-induced improvement in amyloid
639 pathology (**Fig 6p-q**), suggesting that EGFR signaling activation and its related anabolic stimulation
640 are required for exercise-induced amyloid clearance and cognitive benefits.
641
642 We sought to identify the upstream signals for exercise-induced growth factor signaling. Among
643 ligands for the EGFR family[34], heparin-binding EGF-like growth factor (HB-EGF) stood out with
644 significant gene expression upregulation in mouse muscles[35] and human muscles after physical
645 exercise[36] (**Fig 7a**). HB-EGF gene expression is also upregulated in blood cells after exercise in
646 human blood[37]. In a proteomics analysis of human blood samples, HB-EGF is the only detectable
647 EGFR ligand upregulated by exercise[38] (**Fig 7b**). HB-EGF is known to cross the blood-brain
648 barrier[39]. We confirmed that the blood HB-EGF levels were elevated after chronic exercise in APP[NL-
649 G-F] mice (**Fig 7c**). To address whether HB-EGF has cognitive benefits, we intranasally administered
650 HB-EGF in sedentary APP[NL-G-F] mice at 3 ug/mouse once every other day from 4 to 7 months old.
651 HB-EGF did not affect body weight (**Fig 7d**) but improved cognitive functions in the object-in-place
652 test (**Fig 7e-f**), Y-maze test (**Fig 7g-f**), and social memory (**Fig 7i**) without affecting the sociability
653 (**Fig 7j**) or locomotor activity during the three-chamber test (**Fig 7k**)**.** HB-EGF also reduced escape
654 latency during the multiple-day Morris water maze test (**Fig 7l**) but did not cause significant
655 differences in the probe test on the last day (**Fig 7m**). HB-EGF did not alter anxiety-related
656 behaviors in the open field arena test or light-dark test (**Fig 7n-o**) but reduced the overall beta-
657 amyloid deposition (**Fig 7p-q**). These results suggest chronic intranasal HB-EGF treatment in mice
658 can ameliorate amyloid-induced cognitive decline and reduce amyloid deposition.
659
660 **DISCUSSION**
661
662 Our results offer a comprehensive overview of transcriptomic and chromatin accessibility changes
663 across different cell types within the mouse hippocampus in response to chronic voluntary exercise.
664 Two recent publications present snRNA-seq analyses conducted on mouse brains. One study
665 examined the whole brain following 12 months of voluntary wheel-running exercise[40], while the other
666 focused on the hippocampus after 4 weeks of wheel-running exercise[41]. These investigations were
667 conducted on wild-type mice without amyloid deposition and did not include snATAC-seq analyses.
668 Our utilization of APP[NL-G-F] mice, coupled with snATAC-seq integration, illuminates the upstream
669 transcriptional factor networks governing hippocampal responses to exercise in the presence of
670 amyloid deposition. Our profiling reveals that exercise reverses amyloid-induced transcriptomic
671 alterations by activating gene transcription. Exercise-induced transcriptional regulatory networks
672 show specificity to cell types for upstream transcription factors. Yet the downstream target genes
673 collectively converge on growth factor signaling pathways, particularly the EGFR/insulin pathway,
674 which is associated with elevated HB-EGF levels in the blood. The cognitive benefits of exercise are
675 blocked by pharmacological inhibition of EGFR/insulin signaling, while chronic intranasal
676 administration of HB-EGF enhances memory function in sedentary APP[NL-G-F] mice. Therefore, the
677 insights gained from single nucleus multi-omics analysis of exercise effects on the brain have
678 opened the door to a potential therapeutic approach for AD by activating growth factor signaling.
679
680 Growth factors, including BDNF, IGF-1, VEGF, and GH, have been implicated in the neurotrophic
681 or synaptogenic effects of exercise[42,43]. Our findings suggest that HB-EGF is a novel growth factor
682 involved in the process. Our results align with prior studies showing HB-EGF administration can
683 enhance the generation of new neurons or oligodendrocytes[44,45]. Consistently, HB-EGF was shown
684 to interact with APP and promote cellular neuritogenesis[46]. These findings do not rule out other
685 signaling pathways in exercise-induced cognitive improvement. Interestingly, EGFR inhibitors were
686 reported to have beneficial effects in AD, with some conflicting results [47,48]. We find that a combined
687 EGFR inhibitor and PI3K inhibitor blocked the effects of exercise training, which may act
688 independently of the baseline effects of the EGFR inhibitor itself. EGFR effects on AD appear to be
689 age-dependent and mediated by glial cells[47]. There is currently no available data on the effects of

EGFR inhibitors in the APP^{NL-G-F} mouse model at the baseline. Although we focus on HB-EGF based on muscle and blood omics datasets available in the literature, many brain cell types can produce EGF factors, which could be an additional source of elevated EGFR signaling. Further research is needed to elucidate the source of HB-EGF and the effects of EGFR inhibitors or agonists on AD progression.

The current study has several limitations. Firstly, the pharmacokinetics and pharmacodynamics of intranasal HB-EGF administration remain unclear. We speculate that intranasal administration may result in brain-enriched distribution, activating EGFR signaling and potentially mimicking the effects of exercise training in promoting neuritogenesis or neurogenesis. Secondly, it is uncertain whether treatment with Gefitinib and Wortmannin reduces EGFR and PI3K signaling in the hippocampus and whether this would negatively impact cognitive functions or amyloid pathology in the baseline condition without exercise training. Hence, it cannot be conclusively stated that the effects of EGFR/PI3K inhibition are attributed explicitly to exercise. Lastly, hippocampal samples from three mice were pooled for the single-nuclei omics analysis to enhance cost efficiency, albeit at the expense of statistical power. Future advancements in techniques may enable more cost-efficient comprehensive profiling. Despite these limitations, the identification of EGFR signaling from non-biased omics datasets, the EGFR/PI3K inhibitors-mediated abrogation, and the intranasal HB-EGF-mediated recapitulation of exercise-induced cognitive improvements and amyloid pathology collectively support a positive role of the EGFR signaling pathway in the cognitive benefits of exercise in the presence amyloid deposition.

