

Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex

Sinisa Hrvatin¹, Daniel R. Hochbaum^{2,3}, M. Aurel Nagy¹, Marcelo Cicconet⁴, Keiramarie Robertson³, Lucas Cheadle¹, Rapolas Zilionis^{5,6}, Alex Ratner⁷, Rebeca Borges-Monroy⁸, Allon M. Klein⁵, Bernardo L. Sabatini^{3*} and Michael E. Greenberg^{1*}

Activity-dependent transcriptional responses shape cortical function. However, a comprehensive understanding of the diversity of these responses across the full range of cortical cell types, and how these changes contribute to neuronal plasticity and disease, is lacking. To investigate the breadth of transcriptional changes that occur across cell types in the mouse visual cortex after exposure to light, we applied high-throughput single-cell RNA sequencing. We identified significant and divergent transcriptional responses to stimulation in each of the 30 cell types characterized, thus revealing 611 stimulus-responsive genes. Excitatory pyramidal neurons exhibited inter- and intralaminar heterogeneity in the induction of stimulus-responsive genes. Non-neuronal cells showed clear transcriptional responses that may regulate experience-dependent changes in neurovascular coupling and myelination. Together, these results reveal the dynamic landscape of the stimulus-dependent transcriptional changes occurring across cell types in the visual cortex; these changes are probably critical for cortical function and may be sites of deregulation in developmental brain disorders.

Neuronal activity shapes brain development and function through multiple mechanisms^{1,2}. For example, post-translational modifications affect neuronal excitability and synapse function in the short term³. These effects are followed by activity-dependent transcriptional programs that lead to long-lasting cellular adaptations necessary for learning and circuit-level homeostasis^{4,5}. Such transcriptional responses are evoked in the brain by a wide range of stimuli, including sensory experience, metabolic changes, circadian rhythm, stress, injury, and pharmacological intervention, and they have been implicated in many biological responses and diseases^{6–10}.

Early work characterizing activity-dependent transcriptional responses in cultured neurons revealed that a depolarizing stimulus rapidly induces calcium-dependent and mitogen-activated protein kinase (MAPK)-dependent early-response genes (ERGs)¹¹. These ERGs encode factors such as *Arc*, which can directly alter synaptic transmission¹², but primarily encode early-response transcription factors (ERTFs). The ERTFs induce a second wave of late-response genes (LRGs) including neuronal modulators and secreted factors (for example, *Bdnf*) that effect changes in circuitry^{13,14}. Until recently, the study of such responses had largely been limited to neurons *in vitro* or bulk tissue^{15,16}. However, newer technologies have extended this analysis to multiple genetically defined but investigator-selected neuronal cell types *in vivo*^{13,17}. From these studies, a model has emerged in which sensory experience induces a common ERTF program across neuronal types, which is followed by distinct LRG programs that regulate synaptic plasticity in a cell-type-specific manner.

Despite this recent progress in understanding sensory-experience-dependent gene expression, major gaps in knowledge still remain. Previous approaches have analyzed a limited number of inhibitory cell types and masked the full diversity of excitatory populations that form functionally and molecularly distinct layers with specialized roles within the local microcircuitry. Furthermore, neuronal activity can induce calcium waves in astrocytes, proliferation and myelination by oligodendrocytes, and structural changes in the neurovasculature^{18–20}. Despite the considerable diversity of stimulus-responsive cell types in the cortex, a comprehensive understanding is lacking regarding how the full complement of cells within a cortical region respond to a sensory stimulus and how this response contributes to neuronal plasticity.

To address this issue, we mapped the transcriptional response to visual stimulation by performing unbiased high-throughput droplet-based single-cell RNA sequencing (scRNA-seq) of 114,601 individual cells in the mouse visual cortex²¹. Remarkably, cells of all 30 transcriptionally identified types (hereafter referred to as ‘cell types’) responded to the visual stimulus with diverse transcriptional changes that diverged immediately across cell populations and resulted in induction of different ERTFs. Unexpectedly, astrocytes, oligodendrocytes, and endothelial cell types were highly responsive to the visual stimulus and initiated transcriptional changes indicative of structural remodeling. Among excitatory neurons, LRGs were enriched in secreted factors that may modulate neuronal-circuit connectivity and demarcate inter- and intralayer cell-type organization. Together, these results indicate that, in response to sensory stimuli, the cortical transcriptome is much more dynamic than previously appreciated, and the widespread changes in experi-

¹Department of Neurobiology, Harvard Medical School, Boston, MA, USA. ²Society of Fellows, Harvard University, Cambridge, MA, USA. ³Department of Neurobiology, Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA. ⁴Image and Data Analysis Core, Harvard Medical School, Boston, MA, USA. ⁵Department of Systems Biology, Harvard Medical School, Boston, MA, USA. ⁶Vilnius University Institute of Biotechnology, Vilnius, Lithuania. ⁷ICCB-L Single Cell Core, Harvard Medical School, Boston, MA, USA. ⁸Program for Bioinformatics and Integrative Genomics, Graduate School of Arts and Science, Division of Medical Sciences, Harvard University, Cambridge, MA, USA. Sinisa Hrvatin, Daniel R. Hochbaum and M. Aurel Nagy contributed equally to this work. *e-mail: bernardo_sabatini@hms.harvard.edu; michael_greenberg@hms.harvard.edu

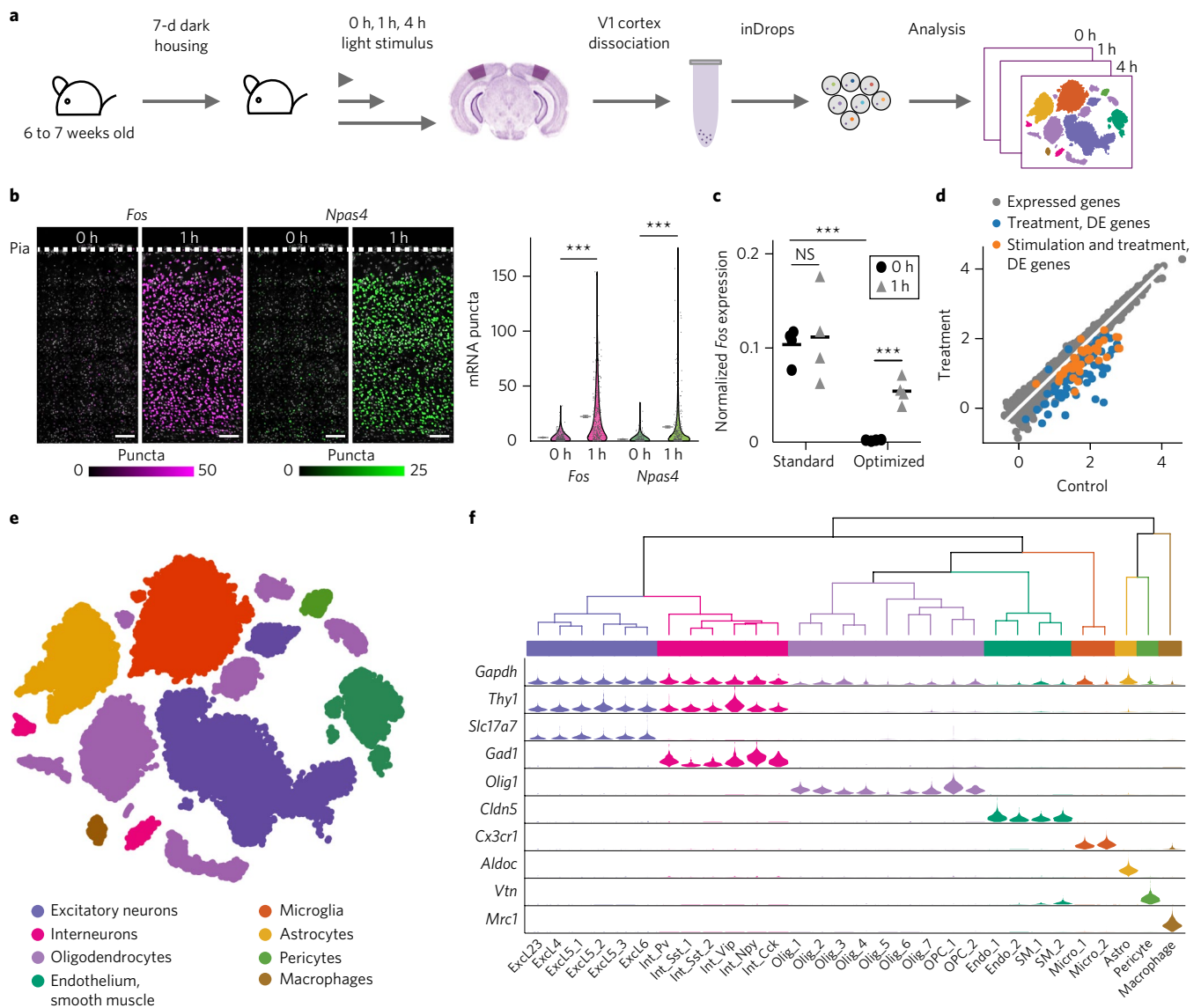


Fig. 1 | Workflow and identification of cell types. **a**, 6- to 7-week-old mice were housed in the dark for 7 d and exposed to light for 0 h (control), 1 h, or 4 h. V1 was dissociated into single cells and subjected to inDrops sequencing. **b**, FISH of the IEGs *Fos* and *Npas4* from mice exposed to light for 0 h or 1 h (left). Nuclei are pseudocolored according to the expression of *Fos* (magenta) or *Npas4* (green) by FISH (Methods). Pia denotes the cortical surface. Scale bars, 100 μ m. Graph shows quantification of FISH signals across time points, with means and 95% confidence intervals denoted by gray lines. A random subset (10%) of the raw data was selected for visualization. For both *Fos* and *Npas4*, $n = 2,667$ cells for 0 h or 2,683 cells for 1 h. *** $P < 10^{-200}$, Mann-Whitney *U* test, two-sided. Experiments were repeated on two cortical slices per time point. **c**, qRT-PCR for *Fos* normalized to *Gapdh*, comparing standard and optimized cell-dissociation protocols designed to limit IEG induction during dissociation. Horizontal lines, means. *** $P = 2.7 \times 10^{-4}$ (standard versus optimized 0 h) or 3×10^{-5} (optimized 0 h versus 1 h); NS, $P = 0.39$ (standard 0 h versus 1 h); $n = 4$ mice, unpaired *t* test, two-sided. **d**, RNA-seq analysis of cocktail-treated and control cells collected from $n = 4$ mice per condition. 114 genes that were significantly differentially expressed (DE) are denoted in blue (FDR < 0.01, |fold change| > 2, limma). 45 of these genes (orange) were also differentially expressed between cocktail-treated light-stimulated and light-unstimulated samples (Supplementary Fig. 2). Axis units are $\log_{10} [(\text{trimmed mean of M values (TMM)} - \text{normalized counts per million}) + 0.1]$. **e**, t-SNE plot of 47,209 cells from the V1 in 23 mice. Colors denote main cell types. **f**, Dendrogram and violin plots showing the distribution of expression of selected marker genes across all 30 analyzed cell types hierarchically clustered on the basis of variable gene expression. Subtypes are described in Methods.

ence-dependent gene expression among the diverse cell types identified here are likely to be essential for proper brain function.

Results

inDrops sequencing to study neuronal activity. We used an in vivo visual stimulus paradigm to study experience-dependent gene expression in the primary visual cortex^{13,16,17}. Mice between 6 and 7 weeks of age were housed in complete darkness for 1 week and then exposed to light for 0 h (control), 1 h, or 4 h (Fig. 1a and

Methods). We used fluorescence in situ hybridization (FISH) to confirm the ability of the light stimulus to induce ERG expression (Fig. 1b). Quantitative reverse-transcription PCR (qRT-PCR) revealed that light induction of ERGs was specific to the visual cortex and did not occur in the somatosensory, motor, or auditory cortices (Supplementary Fig. 1).

Immediately after light stimulation, the visual cortices were dissected and dissociated into single-cell suspensions (Fig. 1a and Methods). In contrast to the stimulus-dependent induction of *Fos*

measured by FISH and qRT-PCR, ERG expression was initially high in control samples of RNA extracted in bulk from single-cell suspensions (Fig. 1c). It has recently been reported that protocols relying on enzymatic digestion induce activity-dependent gene transcription in brain cells²², thus making it difficult to use previously generated scRNA-seq datasets to identify genes whose activity is induced in vivo in response to sensory stimuli. We therefore added to the dissociation solution a cocktail of inhibitors that block neuronal activity, calcium entry, transcription, and translation (Methods). This cocktail decreased the expression of *Fos* in the control but not the stimulated sample (51-fold, $P=3.2 \times 10^{-5}$, unpaired *t* test), thus preserving the pattern of in vivo activity-dependent gene induction detected by FISH (Fig. 1b). Genome-wide analysis showed differential expression of 114 genes (false discovery rate (FDR)-corrected $q < 0.01$; expression change twofold or greater), 110 of which were downregulated after inhibitor treatment (Supplementary Table 1). Of these 110 genes, 45 were also regulated by the light stimulus (FDR-corrected $q < 0.01$; Fig. 1d and Supplementary Fig. 2), thus further supporting the use of these blockers to prevent aberrant transcription during the enzymatic dissociation.

Using this optimized dissociation protocol to preserve the dynamic transcriptional state of ERGs in vivo, we conducted stimulus-dependent scRNA-seq profiling of the visual cortex after light exposure for various durations. After dissociation, single cells were captured, and their mRNAs were barcoded with the inDrops platform²¹ (Methods). In total, 28 preparations from 23 mice were sequenced (Supplementary Fig. 3), yielding 65,539 cells that passed initial quality-control tests (Methods).

