nature research

Corresponding author(s): Manolis Kellis and Li-Huei Tsai

Last updated by author(s): Apr 4, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

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

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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	no software was used
Data analysis	CellRanger software (2.0.0 version) (10x Genomics) was used to align clean reads to the hg38 human genome (GRCh38.p5 (NCBI:GCA_000001405.20). The single-cell analysis package scanpy (version 0.4.4) implemented in python was used for clustering analyses. The R packages scran ((version 1.8.2) and scater ((version 1.8.1) were used for single-cell data manipulations and QC analyses. The R package Seurat (2.3.4) was used to compared reproducibility of markers. Bulk RNA-seq differential analyses we performed by fitting a linear model using the R package limma (version 3.38.3). The R package lme4 (1.1-21) was used to fit Poisson mixed models, the R package RUVseq (1.16.1) to remove unwanted variation from RNA data, and the R package metap (1.1) to compute aggregate p-values. All statistical analyses and visualizations were implemented in R (version 3.4.3). GSEA was applied to identify priori defined gene sets that show statistically significant differences between two given clusters. Gene Ontology (GO) enrichment analyses were performed using Metascape (version 3.0). Raw MPRAGE images were processed to generate total volumes including gray matter, white matter, CSF and intracranial volumes using SPM (version 12). White matter lesions appearing hyperintense in T2-weighted images were segmented based on FLAIR and MPRAGE data using BIANCA. All self-organizing maps were created using the Kohonen R package (version 3.0.8).

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

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The single-nucleus RNA-Sequencing data is available at Synapse (https://www.synapse.org/#!Synapse:syn18485175). The DOI for this dataset is: 10.7303/syn18485175. The data is available under controlled use conditions set by human privacy regulations.

Field-specific reporting

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

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

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	No explicit calculations were performed to determine sample size. Rather, we aimed to analyze brain tissue from an equal number of men and women and at least 12 individuals per group. Therefore we analyzed brain tissue from 24 individuals with low amyloid burden and 24 individuals with high amyloid burden. These numbers of samples was sufficient to perform a confident data analysis.
Data exclusions	Low quality snRNA-seq libraries were excluded and the exclusion criteria are described in the manuscript as follows. Cells with a high ratio of Mitochondrial (MT) relative to endogenous RNAs had low starting amounts of RNA, which might indicate that source cells were dead or stressed, resulting RNA degradation. Outlier cells in these quality metrics were found to cluster together in the tSNE 2D space . Based on these observations and subsequent scatter plot analyses, cells with less than 200 detected genes, and cells with abnormally high ratio of counts mapping to MT genes, relative to the total number of detected genes were removed. Specifically, given a highly skewed empirical distribution of the MT ratio values (i.e., having an elbow shape clearly separating high and low scores) outlier cells were classified in two groups using the k-means clustering algorithm (k=2) on the MT ratio, and subsequently removed.
Replication	Verification of the single-nucleus RNA-seq data was perfomred through validation using RNA in situ hybridization, quantitative RT-PCR, and immunohistochemistry on tissue derived from a subset of the individuals analyzed using snRNA-seq. These experiments validated the findings derived from snRNA-seq.
Randomization	The study participants were allocated into groups based on the overall amyloid level.
Blinding	Investigators were blinded to group allocation for the quantification of the RNAscope data.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Antibodies

Antibodies used	anti-Iba1 (Synaptic Systems; Cat. No. 234 004, Polyclonal Guinea pig antiserum, Lot number: 2-13, 1:500); anti-Human HLA-DP, DQ, DR antigen (Agilent, M077501-2, clone CR3/43, Lot number: 20047190, 1:100; anti-β-Amyloid (Cell Signaling Technology, #8243, D54D2, Lot number: 4, 1:500); anti-OLIG2 (Atlas, HPA003254, rabbit polyclonal, Lot number: C114365, 1:1000);

anti-CRYAB (LSBio, LS-B3696, rabbit polyclonal, Lot number: 124639, 1:200);
anti-QDPR (Atlas Antibodies, HPA065649, rabbit polyclonal, Lot number: R92923, 1:2500);
anti-Rabbit IgG (goat), HRP-labeled (PerkinElmer, NEF812001EA, from goat serum, Lot number: 10311573, Dilution);
Alexa Fluor®488 conjugated anti-NeuN antibody (MilliporeSigma, catalog number MAB377X, clone A60, Lot number: 3101114, 1:500)

