

 Open access • Posted Content • DOI:10.1101/241315

A topographic axis of transcriptional identity across thalamic projection systems

— [Source link](#) 

James W Phillips, Anton Schulmann, Erina Hara, Chenghao Liu ...+7 more authors

Institutions: Brandeis University

Published on: 20 Feb 2018 - bioRxiv (Cold Spring Harbor Laboratory)

Related papers:

- [A single spectrum of neuronal identities across thalamus](#)
- [Distinct subnetworks of the thalamic reticular nucleus.](#)
- [Higher-Order Thalamic Circuits Channel Parallel Streams of Visual Information in Mice.](#)
- [Local connections to specific types of layer 6 neurons in the rat visual cortex](#)
- [Patterning of thalamus](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/a-topographic-axis-of-transcriptional-identity-across-49c1ibg5ht>

A single spectrum of neuronal identities across thalamus

James W. Phillips^{1,3†,*}, Anton Schulmann^{1,2†}, Erina Hara^{1,¥}, Chenghao Liu², Lihua Wang¹,
Brenda C. Shields^{1,∅}, Wyatt Korff¹, Andrew L. Lemire¹, Joshua Dudman¹, Sacha B. Nelson^{1,2,*},
Adam Hantman^{1,*}

¹HHMI Janelia Research Campus, Ashburn, Virginia, USA

²Brandeis University, Waltham Massachusetts, USA

³Department of Physiology, Development and Neuroscience, University of Cambridge, UK

*Correspondence to: James Phillips (phillipsj10@janelia.hhmi.org), Sacha Nelson

(nelson@brandeis.edu), Adam Hantman (hantmana@janelia.hhmi.org)

†Co-first authorship

¥Present Address: Department of Molecular, Cellular & Biomedical Sciences, CUNY School of
Medicine, 160 Convent Ave, New York, NY 10031

∅Present Address: Duke University - Pratt School of Engineering 311 Research Drive, Durham,
NC 27710

ABSTRACT

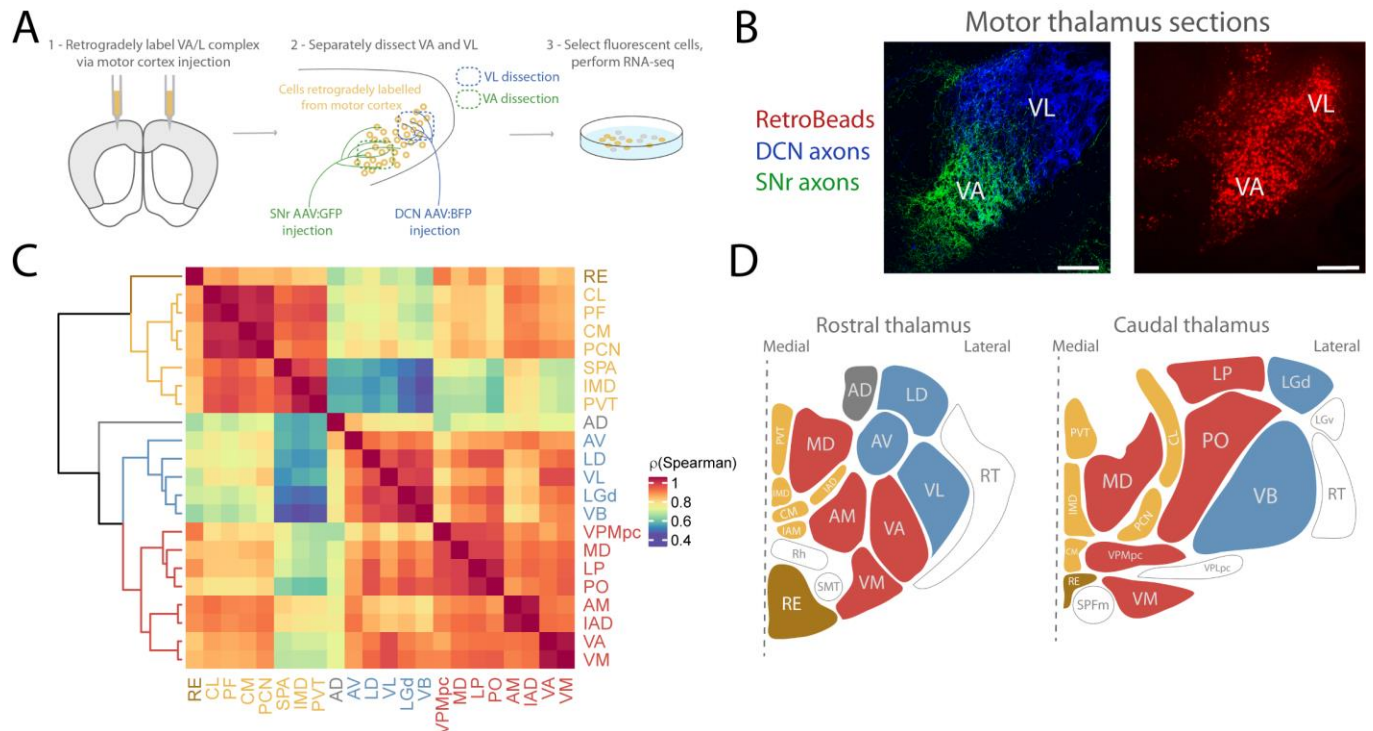
Uncovering common principles by which diverse modalities of information are processed is a fundamental goal in neuroscience. In mammalian brain, thalamus is the central processing station for inputs from sensory systems, subcortical motor systems, and cortex; a function subserved by over 30 defined nuclei^{1,2}. Multiple thalamic nuclei send convergent information to each region of the forebrain, but whether there is a conserved architecture across the set of thalamic pathways projecting to each forebrain area has remained unresolved³⁻⁵. To uncover organizational principles of thalamic pathways, we produced a near-comprehensive transcriptomic atlas of thalamus. This revealed a common logic for thalamic nuclei serving all major cortical modalities. We found that almost all nuclei belong to one of three major profiles, with a given cortical area getting input from each of these profiles. These profiles lie on a single axis of variance aligned with the mediolateral axis of thalamus, and this axis is strongly enriched in genes encoding receptors and ion channels. We further show that each projection profile exhibits different electrophysiological signatures. Single-cell profiling revealed that rather than forming discrete classes, thalamic neurons lie on a spectrum, with intermediate cells existing between profiles. Thus, in contrast to canonical models of thalamus that suggest it is a switchboard primarily concerned with routing distinct modalities of information to distinct cortical regions, we show that the thalamocortical system is more akin to a molecularly-defined ‘filter bank’ repeatedly applied across modality. Together, we reveal striking covariation in the organization of thalamic pathways serving all input modalities and output targets, establishing a simple and comprehensive thalamic functional architecture.

MAIN TEXT

To understand the organization of thalamic pathways, we combined anatomical and genetic approaches to produce a near-comprehensive, projection-specific transcriptomic atlas of murine

47 thalamus. Thalamic nuclei were retrogradely labeled from individual forebrain areas,
 48 microdissected and cells pooled (8 projection targets, 22 nuclei, 120 samples, Extended Data
 49 Tables 1 and 2). Anterograde tracing of inputs to thalamus was used when identification of
 50 nuclear boundaries was ambiguous (Fig. 1, A and B, Extended Data Table 1). We then used
 51 hierarchical clustering to explore the relationship between the transcriptomes of thalamic nuclei
 52 (on the 500 most differentially expressed genes via an ANOVA-like test, see methods, Extended
 53 Data Fig. 1B and Supplementary Table 2), and identified five major subdivisions of nuclei across
 54 thalamus (Fig. 1C). Anterior dorsal nucleus (AD) and nucleus reuniens (RE) each formed
 55 profiles of their own, leaving three major multi-nuclei profiles. These major profiles were not
 56 explained by cortical projection target or modality, since the multiple nuclei projecting to motor,
 57 somatosensory or visual cortices split across different profiles. For example, central medial
 58 (CM), ventral anterior (VA) and ventral lateral (VL) nuclei all project to motor cortex, but are
 59 split across the three profiles. By typically receiving input from each of these profiles, each
 60 cortical region samples from all three genetically defined pathways. Our nuclear subgrouping
 61 also did not obey rules of the prominent core/matrix scheme of thalamus, *Calb1*⁺ neurons which
 62 define ‘matrix’ nuclei are split by the first branch of our hierarchical clustering (Extended Data
 63 Fig. 2)⁵⁻⁷. Rather, our three major profiles were best distinguished by anatomy with nuclei of
 64 each profile occupying a characteristic position along the mediolateral axis of the thalamus (Fig.
 65 1D). We thus find that the architecture of thalamus is dominated by genetic differences that are
 66 organized topographically.

67
 68



69
 70
 71
 72
 73
 74

Fig. 1. A near-comprehensive transcriptomic atlas allows unbiased clustering of thalamic gene expression profiles

A. Schematic of experimental pipeline to obtain transcriptomic atlas of the thalamus. In this example, motor thalamic neurons were retrogradely labelled from their primary

75 projection target (motor cortex), manually dissected and sorted. Viruses expressing green
76 and blue fluorescent proteins (GFP, BFP), respectively, were injected to the deep
77 cerebellar nuclei (DCN) and substantia nigra pars reticulata (SNr) to label motor nuclear
78 subdivisions (ventral lateral (VL) and ventral anterior (VA), respectively) previously
79 identified^{8,9}.

80 B. Example labelling from scheme shown in A. Coronal sections. Scale bars = 200 μ m.

81 C. Hierarchical clustering of thalamic nuclei using Spearman's correlation of top 500 most
82 differentially expressed genes (Extended Data Fig. 1B) across all 22 nuclei. Major
83 clusters defined as the top 5 branches of cluster dendrogram.

84 D. Topographic localization of gene expression profiles in thalamus. Coronal thalamic
85 section schematics with nuclei colored as in Fig. 1C. Unsampled nuclei are left
86 uncolored. (see Fig. 1B).

87

88

89 To understand the pattern of gene expression differences between the thalamic projection
90 profiles, we performed principal component analysis (PCA) on our data. Principal component 1
91 (PC1, 38% explained variance) separated nuclei into the same major profiles identified via
92 hierarchical clustering (Fig. 2A, Extended Data Fig. 3A). Again, position on this first dimension
93 strongly correlated with mediolateral position, demonstrating topogenetic architecture in
94 thalamus (Fig. 2B, Extended Data Fig. 3B). Based on their relative order on this first component,
95 we named the three major profiles primary, secondary, and tertiary. The progressive difference
96 from primary to tertiary nuclei was also evident in the number of genes differentially expressed
97 between the groups, with the primary and tertiary nuclei being most distinct, and the other two
98 comparisons being less so (Fig. 2C). This primary axis was prominently enriched in genes
99 encoding neurotransmitter receptors, ion channels, and signaling molecules (Fig. 2D and E).

100 Thus, nuclei of the major profiles sit continuously along a single axis of genetic variance which
101 is aligned with the mediolateral spatial axis of thalamus.

