Variation in neuronal activity state, axonal projection target, and position principally define the transcriptional identity of individual neocortical projection neurons

Maxime Chevée[1,2], Johanna D. Robertson[3], Gabrielle H. Cannon[4]
Solange P. Brown[2,] and Loyal A. Goff[2,4,*]*

Affiliations:

[1]Biochemistry, Cellular and Molecular Biology Graduate Program, Johns Hopkins University School of Medicine, Baltimore, Maryland, 21205, USA.

[2]Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland, 21205, USA.

[3]Human Genetics Training Program, McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, 21205, USA.

[4]McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, 21205, USA.

[*]Corresponding Authors: spbrown@jhmi.edu (SPB) and loyalgoff@jhmi.edu (LAG)

ABSTRACT

Single-cell RNA sequencing technologies have generated the first catalogs of transcriptionally defined neuronal subtypes of the brain. However, the biologically informative cellular processes that contribute to neuronal subtype specification and transcriptional heterogeneity remain unclear. By comparing the gene expression profiles of single layer 6 corticothalamic neurons in somatosensory cortex, we show that transcriptional subtypes primarily reflect axonal projection pattern, laminar position within the cortex, and neuronal activity state. Pseudotemporal ordering of 1023 cellular responses to manipulation of sensory input demonstrates that changes in expression of activity-induced genes both reinforced cell-type identity and contributed to increased transcriptional heterogeneity within each cell type. This is due to cell-type specific biases in the choice of transcriptional states following manipulation of neuronal activity. These results reveal that axonal projection pattern, laminar position, and activity state define significant axes of variation that contribute both to the transcriptional identity of individual neurons and to the transcriptional heterogeneity within each neuronal subtype.

INTRODUCTION

Single-cell RNA sequencing (scRNA-seq) approaches have been harnessed to reveal previously hidden levels of complexity in cell types and states within a given tissue[1, 2]. Nowhere is this more relevant than in the mammalian central nervous system, an organ system dependent on a remarkable diversity of cell types and cell states for its function. The neocortex contains a wide variety of neuronal cell types organized into the circuits that direct higher brain functions such as perception, memory and cognition. Furthermore, cortical neurons are highly dynamic, undergoing significant changes in gene expression during development and throughout adulthood in response to activity and experience[3-8]. Recent studies have generated a more comprehensive 'parts list' of discrete cell types within the neocortex than previously available[9-16]. While such surveys yield insights into the diversity of cortical cell types, the sources of transcriptional variation both within and across cell types remain poorly understood[17-19].

We compared expression profiles of layer 6 corticothalamic neurons (L6CThNs), a heterogeneous population of related cortical projection neurons defined in anatomical, functional, and gene expression studies, making them ideally suited for investigating the relationships between transcriptional subtypes and other cellular properties[12, 16, 20-29]. By combining scRNA-seq with an enrichment strategy that preserved axonal target information, we identified two transcriptionally distinct L6CThN subtypes that each demonstrates a reproducible bias for long-range projection targets. The transcriptional profiles of L6CThNs also reflect their laminar position within L6. These two L6CThN subtypes exhibit divergent signatures of neuronal activity both in their baseline expression patterns as well as in their responses following manipulation of sensory input. The subtype biases in the choice of response following sensory manipulation increased transcriptional heterogeneity both within and between cell types and

70 reinforced the transcriptional identities of the two L6CThN subtypes. These results demonstrate
71 that scRNA-seq resolves relationships between gene expression and features such as axonal
72 projection pattern, spatial organization, and cell state, and identifies the independent
73 contributions of multiple biological signals that together determine transcriptional
74 heterogeneity within and across neuronal populations.
75
76 **RESULTS**
77 ***Transcriptional profiling of layer 6 corticothalamic neurons reveals two subtypes which reflect***
78 ***axonal projection bias***
79 Studies of primary sensory cortex demonstrate that L6CThNs are heterogeneous[20-29]. In rat
80 barrel cortex, L6CThNs in upper layer 6 project to the ventral posterior medial nucleus (VPM) of
81 the thalamus, while L6CThNs in lower layer 6 project to both VPM and the posterior medial
82 nucleus (POm)[21, 25]. To distinguish between these two projection classes in mouse, we first
83 validated Cre recombinase expression as a reliable marker for L6CThNs in the barrel cortex of
84 *Neurotensin receptor 1-Cre* mice[30-32] (Fig. 1a-e; *Ntsr1-Cre*, Gensat 220). Next, we showed that a
85 subset of Cre-expressing, VPM-projecting L6CThNs in lower layer 6 also projects to POm (Fig. 1f-
86 j).
87 To determine whether L6CThNs subclasses projecting to VPM only or to both VPM and
88 POm are distinguished by their gene expression profiles, we labeled the two subclasses in
89 adolescent mice, microdissected barrel cortex containing retrogradely labeled cells, dissociated
90 the tissue into a single-cell suspension, sorted the differentially labeled L6CThNs using
91 Fluorescence Assisted Cell Sorting (FACS; Supplementary Fig. 1a-b), and collected enriched
92 populations of each subclass for bulk RNA sequencing. We identified 69 genes with significant

93 differential expression (DE) between the two sorted populations (Fig. 1k, Supplementary File 1,
94 Supplementary Table 1; Cuffdiff2; 10% false discovery rate (FDR)). These 69 DE genes
95 demonstrate that subclasses of L6CThNs distinguished by their long-range axonal projection
96 patterns are differentiated by their gene expression profiles.

97 This bulk analysis was predicated on prior knowledge of existing morphological
98 subclasses and may have obscured underlying transcriptional subtypes that comprise each
99 projection class of L6CThNs. To resolve transcriptionally defined subtypes of L6CThNs, we next
100 evaluated the gene expression landscape of single L6CThNs using an unbiased classification
101 approach. We first sorted individual, labeled L6CThNs (Supplementary Fig. 1c) and collected 96
102 VPM-only and 96 VPM/POm L6CThNs from each of two replicate mice, totaling 384 single
103 L6CThNs. Individual cell lysates were subjected to a modified Smart-Seq2 library preparation
104 and scRNA-seq analysis. In total, 346 single L6CThNs passed quality control filters and were used
105 for further analysis (Supplementary Fig. 1d-i; Supplementary File 2). We confirmed the fidelity of
106 our enrichment by assessing each cell for neuronal and non-neuronal markers (Supplementary
107 Fig. 1j).

108 To identify transcriptional subtypes of L6CThNs, we selected genes with the greatest
109 likelihood of contributing to differences across the ensemble of single L6CThNs by identifying
110 genes with high residuals to a mean-variance model fit independently in each of the two
111 replicates and selecting the intersection of these two sets (Supplementary Fig. 2a). These 261
112 genes with high dispersion were therefore depleted for genes associated with technical
113 variation between replicates[33, 34] (Supplementary Table 2). Weights on the first three significant
114 principal components across all cells using this gene set[35] (Supplementary Fig. 2b; permutation
115 parallel analysis; $p < 0.001$) were used for a tSNE dimensionality reduction followed by k-means

116 clustering (Fig. 2a). Single-cell transcriptional profiles of the 346 L6CThNs clustered into at least
117 two distinct subtypes. Fitting the data to three subtypes (k=3) or more did not significantly
118 improve the performance of the clustering (Supplementary Fig. 2c). We compared our
119 classification approach to several recently described single-cell clustering utilities[36, 37] and found
120 a high-degree of agreement (Supplementary Fig. 2d; SC3: 93.77% agreement; CIDR: 90%
121 agreement). These results indicate that independent, unbiased clustering approaches based on
122 genes with higher than expected variance across the L6CThN population identify two major
123 subtypes of L6CThN.

124 To determine the relationship between transcriptional identity and morphological
125 subtypes, we next compared the distribution of VPM-only and VPM/POm projection labels
126 across the two transcriptionally defined subtypes (Fig. 2b). The majority of neurons in subtype 2
127 were labeled VPM-only (79%, 103 of 130 cells, $p < 3.768 \times 10^{-17}$, hypergeometric test) while most
128 neurons in subtype 1 were VPM/POm (65%, 141 of 216 cells, $p < 3.768 \times 10^{-17}$, hypergeometric
129 test), a distribution significantly different from that expected by chance. Furthermore, the
130 axonal projection bias was reproducible across replicates (Fig. 2c) even though the segregation
131 of projection-defined subclasses was incomplete for both transcriptionally defined subtypes.
132 Together our results indicate that each transcriptionally defined subtype of L6CThN is enriched
133 for neurons targeting specific sets of thalamic nuclei.

134

135 ***Transcriptional differences between subtypes of layer 6 corticothalamic neurons***

136 To assess the transcriptional differences between the two identified L6CThN subtypes, all
137 expressed genes were subjected to the Monocle2 differential test, using model formulas that
138 accounted for both batch effects and the number of genes detected in each cell, a proxy

139 measure for efficiency of RNA capture and library synthesis. We identified 286 genes that were
140 significantly differentially expressed between the two subtypes (Fig. 2d; Monocle likelihood
141 ratio test, 0.1% FDR; Supplementary File 2, Supplementary Table 3), only six of which overlapped
142 with the 69 DE genes observed in our bulk RNA-seq analysis of sorted L6CThNs despite the high
143 correlation between our bulk RNA-seq and scRNA-seq data (Supplementary Fig. 2e-g). This
144 result, in conjunction with our previous observation of incomplete label segregation across
145 L6CThN cell types, suggests that this discrepancy is primarily due to significant sample
146 heterogeneity arising from retrograde label inefficiencies. Importantly, parameters such as total
147 mapped fragments (mass), total estimated mRNAs per cell, number of genes detectably
148 expressed per cell, and replicate did not result in biased clustering across the DE gene list (Fig.
149 2d), suggesting a minimal influence of technical variation on the list of DE genes. The significant
150 differential expression of genes, including *Fxyd6* and *Lamp*5, between the two subtypes was
151 consistent with expression patterns seen in the Allen Mouse Brain Atlas[38] (Fig. 2e;
152 http://mouse.brain-map.org, Fxyd6-RP_051017_01_E10-coronal, Lamp5-RP_050725_01_B03-
153 coronal). The two subtypes also shared some transcriptional similarities with two recently
154 defined subtypes in primary visual cortex[12] (Supplementary Fig. 3). Several long non-coding
155 RNAs (lncRNAs) were also specifically enriched in each subtype. For example, linc-Tmem20[8] was
156 significantly enriched in subtype 1 L6CThNs (Fig. 2f) and was preferentially expressed in lower
157 layer 6 (Fig. 2g). Conversely, the lncRNA *Pantr1* was the gene with the greatest predictive power
158 for neurons in subtype 2 (Fig. 2f, AUC = 0.876, power = 0.752, ROC analysis). Using a mouse line
159 in which *LacZ* was knocked into the *Pantr1* locus[39], we confirmed greater *LacZ* expression in
160 VPM-only L6CThNs relative to VPM/POm L6CThNs (Fig. 2h,i). Furthermore, in contrast to our
161 analysis of L6CThNs by anatomical labeling, we found that the total number of genes with

162 detectable expression in each transcriptional subtype was significantly greater for subtype 1
163 relative to subtype 2, a finding consistent across replicates (Fig. 2j), and that the mean pairwise
164 Euclidean distance between cells within each subtype was also greater in subtype 2 than in
165 subtype 1, indicating greater cell-cell variation across subtype 2 (Fig. 2k; subtype 1 μ=72.16,
166 σ=5.80; subtype 2 μ=80.59, σ=7.07; $p < 2.2 \times 10^{-16}$, Welch's two-sample t-test). Importantly, we
167 found only two genes, *Lypd1* and *Calm2*, with a significant combinatorial effect of subtype and
168 label, suggesting that label does not distinguish subpopulations within each subtype. Together,
169 these data identify two L6CThN subtypes and confirm a relationship between transcriptionally
170 defined neuronal subtypes and the projection targets of these neurons.

171 To identify cellular processes that differentiate the two transcriptional subtypes of
172 L6CThNs, we queried the DE gene list for enrichment of annotated gene sets from public
173 databases (Supplementary Fig. 4a-d). Significant gene sets included Gene Ontology and
174 Reactome terms related to general features of neurons such as “Neuronal part” and “Synaptic
175 transmission” ($p < 0.01$; hypergeometric test, Benjamini-Hochberg corrected), highlighting the
176 limited resolution of gene sets in currently available public databases for generating biological
177 insights among neuronal subtypes. To identify more informative biological processes that shape
178 differences in the response properties of L6CThNs, we compared the expression of voltage-
179 gated ion channels, neurotransmitter receptors and neuropeptides, a number of which were
180 differentially expressed (Supplementary Fig. 4e-h). Several G protein-coupled receptors and
181 neuropeptides were differentially expressed. For example, *Adcyap1* (PACAP) and a gene
182 encoding a peptide for processing PACAP, *Pam*, were preferentially expressed in subtype 2
183 (Supplementary Fig. 4h-i). Interestingly, receptors for PACAP are found in primary sensory
184 thalamic nuclei[40] and modulate thalamocortical interactions[41], consistent with the projection

185 pattern bias of subtype 2 L6CThNs. Our data indicate that a focused analysis identifies gene
186 expression differences that reflect relevant functional features.
187
188 ***Distinct cellular processes are coordinately regulated within subtypes of L6CThN***
189 To identify cellular processes that contribute to the heterogeneity of gene expression across the
190 transcriptomes of all L6CThNs analyzed, we performed a weighted gene co-expression network
191 analysis (WGCNA) on all genes expressed in the 346 L6CThNs[42]. This analysis yielded 22 modules
192 of co-regulated genes, which we classified using hierarchical clustering (Fig. 3a, Supplementary
193 Fig. 5a). We tested for significant correlations between subtype identity for each cell and the
194 eigenvalues of each module. Three modules significantly correlated with subtype 1 (Black,
195 Turquoise, and Cyan; Fig. 3b, Supplementary Fig. 5b; $p < 0.01$; Pearson's product moment
196 correlation test) and four with subtype 2 (Red, Purple, Blue and Midnight Blue; Fig. 3b,
197 Supplementary Fig. 5b; $p < 0.01$; Pearson's product moment correlation test). When the same
198 analysis was performed based on cell label, VPM-only or VPM/POm, correlation coefficients and
199 confidence measures were all weakened (Fig. 3b, Supplementary Fig. 5c). These seven modules
200 also showed significant correlation with the first principal component of the PCA on the high-
201 variance gene set (Fig. 3c, Supplementary Fig. 3h), suggesting that subtype identity explains a
202 significant amount of variation in gene expression across these neurons. Five of these modules
203 were enriched for genes identified as significantly differentially expressed between subtype 1
204 and subtype 2 (Fig. 3d, Supplementary Fig. 3d). No module was correlated with replicate or
205 other potentially confounding technical parameters (Supplementary Fig. 5f,g), confirming that
206 the variations we observe are driven by biological differences among neurons rather than
207 technical variation. Taken together, these data demonstrate that the greatest source of

