

# Detecting Activated Cell Populations Using Single-Cell RNA-Seq

## Highlights

- Act-seq minimizes artificial transcriptional changes during tissue dissociation
- Act-seq enables unbiased characterization of cell types and their acute activation
- Application of Act-seq provides the first molecular taxonomy in the amygdala
- Application of Act-seq identifies neuronal subpopulations activated by stress

## Authors

Ye Emily Wu, Lin Pan, Yanning Zuo, Xinmin Li, Weizhe Hong

## Correspondence

whong@ucla.edu

## In Brief

Wu et al. develop Act-seq, which minimizes artificially induced transcriptional changes during single-cell dissociation and thus enables faithful characterization of baseline transcriptional profiles and detection of specific cell populations activated by physiological stimuli using single-cell sequencing.



# Detecting Activated Cell Populations Using Single-Cell RNA-Seq

Ye Emily Wu,<sup>1,3</sup> Lin Pan,<sup>1,3</sup> Yanning Zuo,<sup>1</sup> Xinmin Li,<sup>2</sup> and Weizhe Hong<sup>1,4,\*</sup>

<sup>1</sup>Department of Biological Chemistry and Department of Neurobiology

<sup>2</sup>Department of Pathology

David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA

<sup>3</sup>These authors contributed equally

<sup>4</sup>Lead Contact

\*Correspondence: [whong@ucla.edu](mailto:whong@ucla.edu)

<https://doi.org/10.1016/j.neuron.2017.09.026>

## SUMMARY

Single-cell RNA sequencing offers a promising opportunity for probing cell types mediating specific behavioral functions and the underlying molecular programs. However, this has been hampered by a long-standing issue in transcriptional profiling of dissociated cells, specifically the transcriptional perturbations that are artificially induced during conventional whole-cell dissociation procedures. Here, we develop Act-seq, which minimizes artificially induced transcriptional perturbations and allows for faithful detection of both baseline transcriptional profiles and acute transcriptional changes elicited by behavior/experience-driven activity. Using Act-seq, we provide the first detailed molecular taxonomy of distinct cell types in the amygdala. We further show that Act-seq robustly detects seizure-induced acute gene expression changes in multiple cell types, revealing cell-type-specific activation profiles. Furthermore, we find that acute stress preferentially activates neuronal subpopulations that express the neuropeptide gene *Cck*. Act-seq opens the way for linking physiological stimuli with acute transcriptional dynamics in specific cell types in diverse complex tissues.

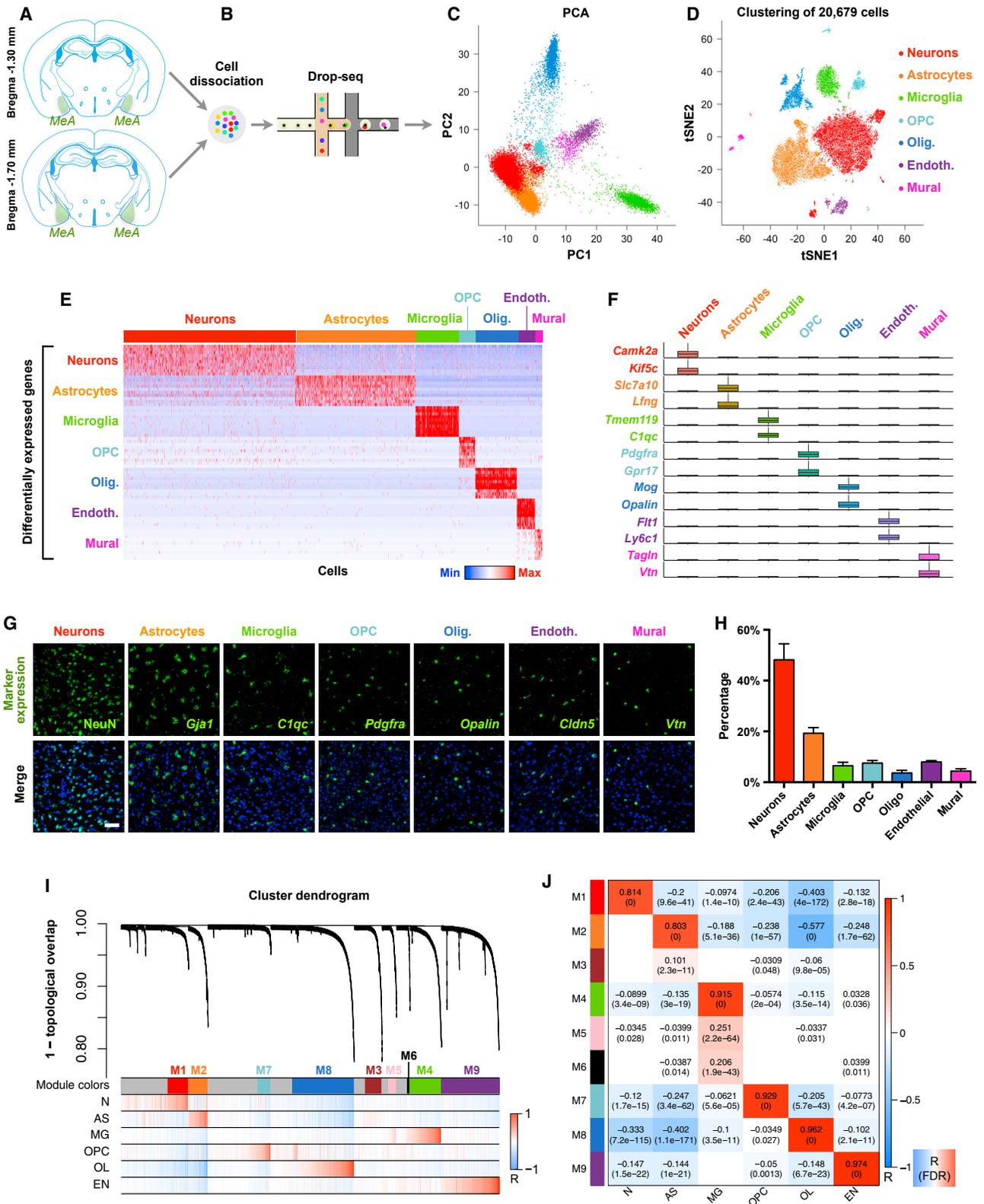
## INTRODUCTION

The mammalian central nervous system supports a multitude of cognitive and behavioral functions through coordinated action of different neural circuits that are composed of diverse sets of differentiated cell types. Essential to the functional dissection of neural circuits is the precise identification and manipulation of neuronal subsets that are activated during a specific behavior or experience. The expression of immediate-early genes (IEGs), which is rapidly and transiently induced by various cellular stimuli including synaptic activity, has been widely used to map neuronal activation patterns and to functionally label ensembles of activated neurons (Kawashima et al., 2014; Tonegawa et al.,

2015). These approaches have been instrumental for understanding the functions of different brain regions and neuronal populations. However, they are mostly limited to the detection of one or a few IEGs, such as *Fos*, *Egr1*, and *Arc*, and the identity of the activated cells are often unclear. Since the activation level and dynamics of different IEGs can vary substantially between brain areas and cell types (Kawashima et al., 2014), pre-selecting one or a few IEGs without a systematic survey may limit the detection of activated cell populations. On the other hand, several techniques for bulk transcriptomic profiling of discrete cell populations based on various molecular tags, such as translating ribosome affinity purification (TRAP) (Heiman et al., 2008), RiboTag (Sanz et al., 2009), and phosphorylated ribosome capture (Knight et al., 2012), allow comprehensive interrogation of the molecular features of different cell populations, including activated neurons. Nonetheless, they do not preserve single-cell identity, and rely on pre-selected “cell-type-specific” markers, which in many cases are not restricted to a single functionally or molecularly distinct cell type. This can confound the interpretation of results from these approaches. For example, it may not be possible to definitively distinguish between activation of neurons and that of non-neuronal cells, or between strong activation of a small fraction of the cells and weak activation of a large fraction of the cells.

Recent advances in single-cell RNA sequencing (scRNA-seq) have offered unprecedented power to reveal fine-scale molecular differences that may underlie functional heterogeneity across cell types (reviewed in Poulin et al., 2016). scRNA-seq can potentially survey transcriptome-wide IEG activation and other transcriptional changes in an unbiased manner, and therefore holds the promise of providing a high-throughput, high-resolution tool for revealing the cell-type identity and molecular features of behaviorally relevant cell populations. Furthermore, single-cell profiling of acute transcriptomic changes will be of broad value to the study of diverse biological questions, such as studies on molecular and cellular responses to tissue injuries and immunological stimuli.

A major, long-standing issue in transcriptional profiling of dissociated cells, including scRNA-seq, has been artificially evoked transcriptional perturbations that occur during conventional whole-cell dissociation procedures. This dissociation process has been shown to cause cell injury, cellular stress, and excitotoxicity, all of which can cause gene expression changes



(legend on next page)

(Kolodziejczyk et al., 2015; Lacar et al., 2016; Poulin et al., 2016). This can not only mask endogenous acute gene expression changes that are truly elicited by physiological stimuli, but may also confound characterization of baseline transcriptional profiles. By dissociating single nuclei instead of whole cells, a recent study showed the feasibility of isolating and profiling nuclei of activated neurons using single-nuclei RNA-seq (snRNA-seq) (Lacar et al., 2016). However, this method relies upon flow sorting of nuclei after immunostaining for one particular IEG (FOS), introducing bias in cell sampling (Lacar et al., 2016). Moreover, compared to sequencing of whole cells, snRNA-seq detects appreciably less mRNAs per cell and biases toward nascent, unprocessed RNA species (Habib et al., 2017; Lake et al., 2017). snRNA-seq therefore may not be suitable when studying cell types with low RNA content and/or when capture of processed, more mature RNA species in the cytoplasm is of importance. A method that enables faithful detection of both baseline and physiologically induced transcriptional features through scRNA-seq of whole cells is much needed.