A fundamental function of growth factor signaling is stimulating anabolic metabolism[49], which may or may not lead to cellular proliferation or organellar growth. Our results suggest that anabolic resistance might be a prevalent feature of the aging brain, contributing to cognitive decline and the pathogenesis of AD, but potentially mitigated by exercise. Thus, the opposing dynamics of growth and senescence could explain the inverse correlation between cancer and AD observed in the elderly human population [50].

REFERENCES

1. McGurran, H., Glenn, J. M., Madero, E. N. & Bott, N. T. Prevention and Treatment of Alzheimer's Disease: Biological Mechanisms of Exercise. *J. Alzheimers Dis. JAD* **69**, 311–338 (2019).
2. Mattson, M. P., Moehl, K., Ghena, N., Schmaedick, M. & Cheng, A. Intermittent metabolic switching, neuroplasticity and brain health. *Nat. Rev. Neurosci.* **19**, 63–80 (2018).
3. Voss, M. W. *et al.* Exercise and Hippocampal Memory Systems. *Trends Cogn. Sci.* (2019) doi:10.1016/j.tics.2019.01.006.
4. Fernandes, J., Arida, R. M. & Gomez-Pinilla, F. Physical exercise as an epigenetic modulator of brain plasticity and cognition. *Neurosci. Biobehav. Rev.* **80**, 443–456 (2017).
5. Prakash, R. S., Voss, M. W., Erickson, K. I. & Kramer, A. F. Physical activity and cognitive vitality. *Annu. Rev. Psychol.* **66**, 769–797 (2015).
6. Hanson, L. R., Fine, J. M., Svitak, A. L. & Faltesek, K. A. Intranasal administration of CNS therapeutics to awake mice. *J. Vis. Exp. JoVE* 4440 (2013) doi:10.3791/4440.
7. Vogel-Ciernia, A. & Wood, M. A. Examining object location and object recognition memory in mice. *Curr. Protoc. Neurosci.* **69**, 8.31.1-17 (2014).
8. Zhou, W. *et al.* Loss of function of NCOR1 and NCOR2 impairs memory through a novel GABAergic hypothalamus–CA3 projection. *Nat. Neurosci.* **22**, 205–217 (2019).
9. Shipton, O. A. *et al.* Left-right dissociation of hippocampal memory processes in mice. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 15238–15243 (2014).
10. Ohno, Y., Murphy, R., Choi, M., Ou, W. & Sumbria, R. K. Full- versus Sub-Regional Quantification of Amyloid-Beta Load on Mouse Brain Sections. *J. Vis. Exp. JoVE* (2022) doi:10.3791/63669.
11. Yao, Z. *et al.* A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184**, 3222-3241.e26 (2021).

744   12. Morabito, S. *et al.* Single-nucleus chromatin accessibility and transcriptomic characterization of
745   Alzheimer's disease. *Nat. Genet.* **53**, 1143–1155 (2021).
746   13. Ge, S., Wang, H., Alavi, A., Xing, E. & Bar-Joseph, Z. Supervised Adversarial Alignment of
747   Single-Cell RNA-seq Data. *J. Comput. Biol. J. Comput. Mol. Cell Biol.* **28**, 501–513 (2021).
748   14. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-
749   regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
750   15. Saito, T. *et al.* Single App knock-in mouse models of Alzheimer's disease. *Nat. Neurosci.* **17**,
751   661–663 (2014).
752   16. Masuda, A. *et al.* Cognitive deficits in single App knock-in mouse models. *Neurobiol. Learn.*
753   *Mem.* **135**, 73–82 (2016).
754   17. Guo, Q., Wang, Z., Li, H., Wiese, M. & Zheng, H. APP physiological and pathophysiological
755   functions: insights from animal models. *Cell Res.* **22**, 78–89 (2012).
756   18. Justice, N. J. *et al.* Posttraumatic stress disorder-like induction elevates β-amyloid levels, which
757   directly activates corticotropin-releasing factor neurons to exacerbate stress responses. *J.*
758   *Neurosci. Off. J. Soc. Neurosci.* **35**, 2612–2623 (2015).
759   19. Duclot, F. & Kabbaj, M. The Role of Early Growth Response 1 (EGR1) in Brain Plasticity and
760   Neuropsychiatric Disorders. *Front. Behav. Neurosci.* **11**, 35 (2017).
761   20. Ramsay, R. G. & Gonda, T. J. MYB function in normal and cancer cells. *Nat. Rev. Cancer* **8**,
762   523–534 (2008).
763   21. Assali, A., Harrington, A. J. & Cowan, C. W. Emerging roles for MEF2 in brain development and
764   mental disorders. *Curr. Opin. Neurobiol.* **59**, 49–58 (2019).
765   22. Dennis, D. J., Han, S. & Schuurmans, C. bHLH transcription factors in neural development,
766   disease, and reprogramming. *Brain Res.* **1705**, 48–65 (2019).
767   23. Imayoshi, I. & Kageyama, R. bHLH factors in self-renewal, multipotency, and fate choice of
768   neural progenitor cells. *Neuron* **82**, 9–23 (2014).
769   24. Hastie, N. D. Wilms' tumour 1 (WT1) in development, homeostasis and disease. *Dev. Camb.*
770   *Engl.* **144**, 2862–2872 (2017).
771   25. Guillemot, F. & Hassan, B. A. Beyond proneural: emerging functions and regulations of
772   proneural proteins. *Curr. Opin. Neurobiol.* **42**, 93–101 (2017).
773   26. Consalez, G. G., Goldowitz, D., Casoni, F. & Hawkes, R. Origins, Development, and
774   Compartmentation of the Granule Cells of the Cerebellum. *Front. Neural Circuits* **14**, 611841
775   (2020).
776   27. Stevanovic, M. *et al.* SOX Transcription Factors as Important Regulators of Neuronal and Glial
777   Differentiation During Nervous System Development and Adult Neurogenesis. *Front. Mol.*
778   *Neurosci.* **14**, 654031 (2021).
779   28. Salta, E. *et al.* Adult hippocampal neurogenesis in Alzheimer's disease: A roadmap to clinical
780   relevance. *Cell Stem Cell* **30**, 120–136 (2023).
781   29. Chen, W.-T. *et al.* Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease.
782   *Cell* **182**, 976-991.e19 (2020).
783   30. Zhou, Y. *et al.* Molecular landscapes of human hippocampal immature neurons across lifespan.
784   *Nature* **607**, 527–533 (2022).
785   31. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).
786   32. Quintela-López, T. *et al.* Aβ oligomers promote oligodendrocyte differentiation and maturation
787   via integrin β1 and Fyn kinase signaling. *Cell Death Dis.* **10**, 445 (2019).
788   33. Ferreira, S. *et al.* Amyloidosis is associated with thicker myelin and increased
789   oligodendrogenesis in the adult mouse brain. *J. Neurosci. Res.* **98**, 1905–1932 (2020).
790   34. Lemmon, M. A., Schlessinger, J. & Ferguson, K. M. The EGFR family: not so prototypical
791   receptor tyrosine kinases. *Cold Spring Harb. Perspect. Biol.* **6**, a020768 (2014).
792   35. Fukatsu, Y. *et al.* Muscle-specific overexpression of heparin-binding epidermal growth factor-like
793   growth factor increases peripheral glucose disposal and insulin sensitivity. *Endocrinology* **150**,
794   2683–2691 (2009).
795   36. Lovrić, A. *et al.* Single-cell sequencing deconvolutes cellular responses to exercise in human
796   skeletal muscle. *Commun. Biol.* **5**, 1121 (2022).
797   37. Glotov, A. S. *et al.* RNA Sequencing of Whole Blood Defines the Signature of High Intensity