Identification of cell types in the adult mouse visual cortex. The inDrops dataset was analyzed through two independent clustering approaches to classify individual cells into cell types according to their patterns of gene expression, thus generating a final dataset containing 47,209 cells (Supplementary Fig. 4 and Methods). On average, we recovered 3,234 transcripts per cell, which represented 1,453 unique genes expressed per cell (Supplementary Fig. 5). The cells were robustly classified into eight main cell types (excitatory neurons; inhibitory neurons; oligodendrocytes and oligodendrocyte precursor cells (OPCs); astrocytes; endothelial and smooth muscle cells; pericytes; microglia; and macrophages) and 30 subtypes (Fig. 1e,f, Supplementary Figs. 6–10 and Supplementary Table 2). The cluster identities (Supplementary Fig. 11) were consistent with recent scRNA-seq data from the mouse visual cortex²³.

Identification of sensory-stimulus-regulated transcriptional programs. To identify visual-experience-regulated genes within each cell type, we analyzed differential gene expression (Methods) across time points, defining differential expression as an FDR < 0.05 and a change in expression of twofold or greater (example results for Exc_L23 in Fig. 2a). Of 25,187 genes, 8,313 were significantly stimulus dependent in at least one cell type, and 611 passed the fold-change threshold (419 upregulated and 192 downregulated; Supplementary Table 3 and Supplementary Fig. 12).

We identified light-induced gene expression changes in all neuronal and, unexpectedly, all non-neuronal cell types in the visual cortex. Across all cell populations, the sensory-experience-regulated genes were grouped into ERGs (362) and LRGs (249) on the basis of the temporal pattern of gene induction (Methods; Fig. 2b and Supplementary Table 3). Gene Ontology (GO) analysis across all 611 genes showed significant enrichment for the terms ‘positive regulation of transcription’ and ‘MAPK signaling’ (additional terms in Supplementary Table 4).

The single-cell resolution of the inDrops data also allowed us to estimate the fraction of cells in each population exhibiting acute transcriptional responses to light stimulation (Methods and Fig. 2c). This analysis revealed that 49–69% of excitatory neurons exhibited

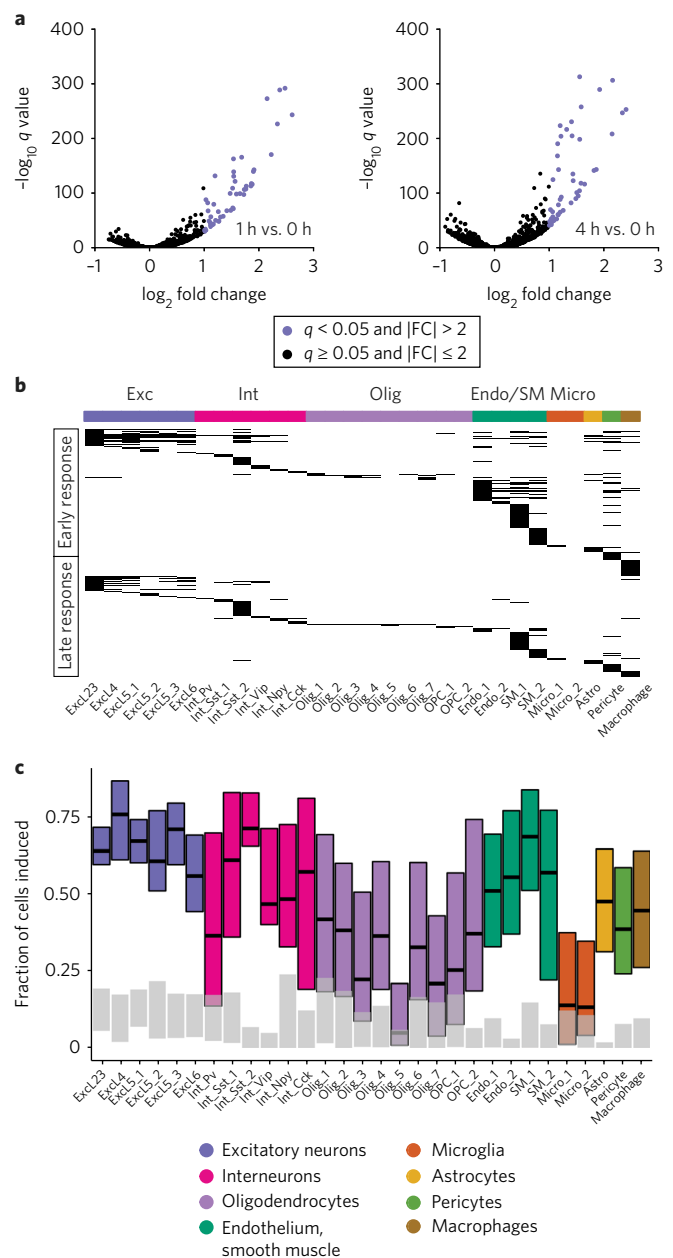


Fig. 2 | Identification of sensory-stimulus-regulated genes. **a**, Sample volcano plots for Excl23 cells indicating genes identified as sensory stimulus regulated ($|\log_2$ fold change) > 1 and FDR < 0.05) for 1h versus 0h (left) and 4h versus 0h (right) comparisons. FC, fold change. Colored dots represent sensory-stimulus-regulated genes. The analysis was performed independently across 30 cell types (Supplementary Fig. 12). **b**, Heat map of all 611 stimulus-regulated genes grouped into ERGs and LRGs by cell type. Exc, excitatory neurons; int, interneurons; olig, oligodendrocytes; endo/SM, endothelium, smooth muscle; micro, microglia. Each horizontal black line represents a stimulus-regulated gene. **c**, Estimation of the percentage of cells with stimulus-regulated transcriptional changes (Methods) at 1h (colored bars) versus 0h (gray bars). We defined induction as requiring two, three, or four induced genes within each cell from a cell-type-specific set; these requirements are represented by the lower, central, and upper lines of the boxes, respectively (Methods).

light-induced transcriptional changes, irrespective of their type or laminar position. In contrast, inhibitory neuronal subtypes showed variable fractions of transcriptionally responsive cells, ranging from

71% in the case of a somatostatin (Sst)-expressing subpopulation (Int_Sst_2) to only 29% of parvalbumin-expressing interneurons (Int_Pv).

Remarkably, a large fraction of non-neuronal cells also exhibited acute light-induced transcriptional responses: endothelial and smooth muscle cell types showed the largest fraction of transcriptionally active cells (49–68%) and were followed by astrocytes (47%), pericytes (38%), and macrophages (38–47%). In contrast, a small fraction of oligodendrocytes and microglia showed light-induced transcriptional responses (4–37% and 11–12%, respectively). Light exposure after housing in the dark thus triggered acute transcriptional responses in a large fraction of cells across the full range of V1 cell types, including an unexpectedly high proportion of non-neuronal cell classes.

Diversity of experience-regulated ERGs. One prevalent model for stimulus-dependent gene expression in the brain suggests that neuronal activity triggers several calcium-dependent signaling pathways, including the Ras/MAPK pathway, which drive the ubiquitous expression of a common cohort of ERGs across many cell types²⁴. Many of these ERGs encode transcription factors (TFs) that have been hypothesized to trigger unique patterns of LRG expression through differential binding to cell-type-specific cis regulatory elements^{13,17}. This hypothesis, however, is based on findings from a limited number of cell types. In an alternative model, the induction of ERTFs themselves is at least somewhat cell-type specific and leads to the activation of distinct sets of LRGs in different cell types.

We tested these two models by examining the ERTF and LRG programs across cell types. In agreement with findings from previous reports, LRGs were shared across fewer cell types than were ERTFs (Mann–Whitney U test, $P=2 \times 10^{-3}$; Fig. 3a). However, of the 38 ERTFs identified, only 19 were induced in three or more cell types, thus suggesting considerable divergence within the early stimulus-responsive gene expression program. To gain further insight into the 19 shared ERTFs, we classified them according to their expression patterns across the 30 cell types and identified four distinct gene sets (Fig. 3b). Two of these gene sets showed sensory-experience-induced expression across most neuronal and non-neuronal cell types (between 3 and 14 cell types). These sets contained canonical immediate-early genes (IEGs) known to regulate the late phases of gene expression (e.g., *Nr4a1*, *Nr4a2*, *Nr4a3*, *Fos*, *Fosl2*, and *Egr1*). We confirmed the induction of a subset of these IEGs across multiple cell types (FISH visual presentation and quantification in Fig. 3c and Supplementary Fig. 13). In contrast, the final two gene sets were specific to either neuronal (*Egr2*, *Egr4*, *Fosb*, *Junb*, and *Npas4*) or non-neuronal cell populations (*Atf3*, *Klf2*, *Klf4*, *Klf10*, and *Maff*).

We investigated the extent to which these ERTFs were coexpressed within individual cells of a given type to determine whether their coexpression was stochastic or reflected a deeper structure of ERTF regulation. To this end, we correlated the expression of ERTFs across individual excitatory neurons, focusing on this cell population because it comprised the largest cohort of transcriptionally responsive cells. We found a significantly higher pairwise correlation between ERTFs within individual cells (Pearson $r=0.23 \pm 0.13$) than between ERTFs shuffled across cells or between expression-matched noninduced genes (Pearson $r=0.002 \pm 0.017$, $r=0.002 \pm 0.029$, $P=0$, $P=0$, Mann–Whitney U test, respectively) (Methods and Fig. 3d, e). Three-color FISH confirmed these conclusions for *Fos* and *Egr1* or *Fos* and *Nr4a1* in excitatory neurons (*Vglut1*⁺; official symbol *Slc17a7*), which demonstrated a high degree of pairwise correlation (Pearson $r=0.74$ and 0.76 , respectively) within individual excitatory neurons (Fig. 3f).

Together, our findings provide a nuanced view of the diversity of experience-regulated gene expression across cell types. Although our data confirm the existence of a core set of ERTFs induced in

both neuronal and non-neuronal cell types, we also found a substantial divergence in ERTF expression that may contribute to LRG diversity among cell types. Moreover, our results suggest that, within individual excitatory neurons, the expression of ERTFs is tightly co-regulated, thus allowing these factors to act in concert in regulating LRG programs.

Excitatory neuronal LRGs. Excitatory neurons throughout cortical laminae are molecularly, physiologically, and hodologically diverse²⁵; however, the diversity of LRG programs among excitatory neuronal subtypes remains to be systematically explored. We identified 19 cell-type-enriched LRGs (of 55 LRGs induced in excitatory neurons), including several secreted modulators of synaptic plasticity¹⁵. For example, *Cbln4*, which encodes cerebellin 4, a secreted factor implicated in inhibitory synapse recruitment²⁶, was enriched in layer 4 excitatory neurons, as confirmed by FISH (Fig. 4a). GO analysis of excitatory LRGs revealed an overrepresentation of secreted factors and genes regulating synapse formation (Supplementary Table 4), in agreement with experience-dependent transcriptional regulation of neuronal connectivity. These findings suggest that long-term adaptations to visual stimuli differ across excitatory neuronal subtypes in the visual cortex and may enable distinct functions in visual processing.

We sought to assess intralayer differences in sensory-experience-dependent transcriptional responses and focused first on layer 5 excitatory neurons, which have been classified into subtypes defined according to their axon-projection patterns and electrophysiological properties^{23,25,27}. We mapped three putative layer 5 excitatory neuron populations onto previously defined transcriptional cell types from the mouse visual cortex²³ (Supplementary Figs. 11 and 14). The three populations of excitatory neurons respond to sensory stimulus by activating a similar number of genes (34–55 genes per cell type). However, despite the anatomical proximity of these neuronal populations, several of their stimulus-dependent genes were subset enriched (Fig. 4b). For example, stimulus-dependent induction of *Pdlim1*, which encodes a protein involved in AMPA-receptor trafficking²⁸, was restricted to predicted corticofugal-projecting *ExcL5_2* neurons. Thus, even within a single cortical layer, functionally heterogeneous cell types exhibited distinct experience-dependent transcriptional responses.

We next asked whether other cortical laminae had subpopulations that transcriptionally diverged after exposure to a sensory stimulus. Further subclustering of excitatory populations revealed two layer 2/3 and three layer 4 excitatory neuron subtypes, in agreement with previous transcriptionally defined cell types (Methods, Fig. 4c,d and Supplementary Figs. 15–18). We found that genes both regulated and not regulated by sensory experience were distinct across these cell types. Noninduced gene markers included *Cdh13* for layer 2/3 and *Ctxn3*, *Calb1*, and *Hsd11b1* for layer 4 (Fig. 4c,d and Supplementary Figs. 19 and 20). These layer 2/3 and layer 4 excitatory neuronal cell types displayed marked differences in experience-regulated transcriptional responses (Fig. 4e). For example, *Cbln4* was differentially expressed across layer 4 subtypes, and the largest light-stimulus induction occurred within the *Calb1*⁺ subtype (Fig. 4f), as corroborated via FISH (Fig. 4g).

Finally, we asked whether these cell types exhibiting distinct experience-induced transcriptional responses were further organized into specific spatial arrangements within their respective layers. FISH revealed that the transcriptionally less responsive *Cdh13*⁺ excitatory cells were enriched superficially in layer 2/3 (Supplementary Fig. 21), in agreement with previous observations of sparser firing in layer 2 than layer 3 in response to sensory stimuli^{25,29,30}. In layer 4, we first examined the two most transcriptionally responsive layer 4 subpopulations, *Calb1*⁺ and *Hsd11b1*⁺ neurons (Fig. 4e), which differed in 44 of the 90 stimulus-regulated genes identified within these subtypes. Whereas *Calb1*⁺ neurons were

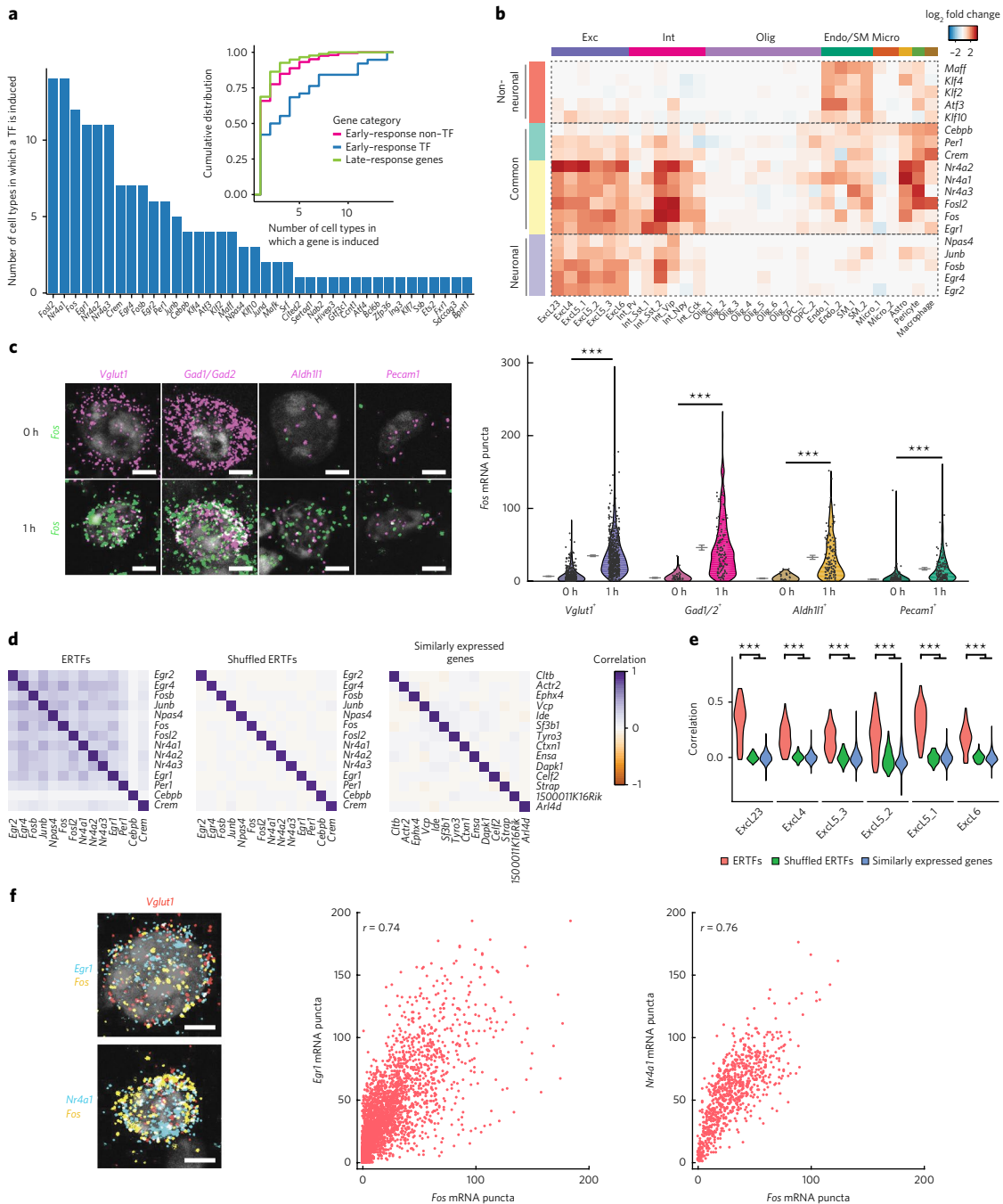


Fig. 3 | Diversity of experience-regulated ERGs. **a**, Number of cell types in which each ERTF is sensory-experience regulated. Inset, cumulative distribution of the number of cell types in which each gene is experience regulated. Stimulus-regulated LRGs were shared across fewer cell types than were ERTFs ($P=2 \times 10^{-5}$, Mann-Whitney U test, two-sided, $n_{\text{ERTFs}}=38$ genes, $n_{\text{LRGs}}=176$ genes, $n_{\text{ERnon-TF}}=205$ genes). **b**, Heat map of \log_2 fold changes between 1h and 0h of the 19 ERTFs shared across at least three cell types. ERTFs are hierarchically clustered into four groups on the basis of gene expression across all cell types. **c**, Left, representative FISH images of *Fos* indicating induction across several cell types. Right, quantification of FISH signals for *Fos* expression across cell types. A random subset (25%) of the raw data is plotted to aid in visualization. For *Vglut1*, $n_{0h}=1,616$ cells, $n_{1h}=2,324$ cells; for *Gad1* and *Gad2*, $n_{0h}=156$ cells, $n_{1h}=497$ cells; for *Aldh1l1*, $n_{0h}=121$ cells, $n_{1h}=477$ cells; for *Pecam1*, $n_{0h}=542$ cells, $n_{1h}=510$ cells. $***P < 10^{-39}$, Mann-Whitney U test, two-sided. Means and 95% confidence intervals are denoted by gray lines. Experiments were repeated on two to eight cortical slices per time point. **d**, Mean pairwise Pearson correlations across individual excitatory neurons at 1h, calculated on the basis of ERTF expression (left, $r=0.23 \pm 0.13$ (mean \pm s.d.)), shuffled ERTF expression (middle, $r=0.002 \pm 0.017$), and expression-matched noninduced genes (right, $r=0.002 \pm 0.029$), $n=91$ pairwise comparisons. Correlations between ERTFs are significantly higher than those between either shuffled ERTFs or expression-matched noninduced genes ($P=0$ and $P=0$, respectively, Mann-Whitney U test, two-sided). **e**, Distributions of pairwise Pearson correlations of ERTFs ($n=91$ pairwise comparisons), ERTFs with shuffled expression, and expression-matched noninduced genes across excitatory neuronal subtypes at 1h. ERTFs are more highly correlated than either shuffled ERTFs or similarly expressed noninduced genes. $***P < 1 \times 10^{-24}$, Mann-Whitney U test, two-sided. **f**, Left, three-color FISH images of *Fos*, *Egr1*, and *Nr4a1* expression at 1h after light stimulation in excitatory neurons (*Vglut1*⁺). Middle and right, quantification of FISH signals. Scatter plots between *Egr1* and *Fos*, and *Nr4a1* and *Fos* coexpression. Expression was highly correlated (Pearson $r=0.74$ and 0.76 , respectively; $n=3,162$ and 778 cells, respectively). Scale bars, $5 \mu\text{m}$. Data were collected from four cortical slices for *Egr1* and *Fos* coexpression, and one cortical slice for *Nr4a1* and *Fos* coexpression.

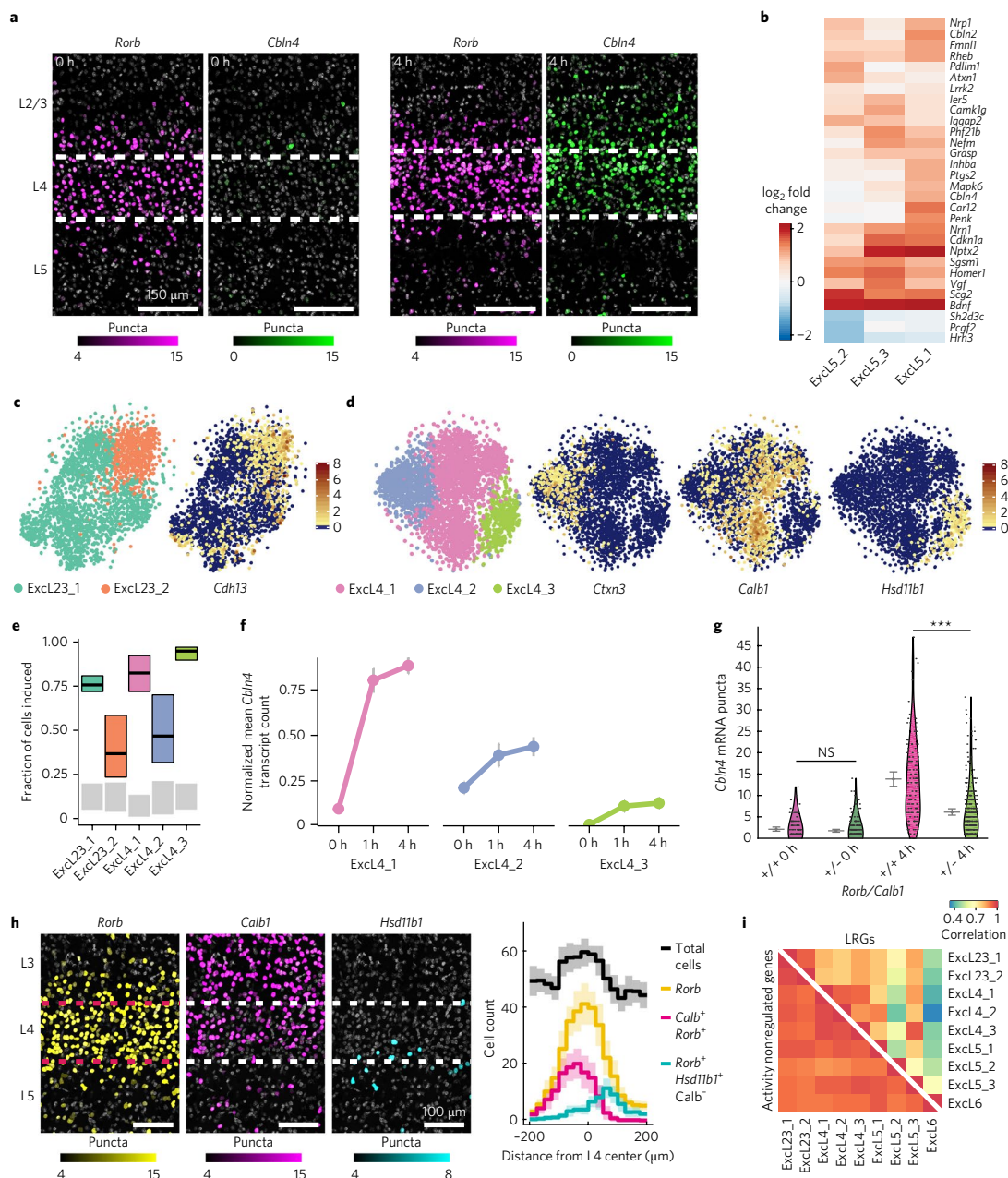


Fig. 4 | Sensory-experience-induced transcriptional responses in V1 excitatory neurons. **a**, FISH of *Cbln4* with *Rorb*, a marker specific for layer (L) 4, at 0 h (control) and 4 h after light stimulus. Nuclei are pseudocolored according to *Cbln4* or *Rorb* expression (Methods). Scale bars, 150 μ m. Experiments were repeated on two cortical slices per condition. **b**, Heat map of \log_2 fold change in expression of LRGs across layer 5 subtypes between 4 h and 0 h. **c**, Left, t-SNE plot showing layer 2/3 excitatory neuronal subtypes. Right, overlay of *Cdh13* expression. Scale indicates normalized expression per cell; $n = 2,941$ cells. **d**, Left, t-SNE plot of layer 4 excitatory neuronal subtypes. Right, overlay of selected marker expression. Scale indicates normalized expression per cell; $n = 3,198$ cells. **e**, Estimation of the percentage of cells with stimulus-regulated transcriptional changes (Methods) at 1 h (colored bars) versus 0 h (gray bars). We defined induction as requiring two, three, or four induced genes within each cell from a cell-type-specific set; these requirements are represented by the lower, central, and upper lines of the box, respectively (Methods). **f**, Mean expression of *Cbln4* across layer 4 subtypes ($n_{\text{Exc4}_1} = 732, 468, \text{ and } 783$ cells; $n_{\text{Exc4}_2} = 343, 214, \text{ and } 210$ cells; $n_{\text{Exc4}_3} = 136, 137, \text{ and } 175$ cells for 0 h, 1 h, and 4 h, respectively). Error bars, s.e.m. **g**, FISH quantification of *Cbln4* across layer 4 subtypes (*Rorb*⁺*Calb1*⁺ (+/+)) and *Rorb*⁺*Calb1*⁻ (+/-). 0 h expression was not significantly different across cell types, $n = 97$ (*Rorb*⁺*Calb1*⁺) or 209 (*Rorb*⁺*Calb1*⁻) cells, $P = 0.062$, Mann-Whitney *U* test, two-sided. The 4-h expression was significantly higher in *Rorb*⁺*Calb1*⁺ than *Rorb*⁺*Calb1*⁻ ($n = 127$ or 286 cells, respectively), $***P < 10^{-17}$, Mann-Whitney *U* test, two-sided. Means and 95% confidence intervals are denoted by gray lines. **h**, Left, *Rorb* (yellow), *Calb1* (magenta), and *Hsd11b1* (cyan) expression in layer 4, as determined by FISH. Scale bars, 100 μ m. Right, quantification of the anatomical distribution of FISH-defined cell types across layer 4 from $n = 13$ cortical slices. Negative/positive values on the x axis correspond to the distance from the center of layer 4 toward the slice surface (negative) or toward deeper cortical layers (positive). The shaded areas around lines indicate 95% confidence intervals around the means. The *Rorb*⁺*Calb1*⁺ ($n = 1,626$ cells) population is enriched superficially within layer 4, as defined by *Rorb* positivity ($n = 4,066$ cells) and an increase in cell density (black). The *Hsd11b1*⁺*Rorb*⁺*Calb1*⁻ ($n = 851$ cells) population is enriched in ventral layer 4 and into superficial layer 5. Further analysis can be found in Supplementary Fig. 22. **i**, Top right, pairwise Pearson correlation between excitatory subtypes calculated on the basis of LRG expression ($n = 55$ genes). Bottom left, same analysis with expression-matched non-stimulus-regulated genes ($n = 55$ genes). Non-stimulus-regulated genes are more correlated across excitatory subtypes than are stimulus-regulated LRGs.