208 variation in the transcriptomes of L6CThNs is the difference between subtypes and reveal
209 several discrete modules of gene expression contributing to this difference.
210
211 ***Projection-dependent and position-dependent gene expression differences contribute to the***
212 ***transcriptional identity of layer 6 neurons***
213 Although the transcriptional signature of L6CThNs reflects projection target bias, axonal
214 projection pattern and sublaminar position are confounded among L6CThNs. VPM/POm
215 neurons are predominantly located in lower layer 6 while VPM-only neurons are biased towards
216 upper layer 6 (Fig. 1h). Thus the gene expression differences we identified may represent
217 differences in sublaminar position within layer 6 rather than axonal projection pattern *per se*. To
218 test this hypothesis, we took advantage of the fact that L6CThNs represent only approximately
219 half of the neurons in layer 6[29, 31]. If a gene's expression reflects axonal projection pattern, its
220 expression should be restricted to L6CThNs in upper or lower layer 6 only. On the other hand, if
221 a gene's expression reflects laminar position, we predict that its expression would be restricted
222 to neurons in either upper or lower layer 6, regardless of their projection pattern.
223 To select target genes to evaluate using single molecule fluorescence *in situ*
224 hybridization (smFISH), we performed a PCA on the mean-centered expression estimates of the
225 high-variance genes across all 346 L6CThNs and used the rotations from this analysis to project
226 all expressed genes into this PCA space to rank order candidate genes (Supplementary Fig. 6a-d).
227 We quantified the expression of target gene mRNAs in individual tdTomato-positive, NeuN-
228 positive neurons identified as L6CThNs and in tdTomato-negative, NeuN-positive neurons in
229 slices of barrel cortex from Ntsr1-Cre;tdTomato mice. These data were then fit to a generalized
230 additive model to test the independent contributions of laminar position and expression of

231 tdTomato (Fig. 4). We found that *Lamp5*, *Serpini1* and *Gabra5* were selectively expressed in a
232 subtype of L6CThNs (Fig. 4b,e-f). In contrast, *Pantr1* varied with laminar position within layer 6
233 in both L6CThNs and non-L6CThNs (Fig. 4c). Our findings reveal that information about position
234 within layer 6 and long range axonal projection pattern is contained in the gene expression
235 differences between the two transcriptional subtypes of L6CThNs.
236
237 ***Neural activity significantly contributes to transcriptionally defined subtype identity***Neuronal
238 activity is known to strongly influence gene expression[4-6]. We therefore hypothesized that
239 neuronal activity state influences the transcriptional profiles of the L6CThNs and may contribute
240 to their transcriptional identity. To test whether any modules reflect activity state, we assessed
241 genes assigned to each module for enrichment of a curated gene set representing the ensemble
242 of genes regulated after induction of neural activity compiled from recently published studies[43-
243 46] (Fig. 3e, Supplementary Fig. 5, Supplementary Table 4). Four of the seven modules correlated
244 with transcriptional subtype (Black, Green, Purple, and Midnight Blue) were significantly
245 enriched for genes induced by neuronal activity[43-46] (Fig. 3e, Supplementary Fig. 5, $p < 0.01$,
246 hypergeometric test). Each of these four modules was also significantly correlated with the first
247 and second principal components of the PCA on the high-variance gene set, suggesting that
248 differences in neuronal activity state explain a significant amount of transcriptional variation
249 across L6CThNs as well. Taken together, our results show that the long range axonal projection
250 pattern, laminar position within layer 6, and the activity state of each neuron are all reflected in
251 the transcriptional profiles of individual L6CThNs and are principal contributors to the identity of
252 the two L6CThN subtypes.
253 Among the four activity-associated modules (Black, Green, Purple, and Midnight Blue),

254 Black was specifically correlated with subtype 1, and Purple and Midnight Blue with subtype 2,
255 suggesting subtype-specific engagement of activity-induced genes in the steady state. The
256 Green module was correlated with PC1-PC3, but demonstrated no significant bias for either cell
257 type. To assess whether these signatures of neuronal activity represent a fundamental aspect of
258 subtype identity, we re-evaluated our classification workflow after regressing out the
259 eigenvalues for each activity-associated module and observed a reduction in the separation of
260 the two L6CThN subtypes in each case, suggesting that steady state differences in neuronal
261 activity genes are a defining characteristic of these two L6CThN subtypes (Supplementary Fig.
262 6e).

263 To further test the influence of neuronal activity on transcriptional identity, we
264 evaluated the enrichment of activity genes along the first two principal components of a gene-
265 centric PCA of the mean-centered expression estimates of high variance genes across all 346
266 L6CThNs. All expressed genes were projected in this PCA space and ranked using their weights
267 along PC1 and PC2 (Supplementary Fig. 6a-d). Both PC1 and PC2 were significantly enriched for
268 genes drawn from our curated list of activity-induced genes[43-46] ($p < 0.01$; Kolmogorov-Smirnov
269 test, Pre-ranked GSEA). The expression levels of the most heavily weighted PC1 genes varied
270 predominantly between subtypes (Fig. 5b-c), while the expression levels of heavily weighted PC2
271 genes additionally varied within each subtype in a pattern similar to classical activity-induced
272 genes like *Fos* and *Bdnf* (Fig. 5d-g), suggesting that activity contributes both to subtype identity
273 and to transcriptional heterogeneity within each subtype. The expression patterns of genes such
274 as *Cdh13*, *Igfbp6*, and *Sla* that strongly contributed to both PC1 and PC2 provide further
275 evidence that axonal projection pattern, laminar location, and activity state are not fully
276 orthogonal axes contributing to cellular identity (orange asterisks, Fig. 5c, Supplementary Fig.

277 6c,d). These data indicate that neuronal activity accounts for significant variation in the gene
278 expression of L6CThNs, contributing to transcriptional heterogeneity both within and between
279 L6CThN subtypes.
280
281 ***Modulation of neuronal activity influences the transcriptional state and identity of L6CThNs***
282 Our data demonstrate that the gene expression profiles of L6CThNs reflect an integration of
283 multiple basis vectors corresponding to discrete, but potentially dependent sources of variation
284 dominated by axonal projection pattern, sublaminar position within layer 6 and neuronal
285 activity. Alterations in the molecular cascades engaged by different patterns of neural activity
286 therefore have the potential to modulate the transcriptional identity of L6CThNs. To test this
287 hypothesis, we unilaterally removed whiskers in a chessboard pattern, a pattern of sensory
288 deprivation shown to engage plasticity mechanisms in the barrel cortex[47, 48], from Ntsr1-
289 Cre;tdTomato mice which were bilaterally injected with a retrograde tracer in POm to label
290 VPM-only and VPM/POm L6CThNs as in our baseline experiments (Fig. 6a). At one and seven
291 days following this sensory manipulation, we collected and sequenced single L6CThNs from the
292 barrel cortex both contralateral and ipsilateral to the manipulation. After preprocessing and
293 quality control, we obtained high-quality transcriptional profiles for 133 of 151 sequenced
294 L6CThNs from 2 replicates one day following whisker removal (Day 1) and for 550 of 563
295 sequenced L6CThNs from 2 replicates seven days following whisker removal (Day 7). When
296 combined with L6CThNs collected under baseline conditions (Day 0), we generated a data set
297 comprised of 1023 sequenced L6CThNs collected from two replicates at each of three time
298 points (Supplementary Fig. 7a-d, Supplementary File 3).
299 To assign each newly acquired L6CThN to its transcriptional subtype, gene expression

300 profiles were first pre-processed and normalized as described for the baseline data set (Day 0).
301 We then performed a PCA using the 286 genes we identified as differentially expressed between
302 transcriptional L6CThN subtypes under baseline conditions and a tSNE analysis on the 1023
303 L6CThNs (Fig. 6b). Neurons were clustered using a Gaussian mixture model[49, 50] (MClust), and
304 cluster assignments were used to designate a subtype identity to each newly profiled neuron.
305 This approach largely recapitulated our original subtype assignments, as only ten neurons
306 classified under baseline conditions were assigned to a different subtype (Fig. 6c). The significant
307 axonal projection bias of the two transcriptional subtypes was maintained as subtype 1 neurons
308 were predominantly labeled VPM/POm (70.7%, 408 of 585 cells, $p < 2.97 \times 10^{-71}$, hypergeometric
309 test) and subtype 2 VPM-only (84.2%, 369 of 438 cells, $p < 2.97 \times 10^{-71}$, hypergeometric test),
310 highlighting the reproducibility of our method and the robustness of our subtype assignment
311 (Fig. 6d).
312 We identified 1134 genes with significant differential expression across the three time
313 points sampled, independent of stable, baseline differences in expression between the two
314 transcriptional subtypes (Fig. 6e, $q \leq 0.001$; Monocle2 test). A k-means clustering analysis of
315 mean expression profiles identified 16 clusters of genes with different temporal expression
316 following deprivation, but failed to identify any clusters with significant subtype-specific effects
317 over time (Fig. 6e). However, the transcriptional changes induced in response to altered sensory
318 input are unlikely to be synchronous across all neurons collected at each time point. To describe
319 the cellular responses to altered sensory input without the confounding effects of neurons in
320 diverse states intermixed at each time point, we established a pseudotemporal ordering for the
321 1023 L6CThNs derived from the 1134 genes with significant differential expression across time
322 points (Fig. 7a). Briefly, using the Monocle2 DDRTree algorithm, cells were arranged in an

323 embedded graph representation in a reduced dimensional space. In this manner, cells with
324 similar expression profiles across the 1134 target genes were positioned next to each other, and
325 a traversal through the graph revealed the sequence of transcriptional changes that reflect
326 progression through a given biological process. Through this projection, we captured a smooth
327 representation of transcriptional modulation to altered sensory input. As expected, the
328 distribution of L6CThNs along pseudotime generally followed the temporal order of collection
329 following whisker removal (Fig. 7a,b). Nonetheless, neurons from each time point were found
330 throughout pseudotime, confirming that the transcriptional response to sensory manipulation is
331 asynchronous across the population of L6CThNs. Interestingly, we observed no bias in the
332 distribution of L6CThNs across pseudotime when the neurons were grouped by hemisphere ipsi-
333 or contralateral to whisker removal, suggesting that longer term transcriptional responses in
334 L6CThNs from both hemispheres are similar (Supplementary Fig. 7e).

335 We identified 1507 genes that were differentially expressed across pseudotime at a
336 much higher stringency than our aggregate analysis independent of transcriptional subtype
337 designation ($q \leq 0.0000001$; Monocle2), indicating that a significantly greater fraction of the
338 transcriptome was identified as modulated by sensory manipulation using the increased
339 resolution of pseudotemporal ordering. To identify cellular processes modulated along
340 pseudotime, we clustered the normalized response curves of the differentially expressed genes
341 (Fig. 7c, Supplementary Table 5) and queried each cluster for enrichment of annotated gene sets
342 from public databases and our curated list of activity-induced genes. The cluster with the
343 earliest changes in expression (Cluster 2) corresponded to significant downregulation of the
344 activity-associated genes ($p < 5.65 \times 10^{-7}$, hypergeometric test) that were enriched in the four
345 activity-associated modules contributing to cellular identity. In contrast, the cluster representing

346 the late response (Cluster 3) corresponded to upregulation of genes associated with chromatin
347 remodeling and reorganization as well as upregulation of lncRNAs, suggesting a slower
348 epigenetic response to sensory manipulation. Interestingly, genes associated with long-term
349 potentiation (LTP) from Cluster 2 (*Calm1*, *Calm2*, *Gria1*, *Gria2*, *Plcb1*, *Plcb4*, *Ppp1ca*, and *Ppp3r1*)
350 were downregulated early in the response while LTP-associated genes in Cluster 3 (*Crebbp*,
351 *Adcy1*, *Grin1*, *Prkcb*, *Ppp3ca*, and *Ppp3cb*) were upregulated towards the end of pseudotime,
352 suggesting that non-overlapping subsets of genes in this single gene set are regulated at distinct
353 phases of the L6CThN response to sensory manipulation. We also identified 75 transcription
354 factors (TFs) with significant differential regulation after sensory manipulation including activity-
355 associated TFs[5, 6] such as *Arc*, *Fos*, *Ier2*, *Junb*, *Mef2c*, and *Nr4a1,* which were expressed early and
356 downregulated over the course of pseudotime. Several TFs involved in neural development[3]
357 including *Neurod6*, *Fezf2*, *Mef2c*, and *Foxp2* were transiently expressed during the response,
358 suggesting a regulatory relationship between activity-dependent plasticity and neural
359 development.
360
361 ***Activity induced changes in gene expression enhance the distinction between subtypes***
362 Our data demonstrate that a transcriptional signature of activity significantly contributes to
363 L6CThN subtype identity under baseline conditions, and that altered sensory inputs result in
364 dramatic transcriptional responses in L6CThNs. These results suggest that the response to
365 sensory manipulation may alter the transcriptional relationship among neurons of the same
366 subtype as well as the distinction between the two transcriptionally defined L6CThN subtypes.
367 To test these hypotheses, we assessed the distribution of pairwise Euclidean distances of the
368 variance stabilized gene expression estimates across all expressed genes for L6CThNs within

369 each subtype for each day following sensory manipulation. In both subtypes, we observed that
370 the response to altered sensory input resulted in an increased mean distance among cells across
371 days (Fig. 8a; $p < 2.2 \times 10^{-16}$, Welch's two-sample t-test). We also identified a significant increase
372 in the variance of the intra-subtype distances across days ($p < 3.59 \times 10^{-5}$, F-test) for all adjacent
373 time points except for subtype 2 Day 1 versus Day 7, indicating that the cell-to-cell variation
374 within both L6CThN subtypes increases in response to altered sensory input. Second, we
375 assessed the distribution of the inter-subtype pairwise Euclidean distances between subtype 1
376 and subtype 2. We found that the mean distance significantly increased from Day 0 to Day 1 as
377 well as from Day 1 to Day 7 (Fig. 8b; $p < 2.2 \times 10^{-16}$, Welch's two-sample t-test) as did the
378 variance of inter-subtype distances ($p < 4.67 \times 10^{-8}$, F-test). These results demonstrate that
379 modulation of neuronal activity increases both the transcriptional heterogeneity within each
380 L6CThN subtype and the relative transcriptional differences between L6CThN subtypes,
381 confirming a dependent relationship between activity level and transcriptional identity.
382
383 ***Subtype-biased responses contribute to transcriptional heterogeneity and subtype identity***
384 These results raise the possibility that subtype-specific responses to sensory manipulation drive
385 the increase in transcriptional heterogeneity and enhance the distinction between L6CThN
386 subtypes. We found that subtype 1 and subtype 2 were differentially distributed along
387 pseudotime at each time point assessed following sensory manipulation (Fig. 8c). Furthermore,
388 pseudotemporal ordering identified a single major branch point in the transcriptional responses,
389 with each of the two subsequent branches exhibiting a significant bias for a specific L6CThN
390 subtype (Branch A and Branch B; Fig. 8d). Consistent with the proportion of the two subtypes in
391 the baseline data, neither subtype 1 nor subtype 2 was significantly enriched in the root state

