

Supplementary materials for

Spatiotemporal characterization of disease-associated neurons in the

Entorhinal Cortex-Hippocampal Circuit during AD progression

Materials and methods

Animals

Male C57BL/6J WT mice were purchased from Charles River Laboratories. Male APP/PS1 mice were originally obtained from the Jackson Laboratory. Animals were bred and housed under specific-pathogen-free conditions with a 12-h dark–light cycle at 21 ± 1 °C and 55–60% humidity, with ad libitum access to water and food. All experimental protocols were approved by the Animal Studies Committee at University of Science and Technology, Hefei, China.

Fluorescence-activated cell sorting (FACS)

WT and APP/PS1 mice were perfused transcardially with PBS, and the entorhinal cortex (EC), hippocampal CA1, and CA3 regions were isolated. Tissues were homogenized in a lysis buffer containing 250 mM sucrose, 25 mM KCL, 5 mM MgCl₂, 10 mM Tris buffer, pH 8.0, 1 μM DTT, protease inhibitors, 0.4U/μl RNase Inhibitor, 0.1% Triton X-100. Brain homogenates were passed through a 74 μm cell strainer to remove cell debris, and nuclei were obtained by centrifugation at 1,000 g for 8 min at 4 °C. Nuclei were resuspended in 500–1,000 μL of staining buffer (RNase-free PBS pH 7.4, 0.5% (wt/vol) RNase-free BSA, 0.2 U/μL of RNase inhibitor). After blocking with PBS containing 0.5% BSA on ice for 15 min, the nuclei were washed with PBS and incubated with an anti-NeuN antibody at 4°C for 1 h, followed by an Alexa Fluor 488–conjugated secondary antibody. Total nuclei were then filtered using a 40 μm cell strainer and sorted by a Beckman Coulter MoFlo Astrios EQ sorter to isolate neuronal nuclei. A gating control sample was prepared for each experiment to set gates for the Alexa 488 channel. The sorted neuronal nuclei were suspended in a lysis buffer (0.2% Triton X-100, 2 U μ/L of RNase inhibitor), snap-frozen in liquid nitrogen, and stored at -80°C until further analysis.

Immunofluorescence staining

Mice were perfused transcardially with PBS, and brains were isolated and fixed in 4% paraformaldehyde (PFA), followed by cryopreservation in 20% for 24 h and 30% sucrose for another 24 h at 4°C. Tissues were embedded in OCT and sectioned at 40 μm thickness using a Leica microtome. Brain sections were washed with PBS for 5 min at RT, and permeabilized with 1% Triton X-100 in PBS for 10 min. After blocking in PBS containing 2% BSA and 0.5% TritonX-100 for 30 min at RT, sections were washed with PBS. Sections were then incubated with primary antibodies overnight at 4°C, followed by incubation with Alexa 488 and/or Alexa 594-labeled secondary antibodies (Invitrogen). Images were acquired using a Leica TCS SPE confocal microscope equipped with a DFC 365 FX Digital Camera.

RNA extraction, Reverse transcription, and Quantitative PCR

Total RNA was extracted from the entorhinal cortex of APP/PS1 or WT mice using Trizol (Invitrogen), followed by DNase I treatment to remove genomic DNA contamination. Reverse transcription was performed using Hiscrypt II Reverse Transcriptase (Vazyme) according to the manufacture's instruction. Quantitative PCR was performed with AceQ™ qPCR SYBR Green Master Mix (Vazyme) on a LightCycler 96 system (Roche) following standard procedures. The measured value for each sample was averaged and compared using CT method, where the amount of target RNA ($2^{-\Delta\Delta CT}$) was normalized to an endogenous reference (ΔCT). Primers for Meg3

detection: CACAGAAGACGAAGAGCTGGA (forward) and GGTAGAGGTGCACAGCAGGT (reverse). Primers for 28S detection: GTTCACCCACTAATAGGGAACGTGA (forward) and GGATTCTGACTTAGAGGCGTTCAGT (reverse).

Data preprocessing

The raw sequencing data were processed for adapter trimming using Soapnuke_v1.1 (Chen et al., 2018), followed by the removal of low-quality reads and reads with a high proportion of N bases. Cleaned reads were aligned to the mouse genome (mm10-GRCm38-V10) or human APP and PSEN1 gene sequences using STAR_v2.5.3a (Dobin et al., 2013). Cells were included based on the following criteria: a unique alignment rate greater than 60%, more than 2,000,000 uniquely mapped reads per cell, and the detection of more than 2,000 genes per cell. This results in the inclusion of 1,663 cells and the exclusion of 47 cells. Post-alignment, the quantified reads were normalized to transcripts per million (TPM) to correct gene length biases in transcript counts.

