nature neuroscience

Corresponding Author: Bosiljka Tasic

Manuscript Number: NN-RS51225-YC

Manuscript Type: Resource

Main Figures: 7

Supplementary Figures: 17

Supplementary Tables: 14

Supplementary Videos: 0

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

Nature Neuroscience: doi:10.1038/nn.4216

	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 #
1b	Computational methods (not stand-alone statistical tests): Principal Component Analysis, WGCNA, and validation by Random forest analysis	Fig. Legend, Methods	1679	Number of cells that pass QC checks	Methods, Fig S2	49 clusters, 1424 core cells, 255 intermediate cells.	Fig S2, Methods	N/A	N/A	N/A	N/A
1c	N/A	N/A	49, 1424	Total number of transcriptomic types and total number of core cells. Also included: number of core cells per each cluster (set of values on top for each cluster).	Fig 1c	Kernel probability densities of RPKM values within each cluster (violin plots), and maximum RPKM for each gene across all clusters	Fig 1c	N/A	N/A	N/A	N/A
2b	N/A	N/A	1424, 255	Total number of core cells, total number of intermediate cells. Also included: number of core cells per each cluster (top), number of core cells per each Cre line/dissection combination (right black column), and number of intermediate cells per each Cre line/ dissection combination (right magenta column).	Fig 2b	Percent of core or intermediate cells in each Cre Line/ dissection combination that are classified into each cluster (size of black or magenta discs at each intersection).	Fig 2b	N/A	N/A	N/A	N/A
3a-c	N/A	N/A	1424	Total number of core cells. For numbers of core cells per each cluster see Fig 2b.	Fig 3 legend, Fig 2b	Bar graphs represent scaled RPKM values for each gene within each cell normalized to the maximum RPKM value of that gene among all cells. The maximum RPKM value is listed for each gene among the cells present in each panel.	Fig 3a-c	N/A	N/A	N/A	N/A

Nature Neuroscience: doi:10.1038/nn.4216

+ -	4a-c	N/A	N/A	Core cells (N = 1424 total, 664 GABAergic, 609 glutamatergic, 151 non-neuronal); Intermediate cells (N = 255 total, 97 GABAergic, 155 glutamatergic, 3 non-neuronal)	Total number of core cells, total number of intermediate cells. For number of core cells per each cluster see Fig 2b.	Fig 4 legend.	Number of core cells per cluster (size of discs), number of intermediate cells between pairs of clusters (thickness of lines).	Fig 4a-c	N/A	N/A	N/A	N/A
+ -	4a	Hypergeometric test	Table S5	Too many values - see Table S5	Total number of cells isolated from upper or lower layer dissections from specific Cre lines and numbers of core cells belonging to specific clusters originating from upper or lower layer dissections from those same Cre lines	Table S5	Too many values - see Table S5	Table S5	Too many values - see Table S5	Table S5	There are no degrees of freedom associated with this test.	Methods
+ -	4d	N/A	N/A	1424; 13,878	Total number of core cells and number of genes used.	Fig 4 legend.	N/A	N/A	N/A	N/A	N/A	N/A
+ -	5a	Limma package in Bioconductor	Fig. Legend	256,430	Total number of exons with 49 transcriptomic cell types	Fig 5a legend, Methods	Too many values - see Table S8	Methods	Too many values - see Table S8, column H (header exon_padj)	Methods	Too many values	N/A
+ -	5b-e	Bayes Factor	Fig. Legend	10 for specific cell types or 20 for broad cell classes	Number of cells sampled for MISO comparisons	Methods	MISO scores and confidence intervals generated by MISO represented by barplots in Fig 5b-e	Methods	Bayes Factor exact values represented by heatmap in Fig 5b-e	Fig 5b-e	Too many values	N/A
+ -	6b	Computational method (not stand-alone statistical test): Random forest analysis	Fig legend	43, 5	Number of cells obtained from retrograde labelling from ipsilateral thalamus or contralateral visual cortex	Fig 6 legend	Classification of 44 cells as core cells beloning to 9 cell types, and 4 cells as intermediate	Fig 6 legend, Methods	N/A	N/A	N/A	N/A
+ -	6b	N/A	N/A	48 single cells, 9 clusters	Number of cells obtained from retrograde labelling from ipsilateral thalamus or contralateral visual cortex, and number of clusters they map to.	Fig 6 legend.	Log10(RPKM+1) of median gene expression in included clusters, Log10(RPKM+1) of single cell gene expression in individual cells.	Fig 6b	N/A	N/A	N/A	N/A