392 (grey arrow; Subtype 1: $p < 0.0715$; Subtype 2: $p < 0.91$, hypergeometric test), indicating that
393 both transcriptional subtypes share a similar early response to sensory manipulation. In
394 contrast, Branch A and Branch B exhibited significant subtype bias: Branch A (red arrow) was
395 biased for subtype 1 (VPM/POM; 80.4%, 148/184 neurons; $p < 7.03 \times 10^{-14}$, hypergeometric test)
396 and Branch B (blue arrow) for subtype 2 (VPM-only; 58.4%, 202/346 neurons; $p < 2.04 \times 10^{-13}$,
397 hypergeometric test). Similar subtype-specific biases were seen at a subsequent branch point
398 (Supplementary Fig. 8). These results indicate that although either cell type may engage the
399 discrete processes represented by each branch in response to sensory manipulation, the cellular
400 decisions to engage a particular response are biased with respect to L6CThN subtypes.
401 Using the Monocle2 branch expression analysis modeling (BEAM) test, we identified
402 1392 genes with significant branch-dependent differential expression (Supplementary Table 6; q
403 ≤ 0.0001; Monocle2 BEAM test); 926 of these overlapped with the 1507 genes with pseudotime-
404 dependent expression suggesting that discrete cellular responses independently contribute to
405 the aggregate transcriptional response to altered sensory manipulation. The branch-dependent
406 genes were organized into seven distinct clusters using hierarchical clustering of the Monocle2
407 branched model fits (Fig. 8e). Hypergeometric testing demonstrated that neurons that progress
408 along Branch A (VPM/POM enriched) were enriched for genes associated with engagement of
409 the proteasome complex, a process involved in synaptic remodeling[51], while Branch B cells
410 (VPM-only enriched) were significantly enriched for genes related to the organization of the
411 postsynaptic density as well as to the regulation of LTP. We identified 79 TFs expressed in a
412 branch-dependent manner, including 31 TFs not identified in our initial analysis of the aggregate
413 response (Fig. 8f), suggesting that much of the aggregate L6CThN response induced by altered
414 sensory inputs is confounded across these two alternative cellular responses. The remaining 27

415 pseudotime-dependent TFs may regulate a uniform response independent of these two
416 subtype-biased responses. Additional genes regulated in a branch-specific manner included
417 thirty-two presynaptic genes (Fig. 8g). Furthermore, neurons committing to a Branch A response
418 demonstrated enhanced expression of GABA receptor subunits (*Gabra5*, *Gabrb1* and *Gabrb2*)
419 while neurons along Branch B induced ionotropic glutamate receptors (Fig. 8h; *Grin1*, *Grin2b*,
420 *Grik5*, and *Grin2a*). Taken together, these subtype-specific biases in transcriptional responses
421 induced by sensory manipulation likely underlie the overall effect we observed on L6CThN
422 identity, enhancing the distinctions between subtype 1 and subtype 2, and significantly
423 increasing both the inter- and intra-subtype heterogeneity of L6CThNs.
424

425 **DISCUSSION**

426 Using scRNA-seq, we identified two transcriptionally defined subtypes of L6CThNs, each
427 exhibiting significant bias for particular long-range axonal projection targets. Analysis of gene
428 expression differences among the profiled neurons determined that variation in axonal
429 projection pattern, laminar position within the cortex, and neuronal activity state all significantly
430 contribute to transcriptional identity, and that these properties are at least partially dependent.
431 Manipulating the activity states of L6CThNs by altering tactile input further showed that each
432 subtype was biased for particular cellular responses to sensory manipulation, revealing
433 functional differences among these neurons. These subtype-biased transcriptional responses
434 not only increased cell-to-cell transcriptional heterogeneity within each subtype, but also
435 enhanced the transcriptional differences between the two subtypes. Together, these data
436 identify the most significant influences on the transcriptional identity of individual cortical
437 projection neurons, and suggest that discrete cellular properties and responses affect the

438 population-level variation and identity of neuronal subtypes.

439 Although it is possible that projection target and transcriptional identity are decoupled
440 for a minority of neurons, the incomplete segregation of retrograde label across subtypes
441 observed here likely represents mislabeling of a subset of L6CThNs, as expected due to the close
442 proximity of VPM and POm and the difficulty of retrogradely labeling all neurons projecting to
443 adjacent targets. Interestingly, the subtypes identified here share some transcriptional
444 similarities with two recently defined subtypes in primary visual cortex[12] (Supplementary Fig. 3),
445 suggesting a conservation of corticothalamic neuron transcriptional identity and axonal
446 projection target bias across functional regions of the cortex[20]. Future studies will be required to
447 determine whether VPM-only and LGN-only projecting neurons as well as VPM/POM and
448 LGN/LP neurons are transcriptionally analogous, and to identify sources of cellular variation
449 across cortical areas. Although the population of L6CThNs we analyzed is comprised of two
450 subtypes, profiling a greater number of neurons or profiling at greater depth may reveal
451 additional rare subtypes of L6CThNs or may show that subtle variations in axonal projection
452 pattern identified in some anatomical studies are not apparent in the expression profiles of
453 cortical neurons[20, 21].

454 Pseudotemporal ordering of the states induced in L6CThN transcriptomes by altered
455 sensory input assessed over seven days revealed that L6CThNs engage at least two molecularly
456 distinct responses in a subtype-biased, but not subtype-specific, manner. Although the distinct
457 transcriptional responses were dominated by neurons collected one and seven days following
458 whisker manipulation, neurons collected in the baseline state were found throughout
459 pseudotime, suggesting that individual cortical neurons may engage these plasticity responses in
460 the non-manipulated state. L6CThNs in the initial phase of cellular response to altered sensory

461 input exhibited similar gene expression changes, indicating that both subtypes engaged a
462 common initial transcriptional response. These results are consistent with previous studies of
463 responses induced by neural activity measured over hours across brain regions and between
464 inhibitory and excitatory neurons which found that common early transcriptional responses lead
465 to cell-type specific late responses[52, 53].

466 The differential response to sensory manipulation resulted in an increase in the
467 transcriptional variation across L6CThNs within each subtype as well as a significant
468 enhancement of the distinction between the subtypes, demonstrating a non-dissociable
469 relationship between neuronal identity and neuronal activity. Because single L6CThNs have the
470 potential to engage either transcriptional response regardless of subtype, our data suggest that
471 extrinsic factors such as distinct activation patterns, due to differences in the circuits in which
472 each subtype is embedded, induce neurons from a given subtype preferentially toward a similar
473 response rather than solely cell-autonomous properties. Interestingly, we found that the
474 transcriptional responses for neurons ipsilateral and contralateral to the sensory manipulation
475 were similar, suggesting that shared cellular mechanisms underlie the previously described
476 functional effects ipsilateral and contralateral to unilateral manipulation at timescales similar to
477 those assessed here[48, 54, 55]. Furthermore, expression of genes that strongly contributed to
478 subtype identity such as *Lamp5* was altered in response to sensory manipulation. These data
479 indicate that factors that change cell state such as plasticity or injury affect our ability to
480 accurately define discrete, stable transcriptional subtypes.

481 Our results indicate that the transcriptional profiles of cortical neurons reflect specific
482 features of these cells, but also that the transcriptional variation across individual neurons is a
483 principal feature of subtype identity with significant functional consequences for individual

484 neuronal responses and subtype function. Subtype 2 L6CThNs were more transcriptionally
485 heterogeneous than subtype 1 neurons under steady-state conditions in part because of
486 baseline differences in gene expression associated with neural activity. In addition, the intra-cell
487 type variation across subtype 2 at day 1 and at day 7 was greater than the inter-cell type
488 variation at the same time points, suggesting that changes in cell-to-cell variation, rather than
489 subtype specific differences, dominate the transcriptional responses to experience dependent
490 plasticity. Together, these data indicate that the contribution of neuronal activity to gene
491 expression differs across distinct neuronal subtypes and that the transcriptional variation due to
492 differences in neuronal activity state plays a central role in defining the identity of cortical
493 projection neurons.

494 **METHODS**

495 **Mice**
496 All procedures were approved by the Johns Hopkins Animal Care and Use Committee and
497 followed the guidelines of the Society for Neuroscience and the National Institutes of Health.
498 Animals used for RNA-sequencing ranged from postnatal day 23 (P23) to P28; animals used for
499 immunohistochemistry and *in situ* hybridization ranged from P23 to P32. Both males and
500 females were used in this study (gender indicated for each replicate below). All animals were
501 kept on a 12 h light/dark cycle, housed with at least two mice per cage and provided with
502 unlimited water and food. The following mouse lines were used:
503 Neurotensin receptor-1 Cre recombinase line[30] (Ntsr1-Cre, GENSAT 220), loxP-STOP-loxP-
504 tdTomato Cre reporter lines[56] (Ai9 and Ai14, Allen Institute for Brain Science), loxP-STOP-loxP-
505 eYFP Cre reporter line[56] (Ai3, Allen Institute for Brain Science), and Pantr1-LacZ[39]. Mice were
506 maintained on mixed backgrounds including C57BL/6 and CD-1.
507
508 **Stereotactic injections**
509 To identify layer 6 corticothalamic neurons (L6CThNs), mice P18 to P23 were anesthetized with
510 ketamine (50 mg/kg), dexmedetomidine (25 µg/kg) and the inhalation anesthetic, isoflurane.
511 Animals were placed in a custom-built stereotactic frame and anesthesia was maintained with
512 isoflurane (1-2.5%). A small craniotomy was performed, and a glass pipette (10–25 µm tip
513 diameter) was lowered into the thalamus. To identify L6CThNs that project to the ventral
514 posterior medial nucleus (VPM) of the thalamus, Ntsr1-Cre;tdTomato mice were injected in VPM
515 (1.1 mm posterior, 1.7 mm lateral , 1.4 mm ventral from bregma) with either Alexa 488 Cholera
516 toxin B (CTB488, ThermoFisher) or green Retrobeads IX (Lumafluor), and Ntsr1-Cre;YFP mice
517 were injected with Alexa 555 Cholera toxin B (CTB555, ThermoFisher). To identify L6CThNs that
518 project to both VPM and to the posterior medial nucleus of the thalamus (POm), Ntsr1-
519 Cre;tdTomato mice were injected in POm (1.35 mm posterior, 1.23 mm lateral , 3.3 mm ventral
520 from bregma) with green Retrobeads IX, and Ntsr1-Cre;YFP mice were injected with either
521 CTB555 or red Retrobeads (Lumafluor). Between 30 and 100 nl of tracer were pressure-injected,
522 and the pipette was kept in position for 5-10 minutes before removal. Following the injection,
523 the incision was sutured and buprenorphine (0.05 mg/kg) was administered to all animals
524 postoperatively. Quantification of the colocalization between retrogradely labeled neurons and
525 Cre expression in Ntsr1-Cre mice (Fig. 1e,j) was performed on mice injected with either CTB or
526 Retrobeads. Preliminary experiments indicated that all tracers used generated similar labeling.
527 However, Retrobeads resulted in increased discrimination of labeled neurons during
528 Fluorescence Assisted Cell Sorting (FACS) and were therefore used for all FACS experiments.
529
530 **Immunohistochemistry**
531 Brains from animals injected with neuronal tracers were removed and placed in a solution of 4%
532 paraformaldehyde (PFA) in 0.01 M phosphate buffered saline (PBS) for 3 h. Coronal sections
533 were cut on a vibratome (50 µm, VT-1000S, Leica). Mice expressing YFP were subjected to
534 immunohistochemistry (1:1000, chicken anti-GFP, GFP-1020, Aves, RRID:AB_10000240, and
535 1:300 AlexaFluor 488-conjugated goat anti-chicken, Life Technologies, A11039,
536 RRID:AB_10563770). Sections were then mounted using Aqua Poly/Mount mounting medium
537 (Polysciences, Inc) and visualized on a confocal microscope (LSM 510, Zeiss) using 10x (0.3 NA),
538 25x (0.8 NA) or 40x (1.3 NA) objectives or on an epifluorescence microscope (AxioObserver.Z1,