Validation

anti-Iba1: reactivity validated by the company for Human. Validated by the company for IHC.
anti-Human HLA-DP, DQ, DR antigen: reactivity validated by the company for Human. The antibody was included in the First International Workshop and Conference on Monoclonal Antibodies to Human MHC Class II Antigens (1983) and its specificity and other characteristics were ascertained by a variety of techniques, including reactivity with isolated antigen, immunoblotting, and labelling of transfected cells.
anti-β-Amyloid: reactivity validated by the company for Human. Validated by the company for IHC on paraffin-embedded human Alzheimer's brain tissue sections.
anti-OLIG2: reactivity validated by the company for Human. Validated by the company for IHC. Has been validated by the Human Protein Atlas in 44 human control brain samples.
anti-CRYAB: reactivity validated by the company for Human. Validated by the company for IHC.
anti-QDPR: reactivity validated by the company for Human. Validated by the company for IHC. Has been validated by the Human Protein Atlas in 44 human control brain samples.
anti-Rabbit IgG (goat), HRP-labeled: Tested by the company to react with rabbit IgG and may recognize other immunoglobulin types that have light chains in common with IgG.
Alexa Fluor®488 conjugated anti-NeuN: reactivity validated by the company for Human.

Human research participants

Policy information about studies involving human research participants

Population characteristics	We selected 24 individuals with elevated β-amyloid pathology and 24 age-matched control individuals with no or very low β-amyloid burden. Individuals were balanced between male and female subjects (12 in each group), and matched for both age (median 86.7 high-amyloid, 87.1 low-amyloid) and years of education (median 19.5 high-amyloid, 18 low-amyloid).
Recruitment	No donors were recruited, the tissue has been obtained from participants in the Religious Order Study.
Ethics oversight	The Religious Orders Study and Rush Memory and Aging Project were approved by an IRB of Rush University Medical Center.

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

Magnetic resonance imaging

Experimental design

Design type	N/A
Design specifications	N/A
Behavioral performance measures	N/A

Acquisition

Imaging type(s)	Structural
Field strength	3T
Sequence & imaging parameters	3D magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence: echo-time (TE)=2.98 ms, repetition time (TR)=2.3 s, inversion time (TI)=900 ms, flip angle=8 degrees, 176 sagittal slices, slice thickness=1 mm, field of view (FOV)=25.6 cm x 25.6 cm, 256x256 acquisition matrix. T2-weighted fluid-attenuated inversion recovery (FLAIR): TE=150 ms, TR=9 s, TI=2.49 s, 35 axial slices, slice thickness=4 mm, FOV=22 cm x 22 cm, 256×256 acquisition matrix.
Area of acquisition	Whole brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	FLIRT (affine registration of T1-weighted MPRAGE to T2-weighted FLAIR), BET (brain extraction), WMLS (automated white matter hyperintensity segmentation based on both T1-weighted MPRAGE and T2-weighted FLAIR), Freesurfer (intracranial volume calculation).
Normalization	T1-weighted MPRAGE data were spatially registered to the T2-weighted FLAIR data using affine registration.
Normalization template	The data were not normalized to a standardized template.

Noise and artifact removal	N/A
Volume censoring	N/A

Statistical modeling & inference

Model type and settings	The total volume of white matter hyperintensities was measured for each participant and then normalized by the corresponding intracranial volume.
Effect(s) tested	N/A

Specify type of analysis: ☒ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016)	N/A
Correction	N/A

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis