

Corresponding Author: T Harkany

Manuscript Number: NN-A54227D

Manuscript Type: Article

Main Figures: 8

Supplementary Figures: 8

Supplementary Tables: 6

Supplementary Videos: 3

Reporting Checklist for Nature Neuroscience

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. For more information, please read [Reporting Life Sciences Research](#).

Please note that in the event of publication, it is mandatory that authors include all relevant methodological and statistical information in the manuscript.

► Statistics reporting, by figure

- Please specify the following information for each panel reporting quantitative data, and where each item is reported (section, e.g. Results, & paragraph number).
- Each figure legend should ideally contain an exact sample size (n) for each experimental group/condition, where n is an exact number and not a range, a clear definition of how n is defined (for example x cells from x slices from x animals from x litters, collected over x days), a description of the statistical test used, the results of the tests, any descriptive statistics and clearly defined error bars if applicable.
- For any experiments using custom statistics, please indicate the test used and stats obtained for each experiment.
- Each figure legend should include a statement of how many times the experiment shown was replicated in the lab; the details of sample collection should be sufficiently clear so that the replicability of the experiment is obvious to the reader.
- For experiments reported in the text but not in the figures, please use the paragraph number instead of the figure number.

Note: Mean and standard deviation are not appropriate on small samples, and plotting independent data points is usually more informative. When technical replicates are reported, error and significance measures reflect the experimental variability and not the variability of the biological process; it is misleading not to state this clearly.

		TEST USED		n			DESCRIPTIVE STATS (AVERAGE, VARIANCE)		P VALUE		DEGREES OF FREEDOM & F/t/z/R/ETC VALUE	
FIGURE NUMBER	WHICH TEST?	SECTION & PARAGRAPH #	EXACT VALUE	DEFINED?	SECTION & PARAGRAPH #	REPORTED?	SECTION & PARAGRAPH #	EXACT VALUE	SECTION & PARAGRAPH #	VALUE	SECTION & PARAGRAPH #	
example 1a	one-way ANOVA	Fig. legend	9, 9, 10, 15	mice from at least 3 litters/group	Methods para 8	error bars are mean +/- SEM	Fig. legend	p = 0.044	Fig. legend	F(3, 36) = 2.97	Fig. legend	
example results, para 6	unpaired t-test	Results para 6	15	slices from 10 mice	Results para 6	error bars are mean +/- SEM	Results para 6	p = 0.0006	Results para 6	t(28) = 2.808	Results para 6	
+ 8d	unpaired t-test	fig. legend	3 + 3	serial slices from 3 mice in each group	fig. legend	no		One-tailed P-value = 0.0332		t = 2.506 with 4 degrees of freedom	not shown	

TEST USED		n			DESCRIPTIVE STATS (AVERAGE, VARIANCE)		P VALUE		DEGREES OF FREEDOM & F/t/z/R/ETC VALUE			
FIGURE NUMBER	WHICH TEST?	SECTION & PARAGRAPH #	EXACT VALUE	DEFINED?	SECTION & PARAGRAPH #	REPORTED?	SECTION & PARAGRAPH #	EXACT VALUE	SECTION & PARAGRAPH #	VALUE	SECTION & PARAGRAPH #	
+ -	S7c	rank sum test	fig. legend	3 animals	109+110+99 cells analysed	fig.	show as box plots	S5c	p<0.001		None reported.	not shown
+ -	8e	Two sample Kolmogorov-Smirnov test	fig. legend	4 + 4 animals	cells from 4 + 4 mice	fig. legend	Shows as CDF	Fig. legend	p=0.0167	figure legend	not reported	not shown
+ -	4, 5c, 6j, 7a,a 2, 8c, S5c S6b,c, S8c	Wilcoxon rank-sum test (p values) with correction for multiple testing (q values)	Fig. legends, Material and methods	Shown in fig 2 a for all 62 clusters		Shown in figures 2a for each cluster	yes	Material and methods, "Error-bar plots."	p values and q values for all genes are present in Supplementary tables 5 and 6	Supplementary tables 5 and 6	not reported	not shown
+ -	5b,	Wilcoxon rank-sum test (p values) with correction for multiple testing (q values)	Fig. legends	Shown in fig 2a for all 62 clusters		Shown in figures 2a for each cluster	averages are shown as a heatmap	Fig. legend	p values and q values for all genes are present in Supplementary tables 5 and 6	Supplementary tables 5 and 6	not reported	not shown
+ -	S5d	Wilcoxon rank-sum test (p values) with correction for multiple testing (q values)	Fig. legends	10+11		Figure legends	averages are shown as a heatmap	Fig. legend	All p values are presented in the figure, Q values < 0.05 for all genes shown	figure S5d	not reported	not shown