Differential Expression Analysis

The differentially gene expression analysis was conducted using the Wilcoxon rank sum test in the R package coin_v1.4-2 (Hothorn et al., 2008), with control for the false discovery rate (FDR). A total of 48,442 genes were analyzed for differential expression between AD and WT samples across different brain regions, age groups, and neuronal subtypes. Genes with adjusted p-values (padj) less than 0.05 were considered differentially expressed. To ensure the comparability, an equal number of cells from different brain regions and age groups were used for differential expression analysis. For analyzing gene expression across neuronal subtypes, a total of 100 cells (50 from WT and 50 from AD brains) were randomly sampled from each subtype. This resampling process was repeated five times, and the average number of differentially expressed genes from these iterations were reported.

Signature genes for each neuronal subtype were selected using the FindAllMarkers function in R package Seurat (Satija et al., 2015). We performed differential expression analysis for 48,442 genes within the indicated neuronal subtype relative to other neuronal types, genes with a logFC > 1 and padj < 0.05 were selected as signature genes.

Clustering analysis

Integrated Clustering analysis was conducted to reduce temporal batch effect using FindIntegrationAnchors in the R package Seurat v4.0.2 (Satija et al., 2015). The top 2,000 most variable genes were selected based on the calculated mean variance using the FindVariableFeatures function. Principal component analysis (PCA) was conducted to reduce the datasets dimensionality using the RunPCA function. The top 22 principal components were selected based on clustering effectiveness to preserve the primary sources of heterogeneity. Clustering analysis was performed using the FindClusters function at a resolution of 0.4. The results were visualized in a two-dimensional plane using tSNE (t-distributed random neighbor embedding). The FindAllMarkers function was used to identify characteristic genes for the clustered neuronal subtypes. Clusters were manually annotated based on the expression of known marker genes.

Pseudo-temporal analysis of cell lineage trajectories

Trajectory analysis was conducted to investigate the molecular characteristic changes in Gfap+ neurons under pathological conditions. Pseudotime analysis was conducted using Monocle v2.18.0 (Trapnell et al., 2014), which employs an unsupervised algorithm to rank individual neurons based on their expression characteristics within cell populations. States were identified by performing hierarchical clustering on differentially expressed genes using the expression values from all ordered neurons. Each neuron was treated as a point in a high-dimensional state space, and Monocle reconstructed the trajectory of state transitions along pseudotime. Cell proliferation and differentiation shall not be taken into accounts in this analysis, as neurons are terminally differentiated cells.

Pathway analysis

Pathway analysis for various neuronal subpopulations was performed using the online platform Metascape (Zhou et al., 2019) (<http://metascape.org/>), with default settings. Signature genes specific to each neuronal subgroup were selected as input for pathway enrichment analysis. Pathways with $p < 0.01$ were considered significantly enriched functional pathways in the indicated neuronal subtype compared to other subtypes, and were classified as function-specific pathways for that neuronal subtype.

Pathway analysis for the DEGs in APP/PS1 relative to WT neurons across different ages, brain regions, and neuronal subgroups was conducted using GSEA_v4.0.3 (Subramanian et al., 2005). Mouse-specific gene sets (GMT) were obtained from the MSigDB (Molecular Signatures Database v7.5.1, available at: <https://www.gsea-msigdb.org/gsea/msigdb/>) and constructed for GSEA. Gene sets with sizes ranging from 15 to 500 genes were selected, with other parameters set to default values. In generalized GSEA, a gene set was considered significantly enriched if the normalized enrichment score (NES) exceeded the threshold (absolute value of $NES > 1$). Specifically, $NES > 1$ indicated upregulation of the functional pathway, while $NES < 1$ suggested downregulation of the pathway.

WGCNA co-expression module analysis

Gene co-expression networks were constructed using the WGCNA_v1.70-3 package in R (Zhang and Horvath, 2005). Gene sets comprising the top 20% of genes with the highest standard deviation were selected for this analysis. The pickSoftThreshold function was employed to determine the optimal soft power value (softpower=6) for calculating the adjacency matrix. The cutreeDynamic function was used for module identification with parameters deepSplit=2 and minModuleSize=30. Modules with a MEDissThres = 0.6 were merged if they were closely correlated. An adjacency matrix was computed using a similarity function based on signed networks and a soft threshold power of 6.