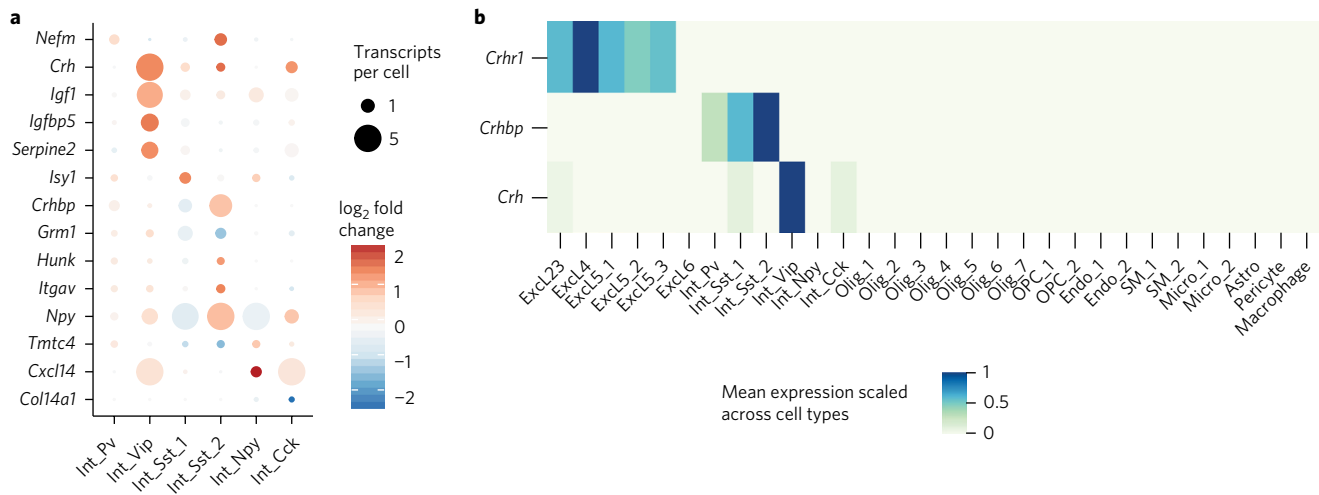


Fig. 5 | Inhibitory neuronal LRGs. a, Cell-type-enriched LRGs across inhibitory cell types. Transcripts per cell represent the mean depth-normalized expression across all cells. Fold change is calculated between 4 h and 0 h. **b**, Scaled mean expression of *Crh*, *Crhbp*, and *Crhr1* across all cell types.

enriched superficially, the *Hsd11b1*⁺ neuronal population was found deep within layer 4 and spanned the boundary between layers 4 and 5 (Fig. 4h and Supplementary Fig. 22). *Ctxn3*⁺ neurons, the least transcriptionally responsive population, were also enriched within superficial layer 4 and were interspersed with the *Calb1*⁺ population (Supplementary Fig. 22). The anatomical organization of these cell types into sublayers, coupled with divergent transcriptional responses to a sensory stimulus, suggests previously unappreciated functional subdivisions located within the laminae of the mouse visual cortex, resembling the cytoarchitecture in higher mammals³¹.

In summary, we observed distinct transcriptional responses both among cortical layers and among neuronal subtypes within individual layers in response to light stimulation. Moreover, we found that the LRG programs activated in response to visual stimulation were highly cell type specific and probably contributed to the cellular features defining the function of each subtype within the circuit (Fig. 4i).

Inhibitory neuronal LRGs. V1 interneurons were classified into six distinct subtypes (Int_Pv, Int_Vip, Int_Cck, Int_Npy, Int_Sst_1, and Int_Sst_2), in agreement with other scRNA-seq-based taxonomies^{23,32} (Fig. 1f and Supplementary Figs. 11 and 23). Analysis of gene expression regulated by sensory stimuli in these cell types showed broad agreement with our prior findings based on cell-type-specific isolation of ribosome-associated mRNA^{13,17}. Specifically, we identified 75 ERGs and 70 LRGs whose expression was light dependent in at least one of the six inhibitory subtypes, and 14 LRGs enriched in a single inhibitory subtype (Fig. 5a).

Of particular interest, in the Int_Vip population we observed selective induction of *Crh* mRNA, which encodes the stress hormone corticotropin-releasing hormone. In the hippocampus, *Crh* signals through the G-protein-coupled receptor *Crhr1* and consequently increases the excitability of pyramidal cells³³. *Crhr1* expression in V1 was enriched in excitatory cell types, thus suggesting that Vip-interneuron-derived *Crh* might directly modulate the excitability of pyramidal neurons in a stimulus-dependent manner (Fig. 5b). Moreover, building upon a recent study in the prefrontal cortex³⁴, in Sst-expressing interneurons we observed stimulus-dependent induction of *Crhbp*, a gene encoding *Crh*-binding protein, a secreted factor that negatively regulates *Crh* signaling. Together, these findings suggest a possible mechanism for the control of cortical microcircuit excitability involving the opposing action of two

activity-regulated signaling peptides derived from distinct inhibitory subtypes.

Experience-dependent transcriptional changes in vasculature-associated cells. Since the discovery over a century ago that neuronal activity rapidly triggers changes in local blood flow³⁵, researchers have recognized that neurons, glia, and associated vasculature (endothelial cells, smooth muscle cells, and pericytes) coordinate their signaling and activity³⁶. In addition, sensory-experience-related neuronal activity restructures cortical vascular networks, although the mechanisms regulating this process remain unknown¹⁸.

We asked whether the vascular transcriptional response was induced by local neuronal activity or by systemic brain-wide changes in blood flow or oxygen levels. To address this question, we carried out qRT-PCR for the endothelial and smooth-muscle-specific ERTF *Klf4* across multiple cortical regions. We found that *Klf4* was induced in V1 but not in other regions of the cortex, thus indicating that the sensory-stimulus-dependent changes in vascular architecture occurred specifically in the region of the brain that is activated by the sensory stimulus (Fig. 6a).

We unexpectedly identified a larger number of activity-regulated genes (257) in endothelial and smooth muscle cells than in excitatory and inhibitory neurons combined (231; Fig. 6b). The large number of induced genes in endothelial and smooth muscle cells suggested that experience-dependent transcriptional programs play an important role in the vascular response. For example, *Angpt2*, which encodes angiotensin 2, a canonical vascular growth factor involved in permeability of the blood-brain barrier and in embryonic and adult angiogenesis^{37,38}, was significantly induced more than twofold in Endo_1 and SM_2 cell types within 4 h of light stimulation (Fig. 6c).

Beyond finding several broadly induced genes, we identified 12 sensory-regulated transcription factors (*Klf2*, *Klf4*, *Klf7*, *Atf3*, *Atf4*, *Bcl6b*, *Jund*, *Maff*, *Mafk*, *Srf*, *Zfp36*, and *Ing3*) that were selectively induced in endothelial and smooth muscle cells but not in neuronal cell types (Fig. 6d–f). A subset of these TFs have previously been implicated in structural changes to brain vasculature³⁹. For example, aberrant overexpression of *Klf2* and *Klf4* in endothelial cells causes cerebral cavernous malformations⁴⁰. Thus, the induction of these ERTFs is likely to be crucial for vascular structural remodeling in response to sensory stimuli.

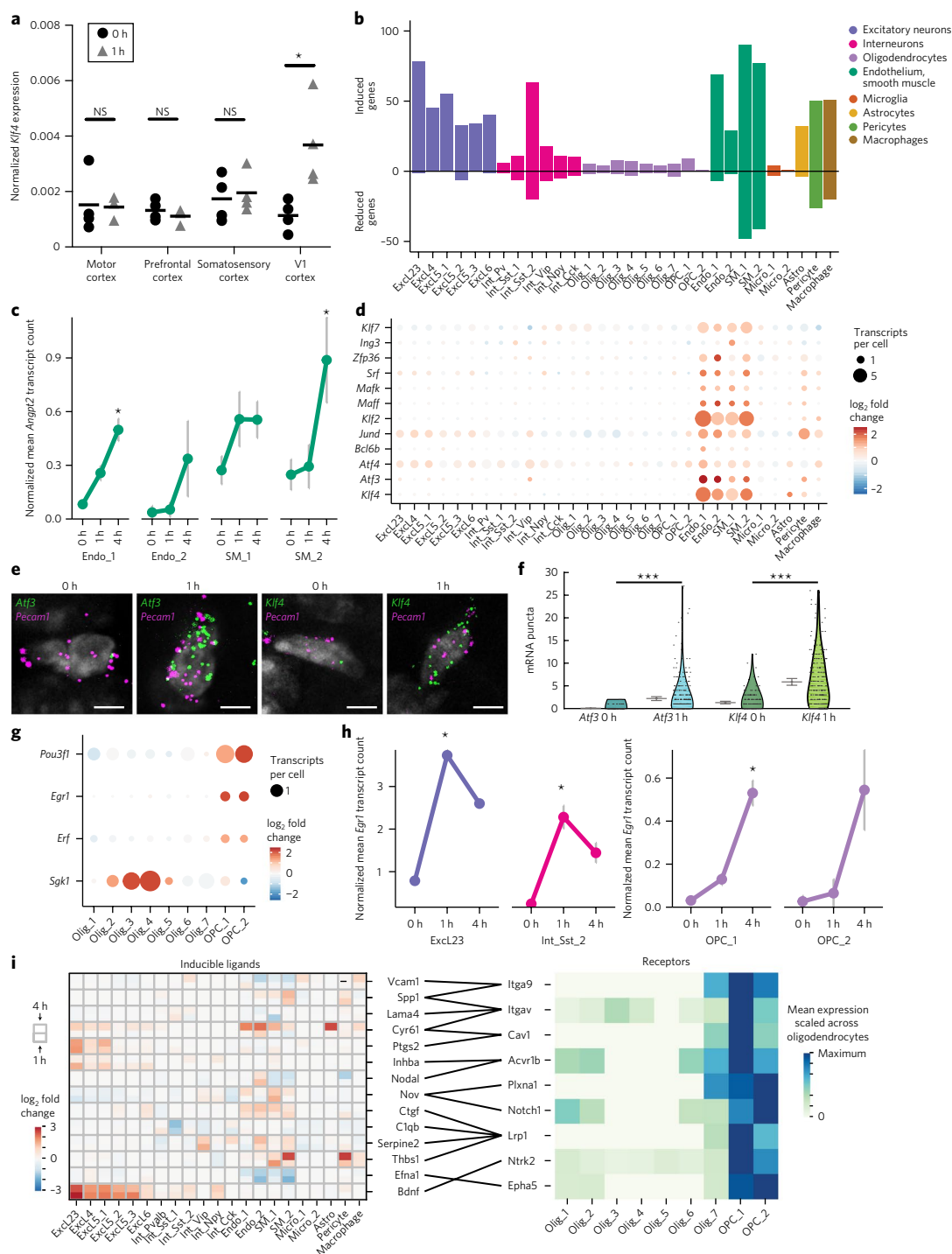


Fig. 6 | Light-induced transcriptional changes in non-neuronal cells. **a**, qRT-PCR of *Klf4* across cortical regions in mice exposed to light for 0 h and 1 h. Horizontal bars, means. *Klf4* induction is significant only in V1; $P = 0.02$, unpaired t test, two-sided. $n = 4$ mice for all 0-h samples; $n = 3$ mice for 1-h motor and prefrontal cortex; $n = 4$ mice for 1-h somatosensory and visual cortex. **b**, Total number of sensory-stimulus-regulated genes across cell types. **c**, Mean expression of the LRG *Angpt2* across endothelial and smooth muscle cells across all three stimulus conditions ($n_{\text{ENDO}_1} = 1,111, 1,107$, and $1,109$ cells; $n_{\text{ENDO}_2} = 33, 65$, and 25 cells; $n_{\text{SM}_1} = 115, 92$, and 116 cells; $n_{\text{SM}_2} = 104, 109$, and 85 cells for 0 h, 1 h, and 4 h, respectively). *FDR < 0.05, fold change > 2, Monocle2. Error bars, s.e.m. **d**, Induction at 1 h of ERTFs enriched in vasculature-associated cells. **e**, FISH of single *Pecam1*⁺ cells at 0 h and 1 h, colabeled with *Atf3* and *Klf4*. Scale bars, 5 μm . Experiments were repeated on two cortical slices per time point. **f**, Quantification of FISH signals for *Atf3* and *Klf4* in *Pecam1*⁺ cells across 0 h and 1 h (for *Atf3*, $n_{0\text{h}} = 275$ cells and $n_{1\text{h}} = 326$ cells; for *Klf4*, $n_{0\text{h}} = 182$ cells and $n_{1\text{h}} = 224$ cells). $P < 10^{-22}$, Mann-Whitney U test, two-sided. Means and 95% confidence intervals are denoted by gray lines. Data are from two cortical slices per time point. **g**, Fold change in expression between 0-h and 4-h samples of light-induced genes in oligodendrocytes and OPCs. **h**, Mean expression of *Egr1* in Excl23, Int_Sst_2, and OPCs across all stimulus conditions ($n_{\text{Excl23}} = 150, 706$, and $1,107$ cells; $n_{\text{Int_Sst}_2} = 59, 52$, and 70 cells; $n_{\text{OPC}_1} = 645, 472$, and 608 cells; $n_{\text{OPC}_2} = 31, 27$, and 43 cells at 0 h, 1 h, and 4 h, respectively). *FDR < 0.05, fold change > 2, Monocle2. Error bars, s.e.m. **i**, Left, heat map of induction of ligands at both 1 h and 4 h in nonoligodendrocyte cell types with receptors enriched in OPCs. Right, scaled mean expression of OPC-enriched receptors. Lines between heat maps denote ligand-receptor pairs.