Here, we developed Act-seq, which combines scRNA-seq with a new single-cell preparation method that eliminates transcriptional perturbations artificially induced by conventional cell preparation procedures. Act-seq enables detailed cell-type classification and faithful detection of both baseline transcriptional profiles and acute gene expression changes in response to physiological stimuli. We applied Act-seq to the medial amygdala (MeA), a brain region important for the modulation of emotional processing and social behavior (Swanson, 2000). Using Act-seq, we generated the first detailed molecular taxonomy of MeA cells, including 16 transcriptionally distinct neuronal subtypes as well as several non-neuronal cell types. We further demonstrated the ability of Act-seq to robustly capture acute gene expression changes in both neurons and non-neuronal cells and found that IEGs were differentially induced in different cell types. Lastly, we applied Act-seq to detecting neuronal activation following acute stress, revealing preferential activation of neuronal subpopulations that expressed the neuropeptide gene *Cck*.

## RESULTS

### Unbiased scRNA-Seq Identifies Distinct MeA Cell Types

Previous studies have indicated heterogeneity in gene expression, electrophysiological property, morphology, connectivity, and function among MeA neurons (Bergan et al., 2014; Bian et al., 2008; Choi et al., 2005; Hong et al., 2014; Keshavarzi

et al., 2014; Unger et al., 2015; Xu et al., 2012; Zirlinger et al., 2001), but the cell type composition within the MeA remains poorly characterized. To dissect cell types in the MeA, we performed high-throughput single-cell transcriptional profiling using Drop-seq (Macosko et al., 2015), wherein single cells are encapsulated into droplets containing single microparticle beads coated with oligonucleotides that carry unique barcodes for identifying individual cells and transcripts (Figures 1A and 1B). We acquired 20,679 single-cell transcriptomes containing a total of 16,233 genes from acutely microdissected and dissociated MeA cells of adult mice (Figures 1A and 1B). Using principle component analysis (PCA), dimensionality reduction by t-distributed stochastic neighbor embedding (tSNE), and clustering with the Louvain-Jaccard algorithm (Blondel et al., 2008; Shekhar et al., 2016), we unbiasedly grouped all the cells into multiple distinct clusters (Figures 1C and 1D, Figure S1). Based on expression of previously known common cell-type markers, we identified seven main cell types in the MeA: neurons, astrocytes, microglia, oligodendrocyte precursor cells (OPCs), oligodendrocytes, vascular endothelial cells, and mural cells (pericytes and vascular smooth muscle cells) (Figures 1C–1F). Differential gene expression analysis revealed numerous genes specifically enriched in each class (Figures 1E and 1F; Table S1, Table S2), many of which have been previously reported as cell-type markers in other brain regions. We further confirmed the expression of markers of each major cell type in the MeA using fluorescent *in situ* hybridization (FISH) or immunohistochemistry (Figure 1G) and observed that the cell-type composition in our Drop-seq data largely corresponded with the proportion of different cell types revealed by staining (Figures 1E and 1H).

In addition to known common cell-type markers, comparison with cell-type specific RNA-seq data in the cortex (Tasic et al., 2016; Zeisel et al., 2015; Zhang et al., 2014) identified several genes that appeared to show cell-type enrichment patterns unique to the MeA. For example, *Gng13* (a G protein subunit) and *Vps8* (an endosomal trafficking protein) were enriched in neurons in the MeA (fold change = 5.7 and 3.5, respectively, false discovery rate [FDR] <  $1.0 \times 10^{-6}$ ) but not in neurons from the previously examined cortical regions (Tasic et al., 2016; Zeisel et al., 2015; Zhang et al., 2014). Similarly, *Kcnt1* (a potassium channel subunit) and *Lgi1* (a leucine-rich repeat protein) showed enrichment in astrocytes only in the MeA (fold change = 7.4 and 6.3, respectively, FDR <  $1.0 \times 10^{-6}$ ), but not in astrocytes in the previously examined cortical regions (Tasic et al., 2016; Zeisel et al., 2015; Zhang et al., 2014). The amygdala importantly

### Figure 1. Classification of MeA Major Cell Types Using Drop-Seq

(A and B) Schematic showing representative sections containing the dissected MeA region (A) and single-cell isolation and sequencing using Drop-seq (B). (C and D) Separation by principle components (PC) 1 and 2 (C) and two-dimensional tSNE visualization (D) of 20,679 MeA cells. Individual dots correspond to single cells, colored according to major cell types. Olig., oligodendrocytes; Endoth., endothelial cells. (E) Heatmap showing expression level of select top cell-type marker genes (fold change > 3, FDR < 0.05) across individual cells. (F) Boxplots showing the distribution of expression level of representative major cell-type markers across different cell types. (G) Immunohistochemistry (NeuN) or FISH (*Gja1*, *C1qc*, *Pdgfra*, *Opalin*, *Cldn5*, and *Vtn*) for main cell-type markers (green) and fluorescent DAPI stain (blue) in the MeA. (H) Quantification of the percentage of different major cell types revealed by staining in all DAPI-stained cells in the MeA. Mean  $\pm$  SD, n = 3 mice for all cell types. (I) Dendrogram showing gene coexpression modules defined using WGCNA. Color bars below indicate the module assignment and Pearson correlation with each major cell type for each gene. N, neurons; AS, astrocytes; MG, microglia; OL, oligodendrocytes; EN, endothelial cells. (J) Heatmap showing Pearson correlation between module eigengenes and each major cell type. Correlation coefficients (R) and FDR-adjusted p values are shown where FDR < 0.05. See also Figures S1, S2, and S3 and Tables S1 and S2.

regulates sex-specific behaviors (Swanson, 2000). Several genes, including *Ecel1*, *Greb1*, *Esr1*, and *Rps6ka6*, have been reported to display sexually dimorphic expression in the MeA (Xu et al., 2012). By examining the expression of these genes in different cell types, we found that *Ecel1*, *Greb1*, and *Esr1* were enriched in neurons, whereas *Rps6ka6* appeared to be more highly expressed in OPCs (Figure S2A), suggesting that sexually dimorphic gene expression in the MeA may arise from both neuronal and non-neuronal cell types.

To confirm the precision of our MeA dissection, we examined the expression of two genes, *Cyp26b1* and *Bmp3*, which were selectively enriched in the neighboring basolateral (BLA) and central (CeA) amygdalar nuclei, but not MeA (Figures S3E and S3F). We confirmed that there was little expression of these genes in our scRNA-seq data generated from the MeA (Figures S3A and S3B). This lack of expression was not due to an inability to detect these genes by Drop-seq, as they were readily detectable by Drop-seq in cells dissociated from a broader amygdalar region that included the BLA and CeA (Figures S3C and S3D).

To further understand gene coexpression relationships across cells at a systems level, we performed weighted gene coexpression network analysis (WGCNA), which defines distinct coexpression modules of correlated genes independent of our cell-type classification (Langfelder and Horvath, 2008). Notably, this unsupervised analysis identified gene coexpression modules that showed specific, strong upregulation in the individual cell types that we have independently defined, supporting our clustering results (Figures 1I and 1J). In addition, gene ontology analysis confirmed that these modules were enriched for genes involved in functional processes closely related to the particular cell types in which they were highly expressed (Figure S2B).

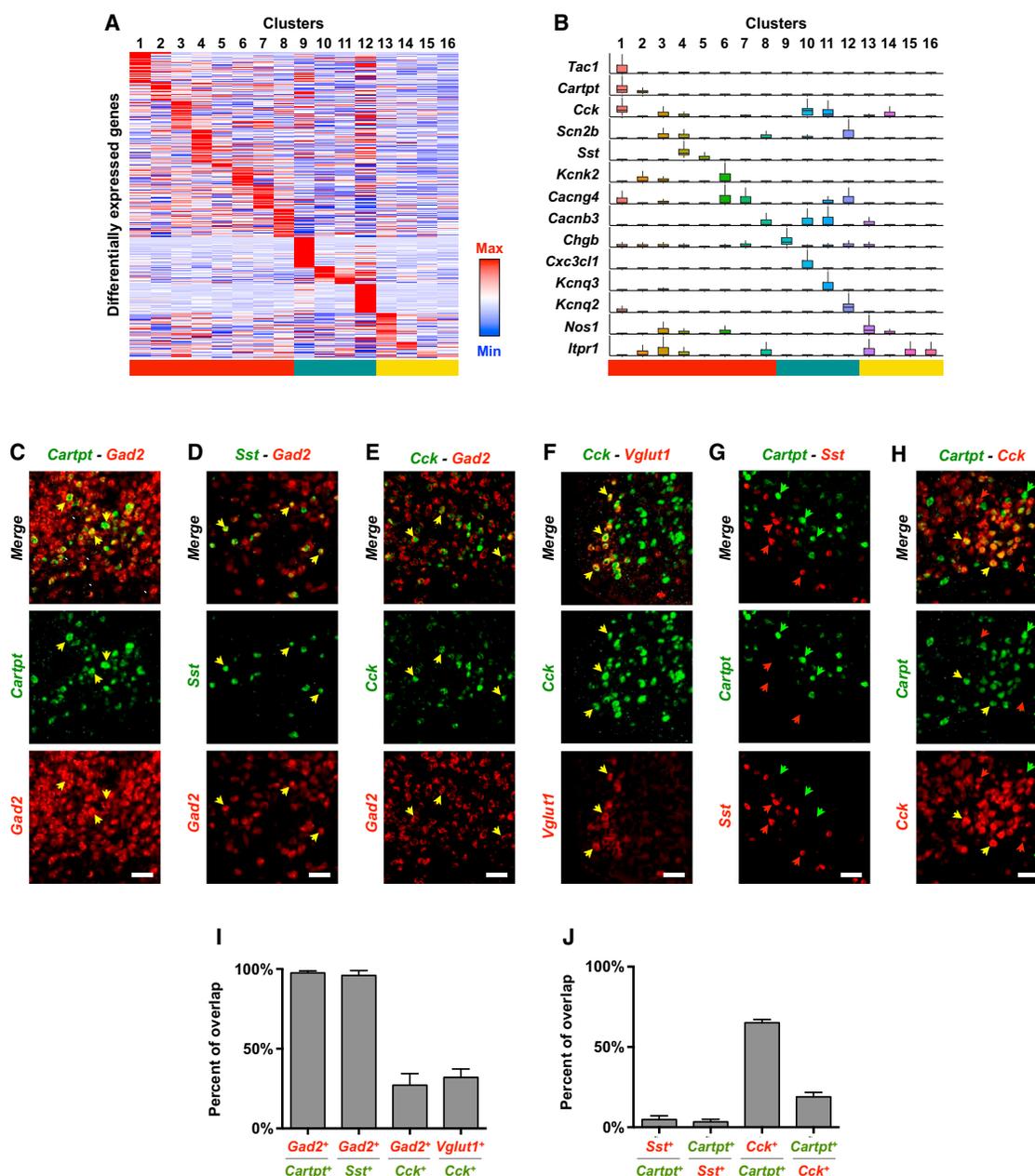
### Neuronal and Astrocyte Heterogeneity in the MeA