► Representative figures

- Are any representative images shown (including Western blots and immunohistochemistry/staining) in the paper?

If so, what figure(s)?

Yes. Histochemical figures in 5d,e,g, 6b-h1, 7b and 8a,d.

- For each representative image, is there a clear statement of how many times this experiment was successfully repeated and a discussion of any limitations in repeatability?

If so, where is this reported (section, paragraph #)?

We have clearly described the number of mice used for each experiment, and for which stainings, in the expanded (online) methods section. The analysis in Fig. S1 is to show the quality of raw data (b,c) and the high level of reproducibility across different biological samples (a). Moreover, a statement is made on male vs. female differences, including a figure (Fig. S2).

► Statistics and general methods

1. Is there a justification of the sample size?

If so, how was it justified?

Where (section, paragraph #)?

Even if no sample size calculation was performed, authors should report why the sample size is adequate to measure their effect size.

1) Here, we used unbiased mathematical approaches to identify cell type variability in the specific brain region. To get a substantial amount of cells per cluster originally we hypothesized (partially based on previous experience as reported by Zeisel et al., 2015) that 3,000 high-quality cells (where only 30% expected to be neurons, thus around 800-1,000 neurons with the predicted number of clusters in the 20-90 range) should be the expected amount of cells for final analysis to reveal genuine group diversity. To avoid "batch"-dependent effects, we controlled (as reported) that any of clusters was not dominated by cells from a particular animal (Fig. S1a).

2) For the statistical analysis presented in Fig. 8d,e we used power calculation to minimize the animal number (to $n = 3 + 3$ and $n = 4 + 4$) in order to obtain significant results with the predicted difference of more than 50% based on our preliminary observations.

3) We constructed a cumulative distribution function (CDF) graph of day/night dependence of phospho-TH levels in >300 cells per from $n = 4$ animals per condition (Fig. 8e). This approach allowed us to minimize the number of animals yet yielding statistically robust data.

2. Are statistical tests justified as appropriate for every figure?

Where (section, paragraph #)?

1) Conventional statistical tests were used for the experiments shown in Fig. 8d,e and Fig. S7c: t-test for Fig. 8d; rank sum test for results shown in Fig. S7c since t-test is not applicable at this point because of the normality test (Shapiro-Wilk) failed; two-sample Kolmogorov-Smirnov test for Fig. 8e.

2) For the data presentation to evaluate the level and specificity of the expression of genes in individual clusters we used the approach described in detail in "Error-bar plots". Briefly, we transformed the data by $\log_2(x+1)$ to calculate the means \pm s.e.m. Likewise, we computed the fraction of positive cells per cluster. We considered meaningful expression if the mean exceeded $2x$ s.e.m. above zero. We show two levels of stringency (bold in figures): i) with power = 0, where there is no correction to the fraction of positive cells and ii) with power = 1, where the average is multiplied by the fraction of positive cells to avoid heterogeneous groups.