Experience-dependent transcriptional changes in oligodendrocytes. Although neuronal activity is known to regulate oligodendrogenesis and myelination, the transcriptional pathways underlying these processes remain unclear^{19,41}. Oligodendrocytes displayed a weak transcriptional response to a light stimulus, and only 33 differentially expressed genes were identified, despite the abundance of these cells in our dataset (10,158 cells). Nevertheless, in multiple oligodendrocyte populations, we did observe several stimulus-induced transcriptional changes of interest, including the induction of *Sgk1*, which encodes serum/glucocorticoid-regulated kinase (Fig. 6g). *Sgk1* overexpression enhances the arborization of oligodendrocytes⁴², thus suggesting that sensory-experience-dependent induction of *Sgk1* may lead to oligodendrocyte remodeling.

We observed the induction of several TF-encoding genes (*Egr1*, *Pou3f1*, and *Erf*) specifically in OPCs and not in immature, premyelinating, or mature oligodendrocytes, thus suggesting that these TFs may play a role in oligodendrogenesis (Fig. 6g). Interestingly, for *Erf* and *Egr1* in OPCs, in contrast to their dynamics in other cell types, we observed a delayed induction that may potentially depend on prior activation of activity-regulated secreted proteins in neurons or other non-neuronal cells (Fig. 6h). Indeed, we found nine receptors⁴³ that were both enriched in OPCs relative to other oligodendrocytes and predicted to bind stimulus-responsive ligands induced in neurons and vasculature-associated cells, including those encoded by *Bdnf*, *Ptgs2*, and *Inhba* (Fig. 6i and Methods). Together, these findings identify sensory-experience-responsive genes that might regulate OPC proliferation, differentiation, and recruitment to neuronal axons.

Discussion

The cerebral cortex is composed of a vast array of cell types whose coordinated activity is necessary for normal brain development and function. Several laboratories have recently applied scRNA-seq to characterize these cell types across brain areas^{8,22,23,32,44–46}; however, that work has largely focused on producing a static cellular taxonomy without revealing the transcriptional changes induced by acute sensory experience and neuronal activity. Moreover, the enzymatic brain-dissociation methods applied in these previous studies are likely to activate stimulus-regulated transcriptional programs and thereby obscure the identification of genes induced in the brain in response to sensory stimuli²². The resource presented here, in contrast, represents an additional dimension of cellular state that must be considered in examining existing cell-type atlases and serves as a critical step toward a holistic understanding of the mechanisms by which experience modulates cortical circuits through the concerted action of diverse cell types.

The degree to which transcriptional responses vary across the panoply of experiences encountered throughout life remains unknown. We restricted our analysis to a single stimulus and cortical region. Intriguingly, despite differences in brain region, stimulus paradigm, and methodology, our 611 experience-regulated genes are highly enriched among those reported by a recent study of activity-dependent chromatin accessibility and bulk gene expression in the dentate gyrus⁴⁷ (Supplementary Fig. 24). Nevertheless, similar analyses of dynamic gene expression in other brain regions in response to a multitude of stimuli will be essential in constructing a complete cell atlas of the central nervous system.

Although droplet-based scRNA-seq has substantial advantages enabling high-throughput and cost-effective analysis of the transcriptomes of thousands of cells, several key limitations remain. First, the anatomical locations of individual cells are lost, thus necessitating the use of FISH or other spatially precise methods to physically map transcriptionally defined cell types. Second, the low transcript-capture efficiency²¹ and shallow depth of sequencing (Supplementary Figs. 3, 5, 9 and 10) resolve only a fraction of transcripts within each cell, thereby requiring large numbers of cells to be analyzed to differentiate cell types. For example, targeted deep

sequencing of inhibitory neurons²³ identified 23 subtypes that were fine-grained subsets of the types described here (Supplementary Figs. 17 and 23). Consequently, sequencing even more cells might uncover additional rare cell types and reveal further heterogeneity of stimulus-dependent transcriptional programs within the mouse visual cortex.

Through the analysis of stimulus-regulated genes, we identified excitatory neuronal cell types that dramatically differed in their response to light stimulus both between and within cortical laminae. Closer examination revealed that cell types residing in the same cortical layer had differentially evoked transcriptional responses and were organized in discrete laminar positions within their respective layers. Thus, despite their proximity, these cell types may have different synaptic connectivity and may serve distinct functions within cortical circuitry. Future studies of these cell types could test whether cells displaying unique experience-dependent transcriptional responses have distinct physiological properties and, within thalamorecipient layer 4, whether these cells receive different thalamocortical inputs, as has been observed in primates⁵¹.

Non-neuronal cells have functions essential to the maintenance of cortical circuitry including regulating synapse development and maturation, neurotransmitter reuptake, and metabolite and oxygen supply. Here, we revealed that, in parallel to the experience-dependent transcriptional response in neurons, a distinct and variable gene network is induced in glial, immune, and vascular cell types. The functions of many of these regulated genes are unknown but are likely to contribute to regulating previously identified activity-dependent processes in non-neuronal cells, including remodeling of vascular networks and structural remodeling of oligodendrocytes. Future studies are necessary to determine how these transcriptional programs act in concert to remodel cortical function and circuitry in response to changes in sensory experience.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41593-017-0029-5>.

Received: 9 March 2017; Accepted: 17 October 2017;
Published online: 11 December 2017

References

- Hensch, T. K. Critical period plasticity in local cortical circuits. *Nat. Rev. Neurosci.* **6**, 877–888 (2005).
- Wiesel, T. N. & Hubel, D. H. Single-cell responses in striate cortex of kittens deprived of vision in one eye. *J. Neurophysiol.* **26**, 1003–1017 (1963).
- Zucker, R. S. & Regehr, W. G. Short-term synaptic plasticity. *Annu. Rev. Physiol.* **64**, 355–405 (2002).
- West, A. E. & Greenberg, M. E. Neuronal activity-regulated gene transcription in synapse development and cognitive function. *Cold Spring Harb. Perspect. Biol.* **3**, a005744 (2011).
- Bading, H. Nuclear calcium signalling in the regulation of brain function. *Nat. Rev. Neurosci.* **14**, 593–608 (2013).
- MacManus, J. P. et al. Translation-state analysis of gene expression in mouse brain after focal ischemia. *J. Cereb. Blood Flow Metab.* **24**, 657–667 (2004).
- Ginty, D. D. et al. Regulation of CREB phosphorylation in the suprachiasmatic nucleus by light and a circadian clock. *Science* **260**, 238–241 (1993).
- Campbell, J. N. et al. A molecular census of arcuate hypothalamus and median eminence cell types. *Nat. Neurosci.* **20**, 484–496 (2017).
- Renier, N. et al. Mapping of brain activity by automated volume analysis of immediate early genes. *Cell* **165**, 1789–1802 (2016).
- Ben-David, E. & Shifman, S. Networks of neuronal genes affected by common and rare variants in autism spectrum disorders. *PLoS Genet.* **8**, e1002556 (2012).
- Thomas, G. M. & Huganir, R. L. MAPK cascade signalling and synaptic plasticity. *Nat. Rev. Neurosci.* **5**, 173–183 (2004).
- Shepherd, J. D. & Bear, M. F. New views of Arc, a master regulator of synaptic plasticity. *Nat. Neurosci.* **14**, 279–284 (2011).
- Mardinly, A. R. et al. Sensory experience regulates cortical inhibition by inducing IGF1 in VIP neurons. *Nature* **531**, 371–375 (2016).

14. Harward, S. C. et al. Autocrine BDNF-TrkB signalling within a single dendritic spine. *Nature* **538**, 99–103 (2016).
15. Lin, Y. et al. Activity-dependent regulation of inhibitory synapse development by Npas4. *Nature* **455**, 1198–1204 (2008).
16. Majdan, M. & Shatz, C. J. Effects of visual activity experience on activity-dependent gene regulation in cortex. *Nat. Neurosci.* **9**, 650–659 (2006).
17. Spiegel, I. et al. Npas4 regulates excitatory-inhibitory balance within neural circuits through cell-type-specific gene programs. *Cell* **157**, 1216–1229 (2014).
18. Lacoste, B. et al. Sensory-related neural activity regulates the structure of vascular networks in the cerebral cortex. *Neuron* **83**, 1117–1130 (2014).
19. Gibson, E. M. et al. Neuronal activity promotes oligodendrogenesis and adaptive myelination in the mammalian brain. *Science* **344**, 1252304 (2014).
20. Wang, X. et al. Astrocytic Ca²⁺ signaling evoked by sensory stimulation *in vivo*. *Nat. Neurosci.* **9**, 816–823 (2006).
21. Klein, A. M. et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
22. Lacar, B. et al. Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat. Commun.* **7**, 11022 (2016).
23. Tasic, B. et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016).
24. Wu, G.-Y., Deisseroth, K. & Tsien, R. W. Activity-dependent CREB phosphorylation: convergence of a fast, sensitive calmodulin kinase pathway and a slow, less sensitive mitogen-activated protein kinase pathway. *Proc. Natl. Acad. Sci. USA* **98**, 2808–2813 (2001).
25. Harris, K. D. & Shepherd, G. M. G. The neocortical circuit: themes and variations. *Nat. Neurosci.* **18**, 170–181 (2015).
26. Chacón, P. J. et al. Cerebellin 4, a synaptic protein, enhances inhibitory activity and resistance of neurons to amyloid- β toxicity. *Neurobiol. Aging* **36**, 1057–1071 (2015).
27. Hattox, A. M. & Nelson, S. B. Layer V neurons in mouse cortex projecting to different targets have distinct physiological properties. *J. Neurophysiol.* **98**, 3330–3340 (2007).
28. Schulz, T. W. et al. Actin/alpha-actinin-dependent transport of AMPA receptors in dendritic spines: role of the PDZ-LIM protein RIL. *J. Neurosci.* **24**, 8584–8594 (2004).
29. Niell, C. M. & Stryker, M. P. Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* **28**, 7520–7536 (2008).
30. de Kock, C. P. J., Bruno, R. M., Spors, H. & Sakmann, B. Layer- and cell-type-specific suprathreshold stimulus representation in rat primary somatosensory cortex. *J. Physiol. (Lond.)* **581**, 139–154 (2007).
31. Nassi, J. J. & Callaway, E. M. Parallel processing strategies of the primate visual system. *Nat. Rev. Neurosci.* **10**, 360–372 (2009).
32. Zeisel, A. et al. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
33. Kratzer, S. et al. Activation of CRH receptor type 1 expressed on glutamatergic neurons increases excitability of CA1 pyramidal neurons by the modulation of voltage-gated ion channels. *Front. Cell. Neurosci.* **7**, 91 (2013).
34. Li, K., Nakajima, M., Ibañez-Tallon, I. & Heintz, N. A cortical circuit for sexually dimorphic oxytocin-dependent anxiety behaviors. *Cell* **167**, 60–72.e11 (2016).
35. Roy, C. S. & Sherrington, C. S. On the regulation of the blood-supply of the brain. *J. Physiol.* **11**, 158–158.17 (1890).
36. Attwell, D. et al. Glial and neuronal control of brain blood flow. *Nature* **468**, 232–243 (2010).
37. Gurnik, S. et al. Angiotensin-2-induced blood-brain barrier compromise and increased stroke size are rescued by VE-PTP-dependent restoration of Tie2 signaling. *Acta Neuropathol.* **131**, 753–773 (2016).
38. Augustin, H. G., Koh, G. Y., Thurston, G. & Alitalo, K. Control of vascular morphogenesis and homeostasis through the angiotensin-Tie system. *Nat. Rev. Mol. Cell Biol.* **10**, 165–177 (2009).
39. Weinl, C. et al. Endothelial SRF/MRTF ablation causes vascular disease phenotypes in murine retinae. *J. Clin. Invest.* **123**, 2193–2206 (2013).
40. Zhou, Z. et al. Cerebral cavernous malformations arise from endothelial gain of MEKK3-KLF2/4 signalling. *Nature* **532**, 122–126 (2016).
41. Hines, J. H., Ravanelli, A. M., Schwindt, R., Scott, E. K. & Appel, B. Neuronal activity biases axon selection for myelination *in vivo*. *Nat. Neurosci.* **18**, 683–689 (2015).
42. Miyata, S., Hattori, T., Shimizu, S., Ito, A. & Tohyama, M. Disturbance of oligodendrocyte function plays a key role in the pathogenesis of schizophrenia and major depressive disorder. *BioMed Res. Int.* **2015**, 492367 (2015).
43. Ramilowski, J. A. et al. A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat. Commun.* **6**, 7866 (2015).
44. Habib, N. et al. Div-Seq: single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science* **353**, 925–928 (2016).
45. Marques, S. et al. Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. *Science* **352**, 1326–1329 (2016).
46. Gokce, O. et al. Cellular taxonomy of the mouse striatum as revealed by single-cell RNA-seq. *Cell Rep.* **16**, 1126–1137 (2016).
47. Su, Y. et al. Neuronal activity modifies the chromatin accessibility landscape in the adult brain. *Nat. Neurosci.* **20**, 476–483 (2017).