To further explore neuronal diversity in the MeA, we performed classification for 2,444 neurons with the highest transcript number in our dataset using the BackSpin algorithm (Zeisel et al., 2015) (STAR Methods). We partitioned the neurons into 16 clusters that could be distinguished by unique sets of markers, including genes encoding neurotransmitter metabolism enzymes and transporters, neuropeptides, receptors, ion channels, and transcription factors (Figures 2A and 2B; Figures S4A and S4B; Table S2). Clusters 1–8 expressed higher levels of GABAergic markers (*Gad1*, *Gad2*, and *Slc32a1*), clusters 9–12 expressed higher levels of glutamatergic markers (*Slc17a6* or *Slc17a7*), and cluster 13 was enriched for *Nos1* (nitric oxide synthase 1) (Figure S4B; Figure 2B). Several neuropeptides, including NPY (neuropeptide Y), CCK (cholecystokinin), substance P, CART peptides, and SST (somatostatin), have been implicated in the regulation of emotional processes by the amygdala (Ebner and Singewald, 2006; Griebel and Holsboer, 2012; Lin and Sibille, 2015; Pérez de la Mora et al., 2007; Reichmann and Holzer, 2016; Rogge et al., 2008; Rotzinger and Vaccarino, 2003). Nonetheless, the distribution of various neuropeptides in different MeA neuronal populations is poorly documented. We found that many neuropeptides exhibited distinct expression profiles across MeA neuronal clusters. To highlight a few, *Tac1* (encoding substance P) was enriched in cluster 1, and *Cartpt*

(encoding CART peptides) was enriched in clusters 1 and 2 (Figure 2B); *Sst* was predominantly expressed in clusters 4 and 5, and was largely non-overlapping with *Tac1* or *Cartpt* (Figure 2B). *Cck* overlapped with *Tac1* and *Cartpt* in cluster 1 and was also expressed in clusters 3, 10, and 11 (Figure 2B). In addition, we found several transcription factors with cluster restricted expression patterns. For example, cluster 9 specifically expressed *Lhx1*, *Lhx5*, and *Nhlh2*, whereas clusters 10 and 11 expressed *Mef2c* and *Bcl11b* (data not shown).

To further validate the observations from Drop-seq, we performed dual-color FISH. Costaining for *Cartpt* or *Sst* and *Gad2* revealed expression of *Cartpt* and *Sst* largely in subsets of *Gad2*<sup>+</sup> neurons (Figures 2C, 2D, and 2I). On the other hand, costaining for *Cck* and *Gad2* or *Slc17a7* confirmed that *Cck* was expressed in subsets of both *Gad2*<sup>+</sup> and *Slc17a7*<sup>+</sup> neurons (Figures 2E, 2F, and 2I). In addition, *Sst* and *Cartpt* were largely mutually exclusive, whereas *Cck* and *Cartpt* signals showed partial overlap (Figures 2G, 2H, and 2J). These results agreed well with our Drop-seq data. We further assessed the spatial distribution of different neuronal subpopulations across different anatomical subdivisions of the MeA using *in situ* hybridization data from the Allen Mouse Brain Atlas (Lein et al., 2007) and observed some degree of spatial separation between several subpopulations. In particular, we found that *Cartpt*<sup>+</sup> cells were enriched in the posterodorsal MeA (MeApd) (Figures S4C–S4F), whereas *Sst*<sup>+</sup> cells showed a reduced density in the MeApd but were broadly distributed in all other MeA subdivisions (Figures S4O–S4R). *Cck*<sup>+</sup> cells were mostly localized in the anteroventral MeA (MeAav) and the MeApd (Figures S4G–S4J), and *Nos1*<sup>+</sup> cells were enriched in the MeAav and the MeApv (Figures S4K–S4N). Taken together, our analysis provides the first systematic transcriptional survey of MeA neurons, revealing aspects of molecular heterogeneity among these neurons that were largely unappreciated before.

Astrocytes play important roles in the regulation of neuronal development and function as well as immune response in the central nervous system (Clarke and Barres, 2013; Sofroniew, 2015). In our dataset, astrocytes comprised the second largest cell population (Figures 1E and 1H). Using Louvain-Jaccard clustering, we further divided the MeA astrocytes into three distinct subclasses (Figure 3A), which we refer to as AS1, AS2, and AS3 hereinafter. While all three clusters shared a number of common astrocyte markers, such as *Gja1*, *Lfng*, and *Agt*, many genes showed higher expression in one population relative to the others (Figures 3B and 3C). For example, *Nkx6-2*, *Sfrp5*, and *Luzp2* were selectively enriched in AS1, while *Myoc*, *Cidea*, and *Gfap* were selectively enriched in AS3 (Figure 3C). AS2 expressed marker genes that partially overlapped with AS1 and AS3 but clearly showed a distinct transcriptional profile from AS1 and AS3. Relative to AS3, AS2 showed the expression of a subset of AS1-expressed genes, such as *Slc7a10*, *Mfge8*, and *Gria2*, but not others, such as *Nkx6-2*, *Sfrp5*, and *Luzp2* (Figures 3B and 3C). AS2 also highly expressed a subset of AS3-expressed genes, such as *Id3*, *Lcat*, and *Id1*, but not others, such as *Myoc* and *Cidea* (Figures 3B and 3C). Consistently, WGCNA performed on astrocytes identified one gene coexpression module (M1) significantly upregulated in AS1 and one module (M2) strongly upregulated in AS3 (Figure 3D). Top enriched gene



### Figure 2. Neuronal Subtypes in the MeA

(A) Heatmap showing expression level of select top genes enriched (fold change > 1.5, FDR < 0.05) in each neuronal subcluster. Expression level is averaged within each subcluster.

(B) Boxplots showing the distribution of expression level of representative subtype markers across 16 neuronal subclusters.

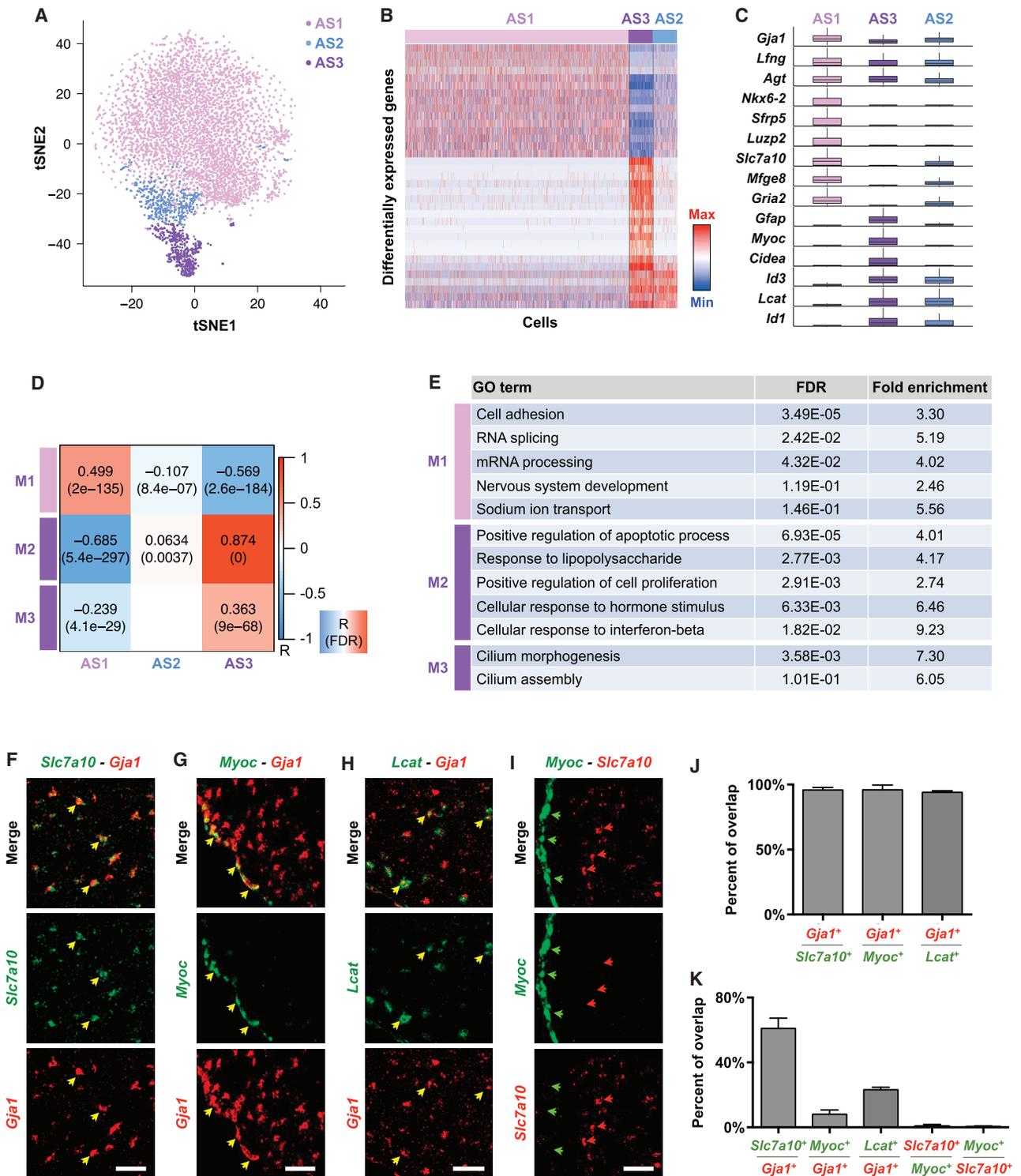
(C–H) Dual-color FISH (dFISH) for select neuropeptide (*Cck*, *Cartpt*, and *Sst*) and neurotransmitter (*Gad2* and *Vglut1*) markers differentially expressed between neuronal subclusters. Yellow, green, and red arrowheads indicate examples of cells positive for both probes, only the green probe, and only the red probe, respectively. Scale bars, 50  $\mu$ m. (C) *Cartpt* and *Gad2*; (D) *Sst* and *Gad2*; (E) *Cck* and *Gad2*; (F) *Cck* and *Vglut1*; (G) *Cartpt* and *Sst*; (H) *Cartpt* and *Cck*.

