Cell Reports, Volume *22*

Supplemental Information

Variation in Activity State, Axonal Projection, and Position Define the Transcriptional Identity of Individual Neocortical Projection Neurons

Maxime Chevée, Johanna De Jong Robertson, Gabrielle Heather Cannon, Solange Pezon Brown, and Loyal Andrew Goff

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Loyal Goff (loyalgoff@jhmi.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice
All procedures were approved by the Johns Hopkins Animal Care and Use Committee and followed the guidelines of the Society for Neuroscience and the National Institutes of Health. Animals used for single cell RNA-seq ranged from postnatal day 23 (P23) to P28, animals used for immunohistochemistry and *in situ* hybridization ranged from P23 to P32. Both males and females were used in this study (gender indicated for each replicate in methods below). All animals were kept on a 12h light/dark cycle, housed with at least two mice per cage and provided with unlimited water and food.

TRANSGENIC ANIMAL MODELS

Information about the generation and genotyping of the transgenic lines used in this study can be found in the corresponding original studies: Ntsr1-Cre (Gong et al., 2007), lox-STOP-lox-tdTomato (Ai9 and Ai14; Madisen et al., 2010), lox-STOP-lox-YFP (Ai3; Madisen et al., 2010) and Pantr1-Lacz (Sauvageau et al., 2013). Mice were maintained on mixed backgrounds including C57BL/6J and CD-1.

METHOD DETAILS

Stereotactic injections
To identify layer 6 corticothalamic neurons (L6CThNs), mice P18 to P23 were anesthetized with ketamine (50 mg/kg), dexmedetomidine (25 µg/kg) and the inhalation anesthetic, isoflurane. Animals were placed in a custom-built stereotactic frame and anesthesia was maintained with isoflurane (1-2.5%). A small craniotomy was performed, and a glass pipette (10–25 µm tip diameter) was lowered into the thalamus. To identify L6CThNs that project to the ventral posterior medial nucleus (VPM) of the thalamus, Ntsr1-Cre;tdTomato mice were injected in VPM (1.1 mm posterior, 1.7 mm lateral , 3.4 mm ventral from bregma) with either Alexa 488 Cholera toxin B (CTB488, ThermoFisher) or green Retrobeads IX (Lumafluor), and Ntsr1-Cre;YFP mice were injected with Alexa 555 Cholera toxin B (CTB555, ThermoFisher). To identify L6CThNs that project to both VPM and to the posterior medial nucleus of the thalamus (POm), Ntsr1-Cre;tdTomato mice were injected in POm (1.35 mm posterior, 1.23 mm lateral , 3.3 mm ventral from bregma) with green Retrobeads IX, and Ntsr1-Cre;YFP mice were injected with either CTB555 or red Retrobeads (Lumafluor). Between 30 and 100 nl of tracer were pressure-injected, and the pipette was kept in position for 5-10 minutes before removal. Following the injection, the incision was sutured and buprenorphine (0.05 mg/kg) was administered to all animals postoperatively. Quantification of the colocalization between retrogradely labeled neurons and Cre expression in Ntsr1-Cre mice (Figure 1E,J) was performed on mice injected with either CTB or Retrobeads. Preliminary experiments indicated that all tracers used generated similar labeling. However, Retrobeads resulted in increased discrimination of labeled neurons during

Fluorescence Assisted Cell Sorting (FACS) and allow for restricted injections relative to other tracers (Katz et al., 1984). Therefore, Retrobeads were used for all FACS experiments.

Immunohistochemistry

Brains from animals injected with neuronal tracers were removed and placed in a solution of 4% paraformaldehyde (PFA) in 0.01 M phosphate buffered saline (PBS) for 3 h. Coronal sections were cut on a vibratome (50 µm, VT-1000S, Leica). Mice expressing YFP were subjected to immunohistochemistry (1:1000, chicken anti-GFP, GFP-1020, Aves, RRID:AB_10000240, and 1:300 AlexaFluor 488-conjugated goat anti-chicken, Life Technologies, A11039, RRID:AB_10563770). Sections were then mounted using Aqua Poly/Mount mounting medium (Polysciences, Inc) and visualized on a confocal microscope (LSM 510, Zeiss) using 10x (0.3 NA), 25x (0.8 NA) or 40x (1.3 NA) objectives or on an epifluorescence microscope (AxioObserver.Z1, Zeiss) using a 5x (0.16 NA) objective. Colocalization of tdTomato, YFP and retrograde tracer was quantified using single-plane confocal images and the Cell Counter plugin in Fiji (Schindelin et al., 2012). For all anatomical comparisons, upper layer 6 and lower layer 6 were defined as the top 50% and bottom 50% of the layer defined by Cre-expressing neurons in Ntsr1-Cre mice.

Cell isolation and enrichment

For bulk sequencing experiments (bulk RNA-seq), POm was injected with green Retrobeads in three Ntsr1-cre;tdTomato mice. For single-cell sequencing experiments (scRNA-seq), POm was injected with red or green Retrobeads in Ntsr1-Cre;YFP mice or in Ntsr1-Cre;tdTomato mice, respectively. Mice were sacrificed 5-7 days after surgery (Bulk RNA-seq: Replicate 1: P23 female; Replicate 2: P23 female; Replicate 3: P26 male; scRNA-seq baseline: Replicate 1: P27 male, Replicate 2: P28 female; scRNA-seq Day 1: Replicate 1: P24 female, Replicate 2: P24 female; scRNA-seq Day 7: Replicate 1: P29 female, Replicate 2: P27 female). Brains were rapidly removed and 300 µm thick somatosensory thalamocortical (Agmon and Connors, 1991) or coronal slices were sectioned (VT-1200s, Leica) in ice-cold sucrose solution containing the following (in mM): 76 NaCl, 25 $NaHCO_3$, 25 glucose, 75 sucrose, 2.5 KCl, 1.25 NaH_2PO_4, 0.5 $CaCl_2$, 7 $MgSO_4$, pH 7.3, 310 mOsm continuously bubbled with 95% O_2/5% CO_2. The slices were then placed in a submersion chamber on an upright microscope in ice-cold sucrose solution continuously bubbled with 95% O_2/5% CO_2. The barrel cortex was identified using differential interference contrast (DIC) and epifluorescence microscopy (Axioskop2 FsPLus, Zeiss) using 4x (0.1 NA) and 40x (water immersion, 0.75 NA) objectives. The location of the retrograde tracer in the lower half of layer 6 was visually confirmed, and then regions of cortex below barrels with appropriate labeling were microdissected and immediately dropped into 2.6 mL equilibrated Papain DNase-I solution. For all experiments, at least three consecutive slices confirmed as clearly labeled were microdissected. For the representative image in Figure 5A, the red and green channels were aligned *post hoc*.

The dissociation protocol was adapted from the trehalose-enhanced neuronal isolation protocol described in Saxena et al., 2012 which is based on the Worthington Papain Dissociation System (Worthington Biochemical Corporation). Briefly, tissue pieces were dropped into a 5 ml tube containing 2.6 mL of a trehalose-Papain-DNase1 solution which included RNase inhibitor, incubated 30-45 min at 37°C with gentle trituration at the half way point using a p1000 pipetter. The reaction was stopped by adding a protease inhibitor solution (for solution compositions, see Saxena et al., 2012. The tissue was then further dissociated and washed by sequential pipetting and low-speed centrifugations. The final pellet was resuspended in 200 µL of media (DMEM, 5% trehalose (w/v), 25 µM AP-V, 0.4 mM kynurenic acid, 6 µL of 40 U/µl RNase inhibitor) at room temperature and introduced into a FACS machine (Beckman Coulter MoFlo Cell Sorter). Neurons

were sorted based on their fluorescence (either tdTomato-positive (red) and Retrobead-positive (green) versus tdTomato-positive only or YFP-positive (green) and Retrobead-positive (red) versus YFP-positive only) directly into lysis buffer either in eppendorf tubes for bulk sequencing (350 µL lysis buffer, Qiagen RNeasy Micro Kit; Replicate 1: tdTomato-only: 1943 cells, tdTomato and tracer: 2287 cells; Replicate 2: tdTomato-only: 386, tdTomato and tracer: 782; Replicate 3: tdTomato-only: 604, tdTomato and tracer: 934) or into individual wells of 96-well plates for single-cell sequencing (2 µL Smart-Seq2 lysis buffer + RNAase inhibitor, 1 µL oligo-dT primer, and 1 µL dNTPs according to Picelli et al., 2014). For baseline experiments, 96 green and 96 read and green neurons were collected from each of two replicates, totaling 384 cells. Upon completion of a sort, the plates were briefly spun in a tabletop microcentrifuge and immediately placed on dry ice. Single-cell lysates were subsequently kept at -80°C until cDNA conversion. Neurons analyzed following sensory deprivation were collected in a similar manner with 151 neurons collected from two replicates at Day 1 and 563 neurons collected from two replicates at Day 7.

Library preparation and amplification

For bulk samples, RNA was extracted with a Qiagen RNeasy Micro Kit. RNA quality was assessed with the Bioanalyzer Pico RNA kit and samples had RIN scores of 8.2-8.8. Library preparation and amplification were performed according to the Smart-Seq2 protocol (Picelli et al., 2014) with some modifications. Principally, all primer and adapter sequences were synthesized with a 5'-biotin to minimize primer dimerization. Briefly, 200 pg of RNA per sample were used as input for template switching cDNA synthesis. Full-length cDNA was amplified by KAPA HiFi mediated PCR for 20 cycles using a 5'-biotinylated ISPCR primer. Then, 500 pg of Ampure XP bead cleaned (1:1) cDNA was used as input for a standard Nextera XT tagmentation reaction, and amplification of adapter-ligated fragments was carried out for 12 cycles during which individual index sequences were added to each distinct sample. Library concentration was assessed with Qubit and library fragment size distribution was assessed on the Agilent Bioanalyzer. Pooled, indexed bulk RNA-seq samples were initially sequenced on the Illumina MiSeq 500 platform, and subsequently the same pool was sequenced on one lane of the Illumina HiSeq 2500 platform to produce 50 bp paired-end reads. Libraries were sequenced to an average depth of $4.2x10^7$ fragments each. Reads from both runs were aggregated by index prior to mapping.

Library preparation and amplification of single-cell samples was performed using a modified version of the Smart-Seq2 protocol. Briefly, 96-well plates of single-cell lysates were thawed to 4°C, heated to 72°C for 3 minutes, and then immediately placed on ice. Template switching first-strand cDNA synthesis was performed as described above using a 5'-biotinylated TSO oligo. cDNAs were amplified using 18 cycles of KAPA HiFi PCR and 5'-biotinylated ISPCR primer. Amplified cDNA was cleaned with a 1:1 ratio of Ampure XP beads and 300 pg was used for a one-half standard-sized Nextera XT tagmentation reaction. Tagmented fragments were amplified for 12 cycles and dual indexes were added to each well to uniquely label each library. Concentrations were assessed with Quant-iT PicoGreen dsDNA Reagent (Invitrogen) and samples were diluted to ~2 nM and pooled. Pooled libraries were sequenced on the Illumina HiSeq 2500 platform to a target mean depth of ~8.0×10^5 50 bp paired end fragments per cell (Mudge and Harrow, 2015) at the Hopkins Genetics Research Core Facility to generate 50 bp paired-end reads.

In situ hybridization and immunohistochemistry

Immunohistochemistry (IHC) for tdTomato and *in situ* hybridization (ISH) for *linc-Tmem20* were combined to show co-localization of the lncRNA transcript and tdTomato protein in L6CThNs. Ntsr1-Cre;tdTomato mice (P23 to P30) were transcardially perfused with 0.1 M PBS followed by

a 4% PFA solution, and brains were post-fixed for 3 h at room temperature in 4% PFA. Then, 30 µm sections were cut on a vibratome (VT-1000S, Leica) and first processed for IHC (1:1000, Rabbit anti-DsRed polyclonal antibody, Living Colors, Clontech Laboratories, RRID: AB_10013483 and 1:300 AlexaFluor 568-conjugated goat anti-rabbit, Life Technologies, A11011, RRID: AB_143157). Next, the sections were processed for ISH following the protocol described in Blackshaw, 2013 without proteinase K treatment. Sections were incubated overnight at 65°C with 5 ng/µL of an anti-sense *linc-Tmem20* DIG labeled probe (DIG RNA labeling Kit (SP6/T7), Roche). Detection of the hybridized DIG probes was performed using an alkaline phosphatase (AP) conjugated antibody (1:1000, Anti-Digoxigenin [21H8] Alkaline Phosphatase, Abcam, Cat. No.: ab119345, RRID:AB_10901703) followed by application of the AP substrate BCIP/NBT (SigmaFast BCIP/NBT, Sigma-Aldrich). The enzymatic reaction was carried out until the desired signal to noise ratio was observed. Images were then taken on an AxioObserver.Z1 (Zeiss) using a combination of epifluorescence and brightfield imaging (10x objective, NA: 0.3). Brightness and contrast were adjusted using Adobe Photoshop (Adobe Systems Incorporated).

For validation of *Pantr1* expression, transgenic mice expressing LacZ driven by the endogenous *Pantr1* promoter locus (Goff et al., 2015; Sauvageau et al., 2013) were injected with CTB488 in VPM and CTB555 in POm. IHC was performed on 50 µm-thick sections (rabbit anti-β galactosidase antibody 1:5000; a kind gift from Dr. Joshua R. Sanes, Harvard University, Cambridge, MA and 1:300 AlexaFluor 647-conjugated donkey anti-rabbit, Life Technologies, A31573, RRID: AB_2536183). Imaging was performed on a confocal microscope (LSM 510, Zeiss) using 10x (0.3 NA), 25x (0.8 NA) or 40x (1.3 NA) objectives. LacZ-positive and retrogradely labeled neurons were manually counted on single-plane confocal images using the Cell Counter plugin in Fiji (Schindelin et al., 2012). Brightness and contrast were adjusted using Adobe Photoshop (Adobe Systems Incorporated).

For detection of single mRNA molecules, we used the commercially available system RNAscope (ACDbio). Ntsr1-Cre;tdTomato mice (P23 to P30) were transcardially perfused with 0.1 M PBS followed by a 4% PFA solution, and brains were post-fixed for 2 h at room temperature in 4% PFA. Fixed brains were then washed in PBS and allowed to equilibrate in a 30% sucrose-PBS solution for 36 h before being embedded and frozen in OCT compound (Scigen, Tissue-Plus #4583). Frozen brains were kept at -80°C for up to 2 months. 20 µm sections were collected on a cryostat (Leica CM3050) and processed according to instructions provided by RNAscope. The following probes were used: tdTomato-C2 (317041-C2), tdTomato-C3 (317041-C3), Mm-Rbfox3 (313311), Rbfox3-C2 (313311-C2), Mm-Lamp5 (451071), Mm-Pantr1 (483711), Mm-Serpini1 (501441), Gabra5-C3 (319481-C3). Imaging was performed on a confocal microscope (LSM 510, Zeiss) using a 63x objective (NA 1.4). Sections from three animals were used for *Pantr1*, *Lamp5* and *Gabra5* and from four animals for *Serpini1*. For quantification of the number of mRNAs per neuron obtained using RNAscope, single plane images were first processed with the open-source software CellProfiler (Broad Institute) to automatically identify individual cells based on DAPI signal and to generate outline files. These outlines were then used as input, in combination with the corresponding confocal images, into FISH-quant (Mueller et al., 2013) Matlab) for quantification of transcript number per cell. Cells with no detectable NeuN (*Rbfox3*) transcript were discounted as non-neuronal cells. A threshold of 20 *tdTomato* puncta was selected as the cutoff for distinguishing tdTomato-positive from tdTomato-negative neurons. This resulted in 164 tdTomato-positive and 188 tdTomato-negative neurons for *Serpini1*, 118 tdTomato-positive and 150 tdTomato-negative neurons for *Lamp5*, 111 tdTomato-positive and 198 tdTomato-negative neurons for *Pantr1*, and 135 tdTomato-positive and 160 tdTomato-negative neurons for *Gabra5*. Individual images were stitched to reconstitute an image spanning layer 6 vertically, and the distance of each neuron from the bottom of the layer

was measured manually using Fiji (Schindelin et al., 2012) and normalized to the height of layer 6. Statistical testing (Likelihood ratio test) and curve fitting (Loess) was performed in R/Bioconductor.

Sensory deprivation

Prior to sensory manipulation, animals were bilaterally injected in POm with green Retrobeads IX (Lumafluor) as described above. To induce experience dependent plasticity, mice (P18 to P23) were anesthetized with the inhalation anesthetic, isoflurane. Once an adequate level of anesthesia was achieved, the animals were placed under a surgical microscope and, using tweezers, individual vibrissae were gently pulled by applying slow, steady tension to the base of the whisker (Li et al., 1995). To produce a chessboard pattern, vibrissae alpha, gamma, A2, A4, B1, B3, C2, C4, C6, D1, D3, D5, D7, E1, E3, E5 and E7 were pulled unilaterally while the rest remained untouched. Following the procedure, the animals were returned to their home cage and housed with at least one other animal until tissue collection. We checked for regrowing vibrissae every two days and removed them if visible.

QUANTIFICATION AND STATISTICAL ANALYSES

Read pre-processing

For both bulk and single-cell libraries, paired-end reads were aligned to the mouse reference genome (mm10) using Hisat2 (Kim et al., 2015) with the default parameters except: -p 8. Aligned reads from individual samples were quantified against a reference transcriptome (GENCODE vM8; Mudge and Harrow, 2015) supplemented with additional lncRNA genes described previously (File S1; Molyneaux et al., 2015) with cuffquant (Trapnell et al., 2013) using default parameters with the following exceptions: --no-update-check –p 8. For bulk RNA-Seq, pre-processed expression estimates were used as input for cuffdiff using default parameters except: -p 16, and the output of cuffdiff was analyzed using the cummeRbund R/Bioconductor package (Trapnell et al., 2012). For single-cell experiments, normalized expression estimates across all samples were obtained using cuffnorm with default parameters. For comparison to results from primary visual cortex (Tasic et al., 2016), raw reads were downloaded from the short read archive (GEO: GSE71585), processed as described above for single-cell samples, and normalized expression estimates were obtained from combining both sources of single-cell reads together via cuffnorm.

Single cell analysis

The normalized FPKM matrix from cuffnorm was used as input for the Monocle2 single-cell RNA-seq framework (Trapnell et al., 2014) in R/Bioconductor (Huber et al., 2015). Relative FPKM values for each cell were converted to estimates of absolute mRNA counts per cell (RPC) using the Monocle2 Census utility (Qiu et al., 2017). A total of 346 individual cells under baseline conditions passed quality filters including: a) total mRNAs between 2000 and 200,000 per cell, and b) >10,000 fragments mapped (total_mass). 12,537 genes were identified as expressed in at least 15 cells across all conditions.

To identify high-variance genes, a generalized additive model (MGCV R package; (Wood, 2011) was fit to the $\log_2$ mean RPC expression versus a cubic spline fit to the $\log_2$ coefficient of variation (BCV) for each replicate dissociation independently (Figure S2A and Figure S7D). The intersection of genes with residuals to this fit greater than 1.0 from each replicate were chosen as 'high-variance' genes and the $\log_2$ expression estimates (with a pseudocount of 1) of these

selected genes were used as input for PCA analysis and t-stochastic nearest neighbor embedding (tSNE) clustering of individual cells (Krijthe, 2015).

To cluster individual cells, we employed a spectral tSNE approach (Macosko et al., 2015), combining non-linear dimensional reduction and density clustering. Briefly, PCA analysis was conducted on high-variance genes and then a subset of components (n = 3, selected by permutation parallel analysis, Figure S2B) were used as input for the tSNE visualization algorithm (perplexity = 30; Krijthe, 2015). K-means clustering was used to partition cells using the two reduced tSNE dimensions into two discrete subtypes. The choice of k was made using 1000 bootstrapped tSNE estimates and evaluating k = 2 through k = 8 for each round. The average silhouette coefficient for each tSNE*k was determined and the mean and standard deviation of these estimates was estimated for each k (Figure S2C). k = 2 was chosen as the k value with the highest average silhouette coefficient and the lowest value of k for which there was no significant improvement in mean silhouette score.

After cluster assignment, differential expression testing was performed across all expressed genes using the Monocle2 VGAM model comparison test (Trapnell et al., 2014). To test for differentially expressed genes between transcriptionally-defined cell types, the following full model was fit to each expressed gene: ~replicate+num_genes_expressed+celltype, and compared to a reduced model in which celltype was removed. The number of genes expressed in each cell was included as an explanatory variable as a proxy for the sensitivity of scRNA-seq in each individual cell. To test for differentially expressed genes between anatomically labeled cells, the following full model was fit to each expressed gene: ~replicate+num_genes_expressed+label, and compared to a reduced model in which label was removed. In both cases, genes were selected as significantly differentially expressed with a 0.1% false discovery rate (Monocle2 test; Benjamini-Hochberg corrected). DE gene lists were tested for gene set enrichment against Gene Ontology and Reactome genesets (Ashburner et al., 2000; Yu and He, 2016) using the clusterProfiler Bioconductor package (Yu et al., 2012). The heatmap of differentially expressed genes was composed using the pheatmap R package (Kolde, 2015) and dendrograms were generated from the hierarchical clustering of the Euclidean distances between points.

To identify gene sets contributing to transcriptional identity in our single-cell dataset, gene-centric PCA was performed on a mean-centered matrix of variance-stabilized expression estimates for high-variance genes across all cells. The resulting rotations were used to project all expressed genes into the same PCA space to identify their weights. These weights were used to rank-order all expressed genes and this ordering was used as input for a pre-ranked GSEA analysis (Subramanian et al., 2005). Gene sets for the GSEA analysis were derived from the Monocle2 differential gene tests described above or from our curated list of neuronal activity genes (Cho et al., 2016; Kim et al., 2010; Lacar et al., 2016; Mardinly et al., 2016) using their adjusted p-value cutoff of $p < 0.01$ (Table S4).

WGCNA

To identify modules of correlated gene expression, we used the Weighted Gene Correlation Network Analysis package WGCNA (Langfelder and Horvath, 2008). Normalized expression data were first batch-corrected with respect to replicates using limma, and the resulting matrix of 12,537 detectably expressed genes and 346 cells was used as input for WGCNA. Cells were hierarchically clustered using the Jensen-Shannon distance with average linkage. A soft threshold power of 11 was used to create a signed Topological overlap matrix directly from the normalized and batch-corrected expression estimates. Modules were then learned with a minimum module size of 30. Modules were merged below a distance threshold of 0.5. Module

eigenvalues were estimated and correlated with cellular traits using the Pearson product moment correlation test. Module gene membership was determined in a similar manner in which individual gene expression estimates were correlated with module eigenvalues and tested for significance (Table S8). To test the effect of each module on the segregation of L6CThN cell types, each module eigenvalue was separately regressed out of the expression matrix using limma, and subsequent values were used as input for a tSNE using identical parameters to the original assay.

Pseudotime and BEAM analyses

Pseudotemporal ordering was performed using the prescribed Monocle2 workflow. Briefly, the 1023 L6CThNs that passed quality controls were used as input (Figure S7A-C). To reconstruct a trajectory that reflected cellular progression in response to altered sensory input, we first performed a differential test to identify genes whose expression changed as a function of collection day, independent of baseline differences between celltype (fullModelFormulaStr = ”~num_genes_expressed+Total_mRNAs+sex+celltype*Day, reducedModelFormulaStr = ” ~num_genes_expressed+Total_mRNAs+sex+celltype”, $q \leq 0.01$). These 1134 genes were used as a filtering set for the DDRTree dimensionality reduction. Cells were tested for differential expression along pseudotime using the following model comparisons with $q \leq 0.0000001$: fullModelFormulaStr = "~num_genes_expressed+celltype+sm.ns(Pseudotime,df=3)", reducedModelFormulaStr = "~num_genes_expressed+celltype". Branch dependent expression was determined using the Monocle2 BEAM test (fullModelFormulaStr = "~num_genes_expressed+celltype+sm.ns(Pseudotime, df = 3)*Branch",reducedModelFormulaStr = "~num_genes_expressed+celltype+sm.ns(Pseudotime, df = 3)", $q \leq 0.0001$). All significant gene lists were tested for gene set enrichment using the hypergeometric test.

DATA AND SOFTWARE AVAILABILITY

The Gene Expression Omnibus Accession Number for the sequencing data reported in this paper is GSE107632. Source code and software tools are available upon request.

RESOURCE TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
chicken polyclonal anti-GFP	Aves labs	GFP-1020; RRID:AB_10000240
AlexaFluor 488-conjugated goat anti-chicken	Life technologies	A11039, RRID:AB_10563770
Rabbit DsRed polyclonal antibody	Living Colors, Clontech Laboratories	RRID: AB_10013483
AlexaFluor 568-conjugated goat anti-rabbit	Life Technologies	A11011, RRID: AB_143157
Anti-Digoxigenin [21H8] Alkaline Phosphatase	Abcam	Cat. No.: ab119345, RRID:AB_10901703
Rabbit anti-β galactosidase antibody	Dr. Joshua Sanes, Harvard	
AlexaFluor 647-conjugated donkey anti-rabbit	Life Technologies	A31573, RRID: AB_2536183
Chemicals, Peptides, and Recombinant Proteins		
Alexa 488 Cholera toxin B	ThermoFisher	C34775
Alexa 555 Cholera toxin B	ThermoFisher	C34776
green Retrobeads IX	Lumafluor Inc.	NA
red Retrobeads IX	Lumafluor Inc.	NA
KAPA HiFi polymerase	KAPA Biosystems	KK2602
Critical Commercial Assays		
Worthington Papain Dissociation System	Worthington Biochemical corporation	LK003150
Qiagen RNeasy Micro Kit	Quiagen	74004
Bioanalyzer Pico RNA kit	Agilent Technologies	5067-1513
Angencourt Ampure XP bead	Beckman Coulter	A63881
SuperScript II RT kit	Thermo Scientific	11904018
Nextera XT	Illumina	FC-131-1096
Quant-iT PicoGreen dsDNA Reagent	Invitrogen	P11496
RNAscope	ACDBio	
RNAscope: Pantr1 probe	ACDBio	mm-Pantr1 483711
RNAscope: Lamp5 probe	ACDBio	mm-Lamp5 451071
RNAscope: Serpini1 probe	ACDBio	mm-Serpini1 501441
RNAscope: Gabra5 probe	ACDBio	mm-Gabra5-C3 319481-C3
Deposited Data: Gene Expression Omnibus Accession Number GSE107632		
Experimental Models: Mouse		

Neurotensin receptor-1 Cre recombinase, GENSAT 220	Mutant Mouse Regional Resource Center 017266-UCD	RRID:IMSR_MMRRC:017266
Rosa-CAG-LSL-tdTomato-WPRE (Ai9)	Allen Brain Institute, Jackson Laboratories	RRID:IMSR_JAX:007905
Rosa-CAG-LSL-EYFP-WPRE (Ai3)	Allen Brain Institute, Jackson Laboratories	RRID:IMSR_JAX:007903
Pantr1-LacZ	Dr. John Rinn, Harvard and Regeneron	(Sauvageau et al., 2013)
Oligonucleotides		
Recombinant DNA		
linc-TMEM20 probe: ATTTAGGTGACACTATAGAATACTCAAGCTATGCATCAAGC TTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTG TGCTGGAATTCGCCCTTCCCTAACTTGCAGGCTCTTGAAAA GATTCCCTGTGTTTTCATCTAAAAGCTTCAGCCTTTCATGTT TTAAATTGAGGTCTTTGACTCATTTCAAATTGATTTTGTGAA TTATTTATCTGATTCCATTTTTTTTTCTCCCAGTAGCTATTCA GTTTTTCTAGCCCCTTTTGTTTAATAGTGTCTTAGTCAGGGT TTTATTGTTGTCAAGAGACATCTTTTTTTTTTTTTTTTTTTGG TTTTTCAAGACAGGGTTTCTCTGTGTAGCCCTGGCTGTCCT GGAACTCACTCTGTAGACCAGGCTGGCCTCGAGCTCAGAA ATCCGCCTGCCTCTGCCTCCTGAGTGCTGGGATTAAAGGCA TGTGCCACCATGCCTGGCACCAAGCAACTCTTATAAAGGCA AACATTTTATTGGGTCTGGCTTACAGATTCAGAGGTTCAGT CCATTATCATCATGGCAGGAAGCATGGCAGCATCCAGGCA GACATGCTGGATAAAGGAGCTGAGAGTTCTACATCTTGAT CTGAAGGTAACCAGGAGGAGAAAGAGGTTGAGCATAGGA CCTTAAAGTCCACACCCACAGTGACAAAGGGCGAATTCTG CAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTA GAGGGCCCAATTCGCCCTATAGTGAGTCGTATTA	This paper	
Software and Algorithms		
CellProfiler	Broad Institute	
FISH-quant	(Mueller et al., 2013)	
Adobe Illustrator	Adobe Systems Incorporated	
Adobe Photoshop	Adobe Systems Incorporated	
Fiji	(Schindelin et al., 2012)	
Other		