Next we used the standard non-parametric Wilcoxon rank-sum test for Figures 4, 5c, 6j, 7a,a2, 8c, S5c,d, S6b,c, S8c. For each gene, we have tested each group vs. the pool of cells from all other groups, a total of 62 comparisons. Secondly, for each gene we have corrected for multiple testing everywhere by using the Benjamini-Hochberg procedure for 62 clusters to control the false discovery rate (FDR). If a test was performed, we reported P and q values and indicate significance at a false discovery rate of 5% (i.e. when $q < 0.05$). Values are presented in Figures, Supplementary tables 4, 5 and 6.

3) Results shown in Fig. 5b were built on the analysis described in "Dendrogram construction and split point listing", the corresponding p and q values are present in Supplementary table 3.

- a. If there is a section summarizing the statistical methods in the methods, is the statistical test for each experiment clearly defined?
- Yes, all methods are described in detail including sections: "Selection of cluster-enriched genes and markers", "Dendrogram construction and split point listing", "Error-bar plots", "Calculating significance using the Wilcoxon rank-sum test", "Statistical analysis of histochemical and imaging data".
- b. Do the data meet the assumptions of the specific statistical test you chose (e.g. normality for a parametric test)?
- Where is this described (section, paragraph #)?
- For the analysis of gene expression we used the non-parametric Wilcoxon rank-sum test (Matlab, "ranksum").
- In case of statistical analysis of histochemical and imaging data, we first checked raw data for normality (Shapiro-Wilk test) and equal variance. If these tests were passed, we used Student's t-test. Otherwise, we analysed our data with the Wilcoxon rank-sum test (another name - Mann-Whitney rank sum test). We analysed two distributions shown in Figure 8e using the two-sample Kolmogorov-Smirnov test. All statistical procedures are described in the Online Methods.
- c. Is there any estimate of variance within each group of data? Is the variance similar between groups that are being statistically compared?
- Where is this described (section, paragraph #)?
- Yes, see above. We do not give a detailed description of this procedure, which is part of the analysis routine provided by SigmaPlot.
- d. Are tests specified as one- or two-sided?
- One-sided for 8d, all rank sum tests are one-sided
- e. Are there adjustments for multiple comparisons?
- Yes, for gene expression data (Benjamini-Hochberg procedure for each gene among 62 clusters).
3. To promote transparency, *Nature Neuroscience* has stopped allowing bar graphs to report statistics in the papers it publishes. If you have bar graphs in your paper, please make sure to switch them to dot-plots (with central and dispersion statistics displayed) or to box-and-whisker plots to show data distributions.
- The submission contains bar graphs for mean values of representative gene expressions with individual data shown as dot density plots in Figs. 1b, 2b and 4. Also, we provide on-line processing service on our website linnarssonlab.com/hypothalamus where graphs for requested genes can be automatically plotted. Bar graphs in Fig. 8 were changed to dot-plots.
4. Are criteria for excluding data points reported?
- Was this criterion established prior to data collection?
- Where is this described (section, paragraph #)?
- Yes, in the "Online Methods" when analyzing the boundaries of neuronal clusters. Criteria for manual cluster analysis, delineation and merging are reported. In addition, false discovery rates were reported.
5. Define the method of randomization used to assign subjects (or samples) to the experimental groups and to collect and process data.
- If no randomization was used, state so.
- Where does this appear (section, paragraph #)?
- For single-cell transcriptome analysis, and in a large number of animals of both sexes, tissues were dissociated and sampled in a blinded fashion. This allowed maximum randomization to preclude bias. In other experiments, tissue samples were randomly processed (e.g. brain slices). Description is reported in Online Methods. We also analysed the "batch-dependent" enrichment of clusters (Fig. S1a), see above.

6. Is a statement of the extent to which investigator knew the group allocation during the experiment and in assessing outcome included?
If no blinding was done, state so.
Where (section, paragraph #)?
- Given the above, all experiments were performed in a blinded fashion, which is clearly reported.
7. For experiments in live vertebrates, is a statement of compliance with ethical guidelines/regulations included?
Where (section, paragraph #)?
- Yes, "Online Methods", 1st paragraph ("Animals, tissue preparation and histochemistry").
8. Is the species of the animals used reported?
Where (section, paragraph #)?
- Yes, same paragraph as above, for mice.
9. Is the strain of the animals (including background strains of KO/transgenic animals used) reported?
Where (section, paragraph #)?
- Yes, C57BL/6N mice were used. Transgenic mice were also described in the "Online Methods" and extensively referenced.
10. Is the sex of the animals/subjects used reported?
Where (section, paragraph #)?
- Yes, all transcriptome data shows X- and Y-chromosome-specific genes. Circuit mapping was performed in male mice only.
11. Is the age of the animals/subjects reported?
Where (section, paragraph #)?
- Yes, for each experiment see "Online Methods".
12. For animals housed in a vivarium, is the light/dark cycle reported?
Where (section, paragraph #)?
- Yes, see the "Online Methods".
13. For animals housed in a vivarium, is the housing group (i.e. number of animals per cage) reported?
Where (section, paragraph #)?
- No. All animals were group housed, usually with littermates to avoid social stress.
14. For behavioral experiments, is the time of day reported (e.g. light or dark cycle)?
Where (section, paragraph #)?
- Yes.
15. Is the previous history of the animals/subjects (e.g. prior drug administration, surgery, behavioral testing) reported?
Where (section, paragraph #)?
- No, particularly since none of the animals were re-used from earlier experiments.
- a. If multiple behavioral tests were conducted in the same group of animals, is this reported?
Where (section, paragraph #)?
- Not applicable.
16. If any animals/subjects were excluded from analysis, is this reported?
Where (section, paragraph #)?
- Not applicable. Non excluded.

a. How were the criteria for exclusion defined?
Where is this described (section, paragraph #)?

Not applicable.

b. Specify reasons for any discrepancy between the number of animals at the beginning and end of the study.
Where is this described (section, paragraph #)?

Not applicable.

► Reagents

1. Have antibodies been validated for use in the system under study (assay and species)?

Yes, in all cases.

a. Is antibody catalog number given?
Where does this appear (section, paragraph #)?

Yes, in Online Methods, "Animals, tissue preparation and histochemistry". 2nd paragraph. For non-commercial antibodies we referenced the original papers and acknowledged our collaborators. Moreover, we reference the Human Protein Atlas project and Synaptic Systems GmbH (who co-authored this study) for specific antibodies that are quality-controlled as shown on their respective websites.

b. Where were the validation data reported (citation, supplementary information, Antibodypedia)?
Where does this appear (section, paragraph #)?

We have extensively referenced the paper for quality controls on both procedures and individual antibodies. These appear primarily in the "Results" section.

2. Cell line identity

a. Are any cell lines used in this paper listed in the database of commonly misidentified cell lines maintained by [ICLAC](#) and [NCBI Biosample](#)?
Where (section, paragraph #)?

None used.

b. If yes, include in the Methods section a scientific justification of their use--indicate here in which section and paragraph the justification can be found.

Not applicable.

c. For each cell line, include in the Methods section a statement that specifies:
- the source of the cell lines
- have the cell lines been authenticated? If so, by which method?
- have the cell lines been tested for mycoplasma contamination?
Where (section, paragraph #)?

Not applicable.

▶ Data availability

Provide a Data availability statement in the Methods section under "Data availability", which should include, where applicable:

- Accession codes for deposited data
- Other unique identifiers (such as DOIs and hyperlinks for any other datasets)
- At a minimum, a statement confirming that all relevant data are available from the authors
- Formal citations of datasets that are assigned DOIs
- A statement regarding data available in the manuscript as source data
- A statement regarding data available with restrictions

See our [data availability and data citations policy page](#) for more information.

Data deposition in a public repository is mandatory for:

- Protein, DNA and RNA sequences
- Macromolecular structures
- Crystallographic data for small molecules
- Microarray data

Deposition is strongly recommended for many other datasets for which structured public repositories exist; more details on our data policy are available [here](#). We encourage the provision of other source data in supplementary information or in unstructured repositories such as [Figshare](#) and [Dryad](#).

We encourage publication of Data Descriptors (see [Scientific Data](#)) to maximize data reuse.

Where is the Data Availability statement provided (section, paragraph #)?

We made data available in multiple ways:

- a search engine is available for individual reprocessing of genes at <http://linnarssonlab.org/hypothalamus>. This platform allows anyone to ask the distribution of individual genes.

- all data, particularly for the definition of junction points, differential expression analysis, and divergence graphs, are provided as Tables S1-S6.

- moreover, all transcriptome data were deposited at: Gene Expression Omnibus GSE74672

▶ Computer code/software

Any custom algorithm/software that is central to the methods must be supplied by the authors in a usable and readable form for readers at the time of publication. However, referees may ask for this information at any time during the review process.

1. Identify all custom software or scripts that were required to conduct the study and where in the procedures each was used.

These include the BaskSpinV2 algorithm developed for cell clustering (as reported by Zeisel et al., Science 2015; Romanov et al., EMBO J 2015, Marques et al., Science 2016). The software code is available as part of the Zeisel et al., Science 2015, which we exhaustively referenced. Moreover, we provided all key variables for our calculations in the "Online Methods".

2. If computer code was used to generate results that are central to the paper's conclusions, include a statement in the Methods section under "**Code availability**" to indicate whether and how the code can be accessed. Include version information as necessary and any restrictions on availability.

Not applicable; see above.

▶ Human subjects

1. Which IRB approved the protocol?
Where is this stated (section, paragraph #)?
TUKEB 84/2014, Hungary and Semmelweis University, Hungary.
2. Is demographic information on all subjects provided?
Where (section, paragraph #)?
Yes, in "Online Methods".
3. Is the number of human subjects, their age and sex clearly defined?
Where (section, paragraph #)?
Yes, in "Online Methods".
4. Are the inclusion and exclusion criteria (if any) clearly specified?
Where (section, paragraph #)?
No subject was excluded.
5. How well were the groups matched?
Where is this information described (section, paragraph #)?
These are proof-of-principle data in $n = 2$ subjects, therefore group matching was deemed non-relevant.
6. Is a statement included confirming that informed consent was obtained from all subjects?
Where (section, paragraph #)?
Not applicable.
7. For publication of patient photos, is a statement included confirming that consent to publish was obtained?
Where (section, paragraph #)?
Not applicable.

► fMRI studies

For papers reporting functional imaging (fMRI) results please ensure that these minimal reporting guidelines are met and that all this information is clearly provided in the methods:

1. Were any subjects scanned but then rejected for the analysis after the data was collected?
Not applicable.
 - a. If yes, is the number rejected and reasons for rejection described?
Where (section, paragraph #)?
Not applicable.
2. Is the number of blocks, trials or experimental units per session and/or subjects specified?
Where (section, paragraph #)?
Not applicable.
3. Is the length of each trial and interval between trials specified?
Not applicable.
4. Is a blocked, event-related, or mixed design being used? If applicable, please specify the block length or how the event-related or mixed design was optimized.
Not applicable.

- a. If so, are the method to account for within subject correlation and the assumptions made about variance clearly stated?
18. If the threshold used for inference and visualization in figures varies, is this clearly stated?
19. Are statistical inferences corrected for multiple comparisons?
- a. If not, is this labeled as uncorrected?
20. Are the results based on an ROI (region of interest) analysis?
- a. If so, is the rationale clearly described?
- b. How were the ROI's defined (functional vs anatomical localization)?
21. Is there correction for multiple comparisons within each voxel?
22. For cluster-wise significance, is the cluster-defining threshold and the corrected significance level defined?

► Additional comments

Additional Comments

None considered relevant.