nature portfolio

Corresponding author(s): Hanchuan Peng

Last updated by author(s): Nov 29, 2024

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	No software was used.
Data analysis	We spatially registered all tracings to CCFv3 using mBrainAligner (https://github.com/Vaa3D/vaa3d_tools/commits/master/hackathon/mBrainAligner, commit number: b10132f). We clustered neuron node coordinates using the Mclust function with default parameters (mclust R package version 5.4.7 (Scrucca, et al, 2016)). We found the minimal volume enclosing groups of neuron nodes with alphashape3d R package version 1.3.1 (Edelsbrunner, et al, 1994). We plotted 2-D sections of the arbor domains overlaid on the mouse brain atlas using the R base plot function (version 4.1.0). To convert surface polygon file format .ply files to 3D masks we used binvox (version 1.35). To assess the relevance of arbor domain connectivity for defining cell types and subtypes, we used a Support Vector Machine (SVM; hyperoverlap R package version 1.1.1). The Wilcoxon tests were performed by rstatix R package (version 0.7.0). With regard to m-score matrix clustering, we applied hierarchical clustering by clustermap function in Seaborn Python package (version 0.11.2). We used umap-learn Python package (version 0.5.1) to implement UMAP decomposition. Optimal number of clusters in ditance-weighted clustering was determined by the Silhouette Score (Rousseeuwński, et al, 1987) (metrics.silhouette_score function from scikit-learn Python package version 0.24.2). For transcriptomics analysis we used the SCANPY python package, version: 1.9.3. For electrophysiological profile analysis, we used the IPFX python package, version: 1.0.7. Plotting was done using ggplot2 R package (version 3.4.0) and Matplotlib (version 3.3.4).

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

Data

Policy information about availability of data

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

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

The local reconstructions of dendrites (DEN-SEU) and key computational materials are available on the Zenodo repository: https://doi.org/10.5281/zenodo.14242860. The Allen Mouse Brain Common Coordinate Framework can be accessed via the Allen Reference Atlas at: https://mouse.brain-map.org/. Neuron full reconstructions from MouseLight data are available at: https://ml-neuronbrowser.janelia.org/. Axonal reconstructions from ION data can be accessed at: https://doi.org/10.12412/BSDC.1667282400.20002. The single-cell transcriptomes data used in this study can be downloaded from: https://portal.brain-map.org/atlases-and-data/rnaseq/mouse-whole-cortex-and-hippocampus-10x. Electrophysiological data of VISp neurons utilized in this study can be accessed at: https://portal.brain-map.org/cell-types/classes/multimodal-characterization.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	n/a
Reporting on race, ethnicity, or other socially relevant groupings	n/a
Population characteristics	n/a
Recruitment	n/a
Ethics oversight	n/a

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

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 formal sample-size calculation was performed. Instead, we used all publicly available full neuron and axonal morphology datasets relevant to this analysis, which included 9298 neurons. These datasets were chosen based on their comprehensive coverage of the morphological diversity and spatial distribution needed for the study.
Data exclusions	No data were excluded from the analyses.
Replication	All attempts at replicating the results were successful using the data and codes we provided. The results were replicated 3 times by a single researcher on separate days.
Randomization	Randomization was not applicable to this study because we did not assign experimental groups
Blinding	Blinding was not applicable in this study because we did not assign experimental groups.

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 and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.