Nature Neuroscience: doi:10.1038/nn.4216

+ -	7a	N/A	N/A	12, 24, 30, 41, 12, 48, 22, 12, 13; 4,21,12,2,4,1,1,3,2	Number of core cells in the indicated clusters; number of core cells with RPKM (Ndnf) ≥ 1	Fig 7a	Kernel probability densities of RPKM values within each cluster (violin plots) for the Ndnf gene.	Fig 7a	N/A	N/A	N/A	N/A
+ -	7f	Mann-Whitney test	Fig legend	10, 6	Number of late and non-late-firing neurons examined	Fig legend	Median resting potential; whiskers are 25th and 75th percentiles	Fig Legend	not significant	Fig Legend	14	Fig Legend
+ -	7f	Mann-Whitney test	Fig legend	10, 5	Number of late and non-late-firing neurons examined	Fig legend	Median input resistance; whiskers are 25th and 75th percentiles	Fig Legend	not significant	Fig Legend	13	Fig Legend
+ -	7f	Mann-Whitney test	Fig legend	12, 6	Number of late and non-late-firing neurons examined	Fig legend	Median sag; whiskers are 25th and 75th percentiles	Fig Legend	4.31x10-4	Fig Legend	16	Fig Legend
+ -	7f	Mann-Whitney test	Fig legend	12, 6	Number of late and non-late-firing neurons examined	Fig legend	Median slope; whiskers are 25th and 75th percentiles	Fig Legend	6.52x10-3	Fig Legend	16	Fig Legend
+ -	7g	N/A	N/A	14	pairs of tdT+ cells	Fig Legend	% of electrically coupled cells; mean intersomatic distance; mean junctional conductance. Errors represent SEM	Fig Legend	N/A	N/A	N/A	N/A
+ -	7h	N/A	N/A	12	synaptically connected tdT+ cells	Fig legend	IPSP mean 10 –90% rise time; IPSP mean tau decay. Errors represent SEM.	Fig Legend	N/A	N/A	N/A	N/A
+ -	S2a	N/A	N/A	425	wells over 39 experiments, with 6-12 wells per experiment	Fig legend	% of wells containing one cell , error bar represents standard deviation	Fig Legend	N/A	N/A	N/A	N/A
+ -	S2b	N/A	N/A	64	Number of FACS experiments	Fig legend	Median DAPI+ cells, whiskers represent 25th and 75th percentile.	Fig and Fig Legend	N/A	N/A	N/A	N/A
+ -	S3b	N/A	N/A	1739	Total number of single cell transcriptomes	Figure	Number of cells that pass QC	Fig and Fig Legend	N/A	N/A	N/A	N/A
+ -	S3c	N/A	N/A	1679	Number of single cell transcriptomes that pass QC	Figure	Final set of clusters, core and transitional cells	Figure	N/A	N/A	N/A	N/A
+ -	S4a	N/A	N/A	92; 38	Number of ERCC RNA species, number of ERCC RNA species present at > 1 molecule per sample	Fig Legend	log10(ERCC RPKM +1), Squared Pearson's correlation coefficients, slope	Fig S4a	N/A	N/A	N/A	N/A