(I–J) Quantification of percentage overlap between cells labeled by different markers in dFISH. (I) Overlap between neurotransmitter and neuropeptide markers; (J) overlap between different neuropeptide markers. Mean  $\pm$  SD,  $n \geq 3$  mice. See also Figure S4 and Table S2.

ontology terms for M1 included cell adhesion, RNA processing, nervous system development, and sodium transport, whereas M2 was enriched for genes implicated in regulation of apoptosis, cell proliferation, and cellular responses to stimuli (Figure 3E),

suggesting possible functional differences between the astrocyte subpopulations.

We further performed dual-color FISH for the subtype-specific markers *Slc7a10*, *Lcat*, or *Myoc* and the common astrocyte



**Figure 3. Astrocyte Heterogeneity in the MeA**

(A) Two-dimensional tSNE visualization of three astrocyte subclusters (AS1, AS2, and AS3) in the MeA. Individual dots correspond to single cells, colored by astrocyte subtypes.

(B) Heatmap showing expression level of select top marker genes (fold change > 2, FDR < 0.05) of three astrocyte subclusters.

(C) Boxplots showing the distribution of expression level of representative common or subtype markers for astrocytes.

(D) Heatmap showing Pearson correlation between astrocyte subtypes and the eigengenes of gene coexpression modules defined by WGCNA in all astrocytes.

(legend continued on next page)

marker *Gja1*. Consistent with the Drop-seq data, *Gja1* was expressed in almost all cells in the *Slc7a10*<sup>+</sup>, *Myoc*<sup>+</sup>, and *Lcat*<sup>+</sup> populations (Figures 3F–3H and 3J). The proportion of *Gja1*<sup>+</sup> cells marked by each of the subtype markers overall matched the proportion in the Drop-seq data (Figures 3B and 3K). We also confirmed that *Slc7a10* (marking AS1 and AS2) and *Myoc* (marking AS3) were largely mutually exclusive (Figures 3I and 3K). Moreover, the subtype markers also showed spatial separations; the AS3 astrocytes (*Myoc*<sup>+</sup>) were restricted to a superficial layer that possibly represents cells that form the glia limitans (Figures 3G and 3I), whereas AS1 and AS2 (*Slc7a10*<sup>+</sup>) were localized in the parenchyma (Figures 3F and 3I). AS2 astrocytes, which were also *Lcat*<sup>+</sup>, appeared to be sparsely distributed in the parenchyma (Figure 3H). Together, these observations suggest astrocyte heterogeneity in the MeA.

### Act-Seq Minimizes Artificially Induced Spurious Gene Expression during Single-Cell Preparation

Identifying specific neuronal populations activated by a certain behavior/experience is a key step in the functional dissection of neural circuits. We attempted to adapt scRNA-seq for profiling acute transcriptional changes (including induction of IEGs) following behaviors/experiences or other physiological stimuli, in order to provide an unbiased, high-throughput tool for the identification of activated cells with single-cell resolution. However, the conventional whole-cell dissociation method with protease treatment has been shown to cause strong IEG upregulation, even to a level comparable to seizure-induced, large-scale neuronal activation (Kolodziejczyk et al., 2015; Lacar et al., 2016; Poulin et al., 2016). This not only precludes detection of physiological stimuli-induced IEG expression, but also could confound analysis of baseline transcriptional profiles.

To develop a dissociation method suitable for detecting acute, endogenous transcriptional alterations, we applied a well-established, general transcription inhibitor, actinomycin D (ActD), during dissociation. ActD inhibits transcription mediated by all three eukaryotic RNA polymerases and provides wide-spectrum, fast transcriptional inhibition with little and slow reversibility (Bensaude, 2011). We performed single-cell dissociation from the amygdala using the conventional or our modified procedure and evaluated the effects of the modified procedure on curbing dissociation-induced IEG upregulation by characterizing single-cell transcriptomes for all major cell types using Drop-seq and unbiased cell clustering. We then compared the expression of a curated collection of 139 previously identified IEGs (Table S3) under different experimental conditions. Consistent with a previous report (Lacar et al., 2016), following conventional dissociation procedures with digestion at 34°C, we detected substantial upregulation of many IEGs, such as *Fos*, *Egr1*, *Fosb*, *Nr4a1*, and *Dusp1*, among all cells in comparison to our

modified methods (Figures 4A–4E and 4J). In addition, we observed a marked increase in the proportion of IEG-positive cells in all major cell types, most prominently microglia, endothelial cells, and mural cells (Figure 4F; Table S4). In contrast, with ActD application, many IEGs showed significantly reduced induction among all cells (Figures 4A–4E and 4J). Moreover, the percentage of IEG-positive cells in all major cell types also significantly decreased (Figures 4F and 4G; Table S4). Nonetheless, an appreciable proportion of IEG-positive endothelial and mural cells still persisted (Figure 4G; Table S4).

We further examined the effect of lowering protease digestion temperatures on the expression of IEGs. Decreasing the digestion temperature from 34°C to 22°C or 11°C in the presence of ActD further helped suppress IEG induction in endothelial and mural cells, leading to minimal IEG induction in all cell types (Figures 4A–4E, 4H, 4I, 4L, and 4N; Table S4). Taken together, these results indicate that the transcription inhibitor ActD alone is effective in inhibiting dissociation-induced IEG expression in all cell types except for endothelial and mural cells and that combining it with lower digestion temperature can block artificial IEG induction in all cell types.

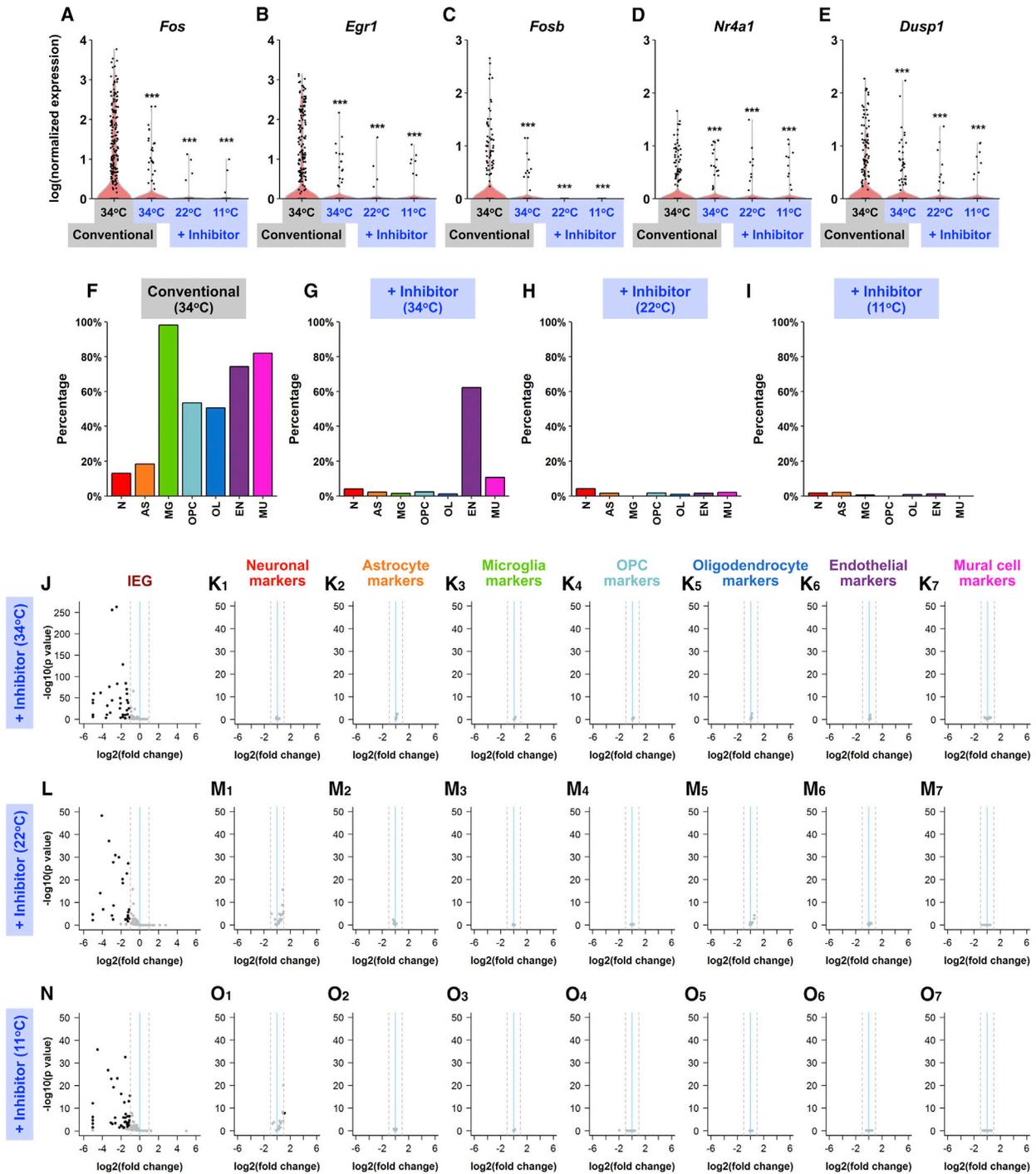
To examine whether our modified dissociation methods affected the detection of genes other than IEGs, we assessed the expression of the top 20 markers for each major cell type. Compared to the conventional method at 34°C, the modified methods did not lead to significant change in any of these markers at 34°C, 22°C, or 11°C, except for a moderate increase in one neuronal marker at 11°C (Figures 4K, 4M, and 4O). Further analysis of all the other ~15,000 genes in our dataset similarly showed minimal changes (Table S4). These results suggested that ActD application and lower digestion temperature caused limited changes in the detection of genes other than the IEGs compared with the conventional dissociation method.