102

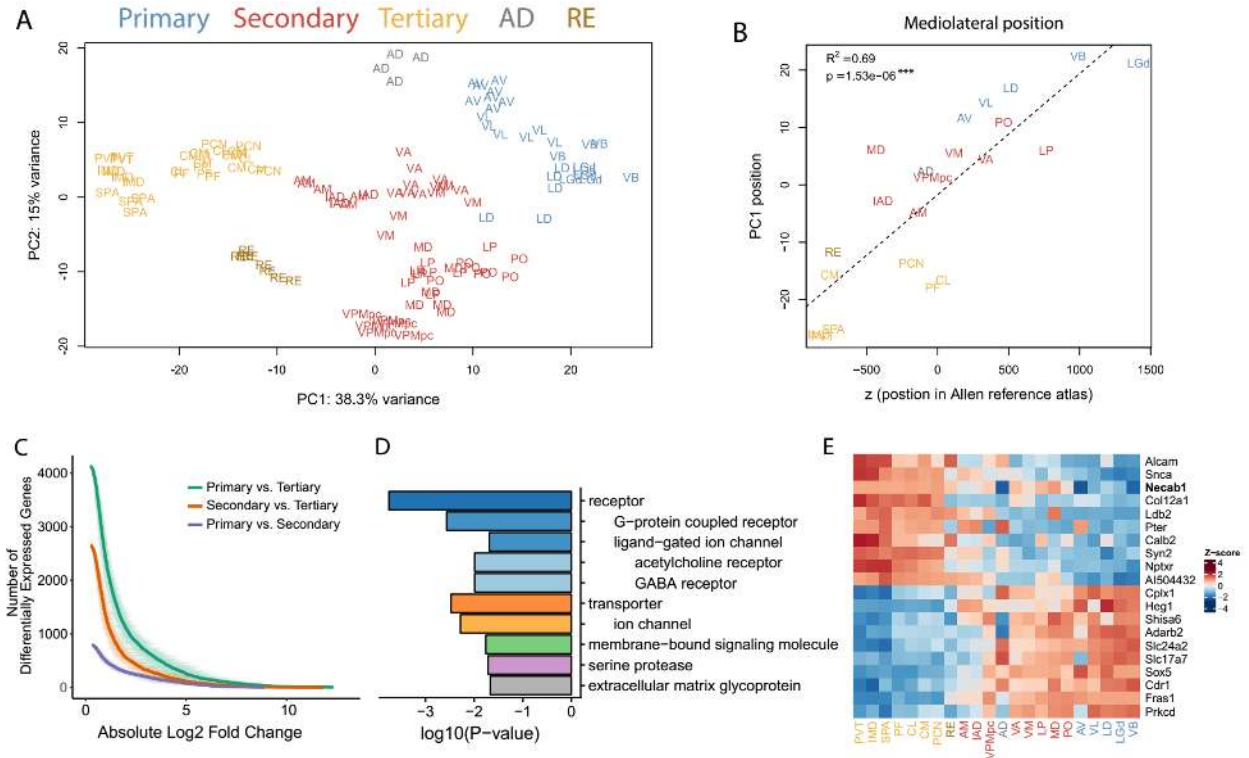
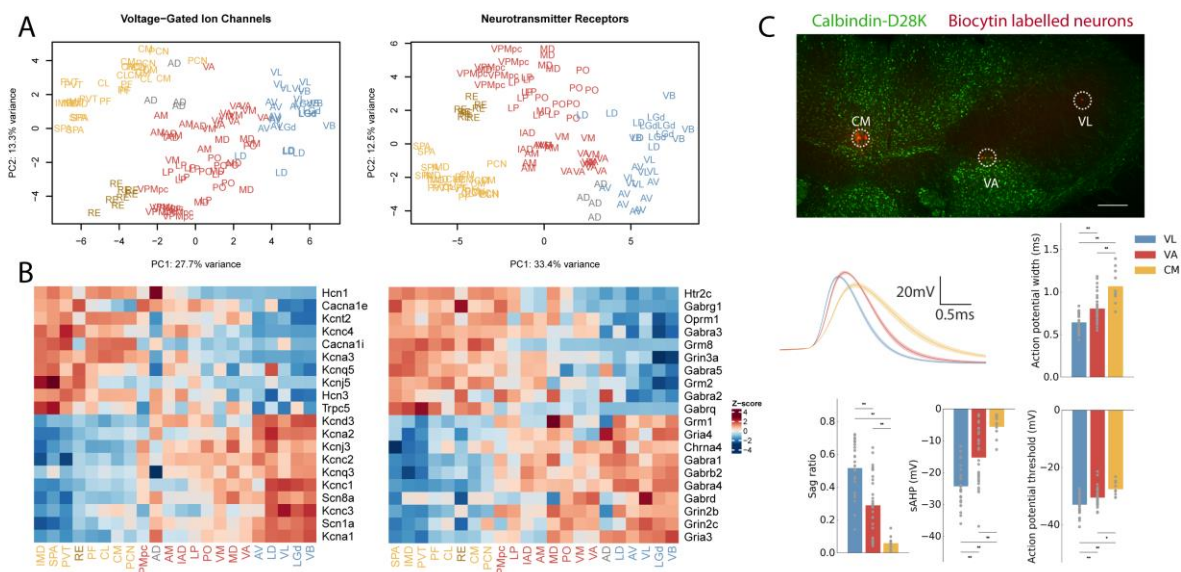


Fig. 2. A common topographic axis of variance separates major thalamic profiles

- A. PCA showing separation of functional nuclear profiles in the first two principal components. The underlying gene set and color scheme are the same as in 1C.
- B. PC1 position is highly correlated with mediolateral position of the nuclei. Mediolateral positions are z voxel coordinates of nuclei centers in the Common Coordinate Framework (CCF) in the Allen Mouse Brain Atlas. 1 voxel = 10 μ m.
- C. Primary nuclei are farthest from tertiary nuclei, with secondary nuclei being intermediate. Plot shows the number of differentially expressed genes at each log fold change level (shown as mean \pm standard deviation) for the three comparisons.
- D. Genes relevant to neurotransmission are overrepresented amongst the top 100 genes with the highest absolute PC1 loadings in our dataset. The ten most highly overrepresented PANTHER protein class terms are shown. P-values based on hypergeometric test. Indentation indicates gene subfamily.
- E. Heatmap of genes with strongest positive and negative loadings on PC1. Nuclei are ordered by their mean position on PC1 of Fig. 2A. Colors represent gene-wise Z-scores.

Given the prominent differences in receptor and ion channel expression between thalamic profiles, we asked whether these profiles correspond to functionally distinct classes of neurons. We first performed PCA on the expression profiles of voltage-gated ion channel or neurotransmitter/modulator receptor encoding genes (Fig. 3A, left and right respectively). Analysis with these limited gene sets reproduced the separation of profiles in PC1 (Fig. 3A), confirming that ion channel and receptor profiles are organized along the same axis identified in Fig. 2. Genes linked to high firing rates via faster channel kinetics, such as Kv3 channels (*Kcnc1*, *Kcnc3*), the *Scn8a* channel, and the *Kcnab3* subunit¹²⁻¹⁴, tended to be progressively elevated

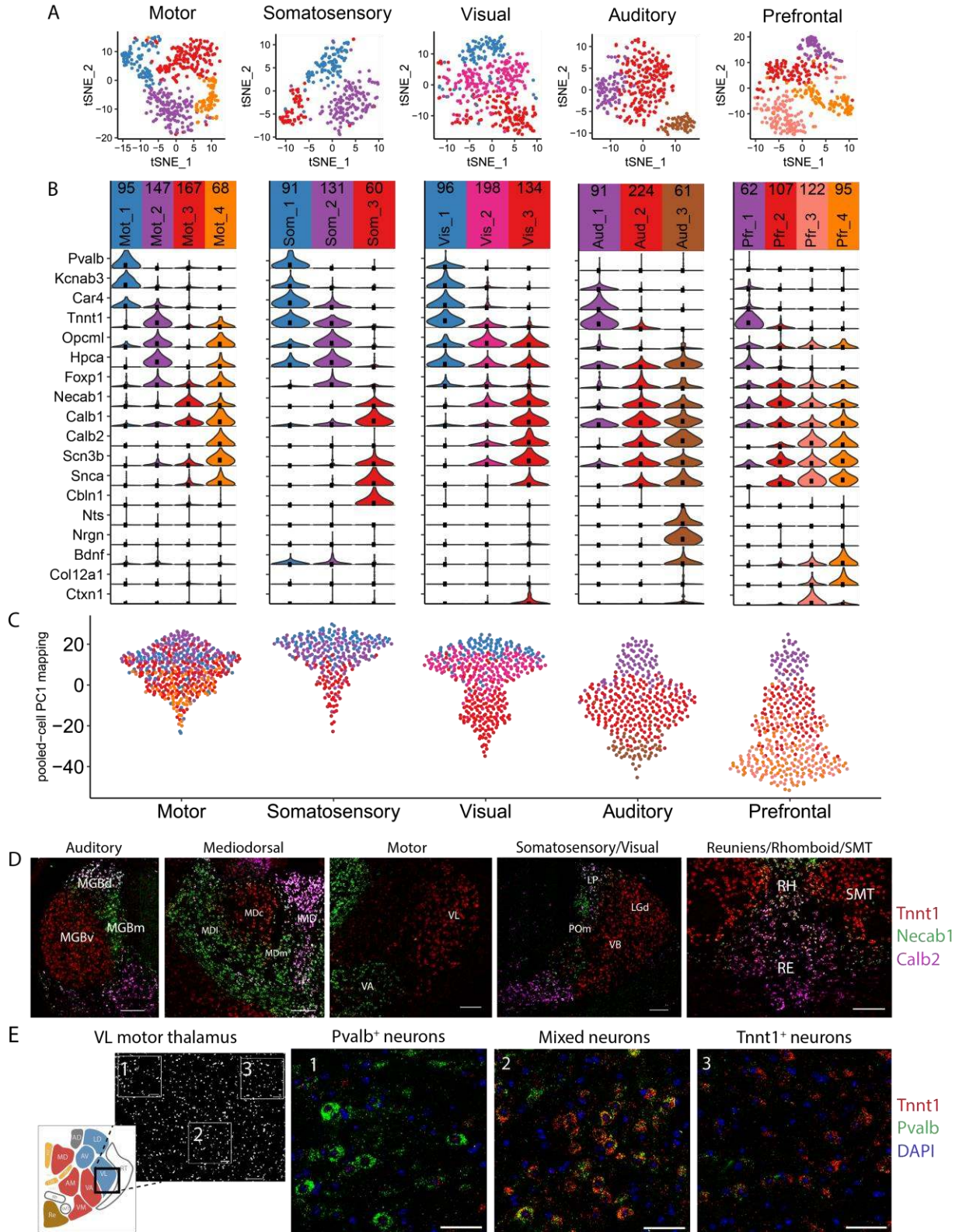
130 toward primary profile nuclei. This raised the possibility that action potential width may
 131 progressively narrow from tertiary to primary nuclei (Fig. 3B). Whole-cell recordings from the
 132 motor-related nuclei CM, VA and VL (representing the three main nuclear profiles; Fig. 2E)
 133 confirmed this prediction (Fig. 3C). Neurons recorded within VL have the narrowest action
 134 potential width and those in CM the widest. In addition, many other electrophysiological
 135 properties showed a systematic gradient ranging from VL through VA to CM (Fig. 3C, and
 136 Extended Data Fig. 4). Prior work has shown substantial differences in electrophysiological
 137 properties between different thalamic nuclei, but to date this has not been incorporated into
 138 thalamic organizational schemes⁹⁻¹¹. Our electrophysiology recordings thus show a systematic
 139 variation of neuronal firing properties from primary to tertiary profiles. The strong enrichment of
 140 neuromodulatory genes, especially among nuclei of the secondary and tertiary type, suggests
 141 further differential modulation of inputs across these profiles.
 142
 143



144
 145
 146 **Fig. 3. Functionally relevant genes and electrophysiological properties vary systematically**
 147 **across nuclear profiles**

- 148 A. PCA including only genes encoding voltage-gated ion channels and
 149 neurotransmitter/neuromodulator receptors. Colors as in Fig. 1C.
 150 B. Heatmap for genes with the highest gene loadings in PC1 from Fig. 3A. Voltage-gated
 151 ion-channels on the left and neurotransmitter receptors on the right. Colors represent
 152 gene-wise Z-scores.
 153 C. Systematic variation of electrophysiological properties across profiles. Whole-cell patch-
 154 clamp recordings from VL (primary thalamus), VA (secondary thalamus) and CM
 155 (tertiary thalamus). Top row: neurons were labelled with biocytin (red) and localized to
 156 individual nuclei with the aid of Calbindin-D28K immunolabelling (green). Scale bar =
 157 100 μ m. Middle row, left shows average action potential shape for VL, VA and CM
 158 neurons (mean \pm SEM). Remaining panels show comparisons for four physiological
 159 measurements across these nuclei (One-way ANOVA with *post-hoc* Tukey HSD test, all
 160 comparisons $P < 0.05$). Sample contained 29 VL neurons, 34 VA neurons and 10 CM
 161 neurons.

162
163



164
165

166 **Fig. 4. A topographic spectrum of thalamic cell identities within and between thalamic**
167 **nuclei**

- 168 A. Overview of single-cell clusters within each projection system visualized via tSNE. Cells
169 colored by cluster identity.
- 170 B. Violin plots for cluster marker genes (inclusion criteria were Likelihood Ratio Test P-
171 value $< 10^{-5}$ and log2 fold change > 0.5 for each cluster).
- 172 C. Projection of single-cell data onto pooled-cell PC1 from Fig. 2A. Each dot is a single cell
173 colored as in A.
- 174 D. Topographic distribution of marker genes within 6 major thalamic modalities. Multi-
175 FISH with probes for *Calb2* (pink), *Tnnt1* (red) and *Necab1* (green). See Extended Data
176 Fig. 7 and 8 for expanded views and quantification. Scale bars = 200 μm .
- 177 E. Multi-FISH within VL thalamus comparing genes marking clusters in single-cell data.
178 Left panel shows field of view (coronal section, scale bar = 100 μm). Right three panels
179 (scale bars = 50 μm) show expansion of the three boxed areas, moving left to right.
180 Middle box shows intermediate cells expressing both single cell cluster markers. (red =
181 *Tnnt1*, green = *Pvalb*, blue = DAPI).

182
183
184 Thus far we analyzed thalamic nuclei by pooling projection neurons from specific anatomical
185 positions. The resulting transcriptional profiles could represent homogenous populations,
186 discrete subtypes, or cells with graded differences. To probe these possibilities, we profiled the
187 transcriptomes of individual neurons from motor, somatosensory, visual, auditory, and prefrontal
188 projection classes (Extended Data Table 3). Analysis of this single-cell RNAseq dataset resulted
189 in multiple clusters for each projection class, and cluster markers included many genes that also
190 distinguished nuclei (Fig. 4B). Single-cell clusters also separated along the first principal
191 component derived from our pooled-cell RNAseq dataset (Fig. 4C, Extended Data Fig. 6).
192 Markers for the single-cell clusters were spatially separated at the single-cell level, but with
193 intermediate cells especially prominent near the anatomical boundaries of thalamic nuclei (Fig.
194 4D, Extended Data Fig. 7). This is consistent with spatially organized heterogeneity or a
195 gradient-like organization rather than intermingled, discrete populations^{15,16}.

196
197 Given the strong relationship between PC1 and anatomical nuclei position (Fig. 2C and Extended
198 Data Fig. 3), the presence of single-cell clusters occupying similar PC1 positions (e.g. clusters 1
199 and 2 of the motor projections neurons, or clusters 1 and 2 of the somatosensory nuclei,
200 Extended Data Fig. 6) suggested that distinct neuron types could also coexist within anatomical
201 boundaries of nuclei. We examined this possibility by performing multi-color fluorescent *in situ*
202 hybridization (multi-FISH) for genes which distinguished amongst the major profiles (e.g. *Pvalb*
203 and *Tnnt1*). Taking motor thalamus as an example, *Pvalb* and *Tnnt1* expressing cells were found
204 within the anatomical boundaries of a single thalamic nucleus (VL; Fig. 4E). Some individual
205 VL cells expressed both *Pvalb* and *Tnnt1*; *Pvalb*-selective, intermediate, and *Tnnt1*-selective
206 cells were distributed along the mediolateral axis (Fig. 4E). Therefore, spatially organized
207 transcriptomic differences can exist even within individual thalamic nuclei.

208
209 A common organizing principle of thalamus divides nuclei into discrete core or matrix subtypes
210 that span modalities^{1,7}. To date the strongest evidence for cross-modal organization is that
211 genetic differences can be larger within a sensory modality than between sensory

212 modalities¹⁷(Extended Data Fig. 9). However, previous studies have focused either on select
213 sensory systems and/or have had limited scope with respect to the molecules investigated¹⁸.
214 Here, using the full transcriptomes of nearly all thalamic pathways, we confirm cross-modal
215 organization but replace the core/matrix dichotomy with a spectrum of profiles that span a single
216 axis of genetic variance. For example, we find that the matrix subtype has substantial diversity,
217 splitting into multiple profiles based upon hierarchical clustering (Extended Data Fig. 2). The
218 axis of variance is dominated by genes that directly shape neuronal properties, leading to
219 conserved, systematic variation of function. This pattern of variation is not only imposed on
220 sensory thalamocortical systems but also diversifies motor, limbic, and cognitive thalamocortical
221 systems. Understanding how the pattern of intra-thalamic molecular variability intersects with
222 input modalities and behavioral relevance will be an important challenge for future work¹⁹⁻²³.

223
224 Our single-cell transcriptomics indicated that a given cortical area samples across a spectrum of
225 thalamic profiles. Some projection neurons exhibited features of more than one profile,
226 suggesting gradient-like transitions between cell types. Using multi-FISH, we mapped these
227 intermediate cell types both to the boundaries between nuclei and within nuclei. Intermediate
228 cells might exhibit hybrid input-output transforms, morphologies, and/or functions.

229
230 Given that neurons spanning the full range of profiles provide input to nearly all cortical areas,
231 our data suggests that thalamus acts as a "filter bank" providing each recipient forebrain area
232 with a broad complement of differentially filtered inputs (Extended Data Fig. 10).

233
234

235 **Materials and Methods**

236
237

237 **EXPERIMENTAL METHODS**

238
239

239 *Animal care*

240
241

241 Experimental procedures were approved by the Institutional Animal Care and Use Committee
242 (IACUC) at the Janelia Research Campus. Mice were housed on a 12 hour light/dark cycle, with
243 *ad libitum* food and water. The majority of mice were 8-12 weeks old (Supplementary Tables 1
244 and 3).

245
246

246 *Acquisition of samples*

247
248

248 Cells were fluorescently labelled to enable manual dissection. This was done through retrograde
249 labelling via either viral or tracer injection into the major projection field of the nucleus of
250 interest. For viral injections, rAAV2-retro expressing cre-dependent GFP or TdTomato under the
251 CAG promoter were injected, with volumes of 50-100nL at 3 depths (see Extended Data Table
252 2)²⁴. Minimum survival time was 3 weeks post-injection. Viruses were prepared by Janelia
253 Virus Services. Non-viral retrograde tracer labelling used the lipophilic tracer DiI (2.5mg/ml in
254 DMSO, injecting volumes of 50-200nL per site, from Molecular probes) or Lumafuor red
255 retrobeads (diluted 3x in PBS, 50-200nL per site). Anterograde labelling of inputs to thalamus
256 was also used in a small number of cases (see Extended Data Tables 1 and 2).

257

258 We referred to the Paxinos and Franklin mouse brain atlas to guide our dissections²⁵. For the
259 majority of regions of thalamus, retrograde tracers labeled populations corresponding to
260 identified thalamic nuclei (Extended Data Tables 1 and 2 for targeting details). However, the
261 caudal intralaminar nuclei (parafascicular complex) were less clearly delineated. This likely
262 reflects additional heterogeneity within this complex beyond that shown in atlases²⁶.

263
264 At no stage were experimenters blinded to sample identity.

265 ***Manual cell sorting and RNAseq***

266 *Sorted pooled-cell RNAseq*

267
268 Fluorescent cells were collected and sequenced as previously described^{15,27}. Briefly, animals were
269 deeply anaesthetized with isoflurane and euthanized. Coronal slices (200-300 μ m) were cut and
270 placed for 1 hour at room temperature with pronases and neural activity blockers in artificial
271 cerebrospinal fluid (ACSF). Relevant regions were then microdissected, and the tissue
272 dissociated. The resulting cell suspensions were diluted with filtered ACSF and washed at least 3
273 times by transferring them to clean dishes. This process produces negligible contamination with
274 non-fluorescent tissue (Extended Data Fig. 1A)²⁷. After the final wash, samples were aspirated in
275 a small volume (3 μ l) and lysed in 47 μ l XB lysis buffer using the Picopure kit (KIT0204,
276 ThermoFisher) in a 200 μ l PCR tube (Axygen), incubated for 30 min at 40 °C on a thermal cycler
277 and stored at -80 °C.

278
279
280 Library preparation and sequencing was performed by the Janelia Quantitative Genomics core.
281 RNA was isolated from each sample using the PicoPure RNA isolation kit (Life technologies)
282 and on-column RNAase-free DNase I treatment (Qiagen). 1 μ L ERCC RNA spike-in mix (Life
283 technologies) was added to each sample. Amplification was then performed using the Ovation
284 RNA-seq v2 kit (NuGEN), yielding 4-8 μ g of cDNA. The Ovation rapid DR multiplexing kit
285 (NuGEN) was used to make libraries for sequencing, which were sequenced on a HiSeq 2500
286 (Illumina).

287 *Sorted single-cell RNAseq*

288
289 Retrogradely labeled cells were isolated as described above, and collected into 8-well strips
290 containing 3 μ L Smart-seq2 lysis buffer, flash-frozen on dry ice, and stored at -80°C until further
291 use²⁸.

292
293 Upon thawing, cells were re-digested with Proteinase K and barcoded RT primers were added.
294 cDNA synthesis was done using the Maxima H Minus RT kit (Thermo Fisher) and E5V6NEXT
295 template switch oligo, followed by heat inactivation reverse transcriptase. PCR amplification
296 using the HiFi PCR kit (Kapa Biosystems) and SINGV6 primer was performed with a modified
297 thermocycling protocol (98°C for 3 min, 20 cycles of 98°C for 20s, 64°C for 15s, 72°C for 4 min,
298 final extension at 72°C for 5 min). Samples were then pooled across strips, purified with Ampure
299 XP beads (Beckman Coulter), washed twice with 70% ethanol and eluted in water. These pooled
300 strips were then combined to create the plate-level cDNA pool for tagmentation, and
301 concentration was determined using Qubit High-Sensitivity DNA kit (Thermo Fisher).

302

303 Tagmentation and library preparation using 600 pg cDNA from each plate of cells was then
304 performed with a modified Nextera XT (Illumina) protocol, but using the P5NEXTPT5 primer
305 and tagmentation time extended to 15 minutes(30). The libraries were then purified following the
306 Nextera XT protocol (at 0.6X dilution) and quantified by qPCR using Kapa Library
307 Quantification (Kapa Biosystems). 6-10 plates were run on a NextSeq 550 flow cell. Read 1
308 contained the cell barcode and unique molecular identifier (UMI). Read 2 was a cDNA fragment
309 from the 3' end of the transcript.

310
311

312 ***Multi-FISH***

313

314 C57Bl/6J mice (~8 weeks old) were anesthetized with isoflurane then fixed via transcathal
315 perfusion with PBS followed by 4% paraformaldehyde in PBS, pH 7.4. Brains were post-fixed at
316 4 °C overnight, washed 3 times with PBS, and cryoprotected in a sucrose series of 10%, 20%
317 then 30% in PBS at 4 °C. All solutions were prepared RNase-free. Brains were sectioned (14µm)
318 on a Leica CM3050S cryostat, mounted onto Fisher SuperFrost Plus slides, and stored at -80 °C .
319 Multi-FISH was performed using the RNAscope Multiplex Fluorescent Assay platform from
320 ACDBio, following the manufacturer's protocol. The probes used were: *Calb2* (ref 313641-C3),
321 *Necab1* (ref 428541 and 428541-C2), and *Tnnt1* (ref 466911-C2). Fluorescent dyes were DAPI,
322 Alexa Fluor 488, Atto 550 and Atto 647. Images were acquired using a Zeiss LSM 880 confocal
323 microscope. Images were acquired with an air 20x (0.8NA) objective unless otherwise specified.

324

325 ***Electrophysiology***

326

327 Acute brain slices were prepared from p20-25 mice. Animals were deeply anesthetized with
328 ketamine/xylazine/acepromizine and transcathal perfused with ice-cold oxygenated cutting
329 solution containing (in mM): 74 NaCl, 3 KCl, 1 NaH₂PO₄, 25 NaHCO₃, 6 MgCl₂, 0.5 CaCl₂, 5
330 Sodium Ascorbate, 75 Sucrose, 10 Glucose. 300 µm coronal slices containing the thalamus were
331 cut on a vibratome (Leica), and then recovered for 15 min at 33 °C and for 15 min at room
332 temperature in oxygenated cutting solution followed by at least another 1 hour at room
333 temperature in oxygenated ACSF containing (in mM): 126 NaCl, 3 KCl, 1 NaH₂PO₄, 25
334 NaHCO₃, 2 MgCl₂, 2 CaCl₂, 10 Glucose. During recordings, slices were perfused with
335 oxygenated 34-35 °C ACSF with 35 mM d,l-2-amino-5-phosphonovaleric acid (APV), 20 mM
336 6,7-dinitroquinoxaline-2,3-dione (DNQX) to block glutamatergic synaptic transmission and 50
337 mM picrotoxin to block GABAergic synaptic transmission. Target neurons in CM, VA and VL
338 were identified based on their distance to the mammillothalamic tract and nuclear borders were
339 confirmed with calbindin immunostaining *post hoc*. Whole-cell recording pipettes (6 – 8 MΩ)
340 were filled with internal solution containing (in mM): 100 K-gluconate, 20 KCl, 10 HEPES, 4
341 Mg-ATP, 0.3 Na-GTP, 10 Na-phosphocreatine, and 0.1% biocytin. Current-clamp recordings
342 were obtained with Multiclamp 700B amplifiers (Molecular Devices) digitized at 10 kHz using
343 IGOR Pro (WaveMetrics). Resting membrane potentials were adjusted to -65 mV and steady
344 state series resistance was compensated. Neurons with series resistance > 30 MΩ or membrane
345 potentials that changed by > 3 mV were excluded.
346 Custom IGOR scripts were used to analyze the data. For each neuron, threshold, amplitude,
347 afterhyperpolarization and half width at half height of the 16th-19th action potentials in trials
348 with 20 to 40 Hz firing rate were averaged.

349

350 *Immunohistochemistry*

351

352 After recordings, slices were fixed with 4% paraformaldehyde and 2.5% sucrose in 5x
353 phosphate-buffered saline (PBS) at 4 °C for 2-10 days. After washing with PBS, slices were
354 blocked in PBS with 0.3% Triton and 3% BSA at 4 °C overnight and then incubated in PBS with
355 0.3% Triton and 3% BSA and rabbit anti-calbindin D-28k (Swant, 1:1000) at 4 °C overnight.
356 After washing, they were incubated in PBS with 0.3% Triton, 3% BSA and 5% goat serum with
357 fluorescent protein conjugated goat anti-rabbit IgG (Invitrogen, 1:1000) and streptavidin
358 (Invitrogen, 1:1000) at 4 °C overnight to label calbindin-expressing neurons and biocytin-filled
359 neurons.

360

361 **ANALYSIS METHODS**

362

363 *Pooled-cell RNAseq analysis*

364

365 *Data processing and quality control*

366 After removing Illumina adapter sequences using cutadapt, reads were mapped to the mouse
367 reference genome (mm10) using STAR with 'ENCODE settings' for RNAseq²⁹. Mean mapping
368 rate was 82.29% with a standard deviation of 2.25%. Unique unambiguous exon-mapping reads
369 were summarized at the gene level using Gencode version M13.

370

371 Contamination with common astrocytic, oligodendrocytic, erythrocytic and microglial transcripts
372 was low, consistent with a lack of substantial contamination by non-fluorescent cells (Extended
373 Data Fig. 1A). To ensure the specificity of our dissections and to control for potential batch
374 effects, we collected several nuclei through multiple independent labelling approaches, and
375 showed that these samples cluster in a highly similar manner (Extended Data Fig. 1C).

376

377 *Differential gene expression*

378 Differential expression was assessed using the Bioconductor package *edgeR*³⁰. Low counts were
379 removed by requiring a Transcripts per million (TPM) > 5 in at least 3 samples. This yielded a
380 list of approximately 17,000 expressed genes. Counts were then fitted to a negative binomial
381 generalized linear model, where each factor level represents a different thalamic nucleus, and a
382 Likelihood Ratio Test was used to assess differential expression between groups. P-values were
383 adjusted for multiple tests using the Benjamini-Hochberg method. Genes with false-discovery
384 rate < 0.05 were considered differentially expressed. For selecting the most differentially
385 expressed genes between any thalamic nuclei, we used an ANOVA-like test (ANODEV test for
386 generalized linear models, as described in edgeR User manual 3.2.6), testing for differences
387 between any of the 22 nuclei, and used the 500 genes with the highest P-value.

388 To avoid bias due to differences in sample number when comparing numbers of differentially
389 expressed genes between different profiles in Fig. 2C, the groups were subsampled (with
390 replacement) to the size of the smallest group. Bootstrapped log₂ fold changes were obtained
391 over 100 iterations.

392 For visualization, clustering and machine learning of gene expression data, we used variance-
393 stabilized counts produced by the variance-stabilizing transformation in the *DESeq2* R
394 package^{31,32}.

395 For assessing the role of modality vs. hierarchical class on distinguishing thalamic nuclei, we
396 used elastic-net regularized logistic regression classifiers. Models were trained with different
397 numbers of randomly selected genes as features over 100 iterations. To avoid bias due to variable
398 group size, groups were subsampled to the size of the smallest group. Model tuning was
399 performed using the *glmnet* and *caret* packages in R, and accuracy of the best model was
400 assessed using 5-fold cross-validation.

401
402 *Unsupervised clustering and functional enrichments*

403 Hierarchical clustering was performed using 1 - Spearman's correlation as a distance metric and
404 complete linkage for agglomeration. Groups were defined by splitting the tree at the level of 5
405 branches. We termed these profiles, not clusters, as we do not mean to imply discreteness
406 between the classifications. PCA was done using the singular value decomposition based *prcomp*
407 function in R. For functional enrichment of differentially expressed genes, we used the
408 PANTHER Protein Class Ontology
409 (http://data.pantherdb.org/PANTHER13/ontology/Protein_Class_13.0), which is a consolidated
410 version of molecular function gene ontology. Over-representation in the top 100 genes with the
411 highest PC1 loadings was assessed via hypergeometric test.

412 For defining voltage-gated ion channels and neurotransmitter receptors, we downloaded the
413 IUPHAR/BPS database (http://www.guidetopharmacology.org/DATA/targets_and_families.csv).
414 Voltage-gated ion channels were the genes defined in the database, while for neurotransmitter
415 receptors we included ionotropic and metabotropic receptors for glutamate, GABA, glycine,
416 acetylcholine, 5-HT, dopamine, trace amine, histamine, and opioids.

417
418 **Single-cell RNAseq analysis**

419
420 *Data processing and quality control*

421 Single-cell RNAseq data was trimmed for adapters using cutadapt and aligned to the mouse
422 genome (mm10) using STAR. To demultiplex cells, collapse UMIs and produce gene-wise
423 counts for each cell, we used a modified version of the *Drop-seq_tools-1.12* pipeline. Briefly,
424 read 1 was tagged based on the cell barcode and UMI, and this information was added to read 2
425 by merging back the reads after mapping, followed by gene-wise tagging of reads that map onto
426 exons and summarization of digital counts.

427 Single cells were required to have more than 20,000 UMIs and more than 2,500 genes detected
428 per cell, which yielded a total of 1,971 cells (Extended Data Fig. 5A). Of these, 22 cells were
429 found to be significantly contaminated with oligodendroglial and vascular cell transcripts
430 (Extended Data Fig. 5C), leaving 1,949 cells for all downstream analyses. Genes were
431 considered expressed if their expression was detected in more than 10 cells.

432 Our single-cell sequencing was not comprehensive, and with improved sequencing approaches
433 further genetic subdivisions may be identified. Single-cell and pooled-cell dissections were not
434 precisely matched, for example motor-projecting midline nuclei were not dissected for single-
435 cell RNAseq. However, pooled-cell and single-cell RNAseq are in close agreement (Extended
436 Data Fig. 4B), indicating that our results are robust to collection method.

437
438 *Single-cell clustering and marker genes*

439 Single-cell clusters were defined using the *Seurat* R package (version 2.0)^{33,34}. Data were log
440 transformed and scaled. For identifying variable genes, genes were divided into 20 bins based on

441 average expression, and genes that were more than 1 standard deviation away from average
442 dispersion within a bin were used for downstream analysis. Single-cell clustering was performed
443 separately for each projection system using shared nearest neighbor (SNN) clustering and
444 limiting the analysis to the top 10 principal components for distance calculation. Clusters were
445 defined by the Louvain algorithm, and clustering resolution set to 0.6. Clusters of cells were
446 visualized using t-distributed stochastic neighbor embedding (tSNE) using the top 10 principal
447 components as input and perplexity set to 30. Marker genes for each cluster were required to be
448 expressed in at least 80% of the cells in the cluster, to have a P-value $<10^{-5}$ (Likelihood Ratio
449 Test), a log₂ fold change > 0.5 . Projection of single-cell data onto pooled-cell principal
450 components was obtained by multiplying (dot product) log-transformed and scaled single-cell
451 data by the pooled-cell principal component loadings.

452 453 454 REFERENCES

- 455
- 456 1. Jones, E. *The Thalamus*. Cambridge University Press (2007).
- 457 2. Sherman, S. M. Thalamus plays a central role in ongoing cortical functioning. *Nature*
458 *Neuroscience* **19**, 533–541 (2016).
- 459 3. Sherman, S. M. & Guillery, R. W. The role of the thalamus in the flow of information to
460 the cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences* **357**,
461 1695–1708 (2002).
- 462 4. Smith, Y. *et al.* The thalamostriatal system in normal and diseased states. *Front Syst*
463 *Neurosci* **8**, 5 (2014).
- 464 5. Jones, E. Viewpoint: The core and matrix of thalamic organization. *Neuroscience* (1998).
- 465 6. Jones, E. G. The thalamic matrix and thalamocortical synchrony. *Trends in Neurosciences*
466 **24**, 595–601 (2001).
- 467 7. Clascá, F., Rubio-Garrido, P. & Jabaudon, D. Unveiling the diversity of thalamocortical
468 neuron subtypes. *European Journal of Neuroscience* **35**, 1524–1532 (2012).
- 469 8. Kuramoto, E. *et al.* Complementary distribution of glutamatergic cerebellar and
470 GABAergic basal ganglia afferents to the rat motor thalamic nuclei. *European Journal of*
471 *Neuroscience* **33**, 95–109 (2010).
- 472 9. Nakamura, K. C., Sharott, A. & Magill, P. J. Temporal Coupling with Cortex
473 Distinguishes Spontaneous Neuronal Activities in Identified Basal Ganglia-Recipient and
474 Cerebellar-Recipient Zones of the Motor Thalamus. *Cerebral Cortex* **24**, 81–97 (2012).
- 475 10. Fogerson, P. M. & Huguenard, J. R. Tapping the Brakes: Cellular and Synaptic
476 Mechanisms that Regulate Thalamic Oscillations. *Neuron* **92**, 687–704 (2016).
- 477 11. Puil, E., Meiri, H. & Yarom, Y. Resonant behavior and frequency preferences of thalamic
478 neurons. *Journal of Neurophysiology* **71**, 575–582 (1994).
- 479 12. Freeman, S. A., Desmazières, A., Fricker, D., Lubetzki, C. & Sol-Foulon, N. Mechanisms
480 of sodium channel clustering and its influence on axonal impulse conduction. *Cellular and*
481 *Molecular Life Sciences* **73**, 723–735 (2015).
- 482 13. Leicher, T., Bähring, R., Isbrandt, D. & Pongs, O. Coexpression of the KCNA3B gene
483 product with Kv1.5 leads to a novel A-type potassium channel. *J. Biol. Chem.* **273**,
484 35095–35101 (1998).
- 485 14. Rudy, B. & McBain, C. J. Kv3 channels: voltage-gated K⁺ channels designed for high-
486 frequency repetitive firing. *Trends in Neurosciences* **24**, 517–526 (2001).

- 487 15. Cembrowski, M. S. *et al.* Spatial Gene-Expression Gradients Underlie Prominent
488 Heterogeneity of CA1 Pyramidal Neurons. *Neuron* **89**, 351–368 (2016).
- 489 16. Gokce, O. *et al.* Cellular Taxonomy of the Mouse Striatum as Revealed by Single-Cell
490 RNA-Seq. *CellReports* **16**, 1126–1137 (2016).
- 491 17. Frangeul, L. *et al.* A cross-modal genetic framework for the development and plasticity of
492 sensory pathways. *Nature* **538**, 96–98 (2016).
- 493 18. Rubio-Garrido, P., Pérez-de-Manzo, F. & Clascá, F. Calcium-binding proteins as markers
494 of layer-I projecting vs. deep layer-projecting thalamocortical neurons: A double-labeling
495 analysis in the rat. *Neuroscience* **149**, 242–250 (2007).
- 496 19. Guo, Z. V. *et al.* Maintenance of persistent activity in a frontal thalamocortical loop.
497 *Nature* **545**, 181–186 (2017).
- 498 20. Schmitt, L. I. *et al.* Thalamic amplification of cortical connectivity sustains attentional
499 control. *Nature* **545**, 219–223 (2017).
- 500 21. Bolkan, S. S. *et al.* Thalamic projections sustain prefrontal activity during working
501 memory maintenance. *Nature Neuroscience* **20**, 987–996 (2017).
- 502 22. Lee, S. *et al.* Bidirectional modulation of fear extinction by mediodorsal thalamic firing in
503 mice. *Nature Publishing Group* **15**, 308–314 (2011).
- 504 23. Llinas, R. R. Bursting of Thalamic Neurons and States of Vigilance. *Journal of*
505 *Neurophysiology* **95**, 3297–3308 (2006).
- 506 24. Tervo, D. G. R. *et al.* A Designer AAV Variant Permits Efficient Retrograde Access to
507 Projection Neurons. *Neuron* **92**, 372–382 (2016).
- 508 25. Paxinos, G. & Franklin, K. The Mouse Brain in Stereotaxic Coordinates, 4th Edition.
509 *Academic Press* (2012).
- 510 26. Kita, T., Shigematsu, N. & Kita, H. Intralaminar and tectal projections to the subthalamus
511 in the rat. *European Journal of Neuroscience* **44**, 2899–2908 (2016).
- 512 27. Hempel, C. M., Sugino, K. & Nelson, S. B. A manual method for the purification of
513 fluorescently labeled neurons from the mammalian brain. *Nat Protoc* **2**, 2924–2929
514 (2007).
- 515 28. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells.
516 *Nat Meth* **10**, 1096–1098 (2013).
- 517 29. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21
518 (2012).
- 519 30. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for
520 differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–
521 140 (2009).
- 522 31. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion
523 for RNA-seq data with DESeq2. *Genome Biol.* **15**, 31–21 (2014).
- 524 32. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome*
525 *Biol.* **11**, R106 (2010).
- 526 33. Butler, A. & Satija, R. Integrated analysis of single cell transcriptomic data across
527 conditions, technologies, and species. *bioRxiv* 1–18 (2017). doi:10.1101/164889
- 528 34. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of
529 single-cell gene expression data. *Nat Biotechnol* **33**, 495–502 (2015).
- 530
531
532

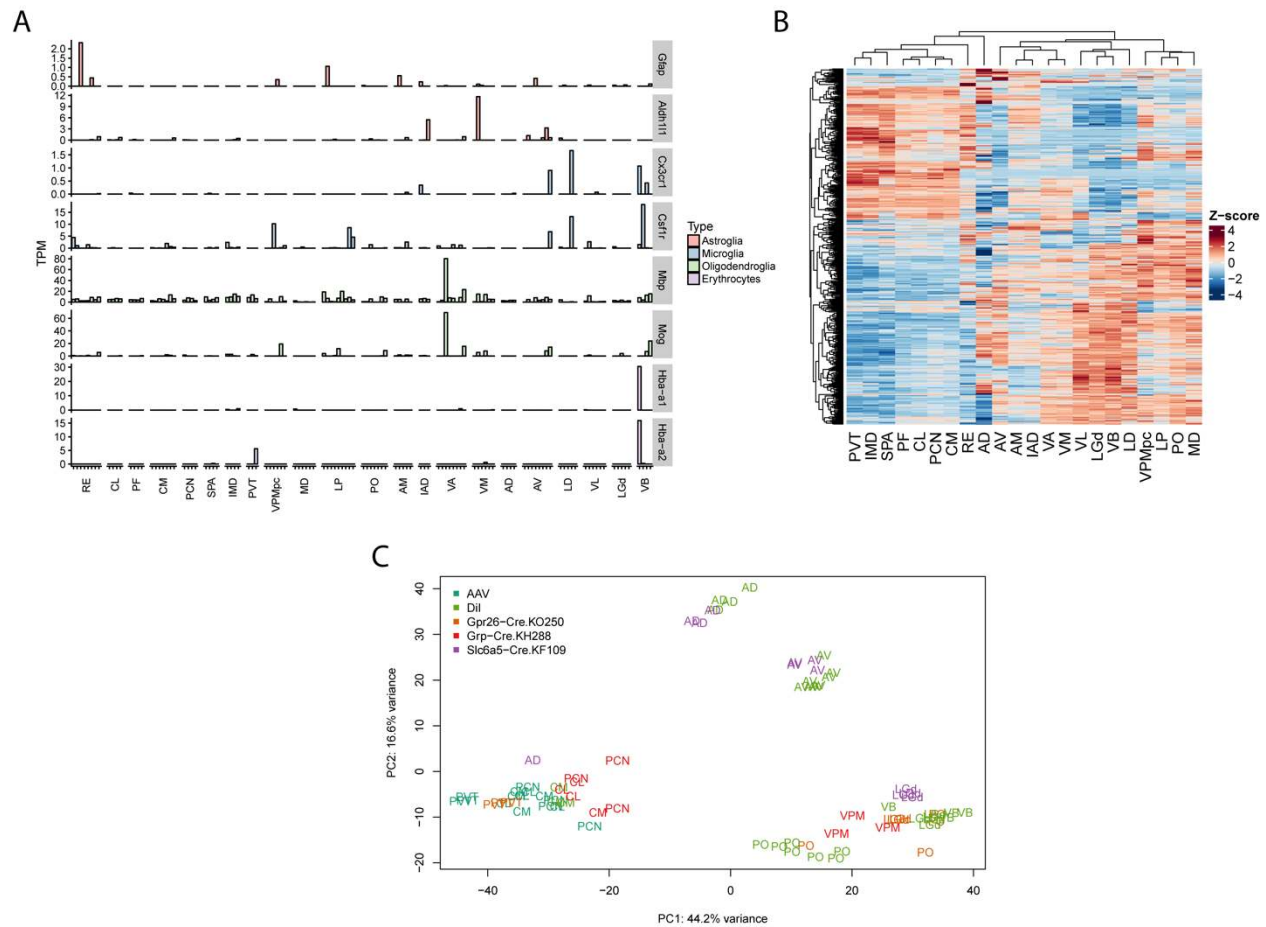
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556

End Notes:

Funding: This project was funded as a small project team (ThalamoSeq) by HHMI at the Janelia Research Campus, following a pilot project in the Dudman/Hantman labs. SN and CL were also supported by grants from NINDS (NS079419) and NIMH (MH105949). AS is funded via the Janelia Visiting Scientists Program. **Contributions:** JP: contributed to all aspects of this project. AS: Analyzed and collected data, planned project and wrote the paper. EH: Planned project and collected data. CL: Collected and analyzed electrophysiology data. LW: Collected data. BS: Collected data. WK: Supervised project, AL: Collected data and developed methods JD: Supervised project and wrote the paper SN: Supervised project. AH: Initiated and supervised project, wrote the paper. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All transcriptomic data used will be made publicly available via the Gene Expression Omnibus. **Further acknowledgements:** We thank Karel Svoboda, Albert Lee and Amy Chuong for critical input throughout the project. We thank Matthew Phillips, Kirandeep Ghataorhe, Brett Mensh and Yves Weissenberger for comments on the manuscript. We thank Vilas Menon, Damian Kao and Mark Cembrowski for help with single-cell RNAseq analysis. We thank Kim Ritola and the Janelia Virology and histology cores for production of viruses and histology. We thank Daniel Morozoff, Yajie Llang, Justin Little, Amy Chuong and Na Ji for surgical protocols and assistance identifying nuclei for dissection, and the Janelia vivarium services for animal care and surgeries.

557
558
559

Extended data figures

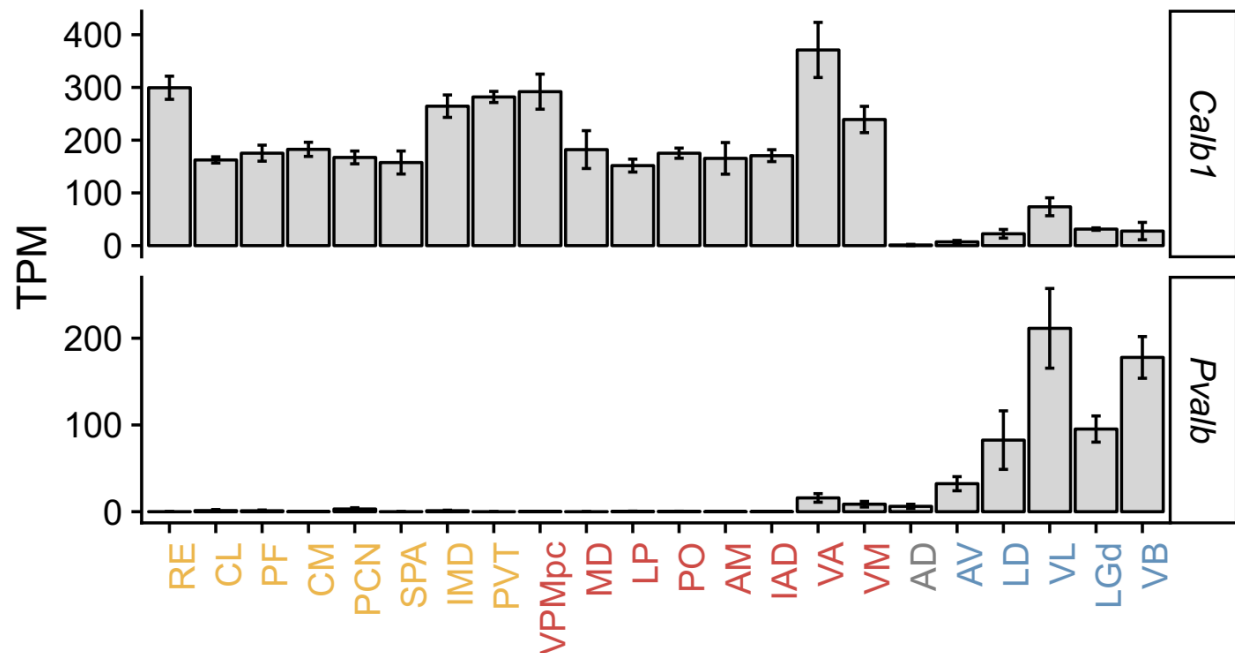


560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577

Extended Data Fig. 1 – Pooled-cell RNaseq quality control and additional analyses

- A) Markers of non-neuronal sample contamination are low across our dataset. Expression (TPM) in pooled-cell samples shown for 8 genes marking astrocytes, microglia, oligodendrocytes and erythrocytes. Only a small number of samples showed expression of contamination markers.
- B) Heatmap of the top 500 differentially expressed genes. Rows and columns are ordered by hierarchical clustering with Euclidean distance metric. Colors represent gene-wise Z-scores.
- C) Samples of the same nucleus obtained via different labelling methods cluster similarly. Principal components analysis of those samples, for which multiple collection methods were used (i.e. GENSAT lines in addition to retrograde labeling) using the top 500 genes with highest variance. Samples are colored by collection approach or transgenic line used.

578



579

580

Extended Data Fig. 2 – Existing thalamic marker genes do not cleanly mark nuclear profiles

581

Expression of *Calb1* and *Pvalb* (mean ± SEM) for each nucleus, with nuclei on the x-axis

582

colored by their profile from Fig. 1. The core/matrix organizational theory proposes that

583

thalamus is divided into two discrete groups, expressing calbindin or parvalbumin. However, the

584

first major branch splits the secondary and tertiary groups, both of which are marked by *Calb1*

585

and would thus both be considered ‘matrix’ nuclei in this theory. Thus existing markers for

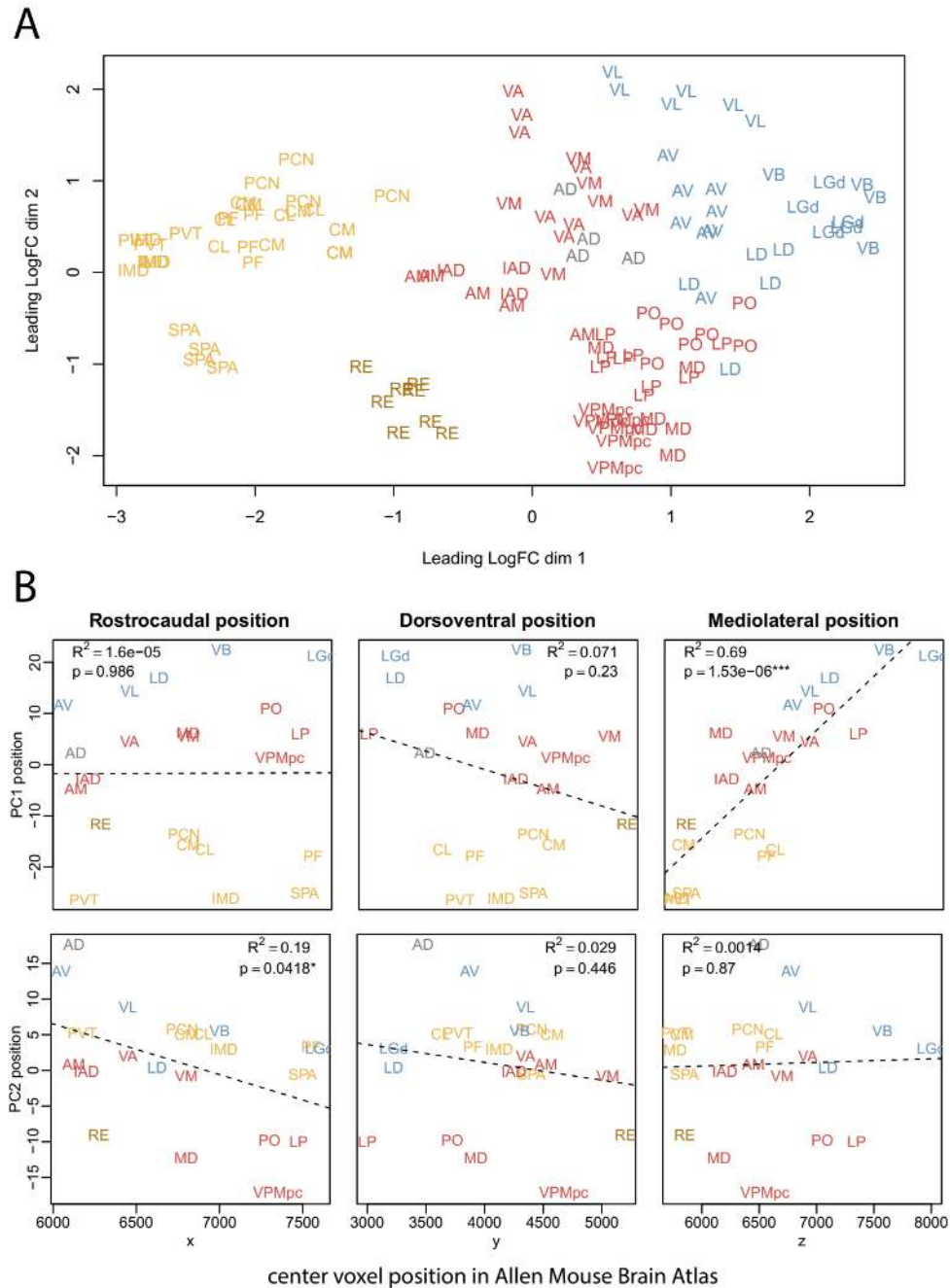
586

thalamic nuclei subgroups do not adequately reflect thalamic organizational structure.

587

588

589

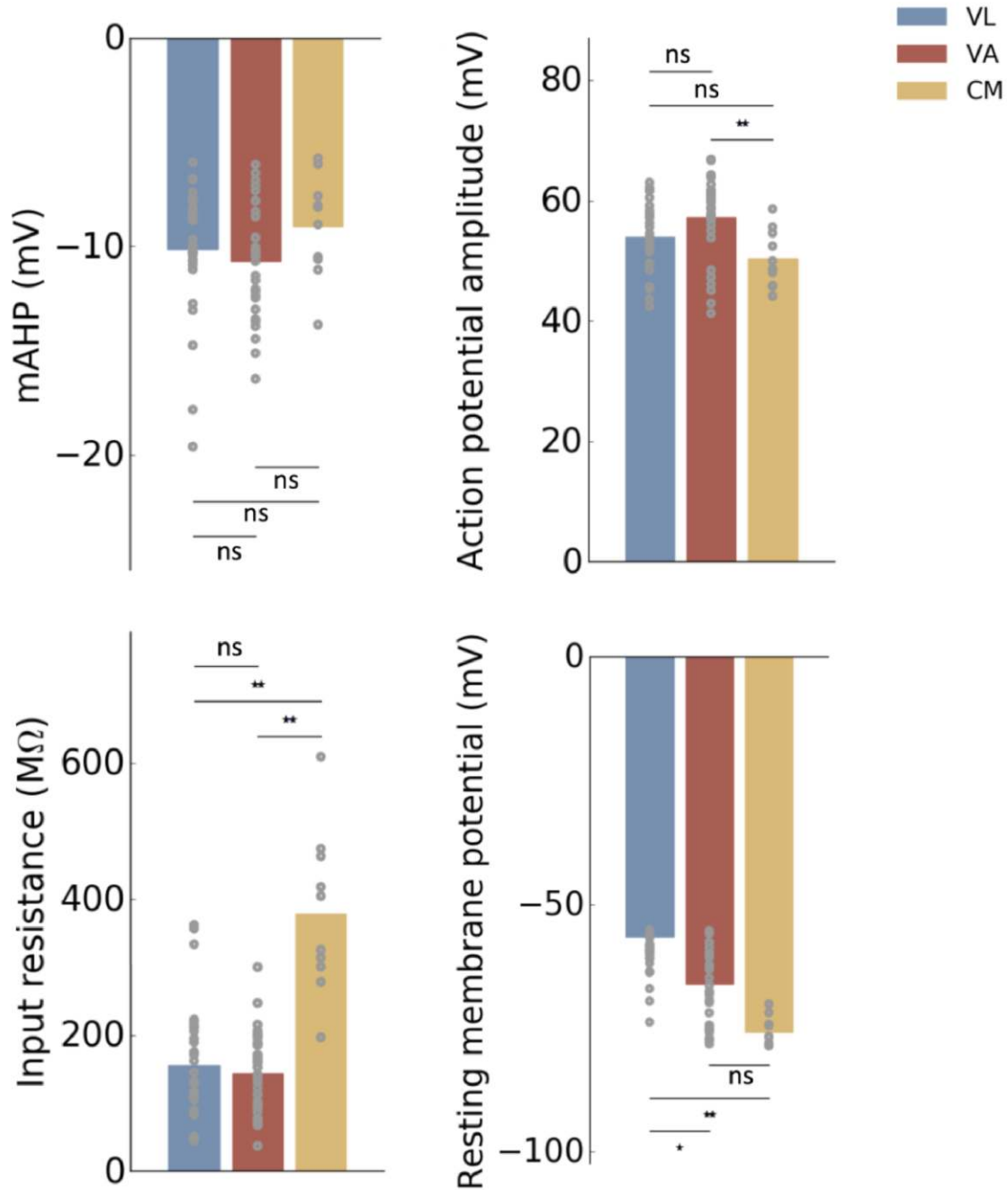


590
591
592
593
594
595
596
597
598
599
600

Extended Data Fig. 3 – Additional analysis of topogenetic axis

- A) Multidimensional scaling using an alternative distance metric also identifies the same leading axis of variance. Distance was defined as the quadratic mean of the log2 fold changes of the top 500 differentially expressed genes between any two samples (meaning that the gene set used for the distance comparison varies between each sample comparison made).
- B) Relationship of PC1 and 2 with topographical position of nuclei. Rostrocaudal, dorsoventral, and mediolateral positions are the x, y, and z voxel coordinates, respectively, in the Allen Mouse Brain Atlas. 1 voxel = 10 μ m.

601
602
603

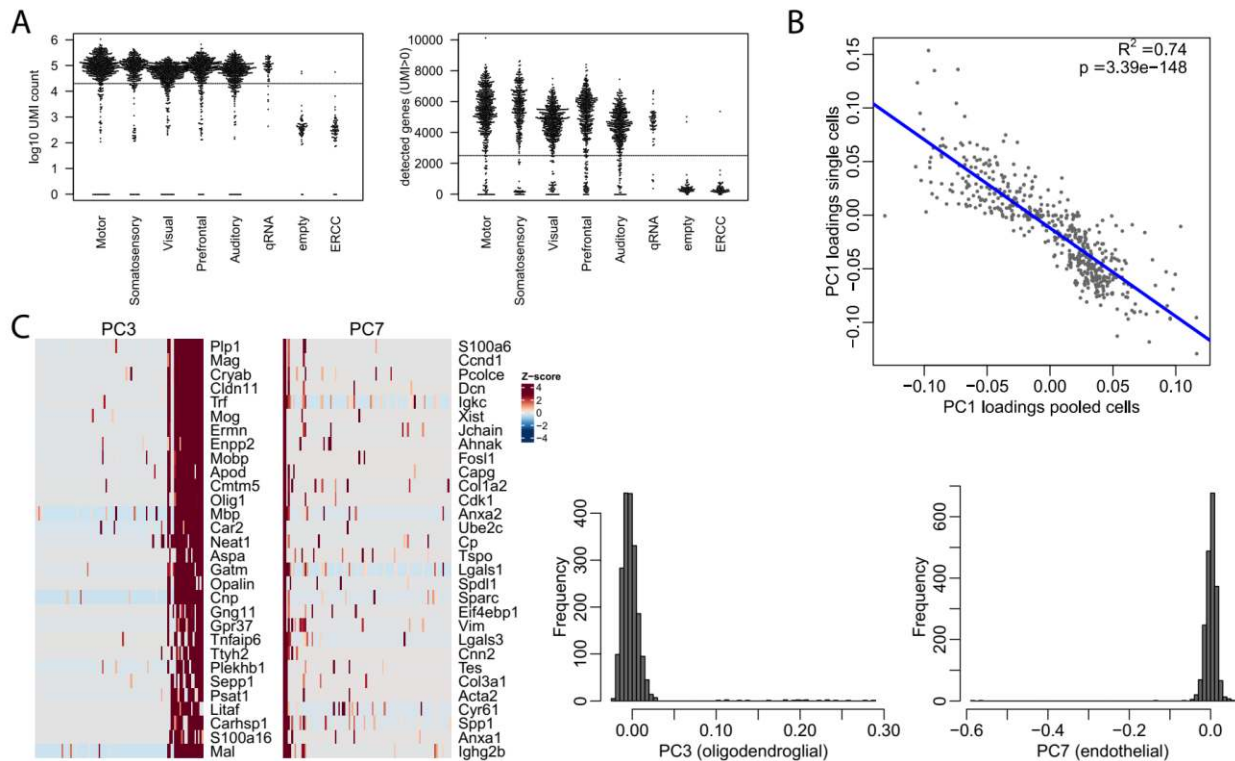


604
605
606
607
608

Extended Data Fig. 4 - Additional electrophysiological properties between thalamic nuclear profiles.

All statistical tests and experimental details are the same as in figure 3C.

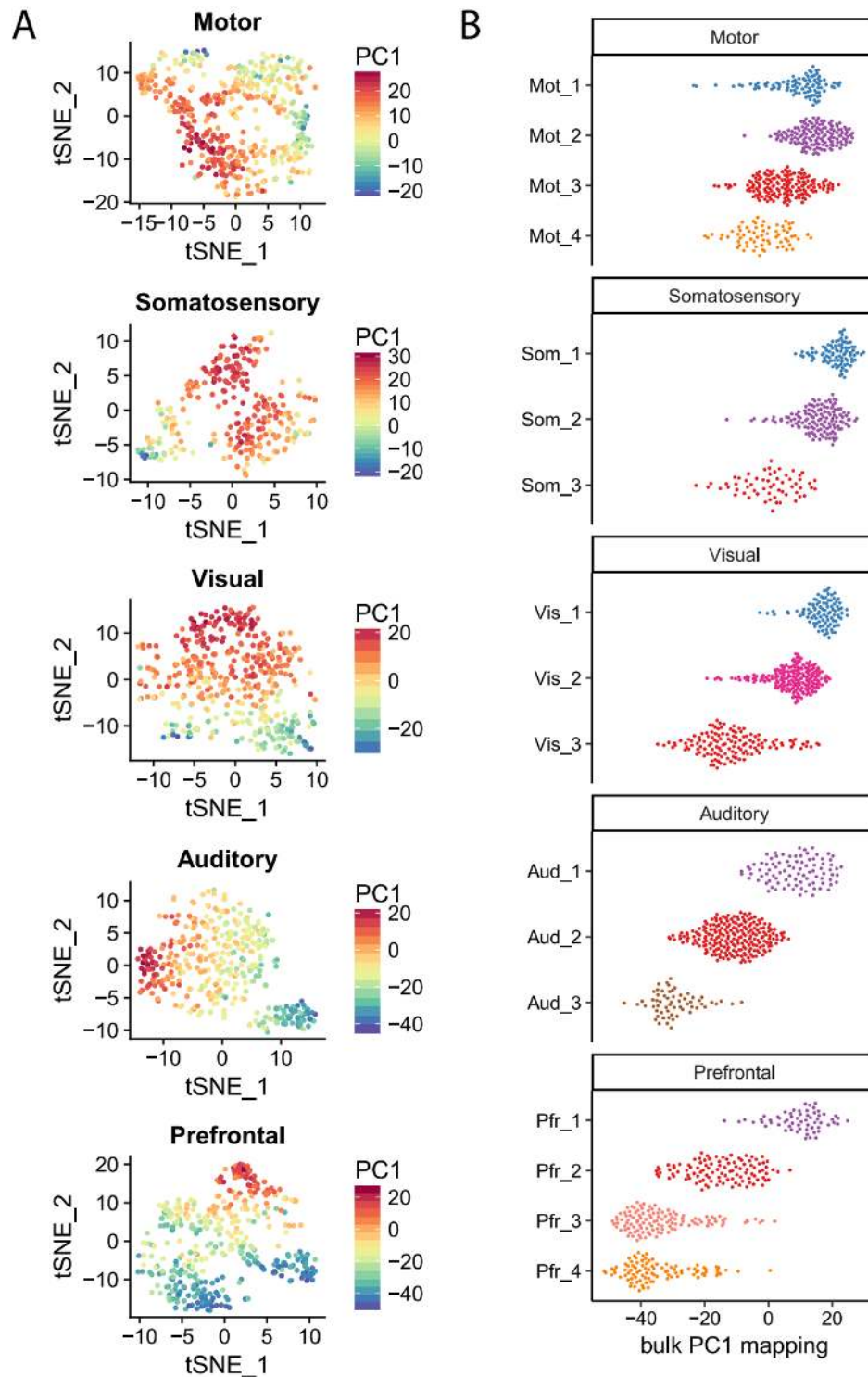
609



610
611
612
613
614
615
616
617
618
619
620
621
622
623

Extended Data Fig. 5 – Quality control for single-cell RNAseq data

- A) UMI count (left) and gene detection rate (right) for all collected single cells. Cutoffs for downstream use are indicated by dashed lines.
- B) PC1 loadings for the most differentially expressed genes between nuclei (gene set as in Fig.1) are highly similar in pooled-cell and single-cell RNAseq data.
- C) PCA on the single-cell RNAseq data revealed that principal components 3 and 7 represented non-neuronal contamination from oligodendrocytes and endothelial cells. Cells with PC3 position >0.05 and PC7 position <0.1 were removed.