Nature Neuroscience: doi:10.1038/nn.4216

+ -	S4b	N/A	N/A	1679; 92; 38	Number of cells that pass QC checks; Number of ERCC RNA species; Number of ERCC RNA species present at > 1 molecule per sample	Fig Legend	Mean log10(ERCC RPKM+1), error bars represent SEM, Squared Pearson's correlation coefficients, slope	Fig S4b	N/A	N/A	N/A	N/A
+ -	S4c	N/A	N/A	1679, 92	Number of single cell transcriptomes that pass QC; number of ERCC RNA species	Fig legend	Detection percentage	Fig S4c	N/A	N/A	N/A	N/A
+ -	S4d-f	N/A	N/A	1679, 92, 24057	Number of single cell transcriptomes that pass QC; number of ERCC RNA species; number of cellular genes	Fig legend	Pearson's correlation coefficient distributions	Fig S4d-e	N/A	N/A	N/A	N/A
+ -	S5a	N/A	N/A	1679; 6; 3	Number of single cell transcriptomes that pass QC; 10 pg cortex transcriptome replicates; 250 ng unamplified cortex replicates	Fig S5a	Median percent of reads that map to various categories, whiskers represent 25th and 75th percentiles	Fig S5a	N/A	N/A	N/A	N/A
+ -	S5b	N/A	N/A	1679	Number of single cell transcriptomes that pass QC	Fig S5b	Percent of reads that map to various categories for each single cell transcriptome	Fig S5b	N/A	N/A	N/A	N/A
+ -	S5c	N/A	N/A	1424, 49	Number of core cells; Number of transcriptomic types	Fig S5c	Mean percent of reads that map to various categories for each transcriptomic type	Fig S5c	N/A	N/A	N/A	N/A
+ -	S5d	N/A	N/A	1424, 49	Number of core cells; Number of transcriptomic types	Fig S5d	Median percent of reads that map to transcriptome for each transcriptomic type; whiskers represent 25th and 75th percentiles	Fig S5d	N/A	N/A	N/A	N/A
+ -	S6a-c	N/A	N/A	23, 19, 7	Number of transcriptomic types (GABAergic, glutamatergic and non-neuronal)	Legend	Mean number of genes detected over all the cells in each group; error bars represent SEM.	Fig S6a-c	N/A	N/A	N/A	N/A
+ -	S6d	N/A	N/A	2,2	Number of transcriptomic datasets: subsampled from raw reads and subsampled from aligned reads.	Legend	Mean number of genes detected over all the cells in each group by subsampling raw reads or post-alignment reads; error bars represent SEM.	Fig S6d	N/A	N/A	N/A	N/A
+ -	S7a-d	N/A	N/A	24,057; 2; 2; 2; 2	Number of genes; number of transcriptomes	Legend	Pearson's correlation coefficients	Fig S7a-d	N/A	N/A	N/A	N/A

Nature Neuroscience: doi:10.1038/nn.4216

+ -	S7e	Two-tailed Mann-Whitney tests with Bonferroni correction	Fig. Legend	1275; 510; 153; 45; 100; 30; 3; 3	Number of comparisons between all samples in cell groups (from Left to Right in figure)	Fig S7e	Median Pearson's correlation coefficients, whiskers represent 25th and 75th percentiles.	Fig S7e	Too many to report: Values represented by inset heatmap in Fig S7e	Fig S7e	Too many to report.	N/A
+ -	S7f	Two-tailed Mann-Whitney tests with Bonferroni correction	Fig. Legend	51; 10; 10; 1; 3	Number of transcriptomic datasets (from Left to Right in figure)	Fig S7f	Median number of genes, whiskers represent 25th and 75th percentiles.	Fig S7f	Too many to report: Values represented by inset heatmap in Fig S7f	Fig S7f	Too many to report.	N/A
+ -	S11a, b	N/A	N/A	480	Number of cells analyzed by qRT-PCR	Fig legend	Gene expression displayed as 30-Ct values for each cell	Fig S11b	N/A	N/A	N/A	N/A
+ -	S11c	N/A	N/A	23	Number of GABAergic cell types	Fig legend	Mean gene expression displayed as log10(RPKM+1)	Fig S11c	N/A	N/A	N/A	N/A
+ -	S12	N/A	N/A	49	Number of transcriptomic cell types	Fig legend	Log10 of 25% trimmed mean RPKM values for each gene among cells in each cluster	Fig S12	N/A	N/A	N/A	N/A
+ -	S13a	Two-tailed Mann-Whitney tests with Bonferroni correction	Fig. Legend	1525; 761; 764; 154	All Neurons; GABAergic neurons; Glutamatergic neurons; Non-neuronal cells	Fig S13a	Median total RNA amount values, 25th and 75th percentiles	Fig S13a	1.34E-82; 7.03E-75; 1.49E-76	Figure Legend	1677; 913; 916	Figure Legend
+ -	S13b	Two-tailed Mann-Whitney tests with Bonferroni correction	Fig. Legend	Displayed for each group in Fig S13b	Number of cells in each core cluster	Fig S13b	Median total RNA amount values, 25th and 75th percentiles	Fig S13b	Values represented by heatmap in Fig S13b	FigS13b	Can be inferred from sum of cells in each compared group - 2	N/A
+ -	S13c	Student's t-Test with Bonferroni correction	Fig. Legend	1463; 721; 742; 147	All Neurons; GABAergic Neurons; Glutamatergic Neurons; Non-Neuronal cells subsampled to 5M total reads	Fig S13c	Mean number of genes detected, error bars represent standard deviations	Fig S13c	8.09E-89; 4.14E-89; 3.02E-98; 1.80E-21	Figure Legend	1608; 866; 887; 1461	Figure Legend
+ -	S13d	N/A	N/A	Displayed for each group in Fig S13d	Number of core cells in each cell type	Fig S13d	Mean number of genes detected, error bars represent standard deviations	Fig S13d	N/A	N/A	N/A	N/A
+ -	S13e	N/A	N/A	49	Number of transcriptomic cell types	Fig legend	Mean number of genes detected (lines), shaded regions represent SEM	Fig S13e	N/A	N/A	N/A	N/A
+ -	S13f	Mann-Whitney test with Bonferroni correction	Fig. Legend	1525; 154	Number of neurons; Number of non-neuronal cells	Fig Legend	Mean number of genes detected (lines), shaded regions represent SEM	Fig S13f	Left to Right: 1.8E-02, 2.7E-86, 1.5E-83, 3.0E-05, 2.1E-14	Figure Legend	1677	Figure Legend