Acknowledgements

We thank E. Griffith and J. Gray for critical reading of the manuscript; the HMS Single Cell Core for sample processing; and D. Harmin, A. Veres, D. Kotliar, H. Finucane, Y. Reshef, B. Hall, T. Otis, and D. Malhotra for discussions. This work was supported by funding from National Institutes of Health (NIH) grant R01 NS028829 and the ROADS Program funded by F. Hoffmann-La Roche Ltd. to M.E.G.; NIH grant R01 NS046579 to B.L.S.; the William F. Milton Fund to D.R.H.; NIH grant T32GM007753 to M.A.N.; NIH Training Grant in the Molecular Biology of Neurodegeneration 5T32AG000222-23 to S.H.; and a Burroughs Wellcome Fund Career award at the scientific interface, an Edward J. Mallinckrodt Scholarship, and NCI grant R33CA212697 to A.M.K.

Author contributions

S.H., D.R.H., and M.A.N. conceived the study and designed all experiments. S.H. and D.R.H. developed the optimized dissociation protocol. S.H. and D.R.H. performed single-cell experiments with assistance from R.Z., A.M.K., and A.R. M.A.N. processed all sequencing data. S.H., M.A.N., and D.R.H. analyzed data with assistance from R.B.-M. and A.M.K. D.R.H., K.R., and L.C. carried out FISH experiments and imaging, D.R.H. analyzed the FISH data. M.C. and D.R.H. developed automated software for FISH analysis. S.H., D.R.H., M.A.N., and M.E.G. wrote the manuscript. M.E.G. and B.L.S. advised on all aspects of the study.

Competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41593-017-0029-5>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to B.L.S. or M.E.G.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Mice. Animal experiments were approved by the National Institutes of Health and the Harvard Medical School Institutional Animal Care and Use Committee and followed the ethical guidelines described in the US National Institutes of Health *Guide for the Care and Use of Laboratory Animals* (<https://grants.nih.gov/grants/olaw/guide-for-the-care-and-use-of-laboratory-animals.pdf>). The experiments used adult (6- to 8-week-old) C57BL/6J virgin male mice (Jackson Laboratory).

Visual stimulation and brain dissection. Mice were housed under a standard light cycle (6:00–18:00) before being placed in constant darkness for 7 d. Mice were either euthanized in the dark (0 h, control condition) or exposed to light for 1 h or 4 h before euthanasia. After isoflurane anesthetization, the eyes were enucleated, the mice were euthanized, the brains were isolated, and the desired cortical regions were microdissected according to the protocols described below.

Fluorescence in situ hybridization (FISH). For sample preparation, mice were light exposed and euthanized as described above. Brains were immediately frozen on dry ice in tissue-freezing medium. Brains were sliced on a cryostat (Leica CM 1950) into 20- μ m sections, adhered to SuperFrost Plus slides (VWR), and immediately stored at -80°C until use. Samples were processed according to the ACD RNAscope Fluorescent Multiplex Assay manual.

For sample imaging, sections containing V1 were imaged on a Leica SP8 X confocal microscope with a 63 \times , 1.4-NA oil-immersion objective (Harvard NeuroDiscovery Center). Tiled V1 cortical areas of $\sim 1.1\text{ mm} \times \sim 0.5\text{ mm}$, containing all cortical layers, were imaged with optical sectioning of 0.5 μm . Channels were imaged sequentially to avoid any optical cross-talk.

For the image-analysis pipeline, we developed automated software to segment DAPI-stained nuclei in FISH images and to count the fluorescent puncta from hybridized probes in each nucleus. A labeled set of 23 images containing a total of $\sim 2,000$ nuclei was used to train a stacked random forest (RF) classifier by using contextual ‘offset’ features adapted from refs ⁴⁸ and ⁴⁹, as well as newly developed circularity features. The RF was trained on three labels: background, nuclei contour, and nuclei, and thus produced three associated probability maps. We used a watershed algorithm on the output probability maps to split neighboring nuclei and create masks. Finally, we eroded the nuclei masks to 80% of the original radius for conservative analysis of fluorescent puncta. Puncta detection was performed as described in ref. ⁵⁰. Code for the image-analysis pipeline is available upon request and has been described in detail in ref. ⁵¹.

Typically, one or two channels of a FISH experiment were devoted to cell-type markers. We used a model-based minimum-error thresholding method⁵² over all cells in an image to determine the fewest number of marker puncta necessary to classify the cell as positive for that marker. To mitigate the risk of false positives, a lower limit of four puncta was used if the model yielded a smaller number of puncta. A lower threshold (two puncta) was used for the layer 4 excitatory neuron cell-type marker *Cttn3*, which was detected at very low abundance (Supplementary Fig. 22). The following cell-type markers were used: *Vglut1* for excitatory neurons; *Pecam1* for endothelial and smooth muscle cells; *Aldh11l1* for astrocytes; and a combination of *Gad1* and *Gad2* for inhibitory neurons. Layer 4 was marked by *Rorb* expression and a peak in the density of DAPI-stained nuclei, and layer 6 was marked by *Foxp2* expression. Subtypes of excitatory neurons were marked by *Cdh13*, *Calb1*, *Cttn3*, *Hsd11b1* (Fig. 4h, Supplementary Figs. 20–22), *Gpr88*, *Bcl6*, and *Nnat* (Supplementary Fig. 14).

For image presentation, for large-area images (Figs. 1b and 4a,h, and Supplementary Figs. 14, 21 and 22), nuclei masks created in the analysis pipeline are displayed and pseudocolored according to the number of puncta contained within the mask. For single-cell images (Figs. 3c,f and 6e), a Gaussian filter was used to decrease uncorrelated noise for visualization purposes only.

Generation of single-cell suspensions. In the standard protocol, V1 was dissected bilaterally in ice-cold choline solution containing 2.1 g NaHCO_3 per liter, 2.16 g glucose per liter, 0.172 g $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ per liter, 7.5 mM $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$, 2.5 mM KCl, 10 mM HEPES, 15.36 g choline chloride per liter, 2.3 g ascorbic acid per liter, and 0.34 g pyruvic acid per liter. The tissue was cut into 300- μ m slices and dissociated with a papain dissociation system (Worthington) according to the manufacturer’s instructions with the following modifications. The EBSS solution was replaced with our dissociation solution (HBSS (Life Technologies), 10 mM HEPES (Sigma), 172 mg kynurenic acid (Sigma) per liter, 0.86 g $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ (Sigma) per liter, and 6.3 g D-glucose (Sigma) per liter, pH 7.35), which was saturated with 95% O_2 and 5% CO_2 . Dissociation was carried out at 37°C for 1 h with 20 U papain per milliliter.

In the protocol optimized for preserving transcriptional state, after the anesthetized animals were transcardially perfused for 5 min with ice-cold choline solution (described above) containing small-molecule cocktail consisting of 1 μ M TTX (Sigma), 100 μ M AP-V (Thermo Fisher Scientific), 5 μ g actinomycin D (Sigma) per milliliter and 10 μ M triptolide (Sigma). V1 was then microdissected, cut into 300- μ m slices, and incubated on ice for 30 min in dissociation solution containing 1 μ M TTX, 100 μ M AP-V (Thermo Fisher Scientific), 5 μ g actinomycin D (Sigma) per milliliter, 10 μ M triptolide (Sigma), and 10 μ g anisomycin (Sigma) per milliliter. Papain was added to a final concentration of 20 U/ml, and the tissue

was dissociated for 1 h at 37°C with gentle agitation in a total volume of 3.2 ml. The remaining procedures were performed per the manufacturer’s instructions without the small-molecule cocktail. After gradient centrifugation, the cells were washed in dissociation solution containing 0.04% BSA and were resuspended in dissociation solution containing 0.04% BSA and 15% Optiprep (Sigma) for scRNA-seq.

RNA isolation, reverse transcription, and qRT-PCR analysis. For qRT-PCR experiments across brain regions (Fig. 6a and Supplementary Fig. 1), the dissected regions from three or four mice were immediately frozen for RNA isolation. For experiments comparing the expression of activity-regulated genes across dissociation procedures, the cells were dissociated with the standard or the optimized dissociation method (described below) from four mice per condition, then were immediately frozen. The RNeasy Mini (Qiagen) isolation procedure was carried out according to the manufacturer’s instructions and included a DNase I digestion. For qRT-PCR analysis, the RNA was reverse transcribed with a High Capacity cDNA Reverse Transcription kit (Life Technologies). qRT-PCR was performed with LightCycler 480 SYBR Green I Master (Roche) on a LightCycler 480 system (Roche). Reactions were run with technical duplicates, which were averaged before subsequent analysis. *Gapdh*, encoding glyceraldehyde 3-phosphate dehydrogenase, was used as a normalization control for the $\Delta\Delta\text{Ct}$ -based quantification. The sequences of real-time PCR primers for *Gapdh* were TGTGTCCGTCGTGGATCTGA (forward) and TTGCTGTTGAAGTCGCAGGAG (reverse). For *Fos*, we used CTGGATTGACTGGAGGCTGC (forward) and TTGCTGATGCTCTTGACTGGC (reverse). For *Klf4*, we used GCAGTCAACAAGTCCCCTCTC (forward) and TAGTCAACAAGTGTGGGTGGC (reverse).

RNA-seq analysis of bulk visual cortex. RNA from three experimental conditions with four replicates each was isolated as described in the section above. The three conditions were drug-cocktail treated and visually unstimulated; drug-cocktail untreated and visually unstimulated; and drug-cocktail treated and visually stimulated (1 h). Sequencing libraries were prepared with a SMART-Seq v4 Ultra Low Input RNA kit for Sequencing and a Nextera XT DNA Sample Prep kit according to the manufacturer’s instructions.

Samples were sequenced on the NextSeq 500 (Illumina) platform, and paired-end reads with 38-nt lengths were obtained. Reads were trimmed with Trimmomatic-0.33 with the following parameters: ILLUMINACLIP:NexteraPE-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:30 (ref. ⁵³). A reference transcriptome was built with TopHat 2.1.1 and Bowtie 1.1.1 against the GRCm38.dna_sm.primary_assembly genome with the GTF file constructed above for inDrops mapping^{54,55}. Reads were mapped against this transcriptome in TopHat with the following options: --mate-inner-dist 500 --no-mixed --transcriptome-index [CUSTOM_TRANSCRIPTOME] --bowtie1 --no-novel-juncs. All samples showed $\geq 90\%$ concordant pair-alignment rate. The featureCounts⁵⁶ package was used to obtain gene counts with the following command options: -p -B -C -g gene_name -a [GTF] -s 1.

Count tables were TMM normalized and converted to CPM with the edgeR software analysis package^{57,58}. Any genes that were not expressed in at least three samples with TMM-normalized CPM > 1 were excluded from further analysis. Differential expression (DE) analyses were conducted with the voom/limma analysis software packages (requiring FDR-corrected $q < .01$) to identify drug-dependent genes (DE between visually unstimulated/cocktail treated and visually unstimulated/not cocktail treated; Fig. 1d) and bulk visual stimulus-dependent genes (DE between visually unstimulated/cocktail treated and visually stimulated/cocktail treated)^{59,60} (Supplementary Fig. 2a).

Single-cell RNA sequencing (inDrops). One or two libraries of approximately 3,000 cells were collected from each animal. inDrops was performed as previously described^{21,61}, generating indexed libraries that were then pooled and sequenced across 15 runs on the NextSeq 500 (Illumina) platform. Four libraries were downsampled to match sequencing depths across samples.

inDrops sequencing-data processing. Transcripts were processed according to a previously published pipeline²¹. Briefly, this pipeline was used to build a custom transcriptome from the Ensembl GRCm38 genome and GRCm38.84 annotation with Bowtie 1.1.1 (ref. ⁵⁵, after filtering the annotation gtf file (gencode.v17.annotation.gtf filtered for feature_type = ‘gene’, gene_type = ‘protein_coding’ and gene_status = ‘KNOWN’). Read quality control and mapping against this transcriptome were then performed. Finally, unique molecular identifiers were used to reference sequence reads back to individual captured molecules, thus yielding values denoted UMIFM counts⁶². All steps of the pipeline were run with default parameters unless explicitly specified.

Quality control for cell inclusion. Our initial dataset contained 114,601 cells with more than 1,000 reads assigned to each cell. All mitochondrially encoded genes were removed from the dataset. Cells with fewer than 700 or more than 15,000 unique-molecular-identifier counts were next excluded, thus yielding 65,539 high-quality cells isolated from 23 animals. The average transcript count per cell was 3,045.

Dimensionality reduction and clustering. All 65,539 cells were combined into a single dataset. Two independent approaches (t-SNE based and Seurat based; described below) were used to cluster cells (described in Supplementary Fig. 4).

t-SNE-based approach (approach 1). Raw counts were first linearly normalized such that each cell in the dataset contained the same number of transcripts (3,045). Next, the 4,000 most variable genes in cells derived from 0-h samples were identified as previously described²¹. Briefly, the v statistic for each gene (a corrected Fano factor accounting for noise in method efficiency and variation in cell size) was computed, and the genes with the largest v statistics were chosen as the most variable genes. Principal component analysis (PCA) based on the 4,000 most variable genes in cells derived from 0-h samples was applied to reduce the dimensionality of the dataset, as described in ref. ²¹. The MATLAB implementation of the t-SNE dimensionality-reduction algorithm used in ref. ²¹ was next applied to position cells on a 2D coordinate system by using the principal components generated above. The perplexity was set to 30, and other t-SNE arguments were left as default⁶³. A machine-learning algorithm with minimal user input was used to define clusters of cells on the basis of their proximity to one another in this 2D space, as previously described⁶⁴. The parameters used were $\text{percNeigh} = 0.02$; $\text{kernel} = \text{'Gauss'}$; $\text{minRho} = 10$, $\text{minDelta} = 3$. This method resulted in 91 initial clusters.