To further evaluate the potential confounding effects of artificial IEG induction on cell clustering, we examined microglia, which showed the strongest IEG upregulation among all cell types when dissociated using the conventional method at 34°C. We first performed PCA and tSNE analysis using all genes in the dataset, including the IEGs, and found that microglia dissociated using the conventional protocol separated into two distinct clusters (Figure 5A), with one showing notably higher IEG level than the other (Figure 5D). In contrast, microglia isolated with our modified methods, which showed low IEG expression, intermingled in a single cluster (Figures 5B and 5C). This suggested that the split of microglia into two clusters was caused by the artificial induction of IEGs. Consistent with this notion, excluding IEGs from the clustering analysis removed this effect, leading all microglia to merge into a single cluster (Figures 5E–5G). As the effects of the conventional dissociation

(E) Top gene ontology terms (FDR < 0.15) enriched in each gene coexpression module in astrocytes.

(F–I) dFISH for a common astrocyte marker (*Gja1*) and three astrocyte subtype markers, *Slc7a10* (enriched in AS1 and AS2), *Myoc* (enriched in AS3), and *Lcat* (enriched in AS2 and AS3). (F) *Slc7a10* and *Gja1*; (G) *Myoc* and *Gja1*; (H) *Lcat* and *Gja1*; (I) *Myoc* and *Slc7a10*. Yellow, green, and red arrowheads indicate examples of cells positive for both probes, only the green probe, and only the red probe, respectively. Scale bars, 50 μm.

(J and K) Quantification of percentage overlap between cells labeled by different astrocyte markers in dFISH. (J) The percentage of *Gja1*<sup>+</sup> cells in *Slc7a10*<sup>+</sup>, *Myoc*<sup>+</sup>, and *Lcat*<sup>+</sup> cells; (K) the percentage of *Slc7a10*<sup>+</sup>, *Myoc*<sup>+</sup>, and *Lcat*<sup>+</sup> cells in *Gja1*<sup>+</sup> cells and the overlap between *Myoc*<sup>+</sup> and *Lcat*<sup>+</sup> cells. Mean ± SD, n = 2 mice for *Lcat/Gja1* dFISH, n = 3 mice for all other groups. See also Table S2.

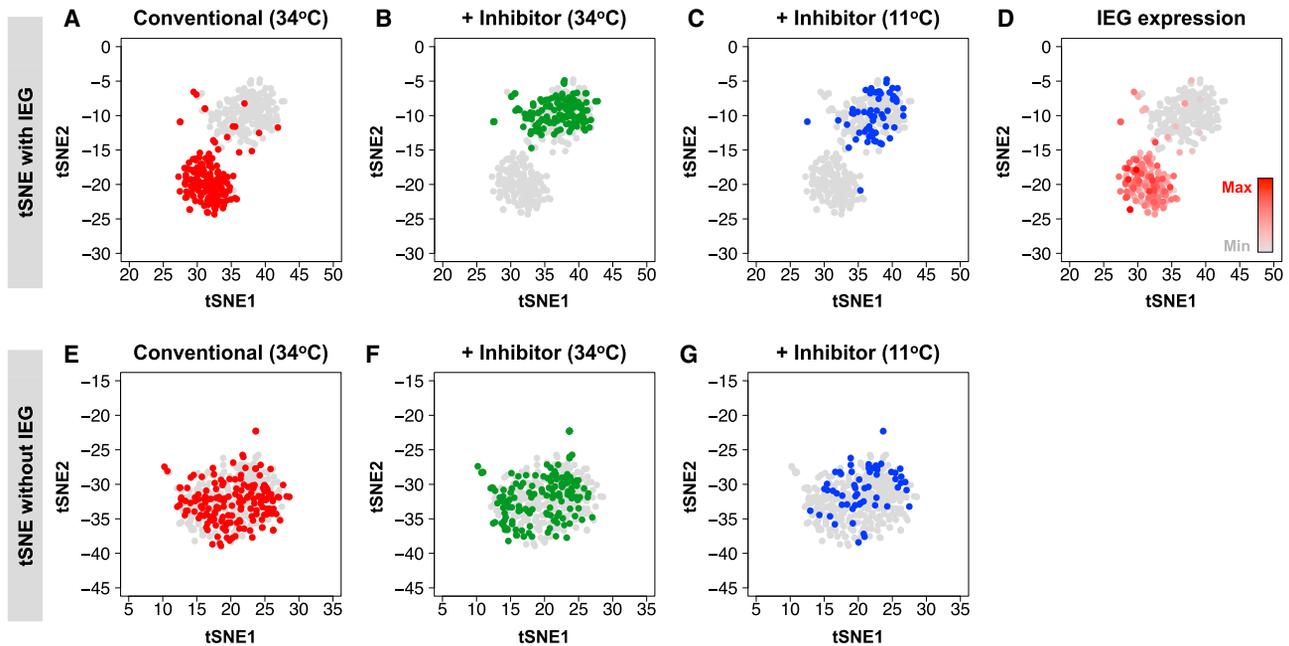


**Figure 4. Act-Seq Suppresses Artificially Induced Gene Expression during Single-Cell Dissociation**

(A–E) Violin plots showing the distribution of the expression level of representative IEGs, including *Fos* (A), *Egr1* (B), *Fosb* (C), *Nr4a1* (D), and *Dusp1* (E), under conventional or modified dissociation conditions. Dots show expression value in individual cells. 400 cells were randomly sampled for each condition to show the same total number of cells across conditions. “+ inhibitor” indicates application of the transcription inhibitor ActD. \*\*\* $p < 10^{-10}$ , one-sided Fisher’s exact test comparing conventional dissociation method with modified methods.

(F–I) Bar plots showing the percentage of IEG-positive cells (STAR Methods) in each major cell type with different single-cell preparation procedures. (F) Conventional dissociation method with 34°C digestion; (G–I) dissociation in the presence of ActD with 34°C (G), 22°C (H), or 11°C (I) digestion.

(J–O) Scatterplots showing fold change in the percentage of cells expressing different IEGs (J, L, and N) or 20 top markers (K, M, and O) for each major cell type along with FDR-adjusted p values obtained from Fisher’s exact test. Comparisons are made between modified and conventional dissociation conditions. Black dots, genes with fold change  $>2$  and  $FDR < 0.05$ . In (J), (L), and (N), fold changes  $<2^{-5}$  are shown as  $2^{-5}$ . See also Tables S3 and S4.



**Figure 5. Confounding Effects of Artificially Induced Gene Expression on Baseline Cell Clustering**

(A–D) Two-dimensional tSNE visualization of microglia. Clustering analysis was performed using all genes in the dataset, including IEGs. Microglia dissociated using the conventional dissociation method split into two distinct clusters (A; red dots), with one showing substantially higher IEG level than the other (D). In contrast, microglia isolated in the presence of ActD either at 34°C or 11°C intermingle in a single cluster (B and C; green and blue dots, respectively) that shows low IEG expression (D).

(E–G) Clustering analysis was performed after excluding IEGs from the dataset. Microglia dissociated using the conventional dissociation method (E) or in the presence of ActD either at 34°C (F) or 11°C (G) merge into a single cluster in the tSNE map.

protocol on cells can be variable and hard to gauge between experiments, it is conceivable that the resultant variability in transcriptional perturbations could create misleading effects similar to what we observed here, which may confound the definition of cell types. Therefore suppressing artificially induced gene expression can also be important for accurately quantifying the baseline transcriptional profile and assigning cell types.