Nature Neuroscience: doi:10.1038/nn.4216

+ -	S14, S15, S16	N/A	N/A	49, 1424	Total number of transcriptomic types and total number of core cells.	Fig Legend	Kernel probability densities of RPKM values within each cluster (violin plots), and maximum RPKM for each gene across all clusters.	Fig S14	N/A	N/A	N/A	N/A
+ -	S17a	N/A	N/A	761; 764; 154; 164; 399; 1128; 175; 939; 249	Number of GABAergic neurons, Glutamatergic neurons, and Non-neuronal cells surveyed in this study and from neocortex and hippocampus in Zeisel et al. study	Fig S17a	Mean number of genes detected within each group	Fig S17a	N/A	N/A	N/A	N/A
+ -	Table S9	N/A	N/A	228	genes	Table S9	Number and percentage of chromogenic ISH experiments that agree with RNA-seq, do not agree RNA-seq for each of 4 different reasons, or are not available	Table S9	N/A	N/A	N/A	N/A
+ -	Table S11	N/A	N/A	72	experimental animals	Table S11	% of C57BL6/J background in each experimental animal and average % of C57Bl6/J background in the complete set of experimental animals	Table S11	N/A	N/A	N/A	N/A

Representative figures

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

   If so, what figure(s)?

   Yes:
   Fig. 1a; Fig. 2a; Fig. 6a; Fig. 7b,c,d,i;
   Fig. S1b, Fig. S2a,b,d; Fig. S9a-c; Fig. S10a-l; Fig. S11a.

2. 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 #)?

   Fig. 1a, S1b and S11a: Yes, Methods, end of Single cell isolation section.
   Fig. 2a: Yes, Figure legend
   Fig. 6a: Yes, end of Fig. 6 legend
   Fig. 7b,c,d,i: Yes, Figure legend
   Fig. S2a,b,d: Yes, Figure legend
   Fig. S9: Yes, Figure legend
   Fig. S10a-l: Yes, Figure legend