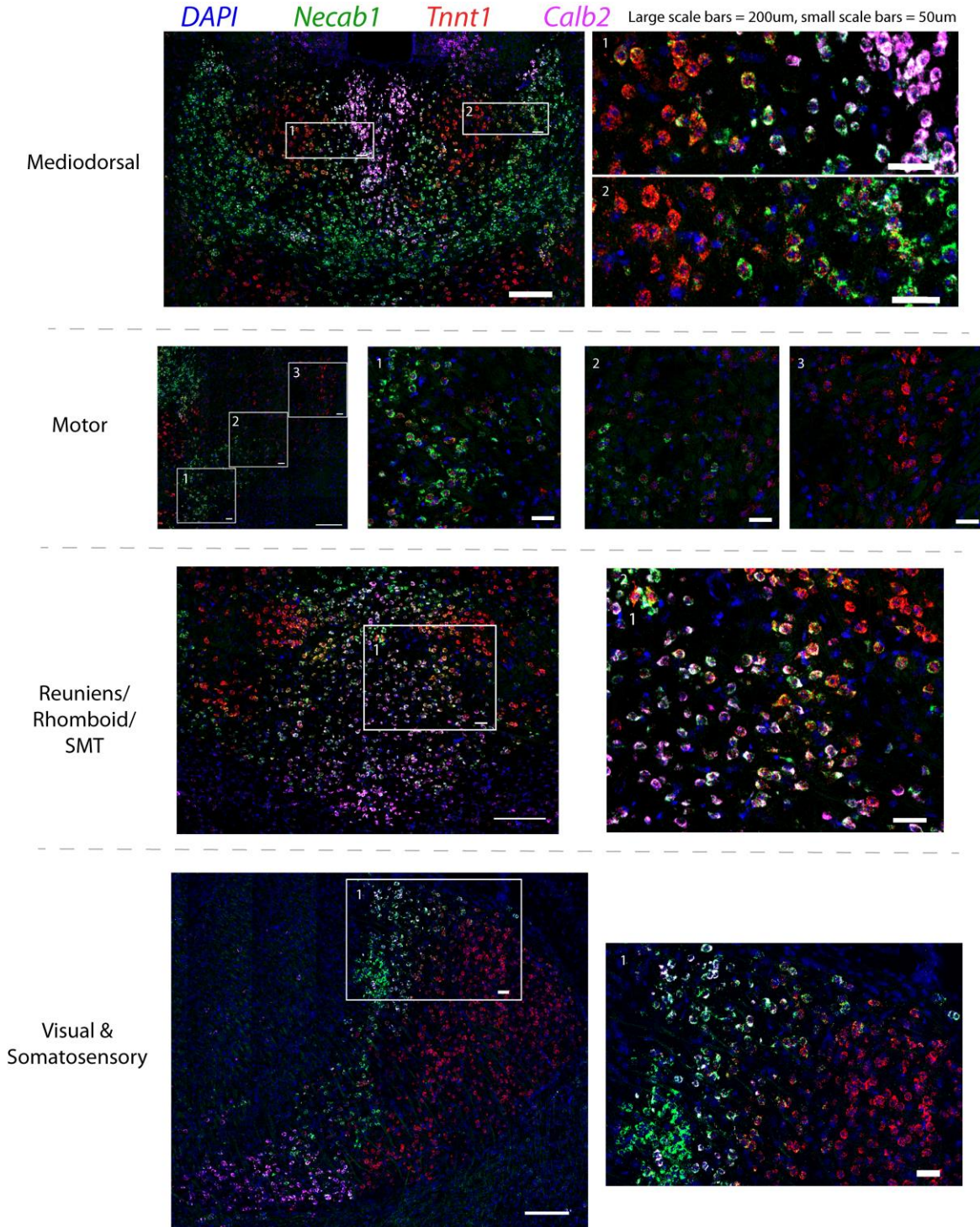


624
625
626
627
628
629
630

Extended Data Fig. 6 – Projection of single cells onto PC1 of pooled-cell RNAseq data

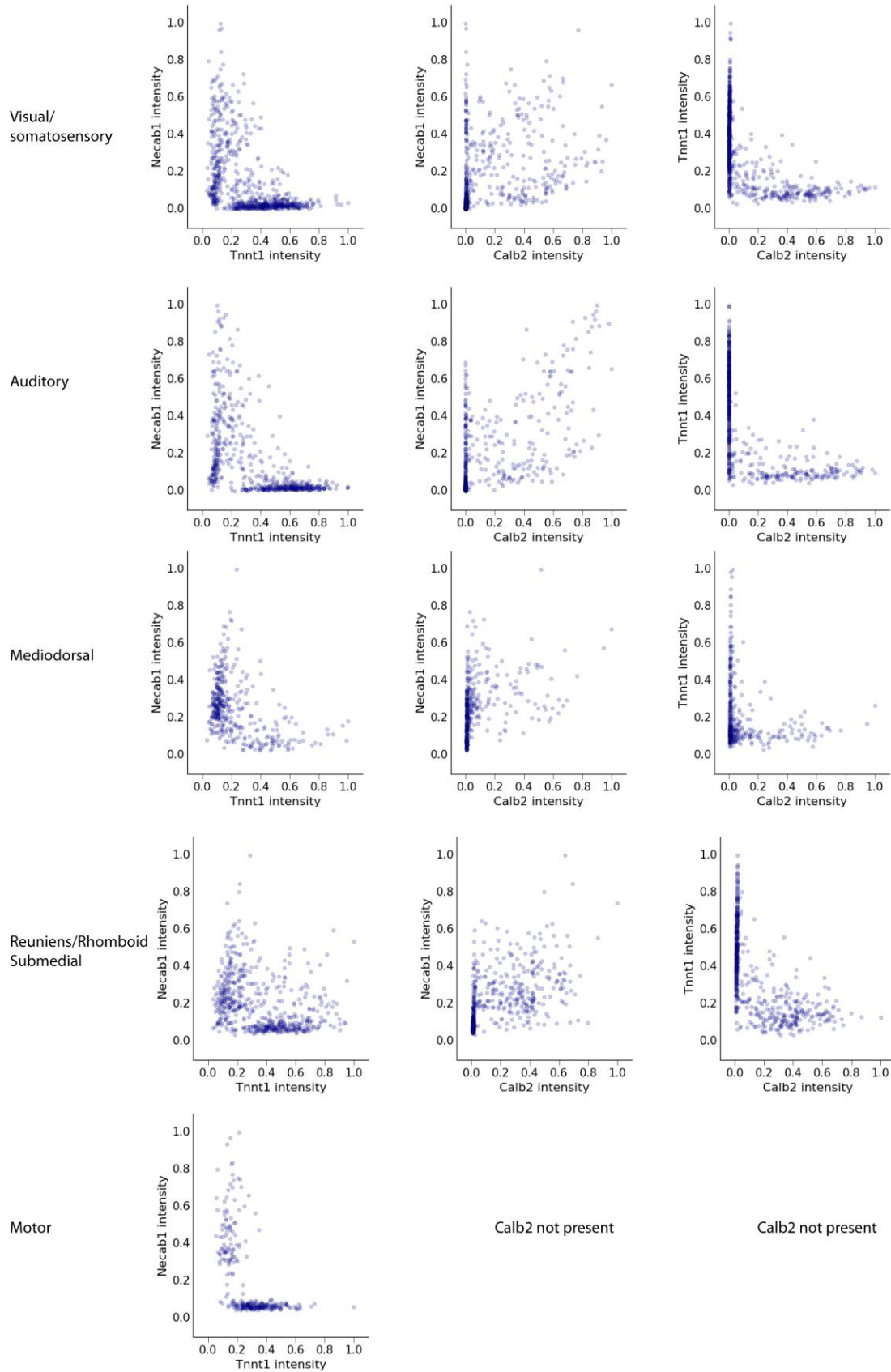
- A) tSNE plots for each projection type with cells from single-cell RNAseq colored by their projection onto pooled-cell PC1.
- B) Positions of single cells projected onto pooled-cell PC1 from Fig. 2, plotted separately for each single-cell cluster.

631
632



633
634
635
636
637
638

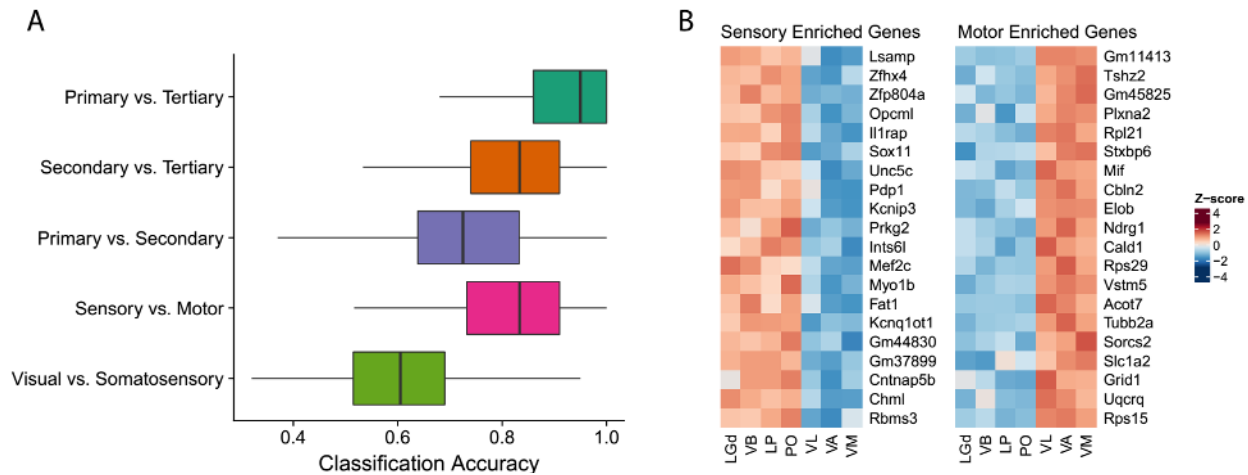
Extended Data Fig. 7 – Multi-FISH show cells with mixed expression of profile marker genes
Expanded views of example regions showing intermediate cells expressing combinations of *Tnnt1*, *Necab1* and *Calb2*, which are preferentially expressed in primary, secondary and tertiary nuclear profiles respectively.



640
641
642
643
644
645
646
647

Extended Data Fig. 8 - Quantification of multi-FISH images shows intermediate cells

Quantification of multi-FISH gene expression images. Regions of interest (ROIs) were drawn in ImageJ. Intensity was normalized first to the ROI size, then divided by the maximum for that channel. Only cells that express at least one of the marker genes were included.