Seurat-based approach (approach 2). The second approach applied the Seurat R package^{65,66}. The data were log normalized and scaled to 10,000 transcripts per cell. Variable genes were identified with the $\text{MeanVarPlot}()$ function, which calculates the average expression and dispersion for each gene, then bins genes and calculates a z score for dispersion within each bin. The following parameters were used to set the minimum and maximum average expression and the minimum dispersion: $\text{x.low.cutoff} = 0.0125$, $\text{x.high.cutoff} = 3$, $\text{y.cutoff} = 0.5$. Next, PCA was carried out, and the top 30 principal components were retained. The clustering resolution was set to 0.6. This method resulted in 32 initial clusters.

Doublet removal and additional clustering. The following analysis was carried out for each approach independently. All clusters containing fewer than 100 cells were discarded. The expression of known marker genes (*Slc17a7*, *Gad1*, *Olig1*, *Aldoc*, *Cldn5*, *Vtn*, *Cx3cr1*, and *Mrc1*) was used to assign each cluster to one of the main cell types: excitatory neurons, inhibitory neurons, oligodendrocytes, astrocytes, endothelial and smooth muscle cells, pericytes, microglia, and macrophages. Clusters with substantial expression of two or more markers were removed because they most probably represented doublet artifacts arising from the cocapture of multiple cells in one droplet.

To resolve additional diversity across these main cell classes identified with either method, all clusters assigned to a single class were then pooled and reanalyzed by reapplication of the same method (t-SNE-based approach or Seurat-based approach). Occasionally, we identified clusters that contained a disproportional fraction of cells from 0 h, 1 h, or 4 h samples regardless of animal of origin, thus suggesting that stimulus-dependent expression was a major contributor to variance in these cell types. In all such cases, the clusters expressing identical cell-type markers were combined into a single cluster. This analysis resulted in 33 final clusters across 55,986 cells for the t-SNE-based approach and 48 final clusters across 56,372 cells for the Seurat-based approach.

Intersecting results obtained from the two approaches. The results from both approaches were intersected to exclude cells that were not consistently assigned. Because there were more clusters generated by approach 2 (48), we took each of those clusters and assessed their degree of overlap with clusters generated by approach 1 (33). First, a 2D matrix was generated with approach 1 clusters along one dimension and approach 2 clusters along the other dimension. The values in the matrix corresponded to the number of cells that were shared between any approach 1 and approach 2 cluster. Among the approach 2 clusters, 3 of the 48 clusters had cells distributed across several approach 1 clusters, and no approach 1 cluster encompassed most (> 50%) of the cells. All the cells from these three clusters were removed from the analysis because the overlap was not deemed sufficient. For each of the remaining 45 approach 2 clusters, we identified a single corresponding approach 1 cluster that encompassed most (> 50%) of the cells. Only the cells shared between each approach 2 cluster and a single majority-corresponding approach 1 cluster were kept for subsequent analysis. This conservative approach filtered out 12% of the cells and generated a dataset containing 48,266 cells robustly classified into nine main cell types (excitatory neurons, inhibitory neurons, oligodendrocytes, astrocytes, endothelial and smooth muscle cells, pericytes, microglia and macrophages) and 33 subtypes.

Classification of the fine-grained cell types. For each of the 33 predicted cell types, we identified a list of enriched genes whose expression in that cell type was threefold greater than the average expression across all other cell types (Supplementary Table 2). These genes were used to assign names to the 33 identified cell types as follows.

Excitatory neurons were marked by the expression of vesicular glutamate transporter 1 (*Slc17a7*) and calcium/calmodulin dependent protein kinase II beta

(*Camk2b*) and separated into layer-specific subtypes: layer 2/3 (ExcL23), layer 4 (ExcL4), three types of layer 5 (ExcL5_1, ExcL5_2, and ExcL5_3) and layer 6 (ExcL6). Layer 6b cells coclustered with layer 6 cells and therefore were analyzed together. We also observed small populations of cells derived from surrounding brain regions including the subiculum, hippocampus, and retrosplenial cortex³², which were removed from subsequent analyses, thus yielding a total of 30 final subtypes and a dataset of 47,209 cells (with an average of 3,234 transcripts per cell).

Inhibitory neurons were identified according to the expression of glutamate decarboxylase 1 (*Gad1*) and were separated into six subtypes according to the expression of previously described neuropeptides. Parvalbumin (*Pvalb*)-positive interneurons (Int_Pv) have previously been described as fast-spiking basket cells. Two types of somatostatin (*Sst*)-expressing interneurons were identified: Int_Sst1 and Int_Sst2. On the basis of in situ hybridization data from the Allen Brain Institute, we observed that Int_Sst_2 cells were distributed throughout the cortex, whereas Int_Sst_1 cells were restricted to layer 5 and layer 6. The remaining three interneuronal cell types were neuropeptide Y (*Npy*)-expressing cells (Int_Npy), layer 2/3 bitufted and bipolar *Vip*-expressing cells (Int_Vip), and upper-layer-enriched cholecystokinin-expressing interneurons (Int_Cck). Although smaller GABAergic subpopulations have previously been reported, we focused our analysis on these six most abundant subtypes.

We identified one major class of astrocytes that expressed aldolase dehydrogenase (*Aldoc*). Although *Gfap* is commonly used as a marker for astrocytes, in agreement with previously published single-cell data³² and in situ hybridization, that *Gfap* expression was restricted to a small subset of astrocytes.

Nine subtypes of *Olig1*-expressing cells were identified⁴⁵. Two subsets, marked by *Pdgfra*, represented OPCs. A large, probably quiescent population of OPCs expressed *C1ql1* (OPC_1), whereas a very small *Pdgfra*⁺ population (102 cells) expressed cell-cycle-associated genes and was likely to be actively cycling (OPC_2)³². Seven additional populations of *Olig1*-expressing cells could be arranged in a continuous progression corresponding to different stages of differentiation: *Bmp4*⁺ immature cells (Olig_7), premyelinating *Tmem2*⁺ cells (Olig_6), and four separate myelinated (*Mag*⁺, *Mog*⁺, and *Mbp*⁺) populations (Olig_1, Olig_2, Olig_3, and Olig_4). We also observed a separate population of mature oligodendrocytes that were marked by *Kif5a* and expressed lower levels of *Mog* and *Mag* but higher levels of *Mbp* (Olig_5) than were found in other oligodendrocyte populations.

In agreement with previous reports, we identified microglia and macrophages as the primary immune cell types in the brain³². Two subtypes of microglia were identified: a more abundant *P2yr12*^{high} (Micro_2) population and a less abundant *P2yr12*^{low}, *Ccl3*⁺/*Ccl4*⁺ population (Micro_1). Because *Ccl3* and *Ccl4* are markers of macrophage activation, it is possible that this second population may represent activated microglia. Macrophages were identified as *Cd36*⁺/*Mrc1*⁺ cells.

Finally, we identified two types of endothelial cells (Endo_1 and Endo_2) and two types of smooth muscle cells (SM_1 and SM_2) expressing the tight-junction protein claudin 5 (*Cldn5*), as well as a population of *Cldn5*⁻ *Vtn*⁺ pericytes. The less abundant of the two endothelial cell types (Endo_2) expressed high levels of hemoglobin alpha (*Hb-a1* and *Hb-a2*) and hemoglobin beta (*Hb-bs* and *Hb-bt*). Although these cells have not been described in previous scRNA-seq datasets, prior work has shown that arterial endothelial cells express hemoglobin alpha, which is enriched at the myoendothelial junction and regulates NO-mediated vascular reactivity⁶⁷. It is therefore likely that our hemoglobin-positive cells are arterial endothelial cells.

One population of smooth muscle cells (SM_2) expressed *Abcc9*, encoding an ATP-binding cassette, and *Kcni8*, encoding an inwardly rectifying voltage-gated potassium channel. These proteins form an ATP-sensitive potassium channel that directly links cellular ATP metabolism with membrane depolarization. The second population of smooth muscle cells (SM_1) expressed *Acta2*, encoding alpha smooth muscle actin, and *Myh11*, encoding myosin heavy chain, thus suggesting that different smooth-muscle-cell types may be subspecialized in their function.

Hierarchical clustering. Hierarchical clustering across all cell types was performed on the depth-normalized dataset (each cell containing 3,234 transcripts) by using the top 4,000 most variable genes (from approach 1). The mean expression for each gene was calculated across each cell type, and the distance between cell types was calculated on the basis of Euclidean distance and hierarchical clustering performed in R with the Ward2 algorithm.

Identification and classification of experience-regulated transcripts. To identify experience-regulated genes for each cell type, we carried out differential gene expression analysis in Monocle2 (ref. ⁶⁸) between cells isolated from the visual cortex in mice exposed to light for 0 h, 1 h, and 4 h. The data were modeled and normalized with a negative binomial distribution consistent with scRNA-seq experimental results. The analysis was performed independently for each of the 30 cell types identified, by separately comparing 1 h to 0 h and 4 h to 0 h. In each analysis, a gene was required to be detected in a minimum of 5% of cells to be included in the differential gene expression test. Genes whose FDR was < 0.05 and whose \log_2 fold change in expression was either > 1 or < -1 were considered activity regulated. The \log_2 fold change was calculated from the depth-normalized

data (with each cell normalized to contain 3,234 transcripts) after 0.12 was added to the expression of each gene: $FC = \log_2(\text{mean1} + 0.12) - \log_2(\text{mean2} + 0.12)$.

For each cell type, activity-regulated genes were classified as either ERGs or LRGs on the basis of their expression patterns. If an induced gene's maximum expression was at 1 h, the gene was classified as an ERG, whereas if the maximum expression was at 4 h, the gene was classified as an LRG. Genes whose expression decreased in response to a stimulus were similarly classified on the basis of the time point of minimum expression. If a gene was classified as an ERG in some cell types and an LRG in others, its final classification was based on the most frequent pattern of induction. If the number of cell types in which a gene was an ERG and a LRG was the same, that gene was classified as an LRG. All induced genes and their classifications are listed in Supplementary Table 3. These classifications are denoted *a* and *b* (early increase and decrease, respectively), and *c* and *d* (late increase and late decrease, respectively).

Determination of induced cells within a population. The top ten genes ranked by fold change between 0 h and 1 h of light stimulation were collected for each cell type. For each gene, the 0-h sample was used to define a ninetieth-percentile expression threshold; i.e., the 10% of cells expressing the gene would be considered positive for that gene. This threshold was meant to represent the expression level present in quiescent cells; we chose 90% to account for the high likelihood that some cells might not be silent even in the unstimulated condition. To classify a single cell as induced, we evaluated the expression of all ten of the most induced genes within that cell type and required at least three of the genes to have an expression level greater than the threshold set by the ninetieth percentile of cells from the 0-h condition. To provide a range for our estimate of induction, we also plotted the results of the same analysis when either two or four genes within each cell met this requirement as the lower and upper bounds of the box.

ERTF hierarchical clustering on the basis of expression patterns across cell types. 19 ERTFs that were induced in at least three cell types were hierarchically clustered on the basis of their \log_2 fold changes between the 0-h and 1-h conditions. The distance metric between genes was Euclidean, and hierarchical clustering was performed with Ward2 algorithms.

Coexpression analysis. Depth-normalized gene expression across single cells at the 1-h time point was correlated (Pearson correlation) between 14 neuronal-induced ERTFs. Expression-matched genes were used as a control and were chosen from the list of all genes whose mean expression in excitatory neurons was between the lowest and highest of the 14 ERTFs. The 14 most closely expressed genes are shown in Fig. 3d. For Fig. 3e, the distribution of similarly expressed genes was generated by random sampling (100 times) of 14 genes whose expression was matched to that of the 14 ERTFs. As an additional control, the expression of each ERTF was shuffled across all cells, such that the expression value for that ERTF was randomly assigned to a different cell. Shuffling across cells retained the average expression and induction of each ERTF but significantly decreased the correlation of expression across ERTFs (Fig. 3e). Statistical differences between the distributions were computed with the Mann–Whitney *U* test.

Identification of excitatory layer 2/3 and layer 4 subtypes. Layer 2/3 and layer 4 excitatory cells were processed independently with the Seurat algorithm. For each set, the data were log normalized and scaled to 10,000 transcripts. Variable genes were identified with the following parameters: *x.low.cutoff*=0.0125, *x.high.cutoff*=3, *y.cutoff*=0.5. The top 30 principal components were chosen, and the clustering resolution was set to 0.6. This analysis initially identified six clusters from Excl23 cells. One cluster containing 22 cells was removed because it was below the 100-cell cutoff. Four of the remaining five clusters contained unbalanced numbers of cells from different time points, thus suggesting that they were separated by stimulus-regulated genes. Consequently, these clusters were merged into a single Excl23_1 population. Excl4 contained four clusters, two of which contained unbalanced numbers of cells from different time points and were merged into the Excl4_1 population.

Correlation between excitatory layer 2/3, layer 4, and layer 5 cell types. In Fig. 4i, left, a union of all LRGs initially identified in Excl23, Excl4, Excl5_1, Excl5_2 and Excl5_3 cell types was created (51 genes). Correlation (Pearson) between cell types was determined on the basis of the mean expression of each of the LRGs in each cell type. In Fig. 4i, right, a random set of expression-matched non-LRGs was generated, and a correlation was computed in the same manner as that for the LRGs. This analysis was repeated 100 times, and the mean pairwise correlation between cell types was plotted.

Correlation between excitatory layer 2/3 and layer 4 marker genes and sequencing depth. We confirmed that the cell-type-specific expression of the marker genes shown in Fig. 4c,d and in Supplementary Fig. 19 was not a function of sequencing depth across layer 2/3 and layer 4 subtypes with the analysis described below (Supplementary Fig. 20).