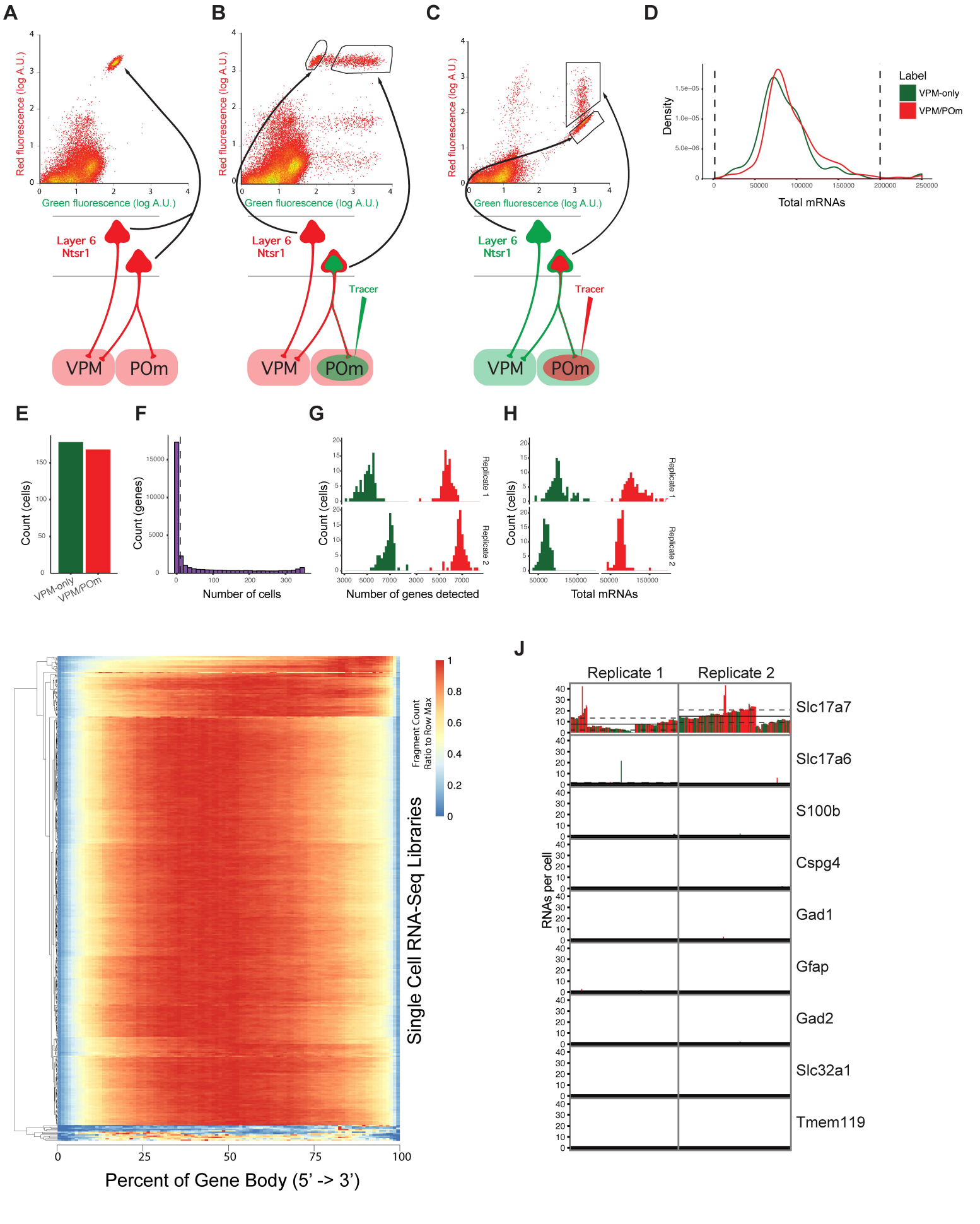


Figure S1
Chevée et al.

Figure S1: Related to Figures 1 and 2. Identification of transcriptional subtypes of layer 6 corticothalamic neurons. (A) Fluorescence-assisted cell sorting (FACS) plot showing a population of layer 6 corticothalamic neurons (L6CThNs) isolated from an Ntsr1-Cre; tdTomato mouse (arrow). (B) FACS plot of L6CThNs distinguished by their long-range projection pattern and collected for bulk RNA sequencing from one of the three replicates. The tdTomato-positive population representing L6CThNs projecting to the ventral posterior medial nucleus of the thalamus (VPM-only) is found in the same location in the FACS plot as the tdTomato-positive L6CThNs in the non-injected control (A). L6CThNs labeled with both the neuronal tracer (green Retrobeads) injected into the posterior medial nucleus of the thalamus (POm) and tdTomato formed a second population shifted to the right (red and green). The range of green fluorescence intensities is likely due to variation in the amount of neuronal tracer taken up by each individual neuron. A population of tracer-positive, tdTomato-negative neurons is likely enriched for corticothalamic neurons, which project to POm only and do not express tdTomato in Ntsr1-Cre;tdTomato mice. (C) FACS plot of L6CThNs isolated for single cell RNA sequencing (scRNA-seq) from one of two replicates, an Ntsr1-Cre;YFP mouse in which a neuronal tracer (red Retrobeads) was injected into POm. Dissociated L6CThNs projecting to VPM only (VPM-only L6CThNs), labeled with YFP only (green), and L6CThNs projecting to both VPM and POm (VPM/POm L6CThNs), labeled with YFP and neuronal tracer (red and green), were individually sorted into two 96-well plates. An additional green-only cell population representing corticothalamic neurons which project to POm-only was not collected. (D) Density plots of total mRNAs detected by scRNA-seq within each cell across the two anatomically defined classes. Dashed lines indicate the upper and lower bound for cell inclusion. (E) Barplot of the number of single cells passing scRNA-seq quality filters for each of the two classes. (F) Histogram of the number of genes with detectable expression by the number of cells in which they were detected (detection threshold = 1 estimated RNA copy per cell). Genes detected in fewer than 15 cells (dashed line) were removed from downstream analyses. (G) Histogram of the number of genes detected per cell in VPM-only and VPM/POm L6CThNs for each replicate. (H) Histogram of the total mRNAs detected per cell in VPM-only and VPM/POm L6CThNs for each replicate. (I) Heatmap of the gene body coverage averaged across all expressed genes for each of the 346 cells shows no significant 3' bias. (J) Expression levels of selected marker genes for major cortical cell types in the cells isolated for scRNA-seq. Bar graphs show the expression levels of 9 marker genes for each individual sorted cell. The markers are: *Slc17a7*, Vesicular Glutamate Transporter 1, a marker of L6CThNs (98.27%; Graziano et al., 2008); *Slc17a6*, Vesicular Glutamate Transporter 2 (0.87%), a marker not associated with L6CThNs; *Gfap* (0.57%), *Cspg4* (Ng2*,* 0.87%), *S100b* (0.87%) and *Tmem119* (0%), markers of glial cells and microglia; and *Gad1* (0.29%), *Gad2* (0.29%) and *Slc32a1* (0%), markers of inhibitory neurons. Each neuron is color-coded by its label identified during FACS (VPM-only: green; VPM/POm: red). The solid black line and dashed lines indicate the mean expression level and the standard deviation, respectively, for each gene by replicate pair.

