## nature research

Last updated by author(s): 08-04-2021	

## **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 our Editorial Policies and the Editorial Policy Checklist.

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

~					
5	tа	ŤΙ	101	h	2

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. <i>F</i> , <i>t</i> , <i>r</i> ) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
x	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
x	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 <i>d</i> , Pearson's <i>r</i> ), indicating how they were calculated
	Our web collection on <b>statistics for biologists</b> contains articles on many of the points above.

## Software and code

Policy information about  $\underline{\text{availability of computer code}}$ 

Data collection

Data were collected using Zen2 Black (Zeiss), Imaris (9.1.2, 64Bit, Bitplane).

Data analysis

We provide an R package, Libra, implementing all methods for DE analysis discussed in this study within a consistent interface. The Libra package is available from GitHub (https://github.com/neurorestore/Libra) and as Supplementary Software 1. In addition, the R source code used to perform data analysis is available from GitHub at https://github.com/neurorestore/DE-analysis.

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 Research guidelines for submitting code & software for further information.

## Data

Policy information about <u>availability of data</u>

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

Raw sequencing data and count matrices have been deposited to the Gene Expression Omnibus under accession code GSE165003 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE165003). The 18 'ground truth' datasets are available from Zenodo at http://dx.doi.org/10.5281/zenodo.5048449.

Field-spe	ecific reporting					
	ne below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.					
X Life sciences	Behavioural & social sciences Ecological, evolutionary & environmental sciences					
For a reference copy of	the document with all sections, see <a href="mailto:nature.com/documents/nr-reporting-summary-flat.pdf">nature.com/documents/nr-reporting-summary-flat.pdf</a>					
Life scier	nces study design					
All studies must di	sclose on these points even when the disclosure is negative.					
Sample size	We assembled a compendium of 46 published single-cell or single-nucleus RNA-seq studies (Supplementary Fig. 3), and performed DE analyses across this compendium to establish the generality of our conclusions. The sample sizes for each of these studies were established by the authors of the original studies.					
Data exclusions	For publications containing more than one comparison, only a single comparison was retained, as described in detail in the Methods section. We retained the comparison involving the greatest number of cells, and used the most fine-grained cell type annotations provided by the authors of the original studies. When count matrices did not use gene symbols, the provided identifiers were mapped to gene symbols, and counts summed across genes mapping to identical symbols. Only cell types with at least three cells in each condition were subjected to DE analysis, and genes detected in less than three cells were removed.					
Replication	Our findings were replicated in 46 published single-cell transcriptomics datasets.					
Randomization	Randomization was not relevant as the study involved re-analysis of published datasets.					
Blinding	Blinding was not relevant as the study involved re-analysis of published datasets.					
We require informat system or method lis  Materials & ex  n/a Involved in tl	ChIP-seq					
<b>x</b> Eukaryotic						
	logy and archaeology  MRI-based neuroimaging and other organisms					
	search participants					
X Clinical da						
Dual use r	esearch of concern					
Antibodies						
Antibodies used	Primary antibodies were: rat anti-Pecam1 (BD Biosciences 550274, 1:200). Secondary antibodies were: Alexa Fluor 555 Goat Anti Rat (1:200, Life Technologies, A21434).					
Validation	We used commercial antibodies. All antibodies have been guaranteed and validated by the manufacturers. We also included positive and negative controls for every histological procedure. The concentration of each antibody was tested before use and confirmed based on the morphology of positive signal.					
Animals and	other organisms					
Policy information	about studies involving animals; ARRIVE guidelines recommended for reporting animal research					
Laboratory animals	Experiments were conducted on adult male or female C57BL/6 mice (15-35 g body weight, 12-30 weeks of age). Vglut2:Cre (Jackson Laboratory 016963) transgapic mice were used and maintained on a mixed genetic background (139/057BL/6).					

Wild animals

Field-collected samples

No wild animals were used in the study.

No field collected samples were used in the study.

Ethics oversight

All procedures and surgeries were approved by the Veterinary Office of the Canton of Geneva (Switzerland).

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