539 Zeiss) using a 5x (0.16 NA) objective. Colocalization of tdTomato, YFP and retrograde tracer was
540 quantified using single-plane confocal images and the Cell Counter plugin in Fiji[57]. For all
541 anatomical comparisons, upper layer 6 and lower layer 6 were defined as the top 50% and
542 bottom 50% of the layer defined by Cre-expressing neurons in Ntsr1-Cre mice.
543
544 **Cell isolation and enrichment**
545 For bulk sequencing experiments (bulk RNA-seq), POm was injected with green Retrobeads in
546 three Ntsr1-cre;tdTomato mice. For single-cell sequencing experiments (scRNA-seq), POm was
547 injected with red or green Retrobeads in Ntsr1-Cre;YFP mice or in Ntsr1-Cre;tdTomato mice,
548 respectively. Mice were sacrificed 5-7 days after surgery (Bulk RNA-seq: Replicate 1: P23 female;
549 Replicate 2: P23 female; Replicate 3: P26 male; scRNA-seq baseline: Replicate 1: P27 male,
550 Replicate 2: P28 female; scRNA-seq Day 1: Replicate 1: P24 female, Replicate 2: P24 female;
551 scRNA-seq Day 7: Replicate 1: P29 female, Replicate 2: P27 female). Brains were rapidly
552 removed and 300 µm thick somatosensory thalamocortical[58] or coronal slices were sectioned
553 (VT-1200s, Leica) in ice-cold sucrose solution containing the following (in mM): 76 NaCl, 25
554 $NaHCO_3$, 25 glucose, 75 sucrose, 2.5 KCl, 1.25 NaH_2PO_4, 0.5 $CaCl_2$, 7 $MgSO_4$, pH 7.3, 310 mOsm
555 continuously bubbled with 95% O_2/5% CO_2. The slices were then placed in a submersion
556 chamber on an upright microscope in ice-cold sucrose solution continuously bubbled with
557 95% O_2/5% CO_2. The barrel cortex was identified using differential interference contrast (DIC)
558 and epifluorescence microscopy (Axioskop2 FsPLus, Zeiss) using 4x (0.1 NA) and 40x (water
559 immersion, 0.75 NA) objectives. The location of the retrograde tracer in the lower half of layer 6
560 was visually confirmed, and then regions of cortex below barrels with appropriate labeling were
561 microdissected and immediately dropped into 2.6 mL equilibrated Papain DNase-I solution. For
562 all experiments, at least three consecutive slices confirmed as clearly labeled were
563 microdissected.
564 The dissociation protocol was adapted from the trehalose-enhanced neuronal isolation
565 protocol described in Saxena et al, 2012[59] which is based on the Worthington Papain
566 Dissociation System (Worthington Biochemical Corporation). Briefly, tissue pieces were dropped
567 into a 5 ml tube containing 2.6 mL of a trehalose-Papain-DNase1 solution which included RNase
568 inhibitor, incubated 30-45 min at 37°C with gentle trituration at the half way point using a p1000
569 pipetter. The reaction was stopped by adding a protease inhibitor solution (for solution
570 compositions, see Saxena et al, 2012[59]. The tissue was then further dissociated and washed by
571 sequential pipetting and low-speed centrifugations. The final pellet was resuspended in 200 µL
572 of media (DMEM, 5% trehalose (w/v), 25 µM AP-V, 0.4 mM kynurenic acid, 6 µL of 40 U/µl
573 RNase inhibitor) at room temperature and introduced into a FACS machine (Beckman Coulter
574 MoFlo Cell Sorter). Neurons were sorted based on their fluorescence (either tdTomato-positive
575 (red) and Retrobead-positive (green) versus tdTomato-positive only or YFP-positive (green) and
576 Retrobead-positive (red) versus YFP-positive only) directly into lysis buffer either in eppendorf
577 tubes for bulk sequencing (350 µL lysis buffer, Qiagen RNeasy Micro Kit; Replicate 1: tdTomato-
578 only: 1943 cells, tdTomato and tracer: 2287 cells; Replicate 2: tdTomato-only: 386, tdTomato
579 and tracer: 782; Replicate 3: tdTomato-only: 604, tdTomato and tracer: 934) or into individual
580 wells of 96-well plates for single-cell sequencing (2 µL Smart-Seq2 lysis buffer + RNAase
581 inhibitor, 1 µL oligo-dT primer, and 1 µL dNTPs according to Picelli et al, 2014[60]. Upon
582 completion of a sort, the plates were briefly spun in a tabletop microcentrifuge and immediately
583 placed on dry ice. Single-cell lysates were subsequently kept at -80°C until cDNA conversion.
584

585 **Library preparation and amplification**
586 For bulk samples, RNA was extracted with a Qiagen RNeasy Micro Kit. RNA quality was assessed
587 with the Bioanalyzer Pico RNA kit and samples had RIN scores of 8.2-8.8. Library preparation and
588 amplification were performed according to the Smart-Seq2 protocol[60] with some modifications.
589 Principally, all primer and adapter sequences were synthesized with a 5'-biotin to minimize
590 primer dimerization. Briefly, 200 pg of RNA per sample were used as input for template
591 switching cDNA synthesis. Full-length cDNA was amplified by KAPA HiFi mediated PCR for 20
592 cycles using a 5'-biotinylated ISPCR primer. Then, 500 pg of Ampure XP bead cleaned (1:1) cDNA
593 was used as input for a standard Nextera XT tagmentation reaction, and amplification of
594 adapter-ligated fragments was carried out for 12 cycles during which individual index sequences
595 were added to each distinct sample. Library concentration was assessed with Qubit and library
596 fragment size distribution was assessed on the Agilent Bioanalyzer. Pooled, indexed bulk RNA-
597 seq samples were initially sequenced on the Illumina MiSeq 500 platform, and subsequently the
598 same pool was sequenced on one lane of the Illumina HiSeq 2500 platform to produce 50 bp
599 paired-end reads. Libraries were sequenced to an average depth of 4.2×10^7 fragments each.
600 Reads from both runs were aggregated by index prior to mapping.
601 Library preparation and amplification of single-cell samples was performed using a
602 modified version of the Smart-Seq2 protocol. Briefly, 96-well plates of single-cell lysates were
603 thawed to 4°C, heated to 72°C for 3 minutes, and then immediately placed on ice. Template
604 switching first-strand cDNA synthesis was performed as described above using a 5'-biotinylated
605 TSO oligo. cDNAs were amplified using 18 cycles of KAPA HiFi PCR and 5'-biotinylated ISPCR
606 primer. Amplified cDNA was cleaned with a 1:1 ratio of Ampure XP beads and 300 pg was used
607 for a one-half standard-sized Nextera XT tagmentation reaction. Tagmented fragments were
608 amplified for 12 cycles and dual indexes were added to each well to uniquely label each library.
609 Concentrations were assessed with Quant-iT PicoGreen dsDNA Reagent (Invitrogen) and
610 samples were diluted to ~2 nM and pooled. Pooled libraries were sequenced on the Illumina
611 HiSeq 2500 platform to a target mean depth of ~8.0×10^5 50 bp paired end fragments per cell[61]
612 at the Hopkins Genetics Research Core Facility to generate 50 bp paired-end reads.
613
614 ***In situ* hybridization and immunohistochemistry**
615 Immunohistochemistry (IHC) for tdTomato and *in situ* hybridization (ISH) for *linc-Tmem20* were
616 combined to show co-localization of the lncRNA transcript and tdTomato protein in L6CThNs.
617 Ntsr1-Cre;tdTomato mice (P23 to P30) were transcardially perfused with 0.1 M PBS followed by
618 a 4% PFA solution, and brains were post-fixed for 3 h at room temperature in 4% PFA. Then, 30
619 µm sections were cut on a vibratome (VT-1000S, Leica) and first processed for IHC (1:1000,
620 Rabbit anti-DsRed polyclonal antibody, Living Colors, Clontech Laboratories, RRID: AB_10013483
621 and 1:300 AlexaFluor 568-conjugated goat anti-rabbit, Life Technologies, A11011, RRID:
622 AB_143157). Next, the sections were processed for ISH following the protocol described in
623 Blackshaw, 2013[62] without proteinase K treatment. Sections were incubated overnight at 65°C
624 with 5 ng/µL of an anti-sense *linc-Tmem20* DIG labeled probe (DIG RNA labeling Kit (SP6/T7),
625 Roche). Detection of the hybridized DIG probes was performed using an alkaline phosphatase
626 (AP) conjugated antibody (1:1000, Anti-Digoxigenin [21H8] Alkaline Phosphatase, Abcam, Cat.
627 No.: ab119345, RRID:AB_10901703) followed by application of the AP substrate BCIP/NBT
628 (SigmaFast BCIP/NBT, Sigma-Aldrich). The enzymatic reaction was carried out until the desired
629 signal to noise ratio was observed. Images were then taken on an AxioObserver.Z1 (Zeiss) using

630 a combination of epifluorescence and brightfield imaging (10x objective, NA: 0.3). Brightness
631 and contrast were adjusted using Adobe Photoshop (Adobe Systems Incorporated).
632 For validation of *Pantr1* expression, transgenic mice expressing LacZ driven by the
633 endogenous *Pantr1* promoter locus[39, 63] were injected with CTB488 in VPM and CTB555 in POm.
634 IHC was performed on 50 µm-thick sections (rabbit anti-β galactosidase antibody 1:5000; a kind
635 gift from Dr. Joshua R. Sanes, Harvard University, Cambridge, MA and 1:300 AlexaFluor 647-
636 conjugated donkey anti-rabbit, Life Technologies, A31573, RRID: AB_2536183). Imaging was
637 performed on a confocal microscope (LSM 510, Zeiss) using 10x (0.3 NA), 25x (0.8 NA) or 40x
638 (1.3 NA) objectives. LacZ-positive and retrogradely labeled neurons were manually counted on
639 single-plane confocal images using the Cell Counter plugin in Fiji[57]. Brightness and contrast were
640 adjusted using Adobe Photoshop (Adobe Systems Incorporated).
641 For detection of single mRNA molecules, we used the commercially available system
642 RNAscope (ACDbio). Ntsr1-Cre;tdTomato mice (P23 to P30) were transcardially perfused with
643 0.1 M PBS followed by a 4% PFA solution, and brains were post-fixed for 2 h at room
644 temperature in 4% PFA. Fixed brains were then washed in PBS and allowed to equilibrate in a
645 30% sucrose-PBS solution for 36 h before being embedded and frozen in OCT compound
646 (Scigen, Tissue-Plus #4583). Frozen brains were kept at -80°C for up to 2 months. 20 µm sections
647 were collected on a cryostat (Leica CM3050) and processed according to instructions provided
648 by RNAscope. The following probes were used: tdTomato-C2 (317041-C2), tdTomato-C3 (
649 317041-C3), Mm-Rbfox3 (313311), Rbfox3-C2 (313311-C2), Mm-Lamp5 (451071), Mm-Pantr1
650 (483711), Mm-Serpini1 (501441), Gabra5-C3 (319481-C3). Imaging was performed on a confocal
651 microscope (LSM 510, Zeiss) using a 63x objective (NA 1.4). Sections from three animals were
652 used for *Pantr1*, *Lamp5* and *Gabra5* and from four animals for *Serpini1*. For quantification of the
653 number of mRNAs per neuron obtained using RNAscope, single plane images were first
654 processed with the open-source software CellProfiler (Broad Institute) to automatically identify
655 individual cells based on DAPI signal and to generate outline files. These outlines were then used
656 as input, in combination with the corresponding confocal images, into FISH-quant[64] (Matlab) for
657 quantification of transcript number per cell. Cells with no detectable NeuN (*Rbfox3*) transcript
658 were discounted as non-neuronal cells. A threshold of 20 *tdTomato* puncta was selected as the
659 cutoff for distinguishing tdTomato-positive from tdTomato-negative neurons. This resulted in
660 164 tdTomato-positive and 188 tdTomato-negative neurons for *Serpini1*, 118 tdTomato-positive
661 and 150 tdTomato-negative neurons for *Lamp5*, 111 tdTomato-positive and 198 tdTomato-
662 negative neurons for *Pantr1*, and 135 tdTomato-positive and 160 tdTomato-negative
663 neurons for *Gabra5*. Individual images were stitched to reconstitute an image spanning layer 6
664 vertically, and the distance of each neuron from the bottom of the layer was measured manually
665 using Fiji[57] and normalized to the height of layer 6. Statistical testing (Likelihood ratio test) and
666 curve fitting (Loess) was performed in R/Bioconductor.
667
668 **Sensory manipulation**
669 Prior to sensory manipulation, animals were bilaterally injected in POm with green Retrobeads
670 IX (Lumafluor) as described above. To induce experience dependent plasticity, mice (P18 to P23)
671 were anesthetized with the inhalation anesthetic, isoflurane. Once an adequate level of
672 anesthesia was achieved, the animals were placed under a surgical microscope and, using
673 tweezers, individual vibrissae were gently pulled by applying slow, steady tension to the base of
674 the whisker[65]. To produce a chessboard pattern, vibrissae alpha, gamma, A2, A4, B1, B3, C2, C4,
675 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.

Read pre-processing

For both bulk and single-cell libraries, paired-end reads were aligned to the mouse reference genome (mm10) using Hisat2[66] with the default parameters except: -p 8. Aligned reads from individual samples were quantified against a reference transcriptome[61] (GENCODE vM8) supplemented with additional lncRNA genes described previously[8] (Supplementary File S1) with cuffquant[67] 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[68]. 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[12], raw reads were downloaded from the short read archive (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[69] in R/Bioconductor[70]. Relative FPKM values for each cell were converted to estimates of absolute mRNA counts per cell (RPC) using the Monocle2 census utility. 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[71] (MGCV R package) 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 (Supplementary Fig. 2a and Fig. 7d). 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[72].

To cluster individual cells, we employed a spectral tSNE approach[73], 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, Supplementary Fig. 2b) were used as input for the tSNE visualization algorithm[72] (perplexity = 30). 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 (Supplementary Fig. 2c). 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[69]. To test for differentially

722 expressed genes between transcriptionally-defined cell types, the following full model was fit to
723 each expressed gene: ~replicate+num_genes_expressed+celltype, and compared to a reduced
724 model in which celltype was removed. The number of genes expressed in each cell was included
725 as an explanatory variable as a proxy for the sensitivity of scRNA-seq in each individual cell. To
726 test for differentially expressed genes between anatomically labeled cells, the following full
727 model was fit to each expressed gene: ~replicate+num_genes_expressed+label, and compared
728 to a reduced model in which label was removed. In both cases, genes were selected as
729 significantly differentially expressed with a 0.1% false discovery rate (Monocle2 test; Benjamini-
730 Hochberg corrected). DE gene lists were tested for gene set enrichment against Gene
731 Ontology and Reactome genesets[74, 75] using the clusterProfiler Bioconductor package[76]. The
732 heatmap of differentially expressed genes was composed using the pheatmap R package[77] and
733 dendrograms were generated from the hierarchical clustering of the Euclidean distances
734 between points.
735 To identify gene sets contributing to transcriptional identity in our single-cell dataset,
736 gene-centric PCA was performed on a mean-centered matrix of variance-stabilized expression
737 estimates for high-variance genes across all cells. The resulting rotations were used to project
738 all expressed genes into the same PCA space to identify their weights. These weights were used
739 to rank-order all expressed genes and this ordering was used as input for a pre-ranked GSEA
740 analysis[78]. Gene sets for the GSEA analysis were derived from the Monocle2 differential gene
741 tests described above or from our curated list of neuronal activity genes[43-46] using their adjusted
742 p-value cutoff of $p < 0.01$ (Supplementary Table 4).
743
744 **WGCNA**
745 To identify modules of correlated gene expression we used the Weighted Gene Correlation
746 Network Analysis package WGCNA[42]. Normalized expression data were first batch-corrected
747 with respect to replicates using limma, and the resulting matrix of 12,537 detectably expressed
748 genes and 346 cells was used as input for WGCNA. Cells were hierarchically clustered using the
749 Jensen-Shannon distance with average linkage. A soft threshold power of 11 was used to create
750 a signed Topological overlap matrix directly from the normalized and batch-corrected
751 expression estimates. Modules were then learned with a minimum module size of 30. Modules
752 were merged below a distance threshold of 0.5. Module eigenvalues were estimated and
753 correlated with cellular traits using the Pearson product moment correlation test. Module gene
754 membership was determined in a similar manner in which individual gene expression estimates
755 were correlated with module eigenvalues and tested for significance. To test the effect of each
756 module on the segregation of L6CThN cell types, each module eigenvalue was separately
757 regressed out of the expression matrix using limma, and subsequent values were used as input
758 for a tSNE using identical parameters to the original assay.
759
760 **Pseudotime and BEAM analyses**
761 Pseudotemporal ordering was performed using the prescribed Monocle2 workflow. Briefly, all
762 1023 L6CThNs passing quality controls were used as input. To reconstruct a trajectory that
763 reflected cellular progression in response to altered sensory input, we first performed a
764 differential test to identify genes whose expression changed as a function of collection day,
765 independent of baseline differences between celltype (fullModelFormulaStr =
766 ”~num_genes_expressed+Total_mRNAs+sex+celltype*Day, reducedModelFormulaStr = ”
767 ~num_genes_expressed+Total_mRNAs+sex+celltype”, $q \leq 0.01$). These 1134 genes were used as

768 a filtering set for the DDRTree dimensionality reduction. Cells were tested for differential
769 expression along pseudotime using the following model comparisons with q ≤ 0.0000001:
770 fullModelFormulaStr = "~num_genes_expressed+celltype+sm.ns(Pseudotime,df=3)",
771 reducedModelFormulaStr = "~num_genes_expressed+celltype". Branch dependent expression
772 was determined using the Monocle2 BEAM test (fullModelFormulaStr =
773 "~num_genes_expressed+celltype+sm.ns(Pseudotime, df =
774 3)*Branch",reducedModelFormulaStr = "~num_genes_expressed+celltype+sm.ns(Pseudotime,
775 df = 3)", q ≤ 0.0001). All significant gene lists were tested for gene set enrichment using the
776 hypergeometric test.
777
778 **Data and software availability**
779 All primary data are archived on the SRA and Gene Expression Omnibus and are available for
780 direct download on our companion website
781 (http://rstudio.gofflab.org:8383/apps/L6CthPN_dash/). Source code and software tools are
782 available upon request.
783
784 **Author Contributions**: MC, SPB and LAG conceived the experiments. MC, JDR, and GHC
785 performed all experiments. MC, SPB and LAG analyzed the data and wrote the paper with input
786 from all authors.
787
788 **Conflict of Interest**: The authors declare no competing financial interests.
789
790 **Acknowledgments**: We thank Hao Zhang and the Johns Hopkins Bloomberg School of Public
791 Health Flow Cytometry and Cell Sorting Core Facility. Sequencing service was provided by the
792 Johns Hopkins Genetic Resources Core Facility. We thank Dr. Joshua R. Sanes, Harvard
793 University, for the anti-LacZ antibody and Dr. Jesse Gray, Harvard Medical School, for the
794 curated list of neuronal activity-induced genes. This work was supported by a Klingenstein-
795 Simons Fellowship (SPB), a Johns Hopkins Science of Learning Grant (SPB, LAG), National Science
796 Foundation Grant 1656592 (SPB, LAG), NIH training grant T32 GM07814 (JDR), a Boerhinger-
797 Ingelheim Fonds Fellowship (MC), and the NINDS (NS050274).
798
799
800
801
802
803 **REFERENCES**

804

805

806 **FIGURE LEGENDS**

807 **Figure 1: Layer 6 corticothalamic neuron subclasses distinguished by their axonal projections**

808 **have distinct gene expression profiles.** (**a**, **f**) The labeling schemes for two classes of layer 6
809 corticothalamic neurons (L6CThNs; VPM: Ventral posterior medial nucleus; POm: Posterior
810 medial nucleus). (**b**) Injection of a retrograde tracer (red, Alexa 555-Cholera toxin B) into VPM in
811 an Ntsr1-Cre;YFP mouse. Low (**c**) and high (**d**) magnification images of layer 6 (L6) of barrel
812 cortex (BC) showing colocalization of the retrograde tracer (red) and YFP (green). (**e**)
813 Quantification of the colocalization (n=4 mice). (**g**) Injection of a retrograde tracer (red) in POm
814 of an Ntsr1-Cre;YFP mouse. Low (**h**) and high (**i**) magnification images of BC showing the
815 colocalization of the tracer to CThNs in lower L6. (**j**) Quantification of the colocalization (n=3
816 mice). (**k**) Matrix showing the 69 genes differentially expressed between pools of VPM/POm and
817 VPM-only L6CThNs from three replicates. Scale bars, (**b**,**g**) 500 µm; (**c**,**h**) 50 µm; and (**d**,**i**) 10 µm.
818
819 **Figure 2: Unbiased clustering of single transcriptomes of layer 6 corticothalamic neurons**
820 **defines two subtypes with strong axonal projection bias.** (**a**) tSNE plot showing two subtypes of
821 layer 6 corticothalamic neurons (L6CThNs) identified through unbiased clustering of single
822 L6CThN transcriptomes. (**b**) Same tSNE plot as in (**a**), with each L6CThN color-coded by its axonal
823 projection label. (**c**) Fraction of VPM-only (green) and VPM/POm (red) L6CThNs in each
824 transcriptionally defined subtype for each replicate. (**d**) Matrix showing the hierarchical
825 clustering of the 346 individual L6CThNs (x-axis) and the 286 genes differentially expressed (DE)
826 between the two subtypes (y-axis, 0.1 % FDR). (**e**) tSNE plots showing the normalized expression
827 levels of two differentially expressed genes enriched in subtype 1 (*Fxyd6*, left) or subtype 2
828 (*Lamp5*, right). (**f**) tSNE plots showing the normalized expression levels of two differentially
829 expressed long-noncoding RNAs enriched in subtype 1 (*linc-Tmem20*, left) or subtype 2 (*Pantr1*,
830 right). (**g**) Low magnification image of *linc-Tmem20* (red) in barrel cortex of an Ntsr1-

831 Cre;tdTomato (green) mouse using a combination of *in situ* hybridization (*linc-Tmem20*) and
832 immunohistochemistry (tdTomato). Insets show higher expression of *linc-Tmem20* in L6CThNs in
833 lower layer 6 (L6, Inset 2) relative to upper L6 (Inset 1). **(h)** LacZ expression in barrel cortex of a
834 heterozygous *Pantr1-LacZ* mouse following injections of green retrograde tracer in VPM and red
835 tracer in POm. Insets show puncta of LacZ in VPM-only L6CThNs (Column 1, green) and not in
836 VPM/POm L6CThNs (Column 2, red and green L6CThNs). **(i)** Boxplot showing the fraction of
837 VPM-only and VPM/POm L6CThNs expressing LacZ (n=3 mice). (**j**) Boxplot of the median number
838 of genes detected across all cells for each subtype by replicate pair. (Replicate 1: Median
839 number of genes detected: Subtype 1: 5582 ± 526.3 (sd), Subtype 2: 5080 ± 650.0 (sd), $p < 2.169$
840 $\times 10^{-10}$, Mann-Whitney test; Replicate 2: Median number of genes detected: Subtype 1: 6950 ±
841 545.4 (sd), Subtype 2: 6569 ± 478.7 (sd), $p < 7.071 \times 10^{-7}$, Mann-Whitney test). (**k**) Cumulative
842 probability distribution of the pairwise Euclidean distances among cells in subtype 1 (gold) and
843 subtype 2 (blue) ($p < 2.2 \times 10^{-16}$, Welch's two-sample t-test). The black line represents the
844 pairwise distances among a random sample of 100 cells drawn from the 346 cells. 95%
845 confidence interval is shown in light grey (Dvoretzky-Kiefer-Wolfowitz inequality). Scale bars, (**g**)
846 100 µm, 20 µm; (**h**) 20 µm, 5 µm.
847
848 **Figure 3: Coordinately regulated gene sets contribute to the transcriptional identities of layer**
849 **6 corticothalamic neuron subtypes.** (**a**) Weighted gene co-expression network analysis
850 (WGCNA) on variance-stabilized gene expression estimates identifies modules of coordinately
851 regulated genes, which are grouped using hierarchical clustering of module eigengenes. (**b**)
852 Pearson correlation of each module eigengene with parameters for both transcriptional subtype
853 and label. Significance of each correlation was determined (*) using the Pearson's product

854 moment test ($p < 0.01$; Benjamini-Hochberg corrected). (**c**) Pearson correlation of each module
855 eigengene with component rotations for principal components 1-5. (**d,e**) Hypergeometric test
856 for enrichment of the 286 genes differentially expressed between transcriptionally defined
857 L6CThN subtypes (**d**) and genes associated with neuronal activity (**e**) within each module ($p <$
858 0.01; Benjamini-Hochberg corrected).
859
860 **Figure 4: Variation in the transcriptional profiles of layer 6 corticothalamic neurons is defined**
861 **by both subtype-specific genes and genes reflecting laminar location within layer 6.** (**a,d**) tSNE
862 plots showing the eigenvalue for each cell for the two WGCNA modules most correlated with
863 the first principal component (**a**, Midnight Blue; **d**, Turquoise). (**b,c,e,f**) Gene expression tSNE
864 plots (*left*) showing the normalized gene expression level in each cell for representative genes
865 with significant weights on PC1. (**b**) and (**c**) are representative members of the Midnight Blue
866 module, and (**e**) and (**f**) belong to the Turquoise module. Single molecule fluorescence *in situ*
867 hybridization (smFISH, *middle*) showing images of mRNAs detected for each gene of interest
868 (magenta), tdTomato (green), NeuN (cyan), and DAPI (blue) in layer 6 of Ntsr1-Cre;tdTomato
869 mice. Quantitative gene expression analysis of smFISH (*right*) showing the number of mRNAs
870 expressed per neuron as a function of normalized vertical position in layer 6 and neuronal cell
871 type (L6CThNs: Ntsr1;tdTomato-positive;NeuN-positive neurons in green and non-L6CThNs:
872 Ntsr1;tdTomato-negative, NeuN-positive neurons in grey). Curves represent Loess fits to the
873 individual data points, grouped by cell type, and shaded areas correspond to 95% confidence
874 intervals. Statistics: (**b**, *Lamp5*) "Subtype specific" $p < 7.3231 \times 10^{-19}$; "CThN+ position specific" p
875 $< 1.175 \times 10^{-17}$; "CThN- position specific" $p < 0.0035$; (**c**, *Pantr1*) "Subtype specific" $p < 0.9931$;
876 "CThN+ position specific" $p < 9.243 \times 10^{-11}$; "CThN- position specific" $p < 1.021 \times 10^{-14}$; (**e**,

877 *Serpini1*) ”Subtype specific” $p < 1.606 \times 10^{-06}$; “CThN+ position specific” $p < 1.045 \times 10^{-06}$; ”CThN-

878 position specific” $p < 0.3342$; (**f**, *Gabra5*) ”Subtype specific” $p < 0.1020$; “CThN+ position

879 specific” $p < 1.994 \times 10^{-3}$; “CThN- position specific” $p < 0.09873$.

880

881 **Figure 5: A signature of neuronal activity significantly contributes to the transcriptional**

882 **identities of layer 6 corticothalamic neuron subtypes.** (**a**) Module eigengenes for the four

883 modules with significant enrichment for genes associated with neuronal activity. (**b,c**) tSNE plots

884 showing the discretized expression of genes with the highest (**b**) and second highest (**c**) weights

885 in each direction of PC1. (**d,e**) tSNE plots showing the discretized expression of genes with the

886 highest (**d**) and second highest (**e**) weights in each direction of PC2. (**f,g**) tSNE plots showing the

887 discretized expression levels of four genes reflecting neuronal activity: *Bdnf* and *Camk2n1* (**f**),

888 and *Fos* and *Plk5* (**g**).

889

890 **Figure 6: Sustained modulation of gene expression after sensory manipulation in layer 6**

891 **corticothalamic neurons.** (**a**) Experimental design for sensory manipulation and single-cell

892 transcriptional analysis of L6CThNs from the barrel cortex at 1 and 7 days post manipulation. (**b**)

893 tSNE plot of all 1023 neurons obtained from baseline (Day 0), 1, and 7 days following sensory

894 manipulation using the 286 genes identified as differentially expressed between L6CThN

895 subtypes at day 0. Day 0 neurons are labeled by transcriptional subtype. Day 1 and Day 7

896 neurons are colored light gray. tSNE positions were fit to a Gaussian mixture model (black lines)

897 which was used to classify Day 1 and Day 7 neurons as members of the previously defined

898 L6CThN transcriptional subtypes. (**c**) tSNE plot colored by transcriptional subtype as assigned in

899 (**b**). 10 of 340 neurons (2.9%) from Day 0 were assigned to a different subtype than in **Fig. 2a**

900 and are labeled light green. (**d**) tSNE plot with neurons colored by projection label. (**e**) k-means
901 clustering analysis of mean-centered gene expression, aggregated by Day and transcriptional
902 subtype (Subtype 1: gold; Subtype 2: blue) for genes with significant differential expression after
903 sensory manipulation. Semi-transparent lines represent individual genes and bold lines
904 represent cluster centroids.
905
906 **Figure 7: Pseudotemporal reconstruction of transcriptional responses to sensory manipulation**
907 **in layer 6 corticothalamic neurons.** (**a**) Discriminative dimensionality reduction projection of
908 1023 L6CThNs using genes identified as significantly differentially expressed after sensory
909 manipulation. Neurons are colored by day relative to manipulation. (**b**) Density distribution of
910 L6CThNs across pseudotime, grouped by day following manipulation. (**c**) Heatmap of normalized
911 response curves for the 1507 genes with significant differential expression across pseudotime,
912 and significantly enriched gene sets identified for each cluster ($p < 1.0 \times 10^{-2}$, hypergeometric
913 test).
914
915 **Figure 8: Sensory manipulation induces distinct cellular responses in layer 6 corticothalamic**
916 **neurons biased with respect to transcriptional subtype.** (**a**) Distribution of the pairwise
917 Euclidean distances within each subtype for subtype 1 (*left*, gold) and subtype 2 (*right*, blue),
918 using variance-stabilized expression estimates for all expressed genes, showing a significant
919 increase in heterogeneity of gene expression for both subtype 1 and subtype 2. (**b**) Distribution
920 of pairwise inter-subtype Euclidean distances between transcriptionally defined L6CThN
921 subtypes across all expressed genes plotted for each day following sensory manipulation. There
922 is a significant divergence between these two subtypes across time points as indicated by a

923 positive shift in the distances after induction of experience-dependent plasticity. (**c**) Density
924 distributions of L6CThNs at each day plotted across pseudotime for the two transcriptional
925 subtypes of L6CThNs. (**d**) Discriminative dimensionality reduction projection of 1023 L6CThNs
926 shown in **Fig. 7a**, now colored by transcriptional subtype. Red and blue arrows indicate the first
927 major cellular response branch point following sensory manipulation. Grey arrow indicates the
928 tree root state and direction of response progression. Pie charts indicate the proportion of each
929 subtype in the population for each branch. (**e**) BEAM analyses of gene sets with significant
930 differential expression dependent on either major branch point showing all significant DE genes.
931 (**f-h**) BEAM heatmap for branch-dependent transcription factors not detected in the aggregate
932 pseudotime response (**f**), presynaptic proteins (**g**), and ligand-gated neurotransmitter receptors
933 (**h**).
934
935 **SUPPLEMENTARY INFORMATION** (provided separately)
936

937 **REFERENCES**

938 1. Junker, J.P. & van Oudenaarden, A. Every cell is special: genome-wide studies add a new
939 dimension to single-cell biology. *Cell* **157**, 8-11 (2014).
940 2. Liu, S. & Trapnell, C. Single-cell transcriptome sequencing: recent advances and remaining
941 challenges. *F1000Res* **5**, 182 (2016).
942 3. Molyneaux, B.J., Arlotta, P., Menezes, J.R. & Macklis, J.D. Neuronal subtype specification in
943 the cerebral cortex. *Nat Rev Neurosci* **8**, 427-437 (2007).
944 4. West, A.E. & Greenberg, M.E. Neuronal activity-regulated gene transcription in synapse
945 development and cognitive function. *Cold Spring Harb Perspect Biol* **3** (2011).
946 5. Lyons, M.R. & West, A.E. Mechanisms of specificity in neuronal activity-regulated gene
947 transcription. *Prog Neurobiol* **94**, 259-295 (2011).
948 6. Flavell, S.W. & Greenberg, M.E. Signaling mechanisms linking neuronal activity to gene
949 expression and plasticity of the nervous system. *Annu Rev Neurosci* **31**, 563-590 (2008).
950 7. Greig, L.C., Woodworth, M.B., Galazo, M.J., Padmanabhan, H. & Macklis, J.D. Molecular logic
951 of neocortical projection neuron specification, development and diversity. *Nat Rev Neurosci*
952 **14**, 755-769 (2013).
953 8. Molyneaux, B.J.*, et al.* DeCoN: genome-wide analysis of in vivo transcriptional dynamics
954 during pyramidal neuron fate selection in neocortex. *Neuron* **85**, 275-288 (2015).
955 9. Darmanis, S.*, et al.* A survey of human brain transcriptome diversity at the single cell level.
956 *Proc Natl Acad Sci U S A* **112**, 7285-7290 (2015).
957 10. Lake, B.B.*, et al.* Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing
958 of the human brain. *Science* **352**, 1586-1590 (2016).
959 11. Pollen, A.A.*, et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity

960 and activated signaling pathways in developing cerebral cortex. *Nat Biotechnol* **32**, 1053-
961 1058 (2014).
962 12. Tasic, B.*, et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics.
963 *Nat Neurosci* **19**, 335-346 (2016).
964 13. Zeisel, A.*, et al.* Brain structure. Cell types in the mouse cortex and hippocampus revealed by
965 single-cell RNA-seq. *Science* **347**, 1138-1142 (2015).
966 14. Hevner, R.F.*, et al.* Beyond laminar fate: toward a molecular classification of cortical
967 projection/pyramidal neurons. *Dev Neurosci* **25**, 139-151 (2003).
968 15. Sugino, K.*, et al.* Molecular taxonomy of major neuronal classes in the adult mouse
969 forebrain. *Nat Neurosci* **9**, 99-107 (2006).
970 16. Sorensen, S.A.*, et al.* Correlated gene expression and target specificity demonstrate
971 excitatory projection neuron diversity. *Cereb Cortex* **25**, 433-449 (2015).
972 17. Sanes, J.R. & Masland, R.H. The types of retinal ganglion cells: current status and
973 implications for neuronal classification. *Annu Rev Neurosci* **38**, 221-246 (2015).
974 18. Wagner, A., Regev, A. & Yosef, N. Revealing the vectors of cellular identity with single-cell
975 genomics. *Nat Biotechnol* **34**, 1145-1160 (2016).
976 19. Dueck, H., Eberwine, J. & Kim, J. Variation is function: Are single cell differences functionally
977 important?: Testing the hypothesis that single cell variation is required for aggregate
978 function. *Bioessays* **38**, 172-180 (2016).
979 20. Bourassa, J. & Deschenes, M. Corticothalamic projections from the primary visual cortex in
980 rats: a single fiber study using biocytin as an anterograde tracer. *Neuroscience* **66**, 253-263
981 (1995).
982 21. Bourassa, J., Pinault, D. & Deschenes, M. Corticothalamic projections from the cortical barrel

983 field to the somatosensory thalamus in rats: a single-fibre study using biocytin as an
984 anterograde tracer. *Eur J Neurosci* **7**, 19-30 (1995).
985 22. Briggs, F., Kiley, C.W., Callaway, E.M. & Usrey, W.M. Morphological substrates for parallel
986 streams of corticogeniculate feedback originating in both V1 and V2 of the macaque
987 monkey. *Neuron* **90**, 388-399 (2016).
988 23. Briggs, F. & Usrey, W.M. Parallel processing in the corticogeniculate pathway of the
989 macaque monkey. *Neuron* **62**, 135-146 (2009).
990 24. Katz, L.C. Local circuitry of identified projection neurons in cat visual cortex brain slices. *J*
991 *Neurosci* **7**, 1223-1249 (1987).
992 25. Killackey, H.P. & Sherman, S.M. Corticothalamic projections from the rat primary
993 somatosensory cortex. *J Neurosci* **23**, 7381-7384 (2003).
994 26. Kwegyir-Afful, E.E. & Simons, D.J. Subthreshold receptive field properties distinguish
995 different classes of corticothalamic neurons in the somatosensory system. *J Neurosci* **29**,
996 964-972 (2009).
997 27. Usrey, W.M. & Fitzpatrick, D. Specificity in the axonal connections of layer VI neurons in tree
998 shrew striate cortex: Evidence for distinct granular and supragranular systems. *J Neurosci*
999 **16**, 1203-1218 (1996).
1000 28. Zhang, Z.W. & Deschenes, M. Intracortical axonal projections of lamina VI cells of the
1001 primary somatosensory cortex in the rat: a single-cell labeling study. *J Neurosci* **17**, 6365-
1002 6379 (1997).
1003 29. Thomson, A.M. Neocortical layer 6, a review. *Front Neuroanat* **4**, 13 (2010).
1004 30. Gong, S.*, et al.* Targeting Cre recombinase to specific neuron populations with bacterial
1005 artificial chromosome constructs. *J Neurosci* **27**, 9817-9823 (2007).

1006 31. Kim, J., Matney, C.J., Blankenship, A., Hestrin, S. & Brown, S.P. Layer 6 corticothalamic
1007 neurons activate a cortical output layer, layer 5a. *J Neurosci* **34**, 9656-9664 (2014).
1008 32. Bortone, D.S., Olsen, S.R. & Scanziani, M. Translaminar inhibitory cells recruited by layer 6
1009 corticothalamic neurons suppress visual cortex. *Neuron* **82**, 474-485 (2014).
1010 33. Brennecke, P.*, et al.* Accounting for technical noise in single-cell RNA-seq experiments. *Nat*
1011 *Methods* **10**, 1093-1095 (2013).
1012 34. Hicks, S., Teng, M. & Irizarry, R.A. On the widespread and critical impact of systematic bias
1013 and batch effects in single-cell RNA-seq data. *BioRxiv* **doi:10.1101/025528** (2015).
1014 35. Chung, N.C. & Storey, J.D. Statistical significance of variables driving systematic variation in
1015 high-dimensional data. *Bioinformatics* **31**, 545-554 (2015).
1016 36. Kiselev, V.Y.*, et al.* SC3: consensus clustering of single-cell RNA-seq data. *Nat Methods* **14**,
1017 483-486 (2017).
1018 37. Lin, P., Troup, M. & Ho, J.W. CIDR: Ultrafast and accurate clustering through imputation for
1019 single-cell RNA-seq data. *Genome Biol* **18**, 59 (2017).
1020 38. Lein, E.S.*, et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**,
1021 168-176 (2007).
1022 39. Sauvageau, M.*, et al.* Multiple knockout mouse models reveal lincRNAs are required for life
1023 and brain development. *Elife* **2**, e01749 (2013).
1024 40. Joo, K.M.*, et al.* Distribution of vasoactive intestinal peptide and pituitary adenylate cyclase-
1025 activating polypeptide receptors (VPAC1, VPAC2, and PAC1 receptor) in the rat brain. *J Comp*
1026 *Neurol* **476**, 388-413 (2004).
1027 41. Sun, Q.Q., Prince, D.A. & Huguenard, J.R. Vasoactive intestinal polypeptide and pituitary
1028 adenylate cyclase-activating polypeptide activate hyperpolarization-activated cationic

1029 current and depolarize thalamocortical neurons in vitro. *J Neurosci* **23**, 2751-2758 (2003).

1030 42. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network

1031 analysis. *BMC Bioinformatics* **9**, 559 (2008).

1032 43. Lacar, B.*, et al.* Nuclear RNA-seq of single neurons reveals molecular signatures of activation.

1033 *Nat Commun* **7**, 11022 (2016).

1034 44. Mardinly, A.R.*, et al.* Sensory experience regulates cortical inhibition by inducing IGF1 in VIP

1035 neurons. *Nature* **531**, 371-375 (2016).

1036 45. Cho, J.H., Huang, B.S. & Gray, J.M. RNA sequencing from neural ensembles activated during

1037 fear conditioning in the mouse temporal association cortex. *Sci Rep* **6**, 31753 (2016).

1038 46. Kim, T.K.*, et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature*

1039 **465**, 182-187 (2010).

1040 47. Wallace, H. & Fox, K. The effect of vibrissa deprivation pattern on the form of plasticity

1041 induced in rat barrel cortex. *Somatosens Mot Res* **16**, 122-138 (1999).

1042 48. Fox, K. *Barrel Cortex* (Cambridge University Press, Cambridge, 2008).

1043 49. Fraley, C., Raftery, A.E., Murphy, T.B. & Scrucca, L. mclust Version 4 for R: Normal Mixture

1044 Modeling for Model-Based Clustering, Classification, and Density Estimation Technical

1045 Report No. 597. *Department of Statistics, University of Washington* (2012).

1046 50. Fraley, C. & Raftery, A.E. Model-based clustering, discriminant analysis and density

1047 estimation. *Journal of the American Statistical Association* **97**, 611-631 (2002).

1048 51. Ehlers, M.D. Activity level controls postsynaptic composition and signaling via the ubiquitin-

1049 proteasome system. *Nat Neurosci* **6**, 231-242 (2003).

1050 52. Spiegel, I.*, et al.* Npas4 regulates excitatory-inhibitory balance within neural circuits through

1051 cell-type-specific gene programs. *Cell* **157**, 1216-1229 (2014).

1052 53. Whitney, O.*, et al.* Core and region-enriched networks of behaviorally regulated genes and
1053 the singing genome. *Science* **346**, 1256780 (2014).
1054 54. Frenkel, M.Y. & Bear, M.F. How monocular deprivation shifts ocular dominance in visual
1055 cortex of young mice. *Neuron* **44**, 917-923 (2004).
1056 55. Glazewski, S., Benedetti, B.L. & Barth, A.L. Ipsilateral whiskers suppress experience-
1057 dependent plasticity in the barrel cortex. *J Neurosci* **27**, 3910-3920 (2007).
1058 56. Madisen, L.*, et al.* A robust and high-throughput Cre reporting and characterization system
1059 for the whole mouse brain. *Nat Neurosci* **13**, 133-140 (2010).
1060 57. Schindelin, J.*, et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods*
1061 **9**, 676-682 (2012).
1062 58. Agmon, A. & Connors, B.W. Thalamocortical responses of mouse somatosensory (barrel)
1063 cortex in vitro. *Neuroscience* **41**, 365-379 (1991).
1064 59. Saxena, A.*, et al.* Trehalose-enhanced isolation of neuronal sub-types from adult mouse
1065 brain. *Biotechniques* **52**, 381-385 (2012).
1066 60. Picelli, S.*, et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* **9**, 171-
1067 181 (2014).
1068 61. Mudge, J.M. & Harrow, J. Creating reference gene annotation for the mouse C57BL6/J
1069 genome assembly. *Mamm Genome* **26**, 366-378 (2015).
1070 62. Blackshaw, S. High-throughput RNA in situ hybridization in mouse retina. *Methods Mol Biol*
1071 **935**, 215-226 (2013).
1072 63. Goff, L.A.*, et al.* Spatiotemporal expression and transcriptional perturbations by long
1073 noncoding RNAs in the mouse brain. *Proc Natl Acad Sci U S A* **112**, 6855-6862 (2015).
1074 64. Mueller, F.*, et al.* FISH-quant: automatic counting of transcripts in 3D FISH images. *Nat*

1075 *Methods* **10**, 277-278 (2013).

1076 65. Li, X., Glazewski, S., Lin, X., Elde, R. & Fox, K. Effect of vibrissae deprivation on follicle
1077 innervation, neuropeptide synthesis in the trigeminal ganglion, and S1 barrel cortex
1078 plasticity. *J Comp Neurol* **357**, 465-481 (1995).

1079 66. Kim, D., Langmead, B. & Salzberg, S.L. HISAT: a fast spliced aligner with low memory
1080 requirements. *Nat Methods* **12**, 357-360 (2015).

1081 67. Trapnell, C.*, et al.* Differential analysis of gene regulation at transcript resolution with RNA-
1082 seq. *Nat Biotechnol* **31**, 46-53 (2013).

1083 68. Trapnell, C.*, et al.* Differential gene and transcript expression analysis of RNA-seq
1084 experiments with TopHat and Cufflinks. *Nat Protoc* **7**, 562-578 (2012).

1085 69. Trapnell, C.*, et al.* The dynamics and regulators of cell fate decisions are revealed by
1086 pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381-386 (2014).

1087 70. Huber, W.*, et al.* Orchestrating high-throughput genomic analysis with Bioconductor. *Nat*
1088 *Methods* **12**, 115-121 (2015).

1089 71. Wood, S.N. Fast stable restricted maximum likelihood and marginal likelihood estimation of
1090 semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B*
1091 *(Statistical Methodology)* **73**, 3-36 (2011).

1092 72. Krijthe, J. Rtsne: T-distributed stochastic neighbor embedding using Barnes-Hus
1093 implementation. R package version 0.10. https://CRAN.R-project.org/package=Rtsne.
1094 (2015).

1095 73. Macosko, E.Z.*, et al.* Highly parallel genome-wide expression profiling of individual cells
1096 using nanoliter droplets. *Cell* **161**, 1202-1214 (2015).

1097 74. Ashburner, M.*, et al.* Gene ontology: tool for the unification of biology. The Gene Ontology

1098 Consortium. *Nat Genet* **25**, 25-29 (2000).

1099 75. Yu, G. & He, Q.Y. ReactomePA: an R/Bioconductor package for reactome pathway analysis

1100 and visualization. *Mol Biosyst* **12**, 477-479 (2016).