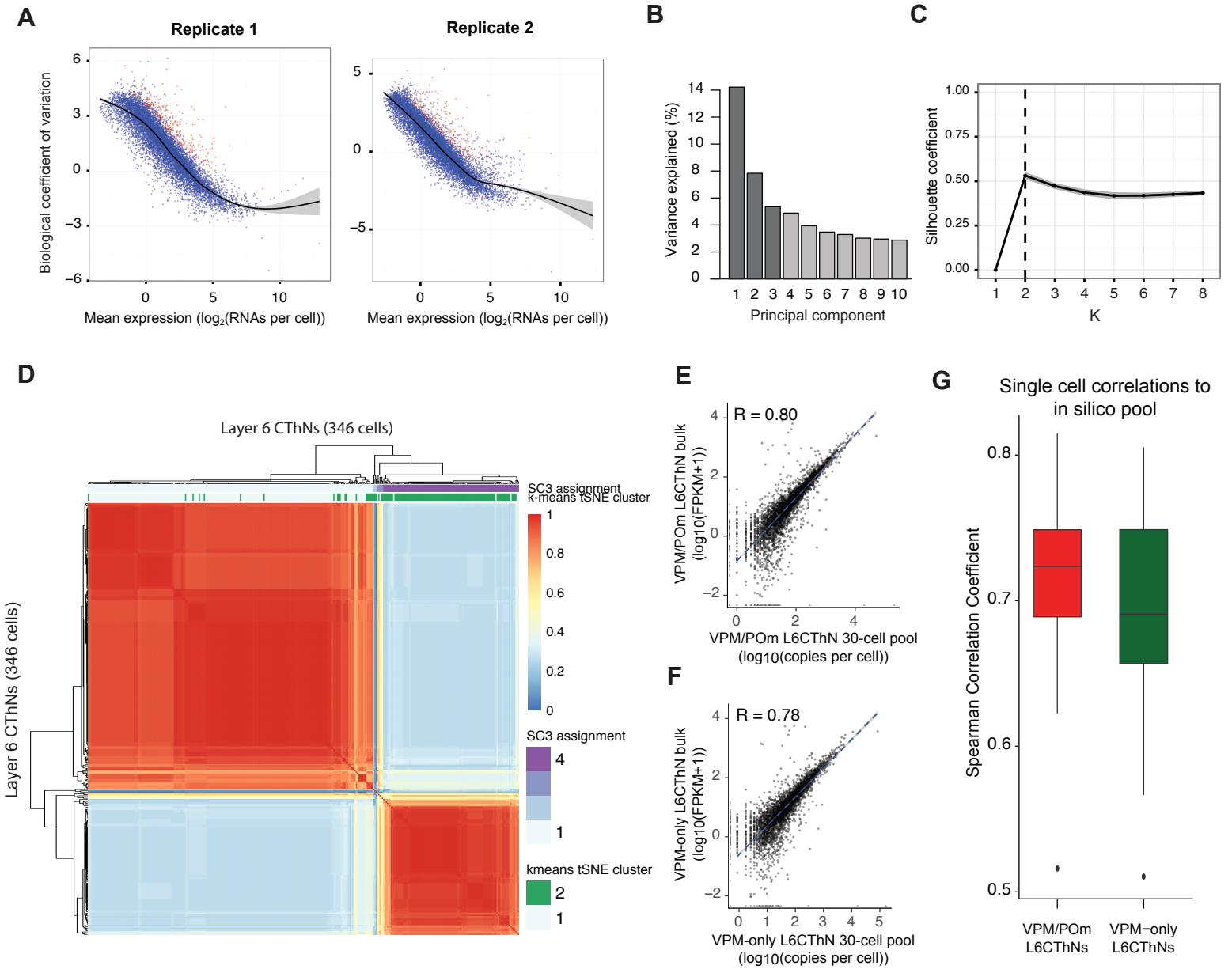


Figure S2: Related to Figure 2. scRNA-seq correlation with bulk RNA-seq and validation of clustering. (A) Scatter plot of the mean-variance relationship for all genes with detectable expression for each replicate animal. The black line represents the GAM fit for each relationship, and the grey ribbon the 95% confidence interval of the model fit. High variance genes selected as the intersection of genes with a residual >1 in both replicates are shown in red. (B) Scree plot showing the weights of each principal component obtained from a principal component analysis of 346 L6CThNs using only the 261 high variance genes common to both replicate 1 and replicate 2 (red genes plotted in (A)). The first three principal components were found to contribute to the variance above noise by permutation parallel analysis ($p<0.001$, shaded dark grey (Chung and Storey, 2015). (C) Summary graph showing the mean (black line) and sd (grey ribbon) of the average silhouette statistic for each silhouette analysis from k=2 to 8, across 1000 tSNE iterations. (D) SC3 heatmap of cell similarities in cluster assignment confirms two major cell type clusters. Color bars show agreement with our established clustering. (E,F) Scatter plots comparing bulk RNA-seq expression estimates (log2(FPKM + 1)) to in silico pools of 30 single L6CThN expression profiles categorized by label. For each subclass of L6CThN we observed high correlations across genes with detectable expression in the single cells (Spearman R = 0.80 for VPM/POm L6CThNs; Spearman R = 0.78 for VPM-only L6CThNs). (G) Boxplot of Spearman rank correlations comparing each single labeled L6CThN to 100 different computational pools of 30 sampled cells identified by their long-range axonal projections (mean Spearman R = 0.72 ± 0.04 (sd) for VPM/POm L6CThNs; mean Spearman R = 0.69 ±0.06 (sd) for VPM-only L6CThNs). These comparisons indicate that our single cell data is consistent with our bulk RNA-seq data, and that the expression patterns in single cells grossly reflect those observed in aggregate.

Figure S2
Chevée et al.

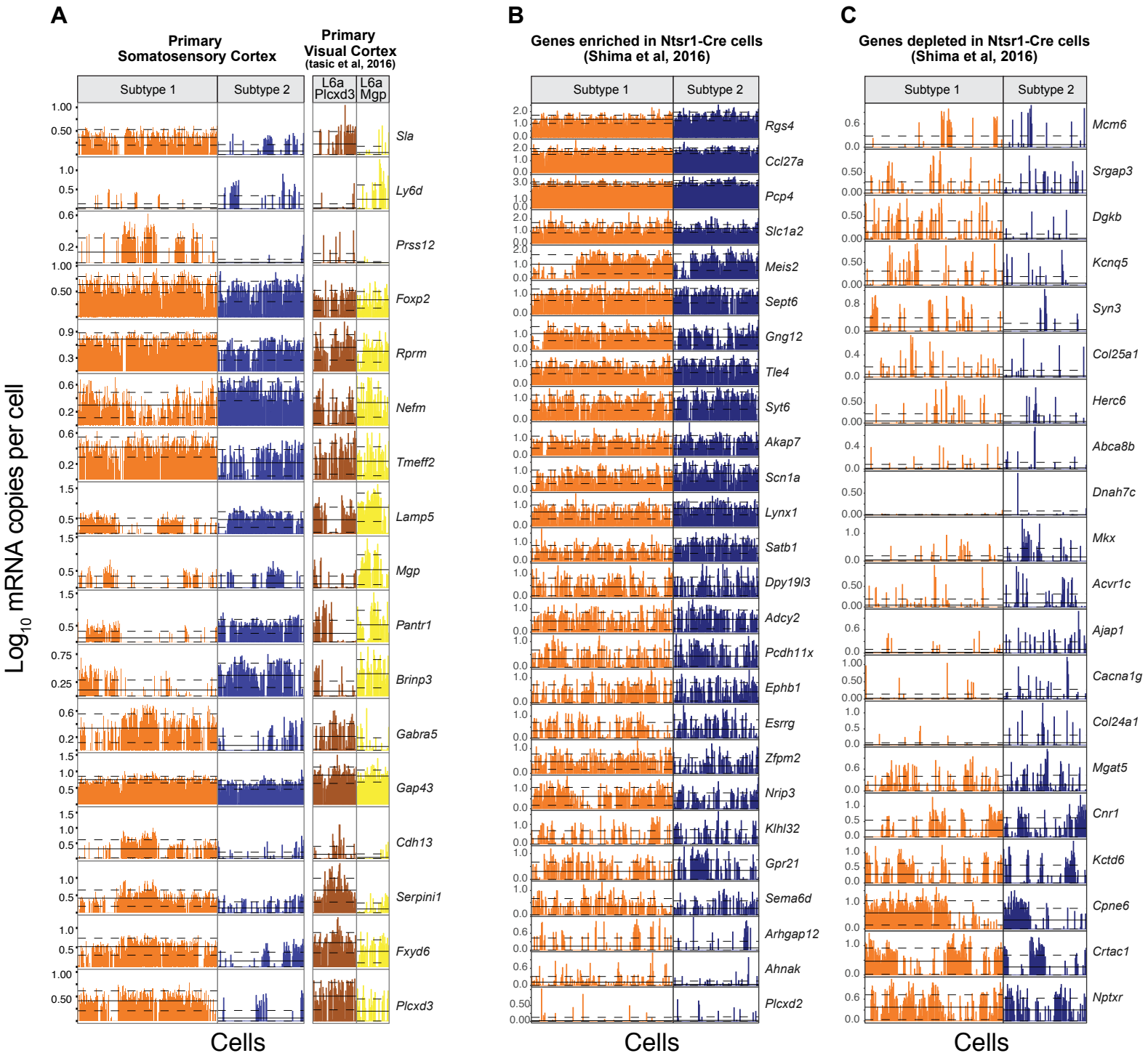


Figure S3
Chevée et al.

Figure S3: Related to Figure 2. Comparison of layer 6 corticothalamic projection neuron subtypes across studies. (A) Barplots comparing the estimated RNA copies per cell for 17 genes highlighted in a study of cortical neurons from primary visual cortex analyzed in (Tasic et al., 2016) across each of 346 single layer corticothalamic neurons (L6CThNs) in the barrel cortex of primary somatosensory cortex (S1) grouped by transcriptionally defined subtype and of 131 single neurons from V1 from Tasic et al., 2016. Cell type assignment for the V1 cells was taken directly from the publicly available annotation for these data. The solid black lines and dashed lines indicate the mean expression level and the standard deviation, respectively, for each gene by subtype and by region. Some genes, such as *Brinp3*, *Cdh13*, and *Fxyd6*, have similar expression profiles across S1 barrel cortex and V1, while other genes, such as *Mgp* and *Prss12*, show clear differences. (B,C) Comparison of genes enriched (B) and depleted (C) in the Ntsr1/P162/P139 bulk samples relative to the 56L bulk samples of corticothalamic neurons from Shima et al., 2016 (see Figure 11B, Shima et al., 2016) with data pooled across analyzed single L6CThNs from barrel cortex (Figure 2). Genes enriched in the Ntsr1/P162/P139 samples are highly expressed across the single L6CThNs from barrel cortex analyzed here (B). Genes depleted in the Ntsr1/P162/P139 bulk samples relative to the 56L bulk samples are lowly expressed across the single cell data. Note the difference in the scale bars in the two gene sets in (B) and (C).