15

798   Exercise at Altitude in Elite Speed Skaters. *Genes* **13**, 574 (2022).
799 38. Mi, M. Y. *et al.* Plasma Proteomic Kinetics in Response to Acute Exercise. *Mol. Cell. Proteomics*
800   *MCP* **22**, 100601 (2023).
801 39. Pan, W. & Kastin, A. J. Entry of EGF into brain is rapid and saturable. *Peptides* **20**, 1091–1098
802   (1999).
803 40. Sun, S. *et al.* A single-cell transcriptomic atlas of exercise-induced anti-inflammatory and
804   geroprotective effects across the body. *Innov. Camb. Mass* **4**, 100380 (2023).
805 41. Methi, A. *et al.* A Single-Cell Transcriptomic Analysis of the Mouse Hippocampus After
806   Voluntary Exercise. *Mol. Neurobiol.* (2024) doi:10.1007/s12035-023-03869-9.
807 42. Voss, M. W., Vivar, C., Kramer, A. F. & van Praag, H. Bridging animal and human models of
808   exercise-induced brain plasticity. *Trends Cogn. Sci.* **17**, 525–544 (2013).
809 43. Vints, W. A. J., Levin, O., Fujiyama, H., Verbunt, J. & Masiulis, N. Exerkines and long-term
810   synaptic potentiation: Mechanisms of exercise-induced neuroplasticity. *Front. Neuroendocrinol.*
811   **66**, 100993 (2022).
812 44. Jin, K. *et al.* Cerebral neurogenesis is induced by intranasal administration of growth factors.
813   *Ann. Neurol.* **53**, 405–409 (2003).
814 45. Scafidi, J. *et al.* Intranasal epidermal growth factor treatment rescues neonatal brain injury.
815   *Nature* **506**, 230–234 (2014).
816 46. da Rocha, J. F. *et al.* APP Binds to the EGFR Ligands HB-EGF and EGF, Acting Synergistically
817   with EGF to Promote ERK Signaling and Neuritogenesis. *Mol. Neurobiol.* **58**, 668–688 (2021).
818 47. Tavassoly, O., Sato, T. & Tavassoly, I. Inhibition of Brain Epidermal Growth Factor Receptor
819   Activation: A Novel Target in Neurodegenerative Diseases and Brain Injuries. *Mol. Pharmacol.*
820   **98**, 13–22 (2020).
821 48. Mansour, H. M., Fawzy, H. M., El-Khatib, A. S. & Khattab, M. M. Repurposed anti-cancer
822   epidermal growth factor receptor inhibitors: mechanisms of neuroprotective effects in
823   Alzheimer's disease. *Neural Regen. Res.* **17**, 1913–1918 (2022).
824 49. Thompson, C. B. & Bielska, A. A. Growth factors stimulate anabolic metabolism by directing
825   nutrient uptake. *J. Biol. Chem.* **294**, 17883–17888 (2019).
826 50. Shafi, O. Inverse relationship between Alzheimer's disease and cancer, and other factors
827   contributing to Alzheimer's disease: a systematic review. *BMC Neurol.* **16**, 236 (2016).
828

ACKNOWLEDGEMENT

AUTHOR CONTRIBUTIONS

XL performed most behavioral assays and amyloid histology analysis; CL performed integrated snRNA-seq/snATAC-seq analysis; WL executed the exercise protocol, isolated nuclei, and performed RNAscope analysis; YD performed the initial snRNA-seq analysis; CG performed the spatial transcriptomics-related analysis; WZ and JL performed some of the behavioral assays; VCC assisted in preparing figures and language editing; HZ provided and advised on the mouse model; UK, DG, ZH, HC assisted or advised on data analyses or data interpretation. ZL guided bioinformatics analysis; YW supervised bioinformatics analysis; ZS and ZL conceived the study and obtained funding; YW and ZS wrote the manuscript with input from other authors.

CONFLICT OF INTEREST
The authors declare no financial conflict of interest.

CONSENT STATEMENT
No human subjects were involved. Consent was not necessary. The publication of the described work is approved by all authors.