Nature Neuroscience: doi:10.1038/nn.4216

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.	Sample size was not calculated a priori. The sample sizes are similar to or higher than those generally employed in the field. Stated in subsection "Statistical analyses and methodology" at the end of Methods section.
2. Are statistical tests justified as appropriate for every figure? Where (section, paragraph #)?	Yes, see subsection "Statistical analyses and methodology" at the end of Methods section.
a. If there is a section summarizing the statistical methods in the methods, is the statistical test for each experiment clearly defined?	Yes, the subsection "Statistical analyses and methodology", at the end of Methods section, summarizes our statistical methods. In addition, the statistical test for each experiment/analysis is defined in the corresponding figure legend.
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 #)?	Yes, as described in the subsection "Statistical analyses and methodology", at the end of Methods section.
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 #)?	Although variances are reported graphically in the figures, we did not compare variance estimates between groups. As a result, when performing statistical tests, we did not make the assumption that variances are equal or that distributions have the same shape. For the differential gene expression tests, we use the DESeq and DESeq2 packages, both of which derive estimates for the underlying distributions (in the form of negative binomial distribution) for the read counts. These statements are included in subsection "Statistical analyses and methodology", at the end of Methods section
d. Are tests specified as one- or two-sided?	All tests are two-sided as stated in subsection "Statistical analyses and methodology", at the end of Methods section.
e. Are there adjustments for multiple comparisons?	Yes, Benjamini-Hochberg correction for FDRs, Bonferroni for p-value-based tests, as stated in subsection "Statistical analyses and methodology", at the end of Methods section. In addition, the adjustments for multiple comparisons, if used, are mentioned next to the name of the corresponding statistical test used in the corresponding figure legend.
3. Are criteria for excluding data points reported? Was this criterion established prior to data collection? Where is this described (section, paragraph #)?	Yes, exclusion of data points was based on several criteria, as mentioned in the Methods section, under Sequencing Data Processing and QC. The criteria were established after data collection.

Nature Neuroscience: doi:10.1038/nn.4216

4. 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 #)?	No randomization was used. This statement is included in subsection "Statistical analyses and methodology", at the end of Methods section.
5. 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 #)?	Data collection and analysis were not performed blind to the conditions of the experiments. This statement is included in subsection "Statistical analyses and methodology", at the end of Methods section.
6. For experiments in live vertebrates, is a statement of compliance with ethical guidelines/regulations included? Where (section, paragraph #)?	Yes, Methods, first paragraph.
7. Is the species of the animals used reported? Where (section, paragraph #)?	Yes, Methods, first paragraph.
8. Is the strain of the animals (including background strains of KO/ transgenic animals used) reported? Where (section, paragraph #)?	Yes, Methods, first paragraph.
9. Is the sex of the animals/subjects used reported? Where (section, paragraph #)?	Yes, Methods, first paragraph.
10. Is the age of the animals/subjects reported? Where (section, paragraph #)?	Yes, Methods. For the age of animals used for cell isolation, see Methods, "Single cell isolation" subsection. For the age of animals treated with tamoxifen or trimethoprim, see Methods, first paragraph.
11. For animals housed in a vivarium, is the light/dark cycle reported? Where (section, paragraph #)?	Yes, Methods, first paragraph.
12. For animals housed in a vivarium, is the housing group (i.e. number of animals per cage) reported? Where (section, paragraph #)?	Yes, Methods, first paragraph.
13. For behavioral experiments, is the time of day reported (e.g. light or dark cycle)? Where (section, paragraph #)?	N/A
14. Is the previous history of the animals/subjects (e.g. prior drug administration, surgery, behavioral testing) reported? Where (section, paragraph #)?	N/A

Nature Neuroscience: doi:10.1038/nn.4216

a. If multiple behavioral tests were conducted in the same group of animals, is this reported?

Where (section, paragraph #)?

N/A

15. If any animals/subjects were excluded from analysis, is this reported?

Where (section, paragraph #)?

Yes.

a. How were the criteria for exclusion defined?

Where is this described (section, paragraph #)?

We used only healthy adult males. We excluded animals with microphthalmia or anophthalmia (Methods, first paragraph).

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 #)?

N/A

Reagents

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

N/A

a. Is antibody catalog number given?

Where does this appear (section, paragraph #)?

N/A

b. Where were the validation data reported (citation, supplementary information, Antibodypedia)?

Where does this appear (section, paragraph #)?

N/A

2. If cell lines were used to reflect the properties of a particular tissue or disease state, is their source identified?

Where (section, paragraph #)?

N/A

a. Were they recently authenticated?

Where is this information reported (section, paragraph #)?

N/A

Nature Neuroscience: doi:10.1038/nn.4216

Data deposition

Data deposition in a public repository is mandatory for:

a. Protein, DNA and RNA sequences
b. Macromolecular structures
c. Crystallographic data for small molecules
d. 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.