1101 76. Yu, G., Wang, L.G., Han, Y. & He, Q.Y. clusterProfiler: an R package for comparing biological

1102 themes among gene clusters. *OMICS* **16**, 284-287 (2012).

1103 77. Kolde, R. pheatmap: Pretty heatmaps. R package verion 1.0.8. https://CRAN.R-

1104 project.org/package=pheatmap. (2015).

1105 78. Subramanian, A.*, et al.* Gene set enrichment analysis: a knowledge-based approach for

1106 interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550

1107 (2005).

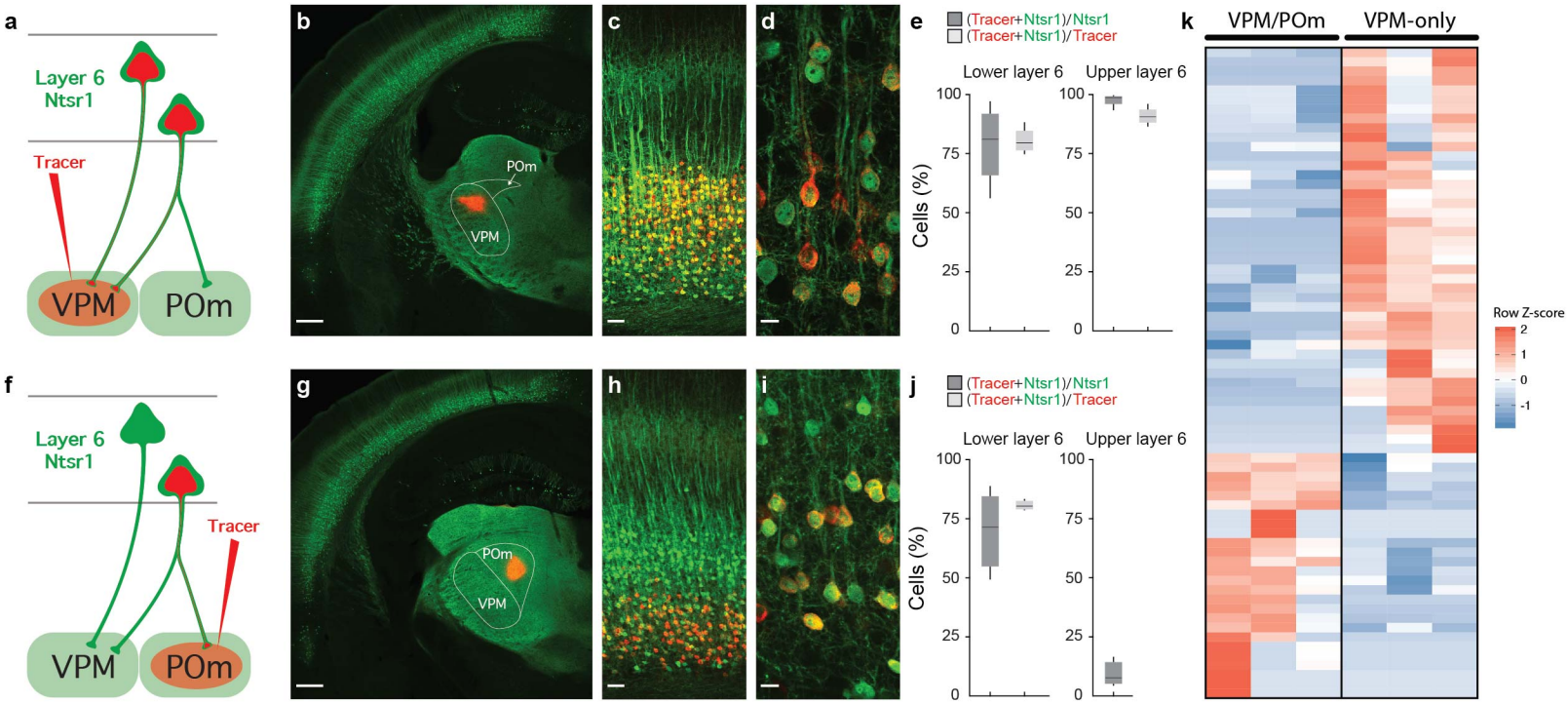

Figure 1
Chevée et al.

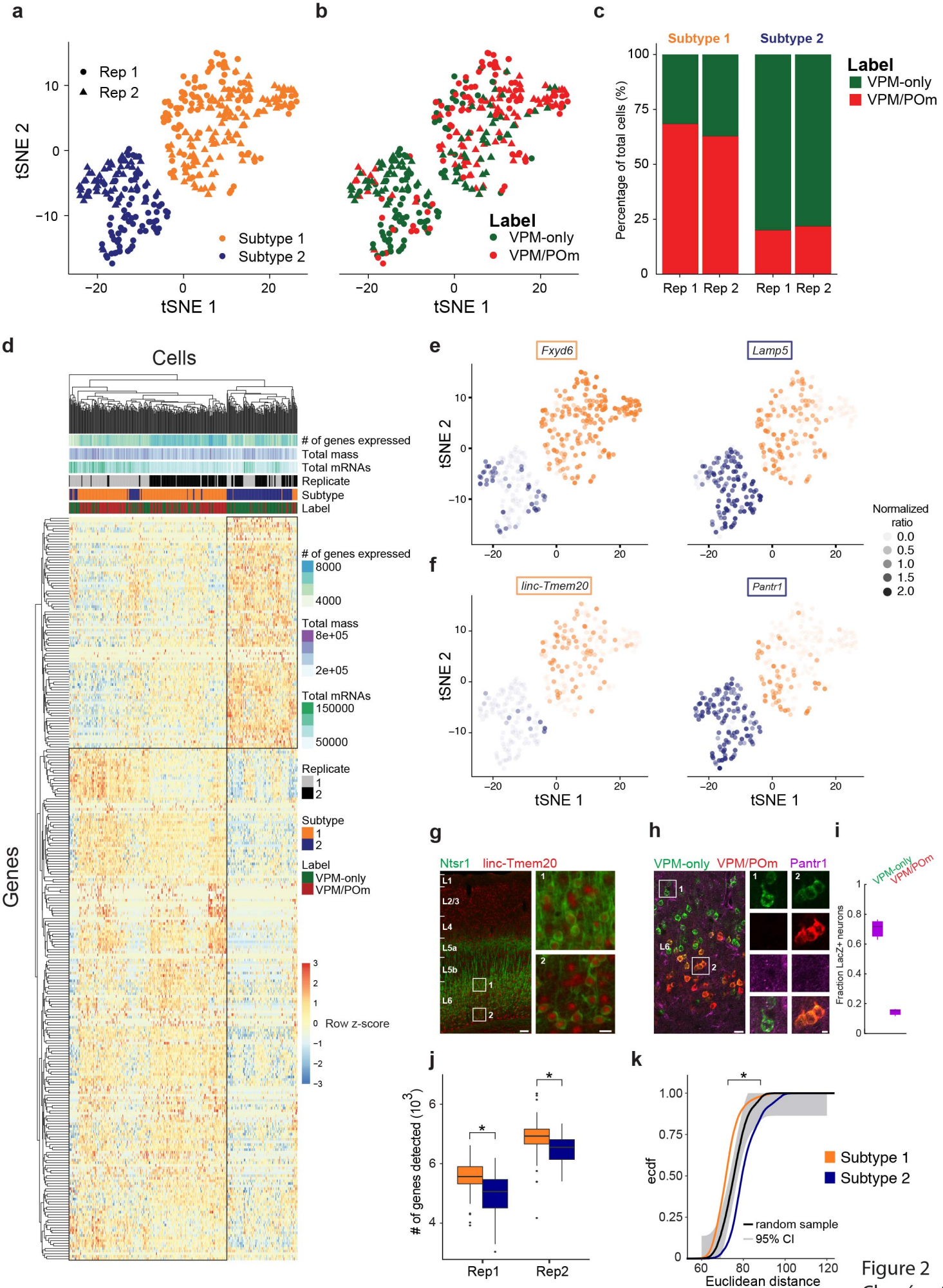

Figure 2
Chevée et al.

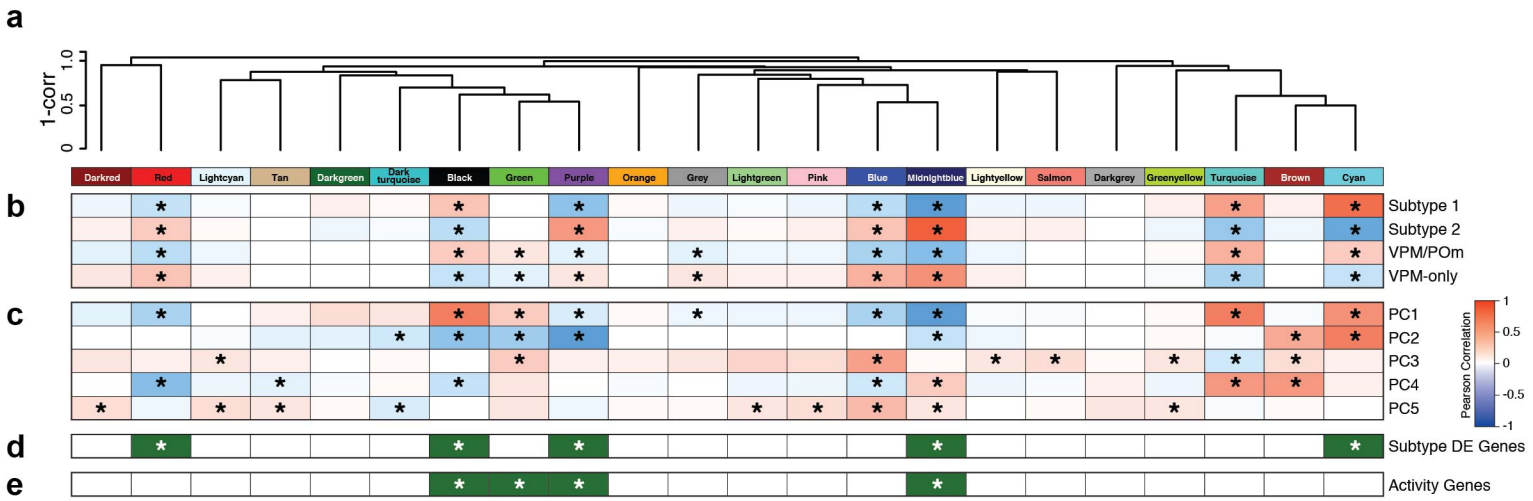

Figure 3
Chevée et al.

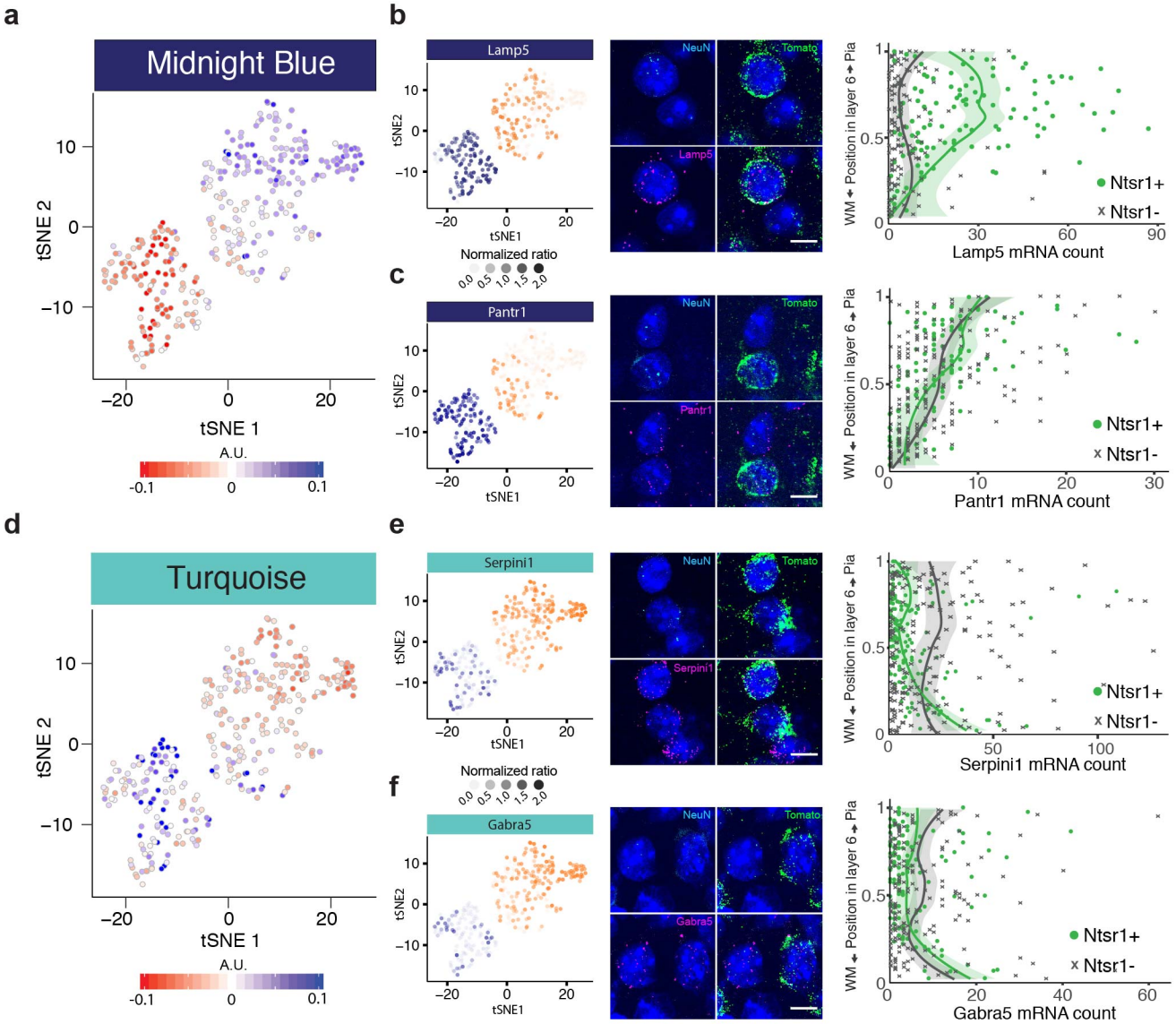

Figure 4
Chevée et al.