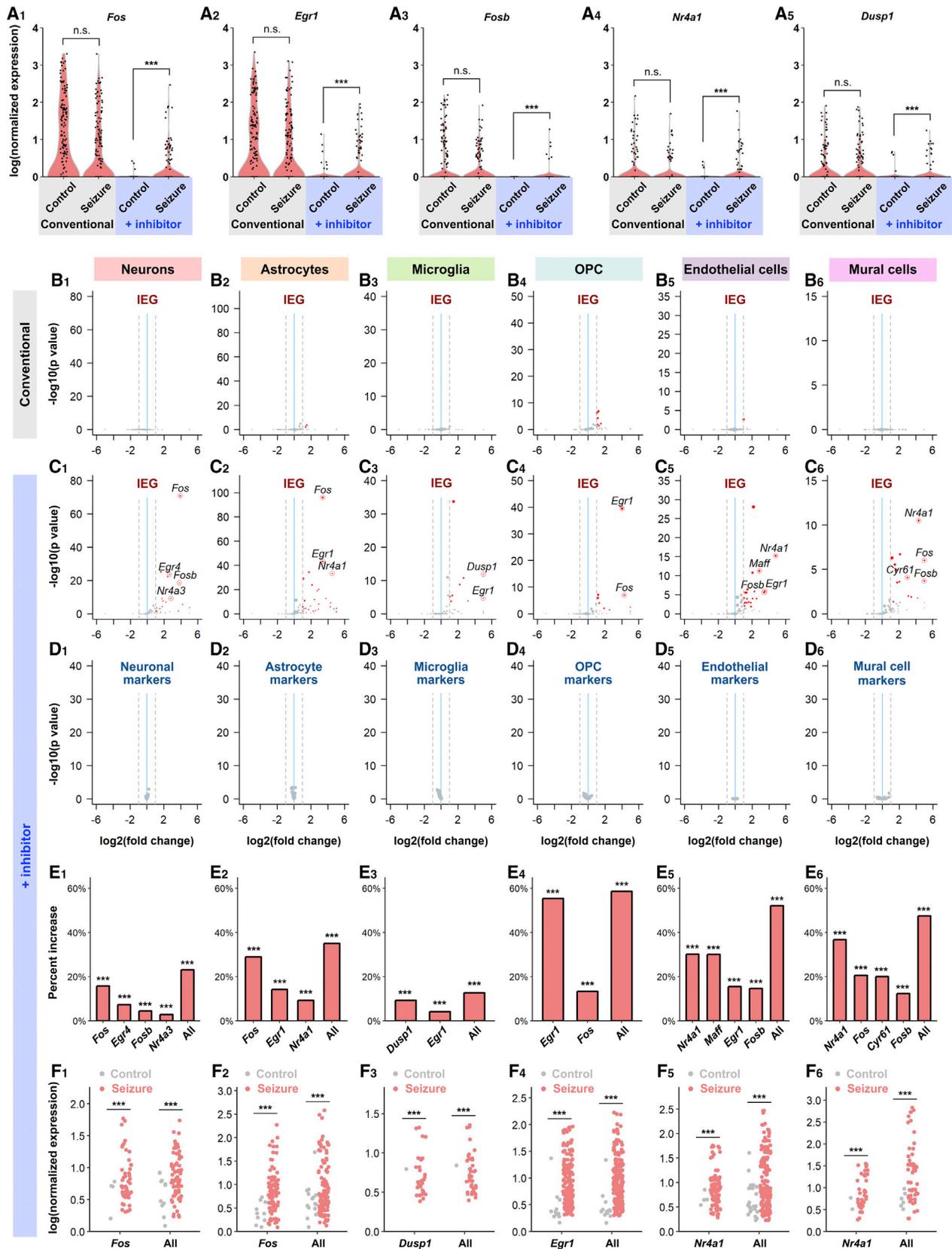
Together, these results demonstrate that our new single-cell preparation method effectively suppresses artificially induced IEG expression during whole-cell isolation. Combining this method with scRNA-seq, which we named Act-seq (activated cell population sequencing), allowed us to perform classification of cell types in the MeA (Figures 1, 2, and 3) and to further evaluate behavior and/or experience-associated acute IEG induction in distinct cell types.

### Act-Seq Robustly Detects Seizure-Induced IEG Expression in the MeA

To further test whether Act-seq improves on the conventional dissociation method in detecting stimulus-induced IEG expression *in vivo*, we compared the ability to detect IEG expression in response to extensive neuronal activation elicited by seizure between these methods. We induced seizure by injecting the convulsant pentylenetetrazole (PTZ), a GABA(A) receptor antagonist (Lacar et al., 2016; Yount et al., 1994). Mice displayed convulsions and freezing ~5 min following the injection. *In situ* hybridization confirmed robust elevation of *Fos* mRNA in the

MeA following seizure compared with control mice (Figures S5A–S5D). We dissociated and sequenced single cells from the MeA using Act-seq or the conventional method and comprehensively surveyed the curated list of 139 previously identified IEGs (Table S3). Since our data indicated that the expression of IEGs may confound cell-type classification (Figure 5), we excluded IEGs from our clustering analysis, which enabled the classification of all major cell types and subtypes independent of activity induced IEG expression. Following clustering, we further analyzed the expression of IEGs after seizure in different cell types.

In agreement with previous findings, the conventional method elicited strong induction of many IEGs among all cells in control animals (Lacar et al., 2016). Under this condition, the control animals showed a comparable level of IEG induction to the seizure-affected animals, thus precluding the detection of IEG changes induced by seizure (Figures 6A). Analysis of expression of all IEGs in individual major cell types identified no or very few IEGs significantly upregulated in each cell type when comparing seizure to control group (Figure 6B). In contrast, IEG expression among all cells in control animals was drastically lower with Act-seq than with the conventional method, thus enabling robust detection of IEG upregulation following seizure (Figures 6A). Furthermore, we found that many IEGs were significantly increased following seizure in multiple major cell types, including neurons, astrocytes, microglia, OPCs, endothelial cells, and mural cells (Figures 6C; Table S5). These results clearly show



(legend on next page)

the capability of Act-seq to detect stimulus-induced IEG expression, which was impossible with the conventional method.

Interestingly, different cell types appeared to display distinct IEG activation profiles following seizure, with different IEGs showing varying magnitude of changes across cell types (Table S5). For example, *Dusp1*, *Nr4a1*, and *Fos* showed the largest fold change among all IEGs in microglia, endothelial cells, and mural cells, respectively (Figures 6C; Table S5). *Cyr61* was among the top up-regulated IEGs in mural cells but did not show any increase in microglia or OPCs (Figures 6C; Table S5). These results demonstrate the utility of our method in systematically screening for the most robust IEG(s) in different cell types.

We next used the most highly induced IEGs in each cell type as activation indicators to assess the level of activation in different cell types. We found that significantly higher proportions of neurons, astrocytes, microglia, OPC, endothelial, and mural cells were IEG positive in seizure-affected mice compared to controls (Figures 6E). In addition to being expressed in larger proportions of cells, these IEGs also showed significantly elevated expression levels after seizure within single cells across different cell types (Figures 6F). Changes in the morphology, proliferation, and/or gene expression of glial cells and endothelial cells following seizure have been implicated in the pathophysiology of epilepsy (Fabene et al., 2008; Luo et al., 2015; Rakhade and Jensen, 2009). However, the acute induction of IEGs in these cells shortly after seizure has not been well characterized. Our results revealed the acute IEG changes that might be important *in vivo* indicators and/or mediators of downstream cellular responses in glial cells and endothelial cells.

We further examined potential expression changes in genes other than IEGs following seizure. Analysis for the top 50 markers for each major cell type revealed no genes significantly differentially expressed, consistent with the absence of overall cell-type alterations following acute seizure (Figures 6D). Furthermore, analysis of all the other ~15,000 genes in the dataset identified only a few genes showing significant differential expression in each major cell type (Figures S6G–S6L). GO analysis did not reveal any functional enrichment among these genes. These results suggested that the most prominent gene expression changes acutely following seizure were mostly in IEGs.

We next examined the activation of different neuronal subtypes. Although different sets of IEGs were induced in different major cell types, we found that a combination of four IEGs, *Fos*, *Fosb*, *Egr4*, and *Nr4a3*, was consistently among the most strongly elevated IEGs across all MeA neuronal subclusters (Table S6) and thus provided a common metric that allowed us to gauge and compare activation across different subclusters. In agreement with large-scale neuronal activation associated with seizure, we observed a significant increase in the proportion of IEG-positive cells and IEG expression levels in the majority of neuronal subtypes (Figure 7A). Consistent with the notion that acute cell activation did not drive cell-type assignment, none of the neuronal subtype markers (Figure 2B) or neuronally enriched genes used for neuronal subclustering were significantly changed following seizure (Figure 7E, Figures S6M and S6N). We also assessed the activation of astrocyte subtypes using the three top IEGs upregulated among all astrocytes following seizure (*Fos*, *Egr1*, and *Nr4a1*; Figure 6C<sub>2</sub>). We found that in all three astrocyte subpopulations, the proportion of cells expressing IEGs and the expression level of IEGs also showed notable increases (Figures 7I and 7K–7M; Table S6). In particular, AS3 displayed a higher percentage of IEG-positive cells and higher IEG expression levels compared to AS1 and AS2 (Figures 7I and 7K–7M). In contrast, key marker genes for the three astrocyte subpopulations (Figures 3B and 3C) exhibited no significant difference between seizure and control groups (Figure 7J). Collectively, these results demonstrate the ability of Act-seq to robustly detect endogenous IEG induction in response to the activation of various brain cell types.