Module membership (kME) was calculated as the average module membership for all genes within a given module based on their adjacency scores. The module membership (MM) was calculated using the signedKME function in WGCNA, which correlated the ME with gene expression values, quantifying how close a gene was to a given module. The correlation between individual genes and biological traits was defined as gene significance (GS). Generally, if GS and MM were highly correlated, it implied that genes were the highly important elements within the modules and were most significantly associated with the biological trait.

AUCscore calculation

AUCCell was used to calculate the AUC scores for genes or gene sets (Aibar et al., 2017). AUCCell is an R package designed for analyzing the state of gene sets in single-cell RNA-seq data to identify cells with active gene signatures. AUC scores were used to quantify gene expression activity.

Homologous gene conversion between human and mouse

Metascape was used for homologous gene conversion between human and mouse gene sets (Zhou et al., 2019). This online tool integrates multiple databases, including Homologene, OrthoMCL, and EggNOG, to facilitate gene comparison across species. For more details, visit the Metascape website at <https://metascape.org/blog/?p=203>.

References

- AIBAR, S., GONZÁLEZ-BLAS, C. B., MOERMAN, T., HUYNH-THU, V. A., IMRICHOVA, H., HULSELMANS, G., RAMBOW, F., MARINE, J.-C., GEURTS, P., AERTS, J., VAN DEN OORD, J., ATA, Z. K., WOUTERS, J. & AERTS, S. 2017. SCENIC: single-cell regulatory network inference and clustering. *Nature Methods*, 14, 1083-1086.
- CHEN, Y., CHEN, Y., SHI, C., HUANG, Z., ZHANG, Y., LI, S., LI, Y., YE, J., YU, C. & LI, Z. 2018. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience*, 7, gix120.
- DOBIN, A., DAVIS, C. A., SCHLESINGER, F., DRENKOW, J., ZALESKI, C., JHA, S., BATUT, P., CHAISSON, M. & GINGERAS, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29, 15-21.
- HOTHORN, T., HORNIK, K., VAN DE WIEL, M. A. & ZEILEIS, A. 2008. Implementing a class of permutation tests: the coin package. *Journal of statistical software*, 28, 1-23.
- SATIJA, R., FARRELL, J. A., GENNERT, D., SCHIER, A. F. & REGEV, A. 2015. Spatial reconstruction of single-cell gene expression data. *Nature biotechnology*, 33, 495-502.
- SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., MUKHERJEE, S., EBERT, B. L., GILLETTE, M. A., PAULOVICH, A., POMEROY, S. L., GOLUB, T. R. & LANDER, E. S. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102, 15545-15550.
- TRAPNELL, C., CACCHIARELLI, D., GRIMSBY, J., POKHAREL, P., LI, S., MORSE, M., LENNON, N. J., LIVAK, K. J., MIKKELSEN, T. S. & RINN, J. L. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature biotechnology*, 32, 381-386.
- ZHANG, B. & HORVATH, S. 2005. A general framework for weighted gene co-expression network analysis. *Statistical applications in genetics and molecular biology*, 4.
- ZHOU, Y., ZHOU, B., PACHE, L., CHANG, M., KHODABAKHSHI, A. H., TANASEICHUK, O., BENNER, C. & CHANDA, S. K. 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications*, 10, 1523.

Extended Data Fig. 1 | Data profiles and molecular dynamics of disease-associated neurons.

(A) Distribution of identified neuronal clusters within different ages (left) or disease status (right), corresponding to Fig 1C.

(B) Up- and down-regulated functional pathways in neurons of APP/PS1 versus WT mice at the indicated ages and brain regions.

(C) Heatmap illustrating module-trait relationship.

(D) The greenyellow module shows a positive correlation with gene expression in EC-stellate cells and a negative correlation with gene expression in *Gfap*⁺ neurons.

(E) *Meg3* expression levels in the EC of WT and APP/PS mice at 6 months, measured by qPCR analysis (n=6 mice per group). Student *t* test ***; $p < 0.001$.

(F) Functional pathway enrichment analysis for top 200 genes negatively correlated with *Meg3* expression within the greenyellow module.

(G) The module yellow is highly correlated with *GFAP*⁺ neurons.