KEYWORDS
Physical exercise, cognition, hippocampus, growth factor, Alzheimer's disease

Figure 1. Exercise improves memory without altering the overall hippocampal cellular composition. Discrimination index and exploration time of the standard object-in-place memory test (**a-c**) and the Y-maze test (**d-f**) in male WT and APP^{NL-G-F} (APP) mice at 10 months old after 6 months wheel-running exercise (EX) or rest (RT) (n = 12 for WT_RT; 8 for WT_EX; 10 for APP_RT; and 9 for APP_EX). Asterisks indicate significant differences with the two-way ANOVA and Fisher's LSD multiple comparisons test. (**g-i**) Discrimination index and exploration time of a modified 5-object in-place test in 4-month-old male WT mice after 2 months wheel-running (n = 12 for RT, 13 for EX). Bar graphs show the mean with S.E.M. Asterisks indicate significant differences by 2-sided t-test. (**j**) Violin plot of cell type-specific marker gene expression levels. (**k**) UMAP of major cell types of the hippocampus based on snRNA-seq and snATAC-seq datasets. Excitatory neurons (EN), inhibitory neurons (IN), microglia (MG), astrocytes (AST), oligodendrocytes (OLG), and oligodendrocyte progenitor cells (OPC). APP/Exercise (APP_EX), APP/Rest (APP_RT), WT/Exercise (WT_EX), WT/Rest (WT_RT).

Figure 2. Exercise counteracts amyloid-dependent transcriptomic changes. (**a**) Scatter plot of gene expression showing a negative correlation between exercise effects (APP_EX vs. APP_RT) and amyloid effects (APP_RT vs. WT_RT) across different cell types. (**b**) Heat map of 833 differentially expressed genes (DEGs) with the reversed pattern in at least one cell type. (**c**) Number of overlapping reversed DEGs across different cell types. (**d**) Overlapping pathways enriched in reversed DEGs across different cell types. (**e-f**) GSEA analysis of the insulin signaling pathway, a top common enriched pathway in reversed genes in different cell types.

Figure 3. Exercise activates the EGFR/insulin signaling. (**a**) Heat map of reversed DEGs within the insulin signaling in different cell clusters. (**b-c**) RNAscope analysis of Stxbp1, a gene of the insulin/EGFR pathway with known function in neurotransmission and a robust reversed expression pattern in EN. Scale bar, 200 µm. n = 3 mice. Bar graphs show the mean with S.E.M. Asterisks indicate significant differences by t-test. ns, non-significant. (**d**) Some of the DEGs in the EGFR pathway. (**e**) Some of the DEGs in the insulin pathway. Images were generated from the KEGG pathway database.

Figure 4. Exercise stimulates cell type-specific transcriptional regulatory networks. (**a**) Heat map of the relative levels of the reversed differentially accessible regions (DARs) in each cluster. DARs upregulated by APP^{NL-G-F} (APP_RT vs. WT_RT) and downregulated by exercise (APP_EX vs. APP_RT) were referred to as "U>D", while those DARs downregulated by APP^{NL-G-F} and upregulated by exercise were referred to as "D>U". (**b**) Top enriched motifs in the DARs in each cell cluster. (**c-h**) Top network showing direction-specific enrichment in EN, IN, and OLG populations.

Figure 5. Exercise impacts excitatory neurons and oligodendrocytes. (**a-b**) UMAP of snRNA-seq data in EN sub-clusters: Prox1+ granule cells and Ociad2+ pyramidal cells. Expression is based on library-size-normalized log values. (**c**) UMAP of hippocampal EN sub-clusters in each individual group. (**d**) The proportion of granule and pyramidal cells within the EN cluster in each group. (**e**) Proportions of granule cells and INs in the DG. (**f**) Box plot of predicted GC immature score (imGC).

17

906 514 of 534 genes in the model were found in our data, and missing genes had only small weights in
907 the model. The log-transformed and max-normalized counts matrix were taken as the input to
908 predict the final imGC score from the logistic regression model. Center line, median; box limits,
909 upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers. Asterisks indicate
910 significant differences. ns, non-significant. (**g**) UMAP of the OLG for trajectory analysis. (**h**) Module
911 score along the pseudotime trajectory of the OLG. (**i**) Distribution of each group on the OLG
912 trajectory.

914 **Figure 6. EGFR and PI3K signaling are required for the cognitive-improving effects of**
915 **exercise.** (**a**) Average actogram profiles of wheel-running activity of 8-month-old male mice treated
916 with EGFR inhibitor Gefitinib and PI3K inhibitor Wortmannin through oral gavage once every other
917 day from 4 to 8 months old. Mice were simultaneously subjected to wheel-running from 4 to 8
918 months old. (**b**) Average daily wheel-running activity (n = 7 cages per group with 2 mice per cage).
919 (**c**) Body weight. (**d-o**) Object-in-place, Y-maze, 3-chamber sociability and social memory, Morris
920 water maze, elevated plus maze, open field arena, and light-dark tests (n = 15 mice for RT, 15 mice
921 for EX, and 16 mice for EX + inhibitor). 2 RT, 2 EX and 1 EX+inhibitor mice were excluding due to
922 discrimination index were greater or smaller than ± 0.7 (**p-q**) Immunostaining of β-amyloid in mice
923 treated with inhibitors. Scale bar: 600 μm. n = 6 mice per group. All bar graphs show the mean with
924 S.E.M. Asterisks indicate significant differences by one-way ANOVA with Fisher's LSD multiple
925 comparisons except Morris water maze was analyzed with repeated-measure 2-way ANOVA with
926 Fisher's LSD multiple comparisons where asterisks indicate differences between RT vs. EX and EX
927 + inhibitor vs. EX.