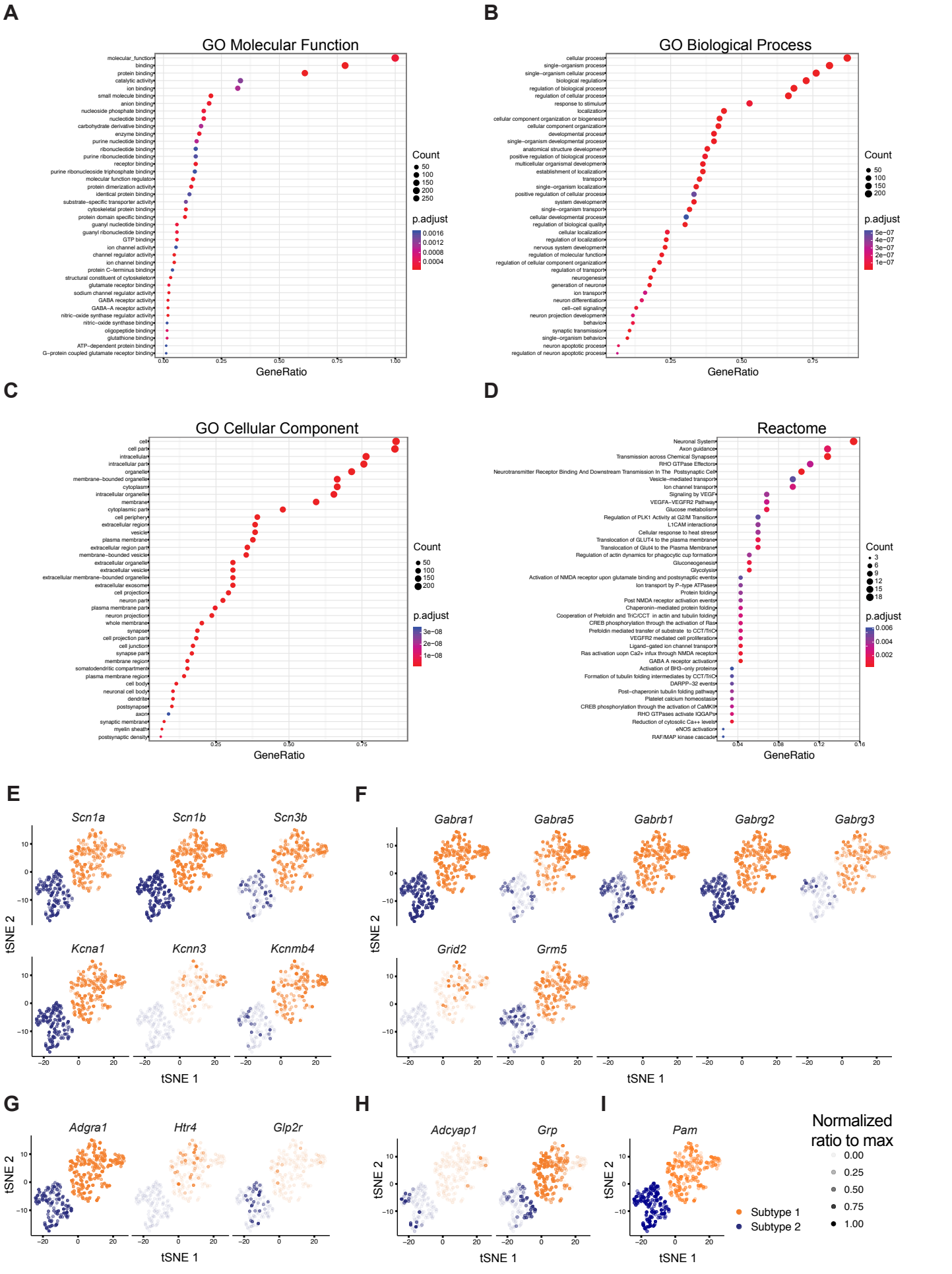
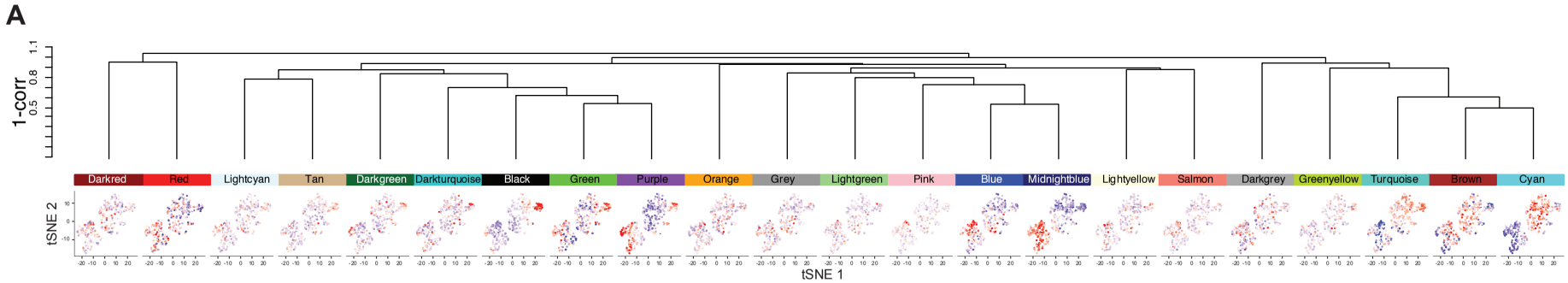


Figure S4
Chevée et al.

Figure S4: Related to Figure 2. Gene sets and example genes differentially recruited by the two transcriptionally defined subtypes of layer 6 corticothalamic neurons. (A-D) Significant Gene Ontology (A-C) and Reactome (D) gene sets associated with the genes that are differentially expressed between subtype 1 and subtype 2 by Benjamini-Hochberg-adjusted hypergeometric tests. (E-H) tSNE plots showing the expression levels of significantly differentially expressed voltage gated ion channels (E), GABA receptors (F,top), glutamate receptors (F,bottom), G protein-coupled receptors (G) and neuropeptides (H) in transcriptionally defined layer 6 corticothalamic projection neurons. (I) tSNE showing the expression levels of *Pam*, which codes for a gene involved in processing the neuropeptide PACAP (*Adcyap1*; Vaudry et al., 2000).



B
C
D
E
F
G
H

−0.1 (0.05)	−0.27 (2e−07)	−0.057 (0.3)	0.00013 (1)	0.097 (0.07)	0.05 (0.4)	0.3 (1e−08)	0.022 (0.7)	−0.51 (7e−24)	0.077 (0.2)	−0.091 (0.09)	−0.079 (0.1)	−0.086 (0.1)	−0.31 (4e−09)	−0.76 (8e−67)	−0.082 (0.1)	−0.1 (0.05)	−0.021 (0.7)	0.099 (0.06)	0.47 (8e−21)	0.12 (0.03)	0.71 (4e−55)	Subtype 1
0.1 (0.05)	0.27 (2e−07)	0.057 (0.3)	−0.00013 (1)	−0.097 (0.07)	−0.05 (0.4)	−0.3 (1e−08)	−0.022 (0.7)	0.51 (7e−24)	−0.077 (0.2)	0.091 (0.09)	0.079 (0.1)	0.086 (0.1)	0.31 (4e−09)	0.76 (8e−67)	0.082 (0.1)	0.1 (0.05)	0.021 (0.7)	−0.099 (0.06)	−0.47 (8e−21)	−0.12 (0.03)	−0.71 (4e−55)	Subtype 2
−0.13 (0.02)	−0.28 (8e−08)	−0.091 (0.09)	0.014 (0.8)	0.04 (0.5)	0.0023 (1)	0.27 (3e−07)	0.14 (0.009)	−0.15 (0.004)	0.021 (0.7)	−0.15 (0.005)	−0.094 (0.08)	−0.11 (0.04)	−0.37 (2e−12)	−0.53 (7e−26)	−0.086 (0.1)	−0.0081 (0.9)	0.033 (0.5)	0.051 (0.3)	0.37 (2e−12)	−0.037 (0.5)	0.28 (1e−07)	VPM/POm
0.13 (0.02)	0.28 (8e−08)	0.091 (0.09)	−0.014 (0.8)	−0.04 (0.5)	−0.0023 (1)	−0.27 (3e−07)	−0.14 (0.009)	0.15 (0.004)	−0.021 (0.7)	0.15 (0.005)	0.094 (0.08)	0.11 (0.04)	0.37 (2e−12)	0.53 (7e−26)	0.086 (0.1)	0.0081 (0.9)	−0.033 (0.5)	−0.051 (0.3)	−0.37 (2e−12)	0.037 (0.5)	−0.28 (1e−07)	VPM-only
0.249	0.003	0.802	0.685	0.045	0.555	2.2e-6	0.999	0.023	0.534	0.756	0.732	0.979	1.0	0.000	0.726	0.878	0.555	0.925	0.999	0.999	0.000	Subtype DE Genes
0.234	0.771	0.072	0.480	0.199	0.020	0.001	0.010	5.0e-7	0.188	0.057	0.303	0.087	0.919	0.000	0.299	0.439	0.199	0.508	1.000	0.300	0.536	Activity Genes
1e−14 (1)	−3.5e−16 (1)	8.8e−15 (1)	1.1e−14 (1)	9.3e−15 (1)	8.7e−15 (1)	7.2e−15 (1)	1.7e−15 (1)	5e−15 (1)	1.3e−14 (1)	8.2e−15 (1)	1e−14 (1)	9.8e−15 (1)	3.4e−15 (1)	2.7e−15 (1)	1.2e−14 (1)	7.5e−15 (1)	7.8e−15 (1)	1.1e−14 (1)	2.9e−15 (1)	2.5e−15 (1)	4.2e−15 (1)	Replicate
1e−14 (1)	−3.5e−16 (1)	8.8e−15 (1)	1.1e−14 (1)	9.3e−15 (1)	8.7e−15 (1)	7.2e−15 (1)	1.7e−15 (1)	5e−15 (1)	1.3e−14 (1)	8.2e−15 (1)	1e−14 (1)	9.8e−15 (1)	3.4e−15 (1)	2.7e−15 (1)	1.2e−14 (1)	7.5e−15 (1)	7.8e−15 (1)	1.1e−14 (1)	2.9e−15 (1)	2.5e−15 (1)	4.2e−15 (1)	Lane
−0.13 (0.02)	−0.37 (2e−12)	−0.029 (0.6)	0.092 (0.09)	0.17 (0.001)	0.14 (0.008)	0.62 (2e−38)	0.24 (6e−06)	−0.2 (2e−04)	0.047 (0.4)	−0.086 (0.1)	−0.099 (0.07)	−0.097 (0.07)	−0.39 (6e−14)	−0.79 (4e−76)	−0.076 (0.2)	−0.049 (0.4)	0.012 (0.8)	0.078 (0.1)	0.63 (1e−39)	−0.055 (0.3)	0.55 (2e−28)	PC1
0.036 (0.5)	−0.003 (1)	−0.078 (0.1)	−0.13 (0.02)	−0.14 (0.01)	−0.21 (1e−04)	−0.5 (6e−23)	−0.42 (5e−16)	−0.78 (8e−72)	−0.0043 (0.9)	−0.0088 (0.9)	0.023 (0.7)	0.035 (0.5)	0.028 (0.6)	−0.26 (9e−07)	−0.099 (0.07)	−0.038 (0.5)	−0.012 (0.8)	0.066 (0.2)	−0.049 (0.4)	0.42 (2e−16)	0.63 (6e−39)	PC2
0.13 (0.02)	0.094 (0.08)	0.16 (0.003)	0.098 (0.07)	−0.022 (0.7)	0.045 (0.4)	−0.016 (0.8)	0.27 (5e−07)	0.097 (0.07)	0.11 (0.04)	0.12 (0.02)	0.21 (1e−04)	0.17 (0.002)	0.47 (1e−20)	0.046 (0.4)	0.15 (0.007)	0.19 (4e−04)	0.05 (0.4)	0.14 (0.008)	−0.24 (8e−06)	0.2 (2e−04)	0.1 (0.06)	PC3
0.0091 (0.9)	−0.53 (3e−26)	−0.1 (0.06)	−0.16 (0.003)	−0.049 (0.4)	0.041 (0.5)	−0.24 (5e−06)	0.14 (0.01)	0.0043 (0.9)	−0.079 (0.1)	0.022 (0.7)	−0.081 (0.1)	−0.087 (0.1)	−0.23 (2e−05)	0.27 (5e−07)	−0.12 (0.03)	−0.071 (0.2)	0.082 (0.1)	−0.089 (0.1)	0.5 (1e−23)	0.49 (1e−22)	0.083 (0.1)	PC4
0.16 (0.003)	−0.096 (0.07)	0.17 (0.002)	0.14 (0.007)	0.052 (0.3)	−0.19 (5e−04)	−0.021 (0.7)	0.13 (0.02)	−0.092 (0.09)	0.053 (0.3)	0.048 (0.4)	0.15 (0.005)	0.19 (5e−04)	0.36 (1e−11)	0.15 (0.005)	0.034 (0.5)	0.06 (0.3)	0.13 (0.02)	0.14 (0.008)	−0.053 (0.3)	0.048 (0.4)	0.015 (0.8)	PC5