### Act-Seq Identifies Neuronal Subpopulations Preferentially Activated by Acute Stress

We next applied Act-seq to evaluating IEG expression changes induced by more moderate neuronal activity in response to a more behaviorally relevant stimulus. A number of previous studies have implicated the MeA in the regulation of stress response and have shown elevated IEG expression in the MeA following acute stressors such as restraint (Crane et al., 2005; Ons et al., 2010; Solomon et al., 2010). However, the identity of the activated cells remains largely unknown. To shed light on this question, we sought to identify MeA neuronal

#### Figure 6. Act-Seq Detects Seizure-Induced IEG Upregulation

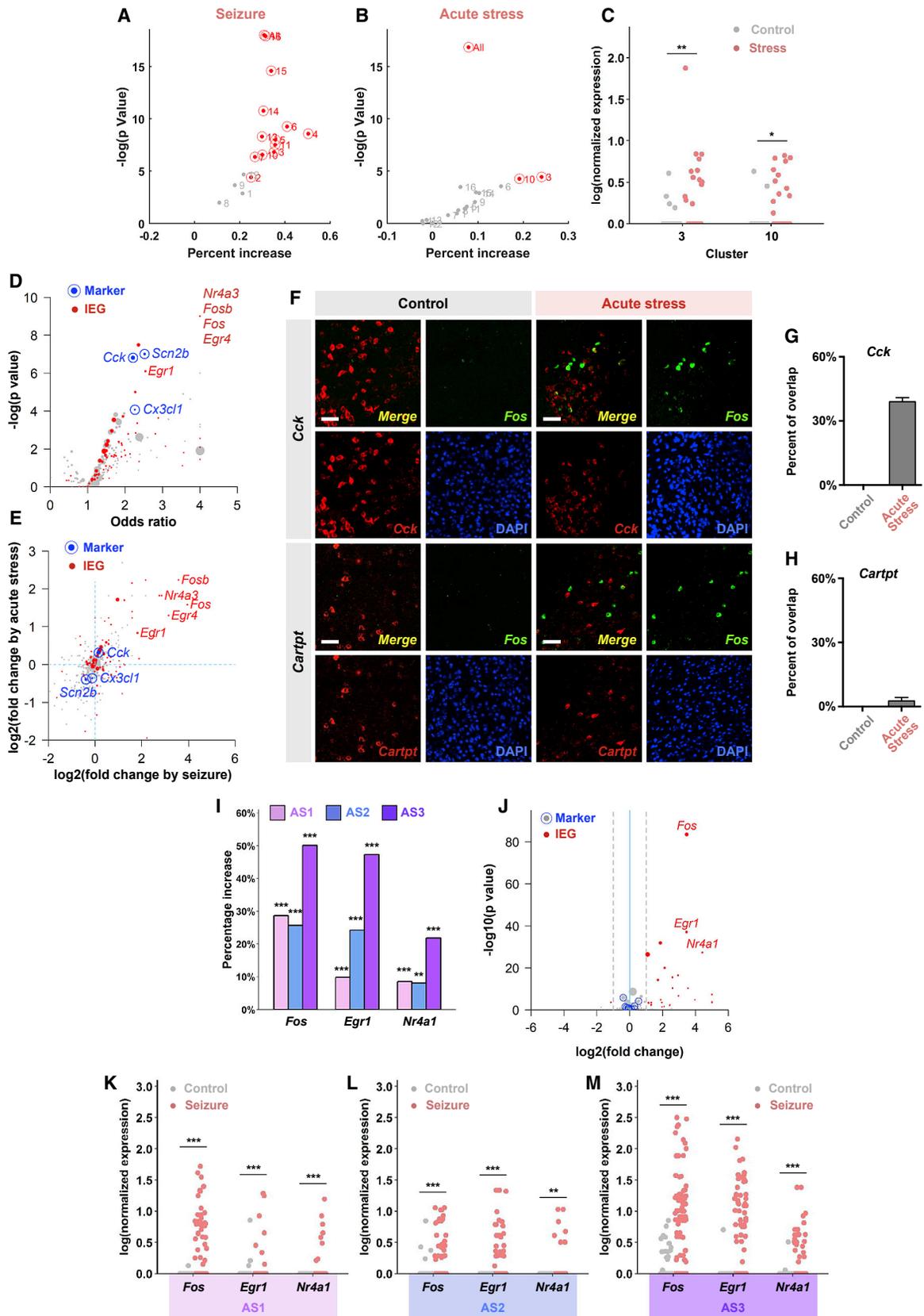
(A<sub>1</sub>–A<sub>6</sub>) Comparison of the performance of the conventional dissociation method and Act-seq in detecting seizure-induced IEG expression. Violin plots show the distribution of the expression level of representative IEGs in cells dissociated from seizure-affected or control animals using the conventional dissociation procedures or Act-seq. Dots show expression value in individual cells. 200 cells were randomly sampled from all cell types for each group to show the same total number of cells across groups. \*\*\* $p < 10^{-10}$ , n.s., non-significant, one-sided Fisher's exact test.

(B<sub>1</sub>–C<sub>6</sub>) Scatterplots showing fold change in the percentage of cells expressing different IEGs in different major cell types in seizure-affected animals versus controls and FDR-adjusted  $p$  value from one-sided Fisher's exact test. Individual dots represent single IEGs. Dot size is proportional to the percentage of cells expressing the corresponding gene in each cell type in seizure-affected animals. Red dots, IEGs with fold change  $>2$  and FDR  $< 0.05$ . In (C<sub>1</sub>)–(C<sub>6</sub>), fold changes  $>2^5$  are shown as  $2^5$ , and top upregulated IEGs are circled and labeled. In (B<sub>1</sub>)–(B<sub>6</sub>), cells were isolated using the conventional dissociation method. In (C<sub>1</sub>)–(C<sub>6</sub>) cells were isolated using Act-seq.

(D<sub>1</sub>–D<sub>6</sub>) Scatterplots showing fold change in the percentage of cells expressing the top 50 markers for each major cell type in seizure-affected animals versus controls and FDR-adjusted  $p$  value from two-sided Fisher's exact test. No marker was significantly differentially expressed with fold change  $>2$  and FDR  $< 0.05$ .

(E<sub>1</sub>–E<sub>6</sub>) Increase in the percentage of cells (percentage in seizure group minus percentage in control group) expressing top upregulated IEGs indicated on the x axis. "All," positive cells defined as expressing any of the individual top upregulated IEGs. \*\*\* $p < 10^{-4}$ , one-sided Fisher's exact test.

(F<sub>1</sub>–F<sub>6</sub>) Distribution of the expression level of top regulated IEGs in control (gray dots) and seizure-affected animals (red dots). "All," the sum of the individual IEGs indicated in (E<sub>1</sub>–E<sub>6</sub>). \*\*\* $p < 10^{-10}$ , likelihood-ratio test (STAR Methods). The plots show 300 (E<sub>1</sub>–E<sub>6</sub>) or 100 (E<sub>6</sub>) cells randomly sampled from each of the control and seizure groups in each cell type in order to show the same total number of cells in each control/seizure comparison. See also Figures S5 and S6 and Table S5.



(legend on next page)

subpopulation(s) activated by acute stress using Act-seq. Consistent with previous reports, *in situ* hybridization confirmed an increase in *Fos* mRNA level in the MeA following 45 min of immobilization stress, albeit to a much lesser extent compared with PTZ treatment (Figures S5E–S5H), as expected. We then performed single-cell isolation and sequencing using Act-seq and observed significant elevation of IEG levels in the whole population of MeA neurons (“All” in Figure 7B; Table S7), demonstrating the sensitivity of Act-seq. On the other hand, no significant alterations were observed for top major cell type markers following acute stress (Figures S7A–S7F). We next compared IEG expression between stressed and control groups for individual neuronal subclusters and found that clusters 3 and 10 showed the most prominent IEG upregulation. Both clusters exhibited significantly increased proportions of neurons expressing IEGs along with elevated IEG expression level within single cells (Figures 7B and 7C), suggesting that these two subpopulations were preferentially activated.

To independently verify the preferential activation of specific neuronal populations, we also examined the enrichment of different genes in all IEG-positive neurons regardless of cell-type classification. This analysis identified the genes *Cck*, *Cx3cl1* (encoding a neuronal transmembrane protein involved in neuron-glia interaction), and *Scn2b* (encoding a neuronal voltage-gated sodium channel subunit) among the most highly enriched non-IEG genes in IEG-positive neurons (Figure 7D). *Cck*, *Cx3cl1*, and *Scn2b* appeared to be the top genes preferentially expressed in clusters 10 (*Cck* and *Cx3cl1*) and 3 (*Cck* and *Scn2b*) (Figure 2B). Since this gene enrichment analysis was independent of any prior classification of neuronal subtypes, identification of *Cck* and two other key marker genes further demonstrated that the corresponding neuronal subtypes were indeed preferentially activated. Of note, expression of neuronal subtype markers (Figure 2B) and neuronally enriched genes used for neuronal subclustering exhibited no significant change in response to immobilization stress, in contrast to the IEGs (Fig-

ure 7E; Figures S7G and S7H), suggesting that the classification of neuronal subtypes was unlikely driven by the experimental treatments.

We further sought to validate the preferential activation of a *Cck*<sup>+</sup> subpopulation by performing dual-color FISH for *Cck* and *Fos* following immobilization stress. We observed a marked increase in the overlap between *Cck* and *Fos* in response to stress; ~40% of *Fos*-positive cells were *Cck* positive, a ratio that was significantly higher than by chance (Figures 7F and 7G; odds ratio = 1.96,  $p = 1.9E-07$ , two-sided Fisher’s exact test). These results were consistent with the activation of cluster 3 neurons, which expressed elevated level of *Fos* following acute stress based on scRNA-seq (Table S7). To test whether the overlap between *Fos* and *Cck* was specific, we also assessed the overlap between *Fos* and *Cartpt*, which showed partial overlap with *Cck* but was enriched in two neuronal subclusters not activated by acute stress according to our scRNA-seq data (Figures 2B and 7B). We found that only ~3% *Fos*-positive cells were *Cartpt* positive, a ratio that was even lower than by chance (odds ratio = 0.70,  $p = 0.049$ , two-sided Fisher’s exact test). These results further supported our finding from scRNA-seq that *Cck*-positive neuronal subpopulations were preferentially activated following acute stress. The observation of preferential activation of *Cck*<sup>+</sup> neurons by acute stress is particularly interesting given the previously reported role of CCK in the modulation of anxiety-related behaviors (Griebel and Holsboer, 2012; Pérez de la Mora et al., 2007; Rotzinger and Vaccarino, 2003). Taken together, these results demonstrate the power of Act-seq to associate acute expression changes in specific cell types with a behavior and/or experience.