929 **Figure 7. Intranasal HB-EGF mimics exercise-induced cognitive improvement.** (**a**) Replot of
930 gene expression levels of EGF family members in human skeletal muscles after long-term exercise
931 training from a published transcriptomic dataset[36]. (**b**) Replot of protein levels of EGF family
932 members in the human blood after long-term exercise training from a published proteomics
933 dataset[38]. (**c**) Serum HB-EGF levels in APP^{NL-G-F} male mice after chronic wheel-running exercise for
934 6 months (n = 11 mice for RT, and 11 mice for EX). (**d**) Body weight gain during intranasal HB-EGF
935 administration in female APP^{NL-G-F} mice. Administration started at 4 months old, with once every
936 other day (n = 13 mice for the vehicle; 13 mice for HB-EGF). (**e-o**) Object-in-place, Y-maze, 3-
937 chamber sociability and social memory, Morris water maze test, open field arena test, and light-dark
938 test in female mice at 7-8-months old after chronic HB-EGF administration (n = 13 mice for the
939 vehicle; 13 mice for HB-EGF; one mouse from the HB-EGF group was excluding from MWM due to
940 low mobility). (**p-q**) Immunostaining of β-amyloid in 8-month-old female mice treated with intranasal
941 HB-EGF for 4 months. Scale bar: 600 μm. n = 6 mice per group. All bar graphs show the mean with
942 S.E.M. Asterisks indicate significant differences by one-way ANOVA with Fisher's LSD multiple
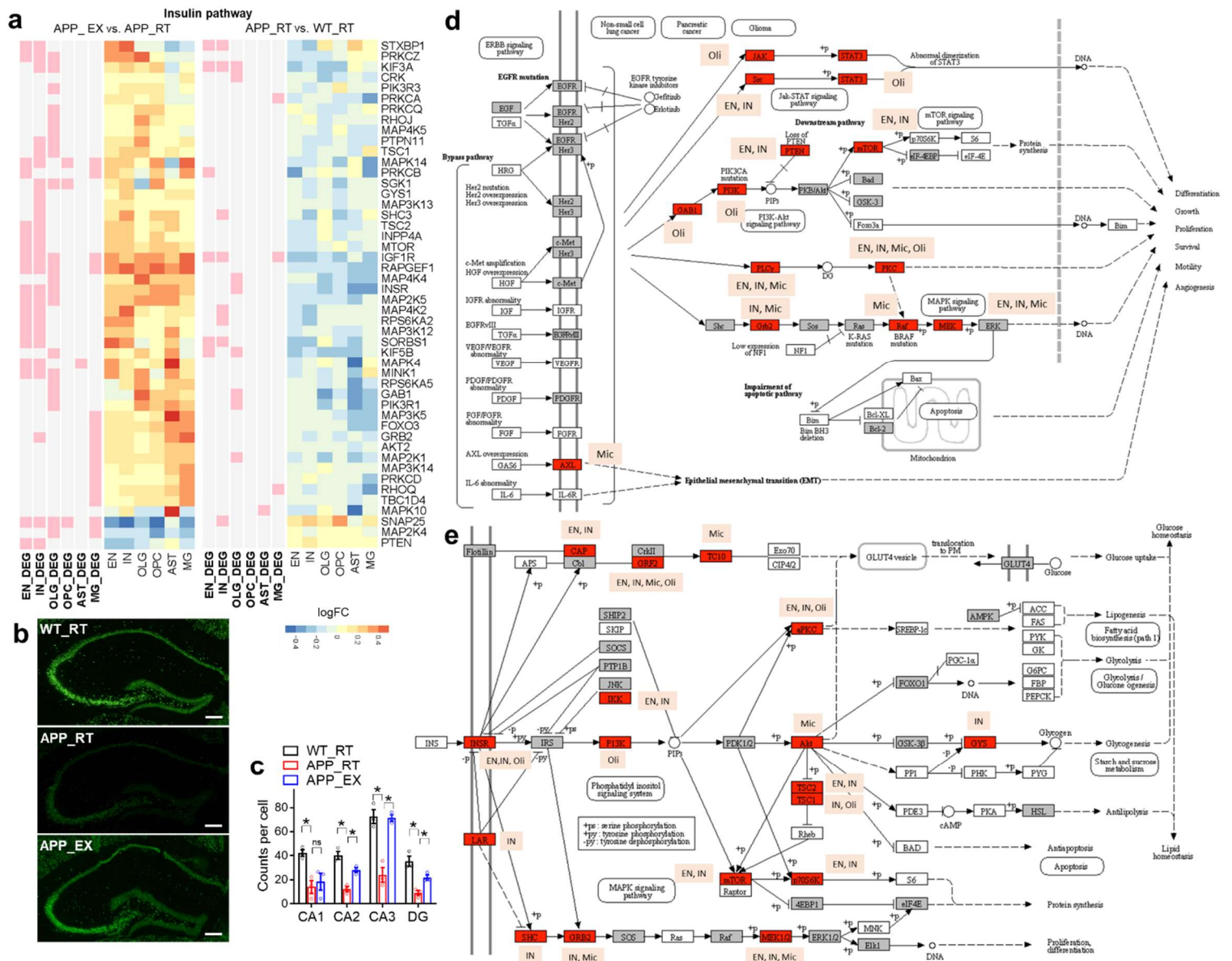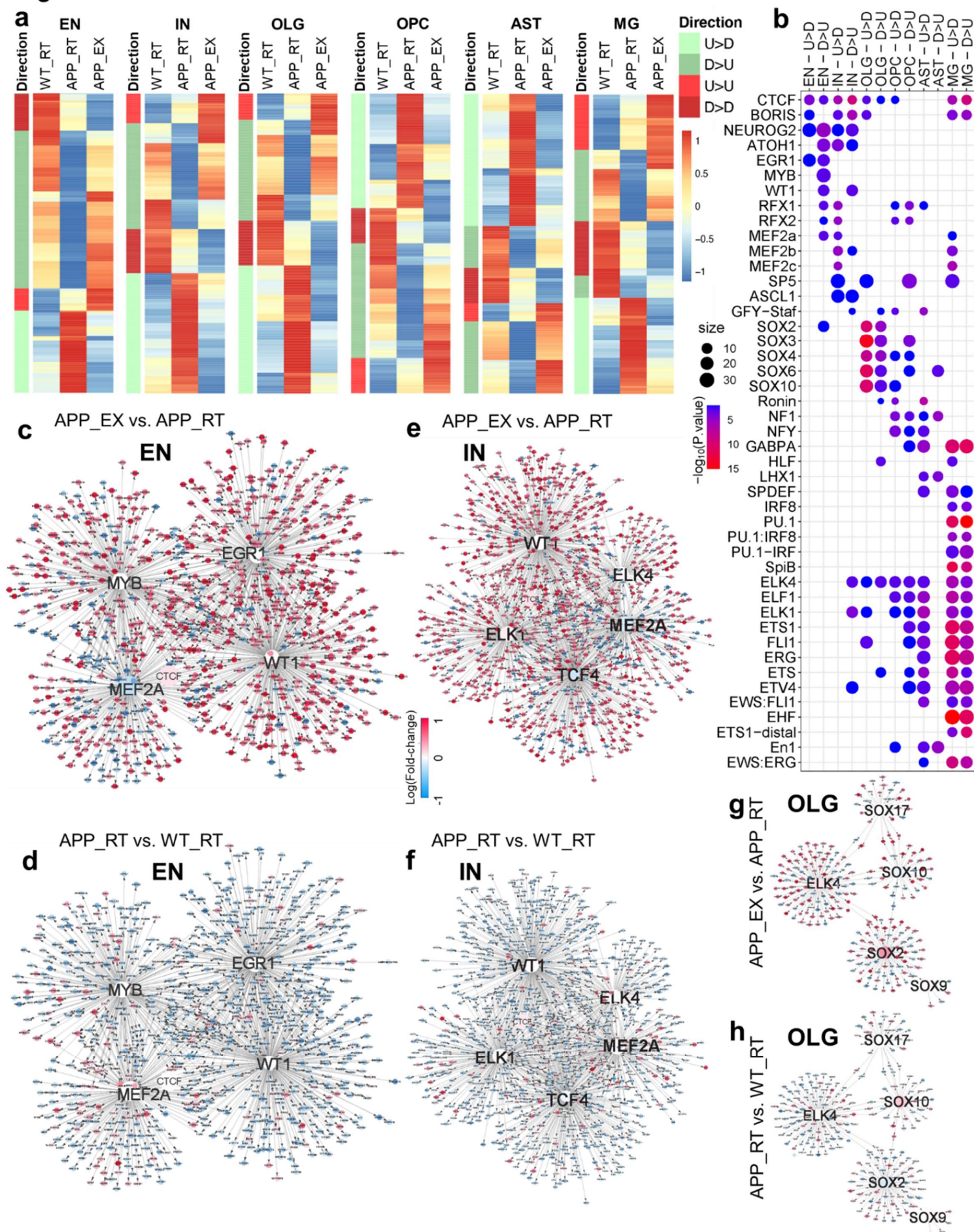943 comparisons or repeated-measure ANOVA with Fisher's LSD multiple comparisons.
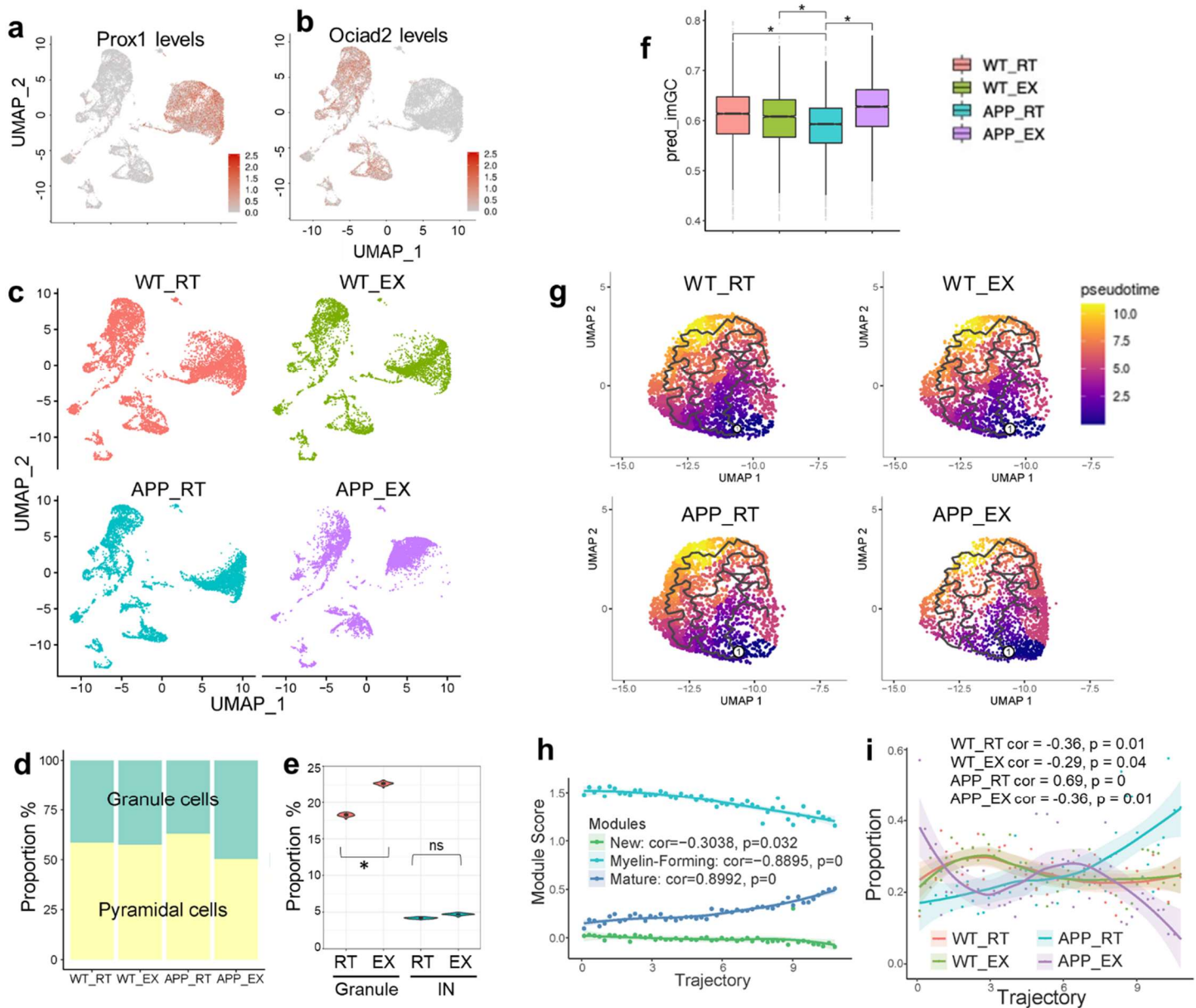
Figure 1
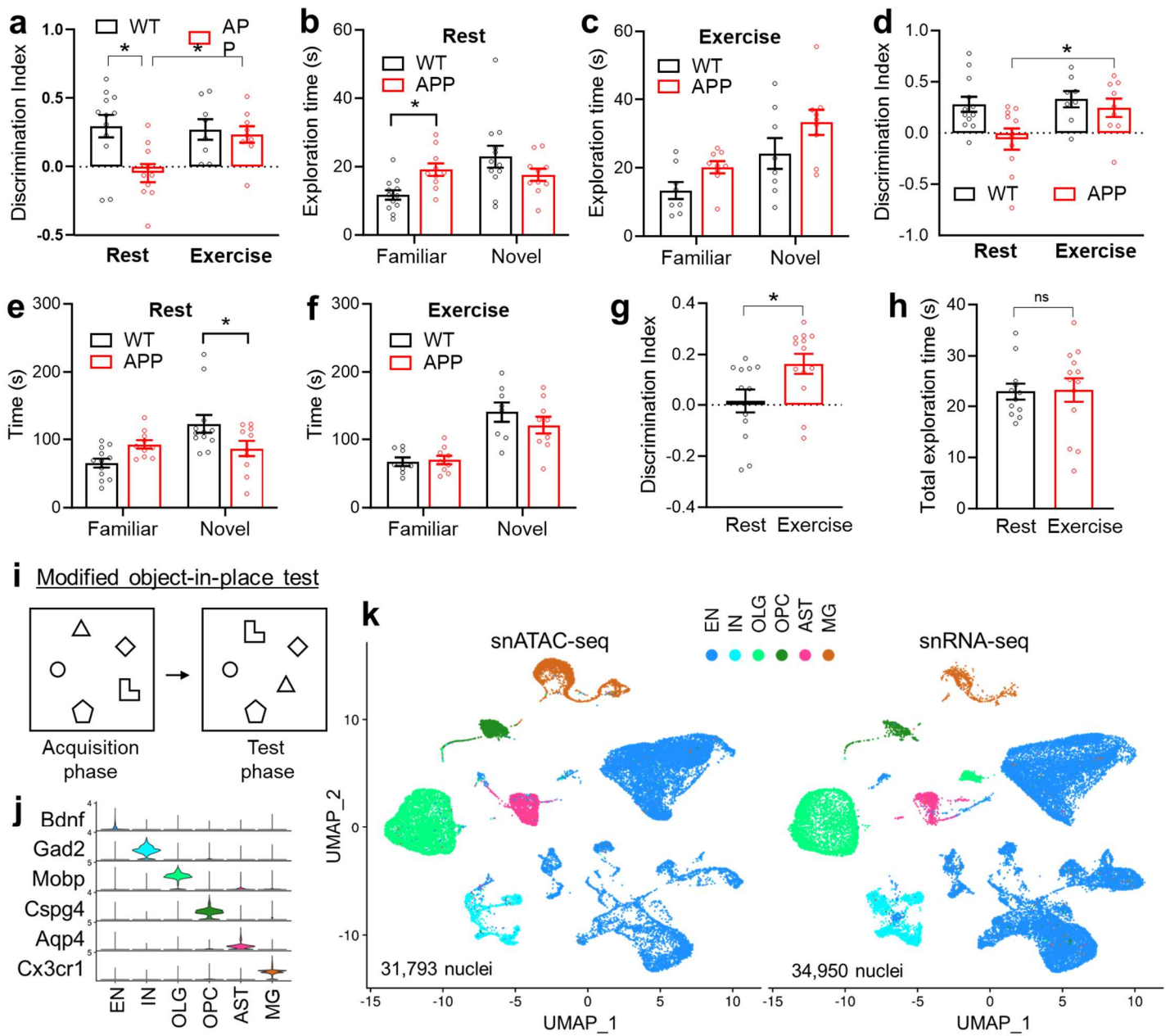
Figure 2
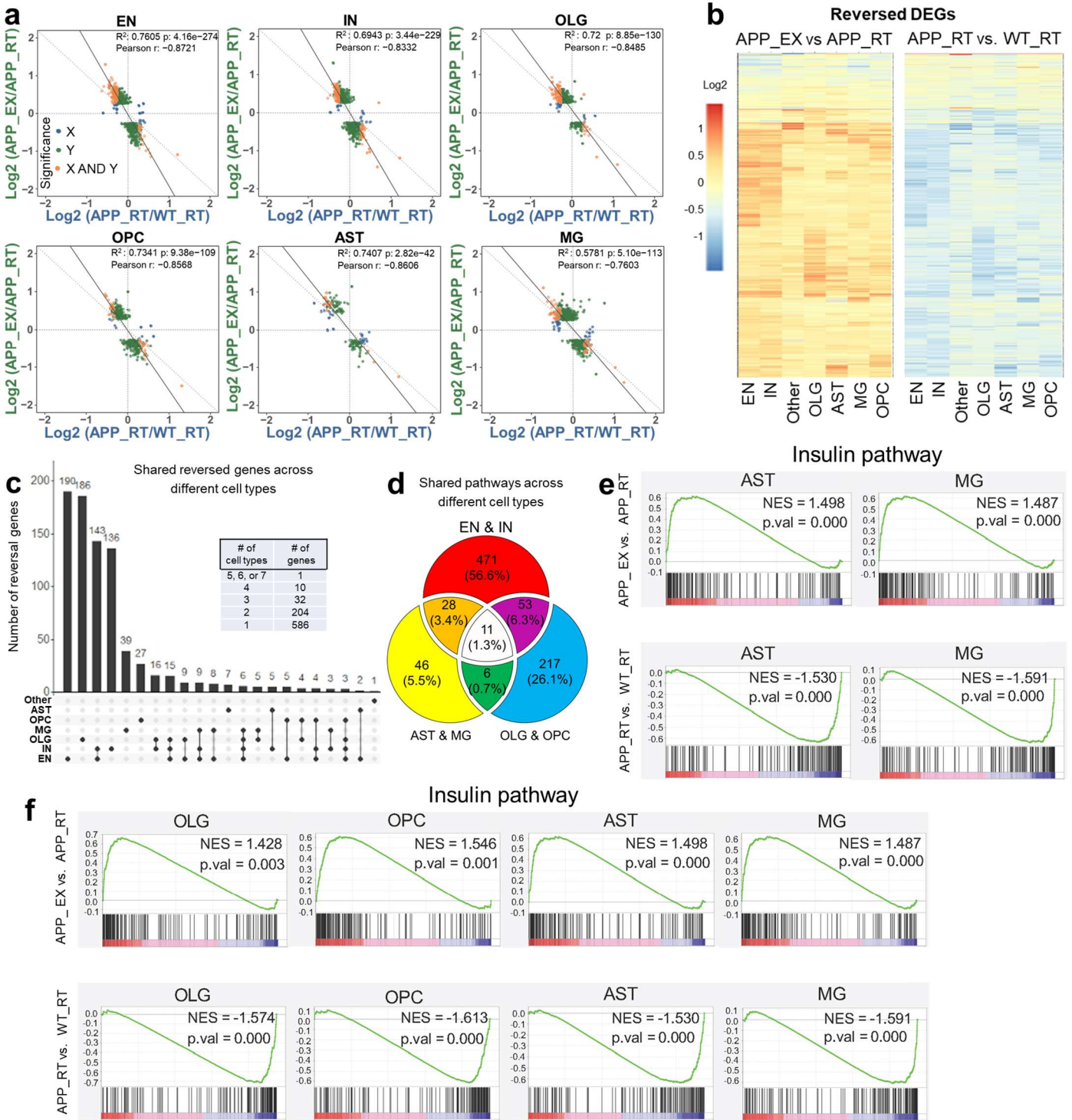
Figure 3