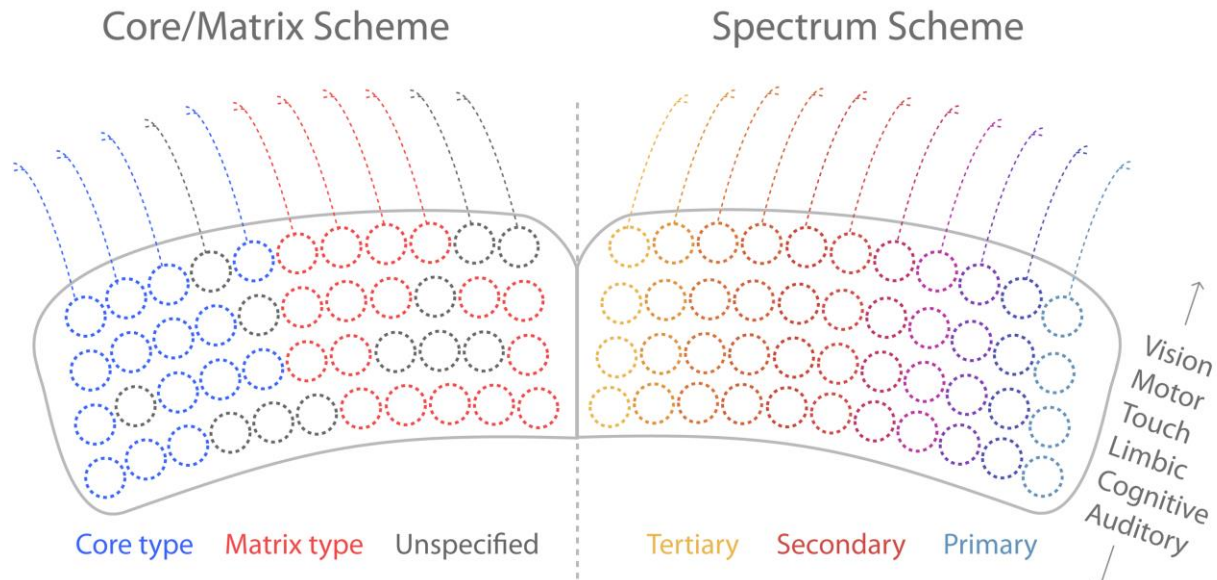


648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664

Extended Data Fig. 9 – Differential gene expression between sensory and motor thalamic nuclei

- A) Classification accuracy for distinguishing Primary, Secondary, and Tertiary type nuclei, as well as for distinguishing motor (VL,VA,VM) vs. sensory (LGd,LP,VB,PO) and visual (LGd,LP) vs. somatosensory (VB,PO) nuclei. Classifiers were obtained using elastic-net logistic regression on 20 random genes over 100 iterations. To prevent bias due to sample size difference, groups were subsampled to the size of the smallest group (n=9) at each iteration. Accuracy was assessed using 5-fold cross-validation. A value of 1 corresponds to perfect classification, while 0.5 corresponds to chance level performance.
- B) Genes that best distinguish motor from sensory nuclei (LGd,VB,LP,PO vs. VL,VA,VM). Plotted are the top 20 genes with false discovery rate <math> < 10^{-3}</math> (Likelihood Ratio Test), fold change > 2, and ordered by highest signal-to-noise ratio (mean log fold change between vs. within group).

Organization of thalamic pathways



665
666
667
668
669
670
671
672
673
674
675
676
677
678
679

Extended Data Fig. 10 – A spectrum of thalamic pathways conserved across modalities

Left: The core/matrix scheme divides thalamus into two discrete cell types based upon their expression of parvalbumin (Core nuclei) and calbindin (Matrix nuclei). However, many nuclei do not express either marker (for example, the anterior nuclei), and calbindin expressing nuclei can have markedly different anatomical and functional characteristics (for example, the rostral intralaminar nuclei and PO both are calbindin⁺).

Right: The spectrum scheme identifies three major profiles of nuclei, but places them on a single spectrum. This spectrum is aligned with the mediolateral axis of thalamus, is based upon the entire transcriptome, and links projection type directly to functional properties. Thalamus therefore consists of a set of filter banks for each modality.

680
681
682
683
684
685
686

Extended Data Tables

Extended Data Table 1 – Sample collection approaches (pooled-cell RNAseq)

[Validation] refers to samples in Extended Data Fig. 1C. For transgenic lines used, see supplementary Extended Data Table 4.

Region name:	Labelling methods:	Replicates:
Anterior dorsal (AD)	Retrograde DiI from retrosplenial cortex. [Validation: Slc6a5-Cre.KF109 transgenic line]	4,[3]
Anterior medial (AM)	Retrograde DiI from cingulate cortex. Retrograde AAV from cingulate cortex.	4,1
Anterior ventral (AV)	Retrograde DiI from retrosplenial cortex. [Validation: Slc6a5-Cre.KF109 transgenic line]	8,[4]
Central lateral (CL)	rAAV2-retro from striatum. [Validation: Grp-Cre.KH288 transgenic line]	4
Central medial (CM)	rAAV2-retro from striatum. Retrograde DiI to striatum. [Validation: Grp-Cre.KH288 transgenic line]	4,3,[1]
Interanterodorsal nucleus (IAD)	Retrograde DiI from cingulate cortex	3
Intermediodorsal nucleus (IMD)	rAAV2-retro from striatum.	4
Laterodorsal nucleus (LD)	Retrograde DiI from retrosplenial cortex.	5
Lateral Geniculate (dorsal) nucleus (LGd)	Retrograde DiI from visual cortex with anterograde CTB-488 tracer injected to retina (Cholera Toxin Subunit B, Alexa Fluor 488™, Molecular Probes). [Validation: Slc6a5-Cre.KF109 and Gpr26-Cre.KO250 transgenic lines]	5,[3],[4]

Lateral posterior nucleus (LP)	Retrograde DiI from visual cortex with anterograde CTB-488 tracer injected to retina (Cholera Toxin Subunit B, Alexa Fluor 488™, Molecular Probes).	9
Medial geniculate body (MGB)	Not collected for pooled-cell RNAseq in this study. Consists of ventral (MGBv), dorsal (MGBd) and medial (MGBm) subdivisions.	na
Mediodorsal nucleus (MD)	Retrograde DiI from frontal cortex. Collected without subdividing the three subdivisions (central: MDc, medial: MDm, lateral: MDl).	6
Paracentral nucleus (PCN)	rAAV2-retro from striatum. [Validation: Grp-Cre.KH288 transgenic line]	4,[3]
Parafascicular nucleus (PF)	rAAV2-retro from striatum. This did not label the entire PF complex, and likely reflects a genetic subset.	4
Posterior medial nucleus (PO)	Retrograde DiI from somatosensory cortex. [Validation: Gpr26-Cre.KO250 transgenic line]	7,[4]
Paraventricular thalamus (PVT)	rAAV2-retro from striatum [Validation: Gpr26-Cre.KO250 transgenic line]	3,[3]
Reuniens nucleus (RE)	Retrograde DiI from cingulate cortex	8
Rhomboid nucleus (RH)	Not collected for pooled-cell sequencing in this study.	na
Subparafascicular nucleus (SPA)	rAAV2-retro from striatum. This did not label the entire parafascicular complex, and likely reflects a genetic subset (medially biased).	4
Submedial thalamus (SMT)	Not collected for pooled-cell RNAseq in this study.	na
Ventral anterior nucleus (VA)	Retrograde DiI from motor cortex. Anterograde labelling of inputs from SNr and DCN with viral tracers (AAV2/1-CAG-GFP and/or AAV2/1-CAG-BFP).	8
Ventral medial nucleus (VM)	Retrograde DiI from motor cortex. Anterograde labelling of inputs from SNr and DCN with viral tracers (AAV2/1-CAG-GFP and/or AAV2/1-	6

	CAG-BFP).	
Ventral lateral nucleus (VL)	Retrograde DiI from motor cortex. Anterograde labelling of inputs from from SNr and DCN with viral tracers (AAV2/1-CAG-GFP and/or AAV2/1-CAG-BFP).	6
Ventrobasal nucleus (VB)	Retrograde DiI from somatosensory cortex [Validation: Grp-Cre.KH288 transgenic line for VPM]	4,[3]
Ventroposteromedial parvocellular (VPMpc)	Retrograde DiI from insular cortex.	6

687
688
689
690
691
692
693
694
695

Extended Data Table 2 – Coordinates for retrograde labelling/trace injections (pooled-cell RNAseq)

All depths relative to brain surface. If depth not stated, injections were made at 300µm and 600µm deep.

Sample area:	Injection coordinates (in millimeters, from bregma, depth from brain surface)	Additional comments
AD	1.7 caudal, 0.25 lateral, 0.4, 1 deep.	
AM	1.35 rostral, 0.2 lateral, 1.25, 1.8 deep.	
AV	1.7 caudal, 0.25 lateral, 0.4, 1 deep.	
CL	0.7 rostral, 1.9 lateral, 2.75, 3.00, 3.25 deep.	
CM	0.7 rostral, 1.9 lateral, 2.75, 3.00, 3.25 deep.	
IAD	1.35 rostral, 0.2 lateral, 1.25, 1.8 deep.	

LD	1.7 caudal, 0.25 lateral, 0.4, 1 deep.	
LGd	2.9 caudal, 2.4 lateral. 2.9 caudal, 1.7 lateral. 3.8 caudal, 3.0 lateral 3.8 caudal, 2.0 lateral	Retina was injected with anterograde tracer (CTB) to distinguish LGd from LP
LP	2.9 caudal, 2.4 lateral. 2.9 caudal, 1.7 lateral. 3.8 caudal, 3.0 lateral. 3.8 caudal, 2.0 lateral.	Retina was injected with anterograde tracer (CTB) to distinguish LGd from LP
MD	2 rostral, 1 lateral, 3.0, 2.5, 2.0 deep. 2 rostral, 1.6 lateral, 3.0, 2.4,1.5 deep. 1.7 rostral, 0.4 lateral, 3.0, 2.4, 1.5 deep.	
PCN	0.7 rostral, 1.9 lateral, 2.75, 3.00, 3.25 deep.	
PF	0.7 rostral, 1.9 lateral, 2.75, 3.00, 3.25 deep.	
PO	0.85, 1.5 caudal,2.75 lateral, 0.4 deep. 1.2 caudal, 2.25 lateral, 0.4 deep.	
RE	1.35 rostral, 0.2 lateral, 1.25+1.8mm deep.	
SPA	0.7 rostral, 1.9 lateral, 2.75, 3.00, 3.25 deep.	
VA	0.3 rostral, 1.6 lateral. 0.8 rostral, 1.6 lateral. 0.5 rostral, 1.4 lateral.	
VM	0.3 rostral, 1.6 lateral	

	0.8 rostral, 1.6 lateral. 0.5 rostral, 1.4 lateral.	
VL	0.3 rostral, 1.6 lateral. 0.8 rostral, 1.6 lateral. 0.5 rostral, 1.4 lateral.	
VB	0.85, 1.5 caudal, 2.75 lateral, 0.4 deep. 1.2 caudal, 2.25 lateral, 0.4 deep.	
VPMpc	0.25 caudal, 4 lateral, 2.4,2.7,2.9 deep.	

696
697
698
699
700
701
702
703
704
705

Extended Data Table 3 – Injection coordinates for retrograde labelling (single-cell RNAseq)

Sample area:	Injection coordinates (in millimeters, from bregma, depth from brain surface)
Auditory Thalamus	3.5 caudal, 4.5 lateral, 2.25 deep. 3.16 caudal, 4.5 lateral, 2.0 deep.
Mediodorsal (Prefrontal) Thalamus	2 rostral, 1 lateral, 3.0, 2.5,2.0 deep. 2 rostral, 1.6 lateral, 3.0,2.4,1.5 deep. 1.7 rostral, 0.4 lateral, 3.0, 2.4, 1.5 deep.
Motor Thalamus	0.3 rostral, 1.6 lateral. 0.8 rostral, 1.6 lateral. 0.5 rostral, 1.4 lateral.
Somatosensory Thalamus	0.85 and 1.5 caudal, 2.75 lateral, 0.4 deep. 1.2 caudal, 2.25 lateral, 0.4 deep.
Visual thalamus	2.9 caudal, 2.4 lateral. 2.9 caudal, 1.7 lateral. 3.8 caudal, 3.0 lateral. 3.8 caudal, 2.0 lateral.

706

707
708
709
710

Extended Data Table 4 – Transgenic mice used in this study

Transgenic line:	Areas collected
Slc6a5-Cre.KF109 (GENSAT)	Anterior dorsal, Anterior ventral, Dorsal lateral geniculate (LGd).
Gpr26-Cre.KO250. (GENSAT)	Paraventricular nucleus (PVT), Posterior medial (PO)
Grp-Cre.KH288 (GENSAT)	Rostral intralaminar nuclei (Central lateral, paracentral), ventral posteromedial nucleus (VPM).

711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726

Captions for Supplementary Tables

Supplementary Table 1 (separate file) – Pooled-cell RNAseq metadata and differential gene expression.

Supplementary Table 2 (separate file) – Pooled-cell RNAseq principal component analysis and PANTHER protein class enrichment.

Supplementary Table 3 (separate file) – Single-cell RNAseq metadata and cluster marker genes.