Pearson Correlation
1
0.5
0
-0.5
-1

Figure S5
Chevée et al.

Figure S5: Related to Figure 3. Visualization of the weighted gene coexpression network analysis (WGCNA) module eigengenes and their associations. (A) WGCNA on variance stabilized gene expression estimates identifies sets of coordinately regulated genes grouped using hierarchical clustering of module eigengenes. The eigengene expression for each module is shown on individual tSNE plots. Pearson correlation of each module eigengene with transcriptional subtype (B), label (C), replicate (F), lane (G), and component rotations for principal components 1-5 from a PCA on high-variance genes (H) are shown. Significance of each correlation was determined using the Pearson's product moment test ($p < 0.01$; Benjamini-Hochberg corrected). (D,E) Significant enrichment for the 286 genes differentially expressed between transcriptionally defined L6CThN subtypes (D) and genes associated with neuronal activity within each module (E) are shown in green ($p < 0.01$, hypergeometric test).

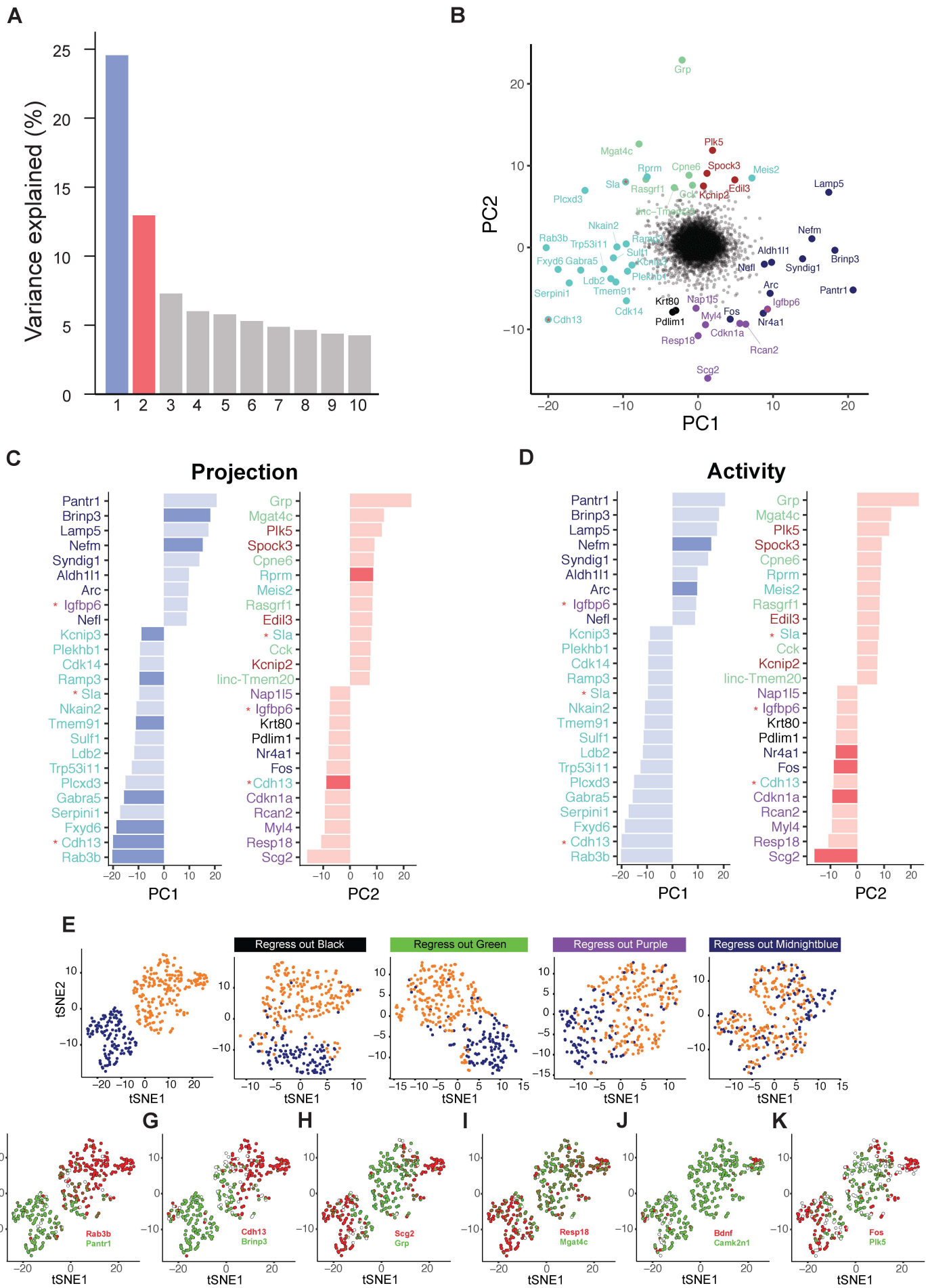


Figure S6
Chevée et al.

Figure S6: Related to Figure 4. Principal component analysis of high variance genes across all layer 6 corticothalamic neurons and influence of activity-induced genes on cellular identity. (A) Scree plot showing the explained variance of the first 10 components of a principal component analysis on the mean centered expression estimates of the high variance genes across all 346 layer 6 corticothalamic neurons (L6CThNs). (B) Projected weights of all expressed genes along PC1 versus PC2 using rotations from the previous PCA analysis. The 25 genes with the highest absolute weights along PC1 and along PC2 are labeled and color-coded by the WGCNA module they belong to (see Figure 3 and Figure S5). All other genes are shown in grey. (C,D) Weights of the top 25 genes along PC1 (blue, *left*) and PC2 (red, *right*). In (C), genes that are differentially expressed between VPM-only and VPM/POm L6CThNs are shown in dark blue (PC1) and dark red (PC2). In (D), genes induced by neuronal activity are shown in dark blue (PC1) and dark red (PC2). Genes that are shared between PC1 and PC2 are marked by an asterisk. (E) tSNE clustering of 346 L6CThNs after regressing out individual activity-associated module eigenvalues suggests that neuronal activity gene expression is a significant contributor to L6CThN transcriptional identity. tSNE plots showing the discretized expression of genes with the highest (F) and second highest (G) weights in each direction of PC1 and the highest (H) and second highest (I) weights in each direction of PC2. (J,K) tSNE plots showing the discretized expression levels of four genes reflecting neuronal activity.