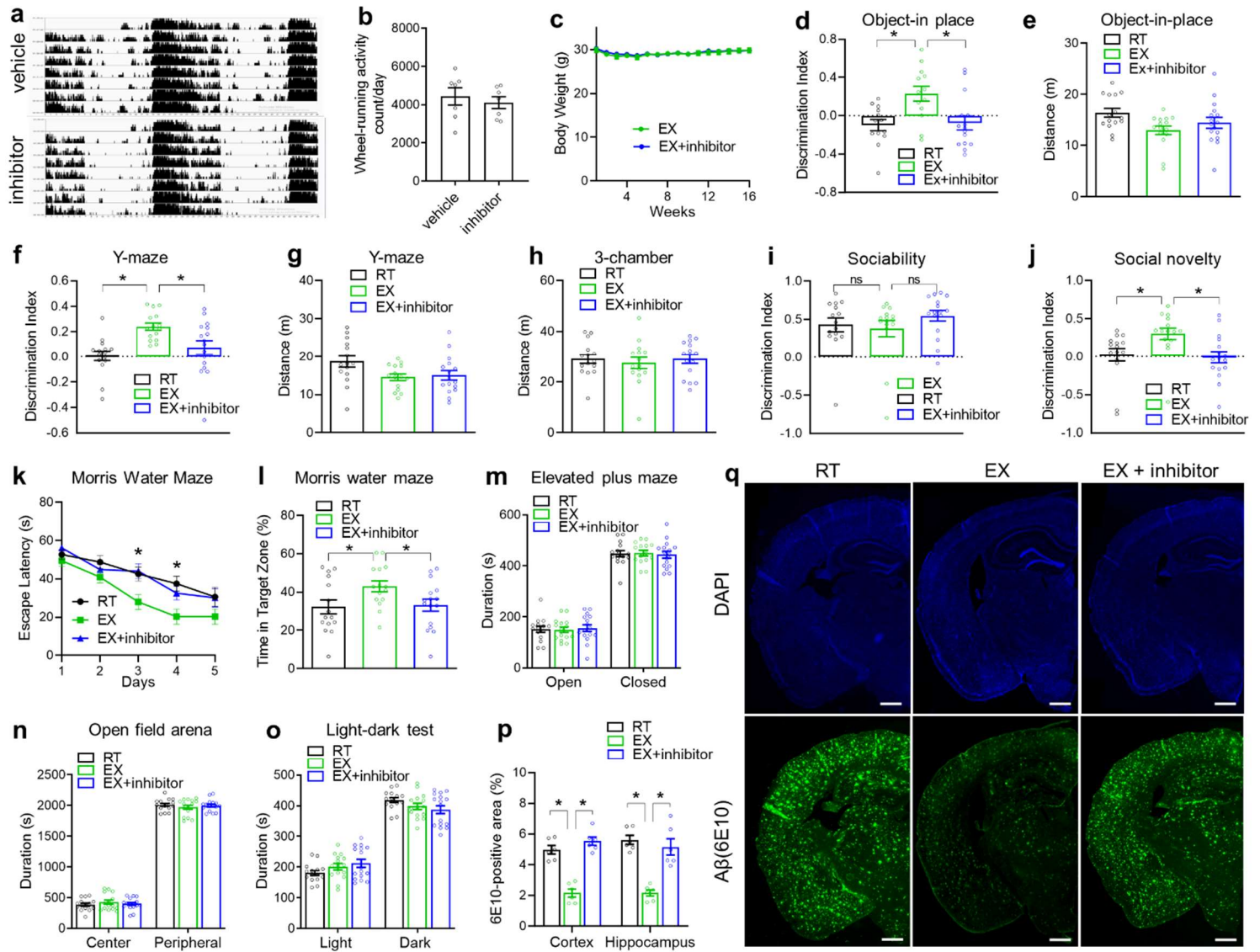# Figure 4

Figure 5

Figure 6

Figure 7