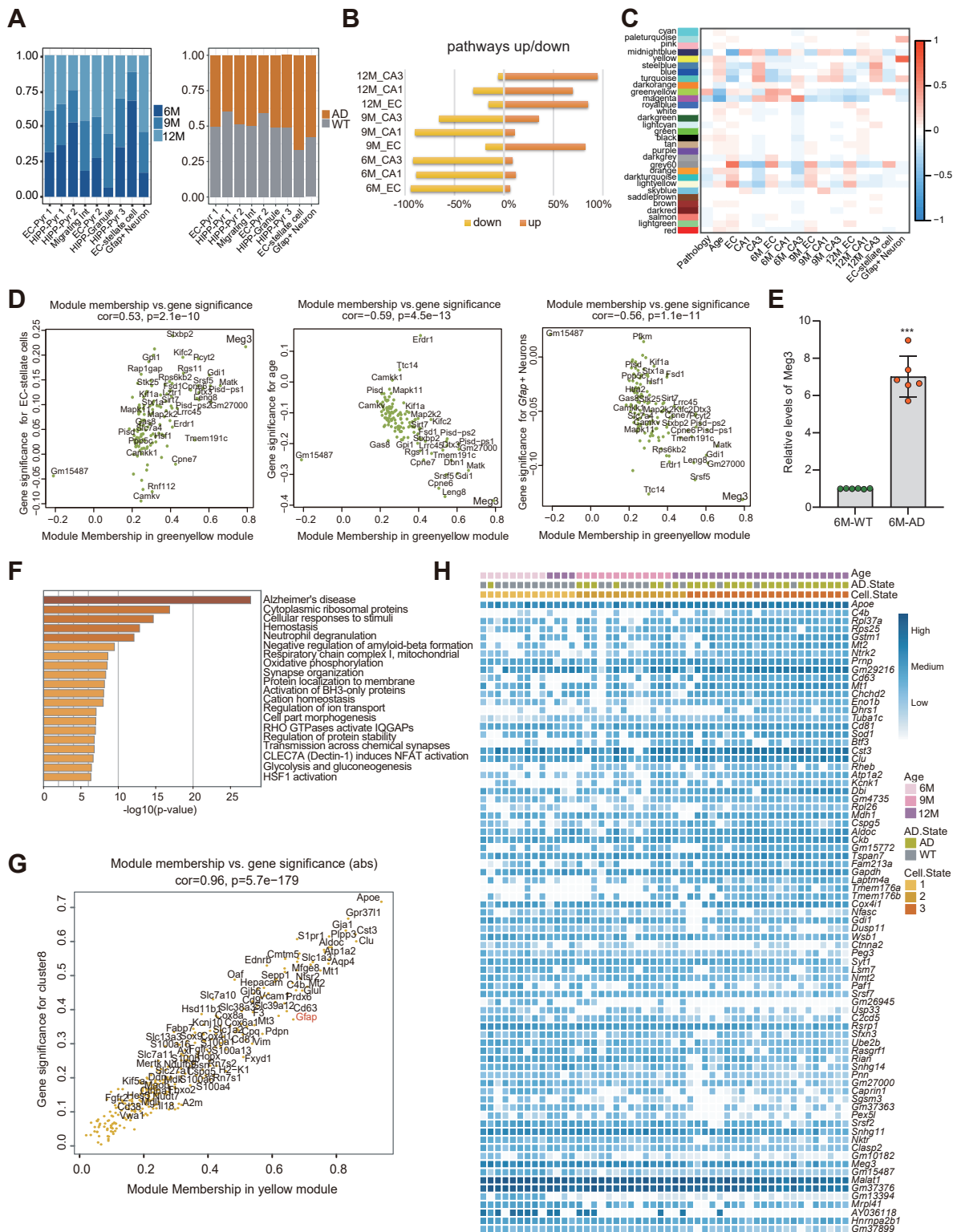
(H) Heatmap showing the top 40 genes associated with disease progression.

Extended Data Fig. 2 | *Meg3* expression across ages, brain regions, and neuronal subtypes in APP/PS1 and WT mice.

(A) *Meg3* expression in the indicated brain regions and ages of APP/PS1 versus WT mice (Wilcoxon signed-rank test; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

(B) *Meg3* expression in the indicated neuronal clusters of APP/PS1 versus WT mice (Wilcoxon signed-rank test; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

Extended Data Fig. 1



Extended Data Fig. 2