The depth-normalized expression level (TPT) of each marker gene was calculated for each cell assigned to a given cell type. 200 cells with $5,000 \leq \text{total}$

UMIFM counts/cell $\leq 10,000$ were sampled from each layer 2/3 or layer 4 subtype at random. The Pearson correlation between marker-gene expression and cell sequencing depth (total UMIFM counts) was then calculated for each layer's marker gene for all subtypes of that layer, by using these 200 depth-matched cells.

Cross-study comparison of transcriptionally defined cell types. Expression levels from each of the two datasets were independently scaled (mean centered, unit variance) by gene within each transcriptionally defined cell type to mitigate batch effects due to different sequencing depths, cell-capture approaches, normalization approaches, and library preparation across this work and ref. ²³. The Pearson correlation was calculated across all pairwise combinations of cell types between the two studies by using only genes expressed in both datasets (normalized expression > 0) in at least one cell type. The cell types were hierarchically clustered (Ward's method) and arranged such that the distance (1 – Pearson correlation) between proximal leaves was minimized. Cell types defined in either study with shared transcriptional identity clustered closely regardless of which study they were taken from (Supplementary Fig. 11), thus suggesting that our classification scheme was consistent with that used in ref. ²³.

Cross-study comparison between similar excitatory neuron subtypes. Using all genes expressed across both studies in at least one cell type was sufficient to establish gross cross-study cell type correlations, but the subtle differences in gene expression across the particularly similar excitatory subtypes were not revealed by this approach. We overcame this limitation by further calculating cross-study correlations for the excitatory cell types alone, using excitatory neuron marker genes identified in either this study (Fig. 3b) or in ref. ²³. Hierarchical clustering was conducted as described in the previous section, and the results are shown in Supplementary Fig. 16.

OPC-specific receptor and induced-ligand identification. We used a human ligand–receptor database⁴³ comprising 2,422 ligand–receptor pairs as the basis for our investigation into protein signaling between OPCs and other cell types.

To identify genes enriched in OPCs relative to other oligodendrocyte populations, we required that (i) the gene's mean expression be higher in OPCs than in any one subtype of oligodendrocytes, (ii) the average expression of the gene in OPCs be at least threefold greater than the mean in oligodendrocytes, and (iii) the gene be in the top 50% most highly expressed genes in OPCs. Intersecting this list with all receptors in the ligand–receptor database yielded 52 OPC-enriched receptors, of which nine had corresponding ligands that were induced in at least one nonoligodendrocyte cell type.

Gene Ontology (GO) analysis. GO analysis was carried out with DAVID 6.8 (refs. ^{69,70}). All expressed genes for the cell type being analyzed were used as background. Expressed genes were defined as genes that were detected in a minimum of 5% of cells. For the GO analysis conducted to generate Supplementary Table 1, all genes expressed with TMM-normalized CPM > 1 in at least three samples were used as background.

Gene set enrichment analyses using the Su et al. dataset. Two gene sets were derived directly from the differential expression analyses conducted in Su et al. (ref. ⁴⁷) on RNA-seq data (rnaseq_1 and rnaseq_4 in Supplementary Fig. 24), with the additional requirement of a minimum twofold change in expression between the unstimulated and electroconvulsively stimulated conditions. The remaining two gene sets were obtained by identifying the transcriptional start sites most proximal to the stimulus-dependent, differentially accessible regions identified in ref. ⁴⁷ (termed *atac_1* and *atac_4* in Supplementary Fig. 24), by using the GRCh37.p11 annotation.

A one-sided Fisher's exact test was used to test the following null hypothesis: the fraction of induced genes identified in the visual cortex in this study (611 total) that was also identified in ref. ⁴⁷ as being stimulus dependent in dentate gyrus < the fraction of stimulus-independent genes identified in this study (12,458 total genes) that was also identified in ref. ⁴⁷ as being stimulus dependent in the dentate gyrus. The Benjamini–Hochberg procedure was applied to correct for multiple hypothesis testing.

Statistical analysis. For Fig. 1c, qRT–PCR for *Fos* expression normalized to *Gapdh* expression, *n* = 4 samples from 4 different mice per condition were analyzed with a two-tailed Student's *t* test. The data distribution was assumed to be normal, although normality was not tested.

For Supplementary Fig. 1 and Fig. 6a, qRT–PCR for *Fos* or *Klf4* expression normalized to *Gapdh* expression. *n* = 3 samples (for motor and prefrontal cortex), and *n* = 4 samples (for somatosensory and visual cortex). Each sample was taken from a different mouse. Analysis was done with a two-tailed Student's *t* test. The data distribution was assumed to be normal, although normality was not tested.

For Fig. 2, to identify experience-regulated genes for each cell type, we carried out differential gene expression analysis in the software package Monocle2. Further details are reported in 'Identification and classification of experience-regulated transcripts' section.

For Fig. 3a,d, nonparametric Mann–Whitney *U* tests were used to test the differences in the distributions of values. For Fig. 3a, $n_{\text{ERTFs}} = 38$ genes and

$n_{\text{L.RGS}} = 176$ genes in each category, and the values tested are the numbers of cell types found to be activity regulated for each gene. For Fig. 3d, genes were selected as indicated in "Coexpression analysis" ($n = 176$ pairwise correlations between two populations), on the basis of a defined set of genes.

For quantification of FISH images in Figs. 1b, 3c, 4g and 6f and Supplementary Fig. 13, all distributions were first tested for normality via the Kolmogorov–Smirnov test. All tested distributions rejected the null hypothesis and were treated nonparametrically. Nonparametric Mann–Whitney U tests were then used to test for differences in the distributions of values.

No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications^{8,71,72}. Mice were randomly assigned to 0-h, 1-h and 4-h time points. No other sample randomization was performed. Experiments were not performed in a blinded fashion.

Life Sciences Reporting Summary. Further information on experimental design is available in the Life Sciences Reporting Summary.

Code availability. Code is available upon request.

Data availability. The data that support the findings of this study are available from the corresponding author. Raw and processed RNA-seq data for both single-cell and bulk experiments are available at Gene Expression Omnibus under accession [GSE102827](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE102827).

To broadly share our data, we have also created an interactive website on which the gene expression of each of the genes in our dataset can be viewed: <http://greenberg.hms.harvard.edu/project/gene-database/>.

References

48. Tu Z. & Bai X. Auto-context and its application to high-level vision tasks and 3D brain image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 1744–1757 (2010).
49. Richmond, D., Kainmueller, D., Glocker, B., Rother, C. & Myers, G. Uncertainty-driven forest predictors for vertebra localization and segmentation. in *International Conference on Medical Image Computing and Computer-Assisted Intervention* 653–660 (Springer, Munich, 2015).
50. Aguet, F., Antonescu, C. N., Mettlen, M., Schmid, S. L. & Danuser, G. Advances in analysis of low signal-to-noise images link dynamin and AP2 to the functions of an endocytic checkpoint. *Dev. Cell* **26**, 279–291 (2013).
51. Cicconet, M., Hochbaum, D. R., Richmond, D. & Sabatini, B. L. Bots for software-assisted analysis of image-based transcriptomics. Preprint at <https://www.biorxiv.org/content/early/2017/08/03/172296> (2017).
52. Fan, J. Notes on Poisson distribution-based minimum error thresholding. *Pattern Recognit. Lett.* **19**, 425–431 (1998).
53. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
54. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
55. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
56. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
57. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
58. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).
59. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
60. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
61. Zilionis, R. et al. Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protoc.* **12**, 44–73 (2017).
62. Islam, S. et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat. Methods* **11**, 163–166 (2014).
63. van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
64. Rodriguez, A. & Laio, A. Clustering by fast search and find of density peaks. *Science* **344**, 1492–1496 (2014).
65. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
66. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
67. Straub, A. C. et al. Endothelial cell expression of haemoglobin α regulates nitric oxide signalling. *Nature* **491**, 473–477 (2012).
68. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
69. Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
70. Huang, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
71. Shekhar, K. et al. Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. *Cell* **166**, 1308–1323.e30 (2016).
72. Baron, M. et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. *Cell Syst.* **3**, 346–360.e4 (2016).

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

▶ Experimental design

1. Sample size

Describe how sample size was determined.

No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications.

2. Data exclusions

Describe any data exclusions.

Four of the 28 sequencing libraries were computationally down-sampled (a random fraction of reads was removed) to match sequencing depths across samples (see methods).
Cell doublets were removed. The criteria is consistent with other papers and is reported in the methods.

3. Replication

Describe whether the experimental findings were reliably reproduced.

The experiment was carried out across 28 independent samples, derived from 23 animals across 20 sequencing runs and 7 batches (separate days on which cells were collected and libraries processed). All data was pooled and analyzed together. We used multiple methods to confirm key findings (scRNAseq, FISH).

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Mice were randomly assigned to the 0h, 1h and 4h cohorts. No other randomization was used.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Only cell encapsulation and transcriptomic alignment were performed in a blinded fashion. Knowledge of experimental conditions was necessary during other parts of data collection. Analysis was conducted using automated scripts.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

Matlab 2016b, R 3.3.1, Limma 3.26.9, edgeR 3.12.1, Python 2.7.6, matplotlib 2.0.2. study. For details see methods.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). [Nature Methods guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

There are no restrictions

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

NA

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

NA

b. Describe the method of cell line authentication used.

NA

c. Report whether the cell lines were tested for mycoplasma contamination.

NA

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

NA

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

Animal experiments were approved by the National Institute of Health and Harvard Medical School Institutional Animal Care and Use Committee. The experiments used adult (6–8 weeks old) C57BL/6J virgin male mice (The Jackson Laboratory).

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

NA

Editorial Policy Checklist

This form is used to ensure compliance with Nature Research editorial policies related to research ethics and reproducibility in the life sciences. For further information, please see our [Authors & Referees](#) site. All questions on the form must be answered.

► Data availability

Policy information about [availability of data](#)

Data availability statement

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated source data
- A description of any restrictions on data availability

A full data availability statement is included in the manuscript.

Required accession codes

Data deposition is mandated for [certain types of data](#).

Confirm that all relevant data have been deposited into a public repository and that all accession codes are provided.

Accession codes will be available before publication No data with mandated deposition All relevant accession codes are provided

► Data presentation

Image integrity

Confirm that all images comply with our [image integrity policy](#).

Unprocessed data must be provided upon request. Please double-check figure assembly to ensure that all panels are accurate (e.g. all labels are correct, no inadvertent duplications have occurred during preparation, etc.).

Data distribution

Data should be presented in a format that shows data distribution (dot-plots or box-and-whisker plots), with all box-plot elements (e.g. center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers) defined. If bar graphs are used, the corresponding dot plots must be overlaid.

Confirm that all data presentation meets these requirements.

Confirm that in all cases where the number of data points is <10, individual data points are shown.

► Structural data

Policy information about [special considerations](#) for specific types of data

If this study did not involve data of these types, check here and skip the rest of this section.

Electron microscopy

For all electron microscopy work, confirm that you have deposited any density maps and coordinate data in [EMDB](#).

Macromolecular structures

For all macromolecular structures studied, confirm that you have provided an official validation report from [wwPDB](#).

► Code availability

Policy information about [availability of computer code](#)

Code availability statement

For all studies using custom code, the Methods section must include a statement under the heading "Code availability" describing how readers can access the code, including any access restrictions.

A full code availability statement is included in the manuscript No custom code used

► Research animals

Policy information about [studies involving animals](#); follow the [ARRIVE guidelines](#) for reporting animal research

If this study did not use animals and/or animal-derived materials for which ethical approval is required, check here and skip the rest of this section.

Ethical compliance

Confirm that you have complied with all relevant ethical regulations and that a statement affirming this is included in the manuscript.

Ethics committee

Confirm that you have stated the name(s) of the board and institution that approved the study protocol in the manuscript.

► Human research participants

Policy information about [studies involving human research participants](#)

If this study did not involve any human research participants, check here and skip the rest of this section.

Ethical compliance

Confirm that you have complied with all relevant ethical regulations and that a statement affirming this is included in the manuscript.

Ethics committee

Confirm that you have stated the name(s) of the board and institution that approved the study protocol in the manuscript.

Informed consent

Confirm that informed consent was obtained from all participants.

Identifiable images

For publication of identifiable images of research participants, confirm that consent to publish was obtained and is noted in the Methods.

Authors must ensure that consent meets the conditions set out in the [Nature Research participant release form](#).

Yes No identifiable images of human research participants

► Clinical studies

Policy information about [clinical studies](#)

If this study was not a clinical trial, check here and skip the rest of this section.

Clinical trial registration

Confirm that you have provided the trial registration number from [ClinicalTrials.gov](#) or an equivalent agency in the manuscript.

Phase 2 and 3 randomized controlled trials

Confirm that you have provided the [CONSORT checklist](#) with your submission.

Yes No Not a phase 2/3 randomized controlled trial

Tumor marker prognostic studies

Did you follow the [REMARK reporting guidelines](#)?

Yes No Not a tumor marker prognostic study

► Methods reporting

Nature Research wishes to improve the reproducibility of the work we publish. As part of this effort, all life science manuscripts require a [reporting summary](#); certain types of research require specialized modules in addition to this form.

Confirm that you have provided a complete and accurate [reporting summary](#).

n/a | Confirmed

For MRI studies, confirm that you have completed the additional [MRI module](#).

For flow cytometry studies, confirm that you have completed the additional [flow cytometry module](#).

For ChIP-seq studies, confirm that you have completed the additional [ChIP-seq module](#).

I certify that all the above information is complete and correct.

Typed signature Sinisa Hrvatin

Date Oct 9, 2017