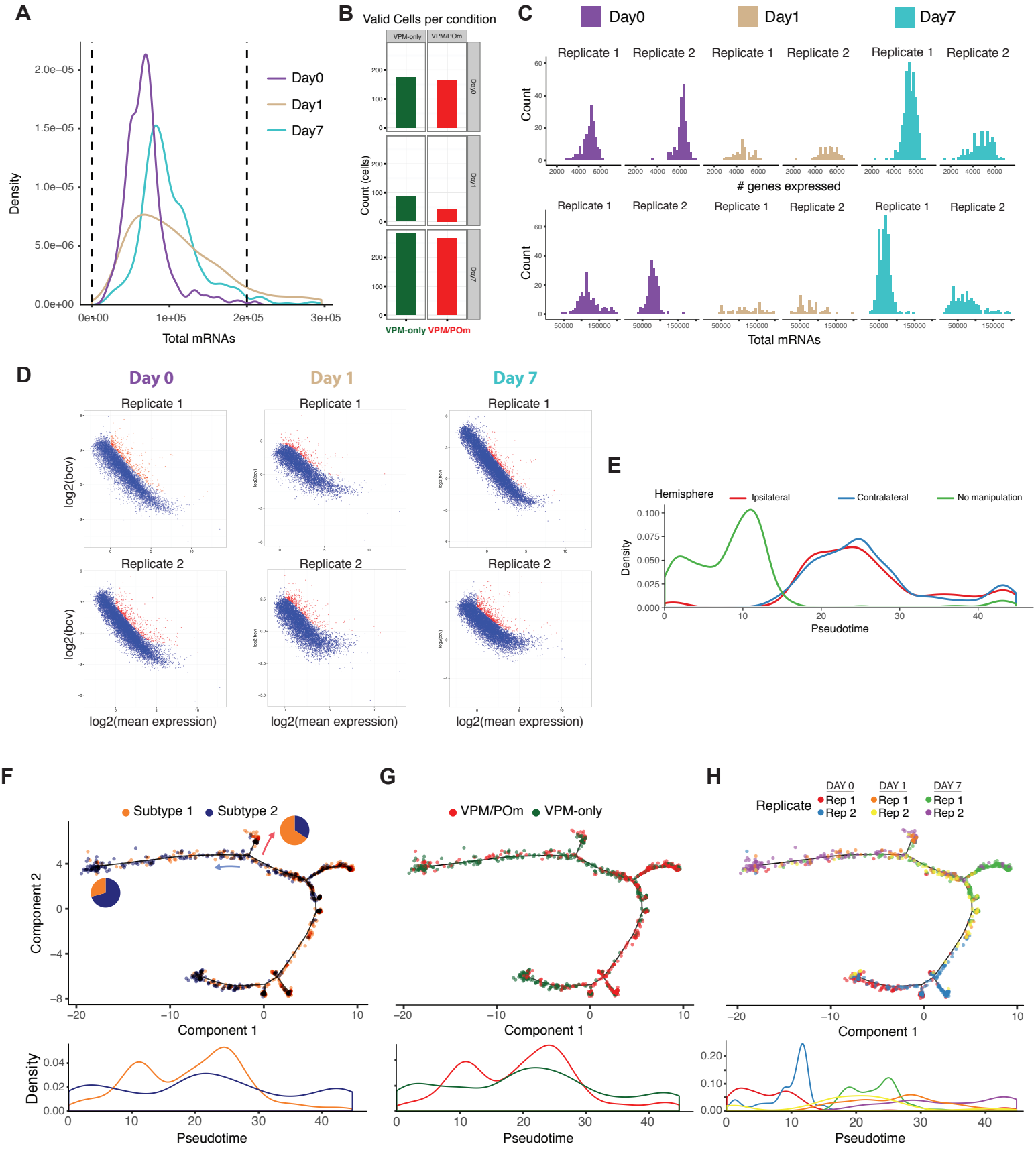


Figure S7
Chevée et al.

Figure S7: Related to Figures 5, 6 and 7. Characterization of the transcriptional responses of layer 6 corticothalamic neurons to manipulation of neuronal activity and contributions of biological and technical effects to the pseudotime trajectory. (A) Density plots of total mRNAs detected by scRNA-seq within each cell across the three conditions (Day 0, Day 1, Day 7). Dashed lines indicate the upper and lower bound for cell inclusion. (B) Barplot of the number of single cells passing scRNA-seq quality filters for each of the two classes at each time point. (C) Histogram of the number of genes detected per cell in L6CThNs at each time point for each replicate (*top*). Histogram of the total mRNAs detected per cell in L6CThNs at each time point for each replicate (*bottom*). (D) Scatter plot of the mean-variance relationship for all genes with detectable expression for each replicate animal at each time point. High variance genes selected as the intersection of genes with a residual >1 in both replicates, for each time point separately, are shown in red. (E) Density distribution of L6CThNs across pseudotime, grouped and color-coded by hemisphere relative to sensory manipulation. Interestingly, we observed no bias in the distribution of L6CThNs across pseudotime when the neurons were grouped by hemisphere ipsi- or contralateral to whisker removal, suggesting that longer term transcriptional responses in L6CThNs from both hemispheres are similar. (F,G,H) Discriminative dimensionality reduction projection (upper panels) of 1023 L6CThNs and density plots (lower panels) colored by transcriptional subtype (F), by retrograde label (G) or by replicate (H). In (F), arrows indicate the second major response branch point following sensory manipulation and pie charts indicate the proportion of each subtype in the population following the branch point. The density plot (lower panel) demonstrates that both transcriptional subtypes are present across the entire pseudotime trajectory (upper panel) including in both terminal branches. In (G), cells are classified instead by their retrograde label assignment. Cells with both labels are present across the entire pseudotime trajectory including in all terminal branches. In (H), neurons are colored by replicate. Neurons are found distributed across pseudotime for each individual replicate contributing to the pseudotime analysis (n = 6). There is no branch of the pseudotime trajectory that is restricted to a single replicate as would be expected from technical variation or batch effect.

References:

Agmon, A., and Connors, B.W. (1991). Thalamocortical responses of mouse somatosensory (barrel) cortex in vitro. Neuroscience *41*, 365-379.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T.*, et al.* (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet *25*, 25-29.

Blackshaw, S. (2013). High-throughput RNA in situ hybridization in mouse retina. Methods Mol Biol *935*, 215-226.

Cho, J.H., Huang, B.S., and Gray, J.M. (2016). RNA sequencing from neural ensembles activated during fear conditioning in the mouse temporal association cortex. Sci Rep *6*, 31753.

Chung, N.C., and Storey, J.D. (2015). Statistical significance of variables driving systematic variation in high-dimensional data. Bioinformatics *31*, 545-554.

Goff, L.A., Groff, A.F., Sauvageau, M., Trayes-Gibson, Z., Sanchez-Gomez, D.B., Morse, M., Martin, R.D., Elcavage, L.E., Liapis, S.C., Gonzalez-Celeiro, M.*, et al.* (2015). Spatiotemporal expression and transcriptional perturbations by long noncoding RNAs in the mouse brain. Proc Natl Acad Sci U S A *112*, 6855-6862.

Gong, S., Doughty, M., Harbaugh, C.R., Cummins, A., Hatten, M.E., Heintz, N., and Gerfen, C.R. (2007). Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. J Neurosci *27*, 9817-9823.

Graziano, A., Liu, X.B., Murray, K.D., and Jones, E.G. (2008). Vesicular glutamate transporters define two sets of glutamatergic afferents to the somatosensory thalamus and two thalamocortical projections in the mouse. J Comp Neurol *507*, 1258-1276.

Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T.*, et al.* (2015). Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods *12*, 115-121.

Katz, L.C., Burkhalter, A., and Dreyer, W.J. (1984). Fluorescent latex microspheres as a retrograde neuronal marker for in vivo and in vitro studies of visual cortex. Nature *310*, 498-500.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Methods *12*, 357-360.

Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S.*, et al.* (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182-187.

Kolde, R. (2015). pheatmap: Pretty heatmaps. R package verion 1.0.8. https://CRAN.R-project.org/package=pheatmap.

Krijthe, J. (2015). Rtsne: T-distributed stochastic neighbor embedding using Barnes-Hus implementation. R package version 0.10. https://CRAN.R-project.org/package=Rtsne.

Lacar, B., Linker, S.B., Jaeger, B.N., Krishnaswami, S., Barron, J., Kelder, M., Parylak, S., Paquola, A., Venepally, P., Novotny, M.*, et al.* (2016). Nuclear RNA-seq of single neurons reveals molecular signatures of activation. Nat Commun *7*, 11022.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics *9*, 559.

Li, X., Glazewski, S., Lin, X., Elde, R., and Fox, K. (1995). Effect of vibrissae deprivation on follicle innervation, neuropeptide synthesis in the trigeminal ganglion, and S1 barrel cortex plasticity. J Comp Neurol *357*, 465-481.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M.*, et al.* (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell *161*, 1202-1214.

Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R.*, et al.* (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat Neurosci *13*, 133-140.

Mardinly, A.R., Spiegel, I., Patrizi, A., Centofante, E., Bazinet, J.E., Tzeng, C.P., Mandel-Brehm, C., Harmin, D.A., Adesnik, H., Fagiolini, M., and Greenberg, M.E. (2016). Sensory experience regulates cortical inhibition by inducing IGF1 in VIP neurons. Nature *531*, 371-375.

Molyneaux, B.J., Goff, L.A., Brettler, A.C., Chen, H.H., Brown, J.R., Hrvatin, S., Rinn, J.L., and Arlotta, P. (2015). DeCoN: genome-wide analysis of in vivo transcriptional dynamics during pyramidal neuron fate selection in neocortex. Neuron *85*, 275-288.

Mudge, J.M., and Harrow, J. (2015). Creating reference gene annotation for the mouse C57BL6/J genome assembly. Mamm Genome *26*, 366-378.

Mueller, F., Senecal, A., Tantale, K., Marie-Nelly, H., Ly, N., Collin, O., Basyuk, E., Bertrand, E., Darzacq, X., and Zimmer, C. (2013). FISH-quant: automatic counting of transcripts in 3D FISH images. Nat Methods *10*, 277-278.

Picelli, S., Faridani, O.R., Bjorklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc *9*, 171-181.

Qiu, X., Hill, A., Packer, J., Lin, D., Ma, Y.A., and Trapnell, C. (2017). Single-cell mRNA quantification and differential analysis with Census. Nat Methods *14*, 309-315.

Sauvageau, M., Goff, L.A., Lodato, S., Bonev, B., Groff, A.F., Gerhardinger, C., Sanchez-Gomez, D.B., Hacisuleyman, E., Li, E., Spence, M.*, et al.* (2013). Multiple knockout mouse models reveal lincRNAs are required for life and brain development. Elife *2*, e01749.

Saxena, A., Wagatsuma, A., Noro, Y., Kuji, T., Asaka-Oba, A., Watahiki, A., Gurnot, C., Fagiolini, M., Hensch, T.K., and Carninci, P. (2012). Trehalose-enhanced isolation of neuronal sub-types from adult mouse brain. Biotechniques *52*, 381-385.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B.*, et al.* (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676-682.

Shima, Y., Sugino, K., Hempel, C.M., Shima, M., Taneja, P., Bullis, J.B., Mehta, S., Lois, C., and Nelson, S.B. (2016). A mammalian enhancer trap resource for discovering and manipulating neuronal cell types. Elife *5*, e13503.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T.*, et al.* (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat Neurosci *19*, 335-346.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol *32*, 381-386.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol *31*, 46-53.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562-578.

Vaudry, D., Gonzalez, B.J., Basille, M., Yon, L., Fournier, A., and Vaudry, H. (2000). Pituitary adenylate cyclase-activating polypeptide and its receptors: from structure to functions. Pharmacol Rev *52*, 269-324.

Wood, S.N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) *73*, 3-36.

Yu, G., and He, Q.Y. (2016). ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Mol Biosyst *12*, 477-479.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284-287.