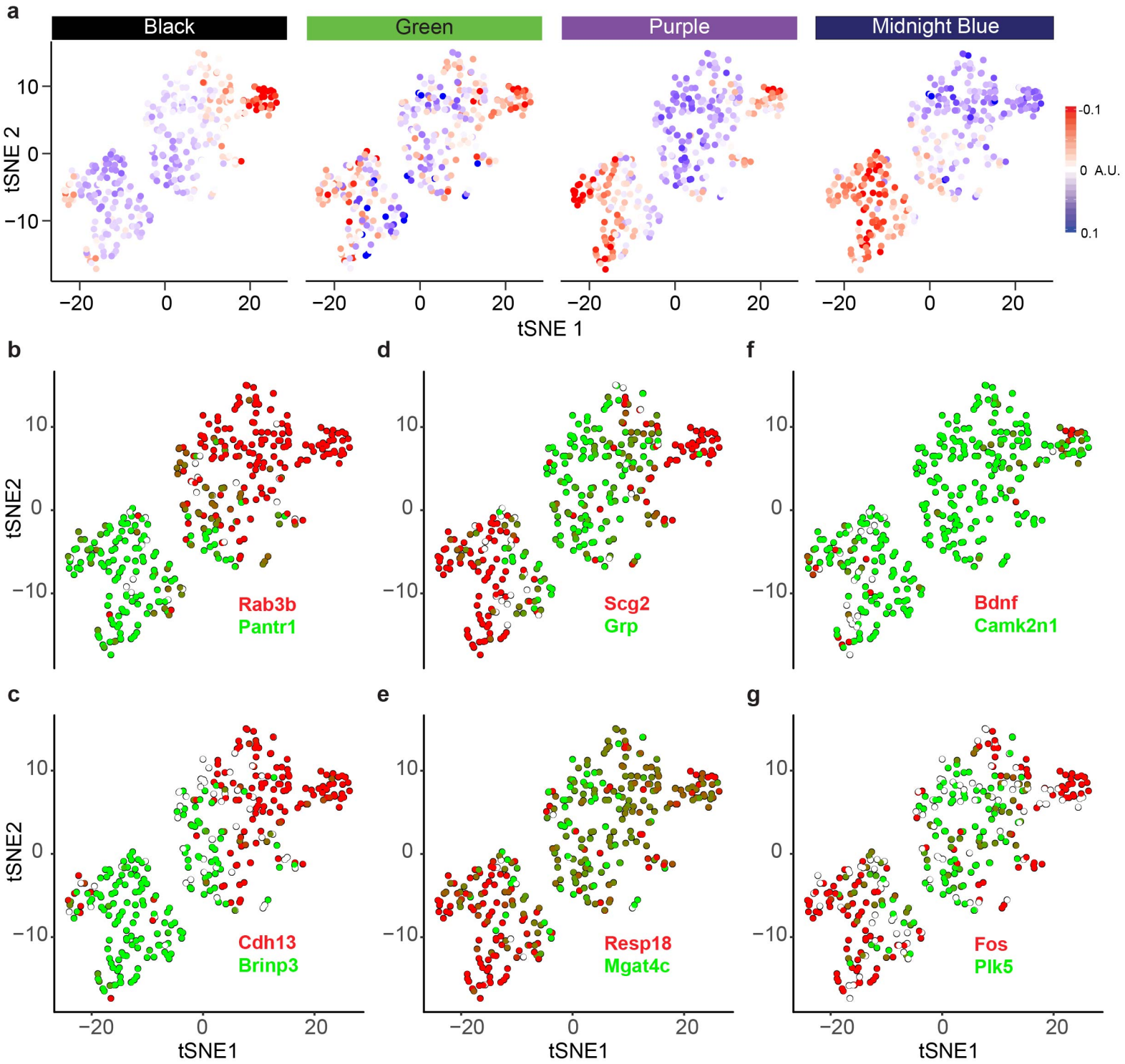

Figure 5
Chevée et al.

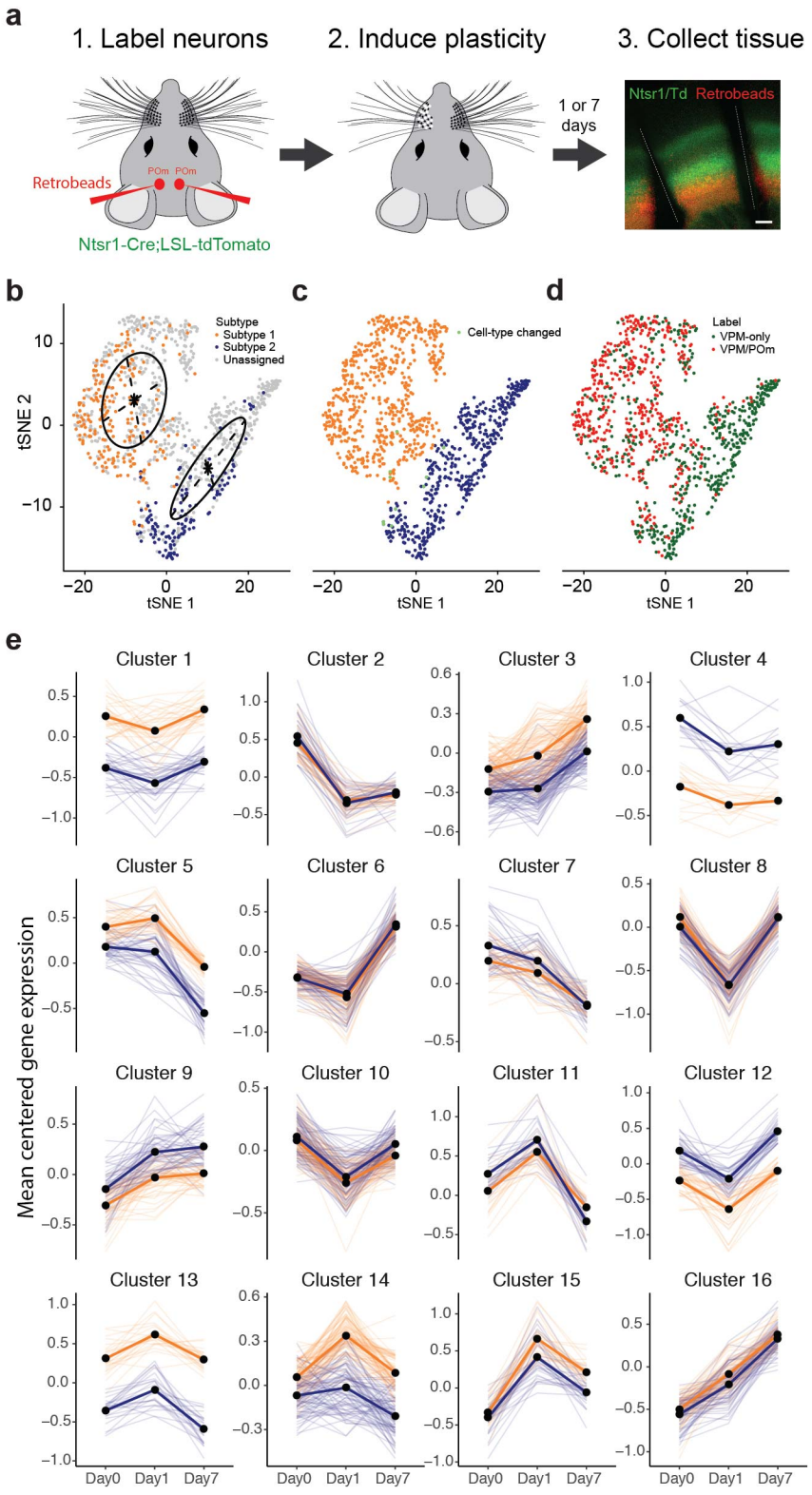

Figure 6
Chevée et al.

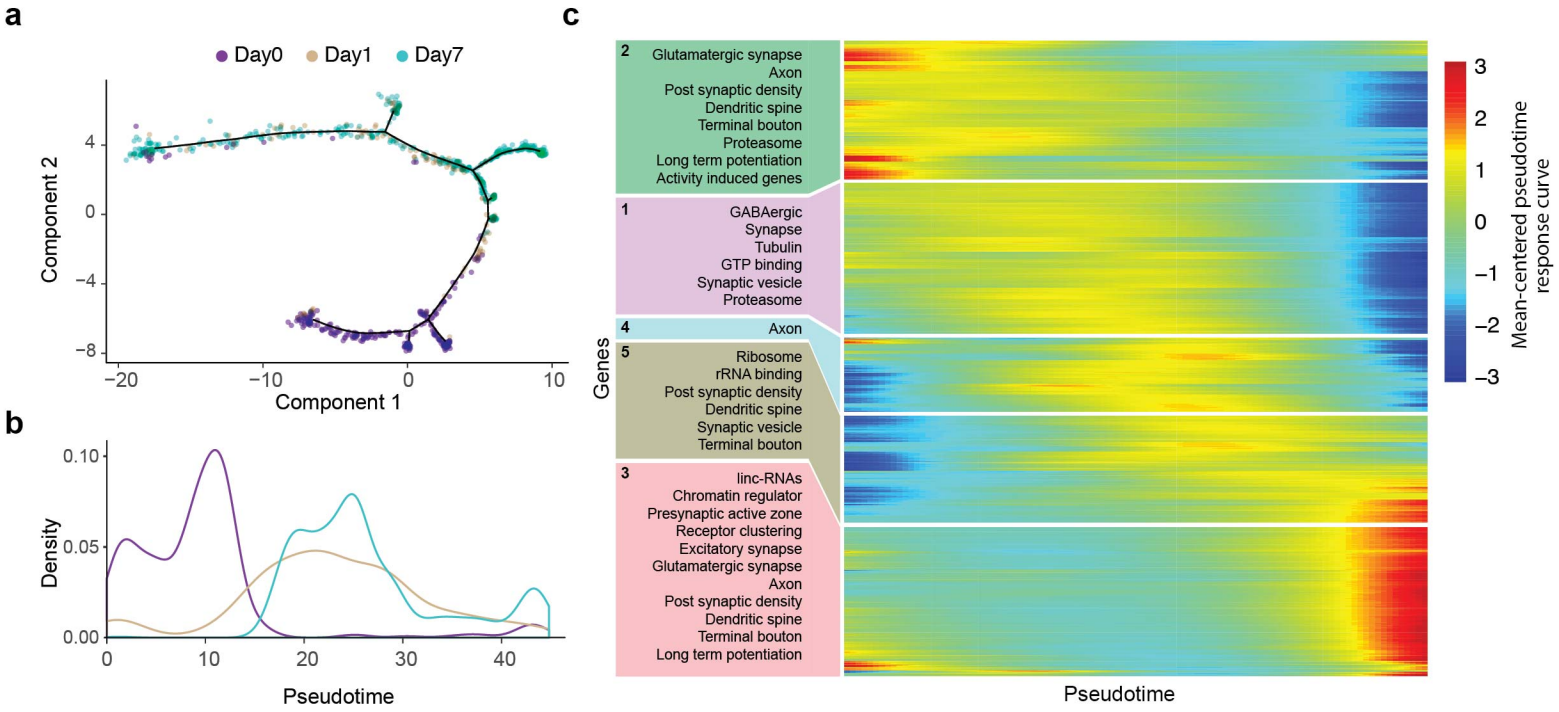

Figure 7
Chevée et al.

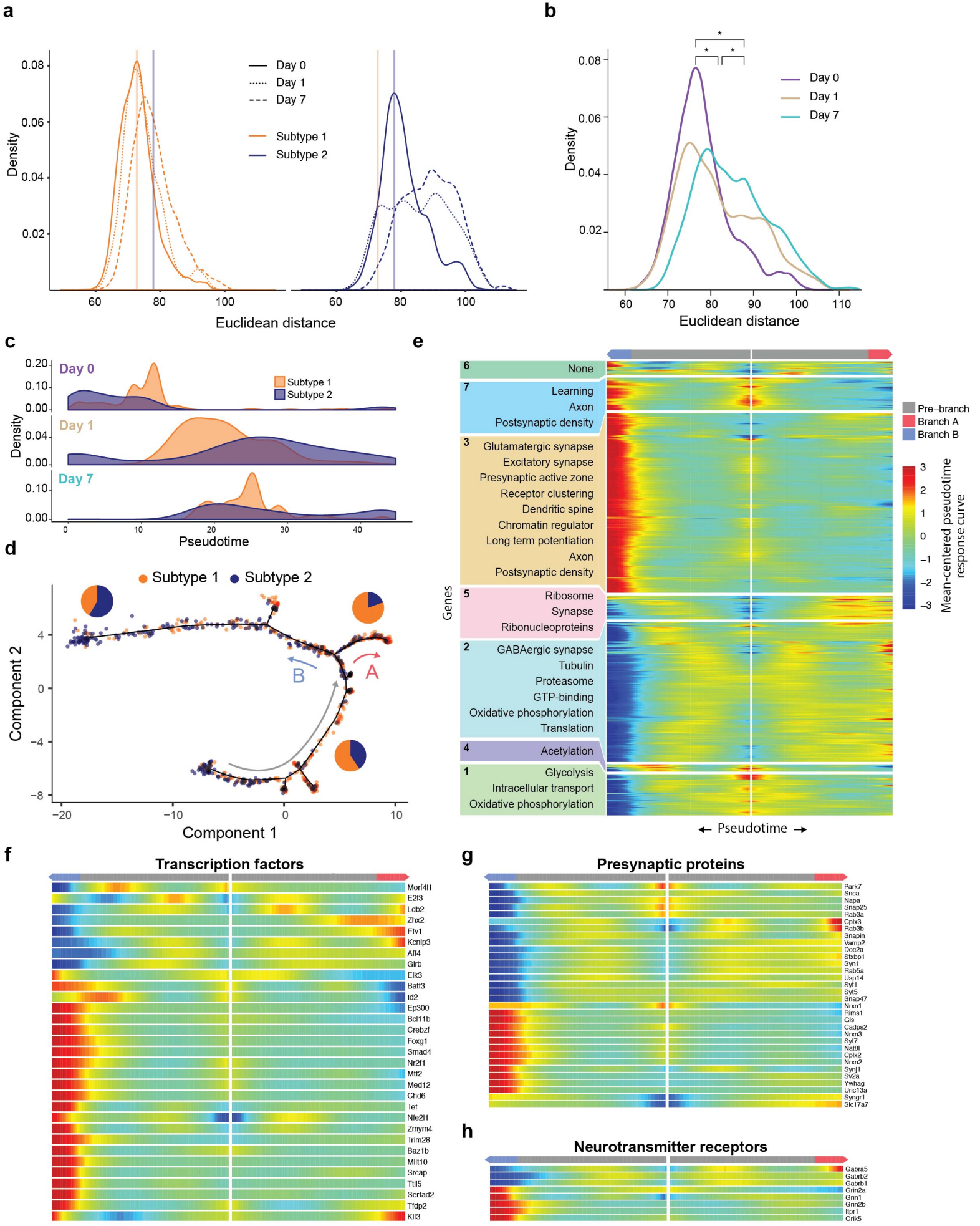

Figure 8
Chevée et al.