1. Are accession codes for deposit dates provided? Where (section, paragraph #)?	Yes, GEO accession GSE71585 for this study is provided in the Data and reagent availability section; GEO accessions for individual cells are provided in Supplementary Table 3.

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.	Each section in the computational methods section has a corresponding custom R script. These include: 1. A script to run iterative clustering algorithm in order to identify cell types 2. A script to identify key differentially expressed genes among different groups of cells, using the DESeq program. 3. A script to run the cross-validation algorithm, which assesses robustness of clusters and assigns"core" and "intermediate" identity to cells
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.	All computer code is made available as a supplemental document.

Human subjects

1. Which IRB approved the protocol? Where is this stated (section, paragraph #)?	N/A
2. Is demographic information on all subjects provided? Where (section, paragraph #)?	N/A
3. Is the number of human subjects, their age and sex clearly defined? Where (section, paragraph #)?	N/A

Nature Neuroscience: doi:10.1038/nn.4216

4. Are the inclusion and exclusion criteria (if any) clearly specified?

   Where (section, paragraph #)?

N/A

5. How well were the groups matched?

   Where is this information described (section, paragraph #)?

N/A

6. Is a statement included confirming that informed consent was obtained from all subjects?

   Where (section, paragraph #)?

N/A

7. For publication of patient photos, is a statement included confirming that consent to publish was obtained?

   Where (section, paragraph #)?

N/A

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?

N/A

   a. If yes, is the number rejected and reasons for rejection described?

      Where (section, paragraph #)?

N/A

2. Is the number of blocks, trials or experimental units per session and/or subjects specified?

   Where (section, paragraph #)?

N/A

3. Is the length of each trial and interval between trials specified?

N/A

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.

N/A

5. Is the task design clearly described?

   Where (section, paragraph #)?

N/A

6. How was behavioral performance measured?

N/A

7. Is an ANOVA or factorial design being used?

N/A

8. For data acquisition, is a whole brain scan used?

   If not, state area of acquisition.

N/A

Nature Neuroscience: doi:10.1038/nn.4216

a. How was this region determined?	N/A
9. Is the field strength (in Tesla) of the MRI system stated?	N/A
a. Is the pulse sequence type (gradient/spin echo, EPI/spiral) stated?	N/A
b. Are the field-of-view, matrix size, slice thickness, and TE/TR/ flip angle clearly stated?	N/A
10. Are the software and specific parameters (model/functions, smoothing kernel size if applicable, etc.) used for data processing and pre-processing clearly stated?	N/A
11. Is the coordinate space for the anatomical/functional imaging data clearly defined as subject/native space or standardized stereotaxic space, e.g., original Talairach, MNI305, ICBM152, etc? Where (section, paragraph #)?	N/A
12. If there was data normalization/standardization to a specific space template, are the type of transformation (linear vs. nonlinear) used and image types being transformed clearly described? Where (section, paragraph #)?	N/A
13. How were anatomical locations determined, e.g., via an automated labeling algorithm (AAL), standardized coordinate database (Talairach daemon), probabilistic atlases, etc.?	N/A
14. Were any additional regressors (behavioral covariates, motion etc) used?	N/A
15. Is the contrast construction clearly defined?	N/A
16. Is a mixed/random effects or fixed inference used?	N/A
a. If fixed effects inference used, is this justified?	N/A
17. Were repeated measures used (multiple measurements per subject)?	N/A
a. If so, are the method to account for within subject correlation and the assumptions made about variance clearly stated?	N/A
18. If the threshold used for inference and visualization in figures varies, is this clearly stated?	N/A
19. Are statistical inferences corrected for multiple comparisons?	N/A
a. If not, is this labeled as uncorrected?	N/A

Nature Neuroscience: doi:10.1038/nn.4216

20. Are the results based on an ROI (region of interest) analysis?	N/A
a. If so, is the rationale clearly described?	N/A
b. How were the ROI's defined (functional vs anatomical localization)?	N/A
21. Is there correction for multiple comparisons within each voxel?	N/A
22. For cluster-wise significance, is the cluster-defining threshold and the corrected significance level defined?	N/A

Additional comments

Additional Comments	N/A

Nature Neuroscience: doi:10.1038/nn.4216