## DISCUSSION

In this study, we have developed Act-seq, which eliminates artificially evoked transcriptional perturbations and enables faithful detection of both baseline transcriptional profiles and acute

### Figure 7. Activation of Neuronal and Astrocyte Subtypes by Physiological Stimuli

(A–H) Activation of MeA neuronal subtypes by seizure and acute immobilization stress. (A and B) Increase in the percentage of IEG-positive cells in all neurons (“All”) or individual neuronal subtypes in seizure- (A) or acute stress-affected (B) animals versus controls along with p value from one-sided Fisher’s exact test. Red dots with circles, neuronal populations with significant increase in both the percentage of IEG-positive cells and expression level of IEGs ( $p < 0.05$ , Fisher’s exact test and  $p < 0.05$ , likelihood-ratio test). (C) Distribution of IEG expression level in cluster 3 and cluster 10 neurons in control (gray dots;  $n = 46$  and  $38$  cells, respectively) and stressed animals (red dots;  $n = 35$  and  $49$  cells, respectively). The plot shows 35 and 49 cells randomly sampled from each of the control and acute stress groups in clusters 3 and 10, respectively, in order to show the same total number of cells in each control/acute stress comparison.  $^{**}p < 0.005$ ,  $^{*}p < 0.05$ , likelihood-ratio test. (D) Overlap between IEG-positive neurons (expressing any of *Fos*, *Fosb*, *Egr4*, or *Nr4a3*) and neurons expressing individual IEGs or neuronally enriched genes used for neuronal subclustering in stressed animals. Odds ratio and p value from Fisher’s exact test are shown. Odds ratios  $> 4$  are shown as 4 and p values  $< 10^{-9}$  are shown as  $10^{-9}$ . Red dots, IEGs. Blue dots with circles, top overlapping non-IEG genes. (E) Fold change in the percentage of neurons expressing IEGs or neuronally enriched genes used for neuronal subclustering in seizure- or acute stress-affected animals versus controls. Fold changes  $> 2^5$  are shown as  $2^5$ . In (D) and (E), dot size is proportional to the percentage of neurons expressing the corresponding gene in stressed animals. (F) dFISH for *Fos* (green) and *Cck* or *Cartpt* (red) in control or stressed animals. (G and H) Quantification of percentage overlap between *Fos*<sup>+</sup> cells and *Cck*<sup>+</sup> (G) or *Cartpt*<sup>+</sup> (H) cells in control or stressed animals. Mean  $\pm$  SD.  $n \geq 4$  mice for all groups. (I–M) Activation of MeA astrocyte subtypes by seizure. (I) Increase in the percentage of AS1, AS2, and AS3 astrocytes expressing the indicated IEGs in seizure-affected animals versus controls.  $^{***}p < 10^{-5}$ ,  $^{**}p < 0.005$ , one-sided Fisher’s exact test. (J) Fold change in the percentage of astrocytes expressing different genes in seizure-affected animals versus controls and FDR-adjusted p value from Fisher’s exact test. Dot size is proportional to the percentage of astrocytes expressing the corresponding gene in seizure-affected animals. Red dots, IEGs significantly upregulated with fold change  $> 2$  and FDR  $< 0.05$ . Fold changes  $> 2^5$  are shown as  $2^5$ . Gray dots, astrocyte subtype markers shown in Figure 3B. Gray dots with blue circles, common and key subtype markers for astrocytes shown in Figure 3C. No astrocyte common or subtype marker was significantly differentially expressed with fold change  $> 2$  and FDR  $< 0.05$ . (K–M) Distribution of IEG expression level in AS1, AS2, and AS3 astrocytes in control (gray dots) and seizure-affected animals (red dots). The plots show 100 cells randomly sampled from each group in order to show the same total number of cells across groups.  $^{***}p < 10^{-5}$ ,  $^{**}p < 0.005$ , likelihood-ratio test. See also Figures S5–S7, Table S6, and Table S7.

gene expression changes, such as induction of IEGs. Act-seq enables systematic, high-resolution identification of activated cell populations and the associated transcriptional responses. Using Act-seq we provide the first detailed molecular taxonomy for the MeA and examine the transcriptional responses of MeA neurons and non-neuronal cells to physiological stimuli, revealing cell-type-specific induction of distinct IEGs following seizure and preferential activation of specific neuronal subtypes in response to acute stress.

### Molecular Taxonomy of the Amygdala

The amygdala is a complex structure anatomically comprised of multiple interconnected nuclei and plays a pivotal role in regulating a variety of cognitive and behavioral functions (Janak and Tye, 2015; Swanson and Petrovich, 1998). Although the medial amygdala has been implicated in innate emotional responses and social behaviors (Swanson, 2000), the molecular identity of MeA cells is still poorly characterized, and a comprehensive transcriptional analysis of different neuronal subtypes is still lacking. Here we generated the first detailed molecular taxonomy for different neuronal and non-neuronal cell types in the MeA and provided a rich list of genetic markers for each major cell type. We identified a number of neuronal subtypes differentially expressing genes closely related to neuronal function. In particular, we found that multiple neuropeptides, important modulators or mediators of neurotransmission, displayed distinct distribution among different neuronal subtypes. We also observed differential distribution of several neuronal subtype markers across different anatomical subdivisions of the MeA, suggesting that transcriptional segregation may partially correlate with spatial separation for some neuronal subpopulations. In addition, we uncovered three astrocyte populations with distinct transcriptional profiles and anatomical localization, providing the first evidence for astrocyte cell-type diversity in the amygdala. Our dataset and findings greatly enrich our knowledge on the cell-type composition of MeA and provide an informative starting point for a comprehensive characterization of molecular diversity among MeA cells. As scRNA-seq technology keeps improving to achieve better detection and higher throughput, the clustering and definition of the MeA cell types and neuronal subtypes will likely continue to be further refined. Additional morphological, connectional, and electrophysiological studies could potentially be coupled with scRNA-seq to further characterize neuronal subtypes and to better understand the relationship between transcriptional heterogeneity and diversity in other cellular properties in the MeA (Zeng and Sanes, 2017).

### Faithful Detection of Acute Transcriptional Changes

A major obstacle in detecting acute changes in transcriptional profile using scRNA-seq has been artificially evoked transcriptional perturbations during conventional whole-cell dissociation procedures using protease treatment (Kolodziejczyk et al., 2015; Lacar et al., 2016; Poulin et al., 2016). This can not only mask endogenous acute gene expression changes that are truly elicited by physiological stimuli, but may also confound characterization of baseline transcriptional profiles. As our data suggested, artificially evoked transcriptional alterations could lead to genera-

tion of misleading cell clusters (Figure 5), thus confounding cell-type classification based on baseline transcriptional profile.

Act-seq effectively minimizes artificially induced transcriptional perturbations during single-cell preparation. This method features the application of the general transcription inhibitor ActD, which, compared to the conventional dissociation method, greatly reduces artificial IEG induction in all cell types (except for endothelial cells) that occurs during cell preparation. The differential responses of different cell types to ActD (endothelial cells compared to other cell types) might be due to cell-type differences in cell permeability to ActD and/or sensitivity of transcription to ActD (Bensaude, 2011; Bosmann, 1971). Performing protease digestion at lower temperatures further blocked spurious IEG expression in endothelial cells, thus minimizing dissociation-induced transcriptional alterations in all cell types. Comparison with the conventional method using seizure-stimulated tissues demonstrated the strength of Act-seq in detecting stimulus-induced transcriptional changes, which was impossible with the conventional protocol. The observation that ActD application alone was able to effectively suppress artificially induced transcriptional changes in most cell types (particularly neurons and glial cells) makes the method straightforward and highly versatile, facilitating its ready incorporation into any dissociation protocols.

In our analysis, computationally excluding the IEGs appeared to suppress the confounding effects of the conventional dissociation method on baseline cell clustering (Figure 5), suggesting that the most prominent transcriptional changes caused by dissociation in these particular experiments were largely limited to IEGs. However, in other scenarios, artificially induced transcriptional changes may not always be restricted to IEGs (for example, when the experiment requires longer dissociation procedures), and determining which genes to exclude from the clustering analysis could be difficult or impractical. In these cases, experimental suppression of dissociation-induced gene expression using Act-seq would be important for minimizing the potential confounding effects. More importantly, experimental removal of artificially caused transcriptional alterations would be necessary for studies that intend to examine any biologically relevant acute gene expression changes.

Together, our improvements can greatly facilitate reliable detection of both baseline transcriptional profiles and acute gene expression changes and can be valuable for various transcriptional profiling techniques requiring dissociation of cells. Although this study has focused on acute transcriptional changes, we envision that the strength of this method in preventing transcriptional artifacts would also greatly facilitate analysis of transcriptional dynamics at different time points following a stimulus, which could help understand downstream effects of IEG expression and gene regulatory relationships.

### Systematic Detection of IEGs in Molecularly Defined Cell Types

Examination of IEGs has been widely used for linking the activity of specific neuronal populations with behavior and/or experience (Kawashima et al., 2014; Knight et al., 2012; Tonegawa et al., 2015). Act-seq enables a systematic, unbiased survey of a large number of IEGs in individual cells and simultaneous

classification of cell types based on transcriptional profiles. These features address two important limitations of conventional methods for IEG detection, such as immunohistochemistry or *in situ* hybridization.

First, with conventional detection methods, usually only one or a small set of pre-selected IEGs are analyzed. However, accumulating evidence suggests that activity-dependent genes are differentially regulated in different brain regions (Kawashima et al., 2014). For example, in the primary somatosensory cortex, both *Fos* and *Arc* were induced by sensory input (Guenther et al., 2013), whereas in the vomeronasal organ, *Egr1*, but not *Arc*, *Fos*, *FosB*, *Jun*, or *Nr4a1*, showed robust elevation after exposure to chemosignals (Isogai et al., 2011). Given the great number of IEGs, the capability of Act-seq to systematically survey different IEGs will be advantageous for sensitive and unbiased identification of activated cell populations.

Second, conventional IEG visualization approaches provide scarce information on the molecular identity of the labeled cells. Costaining with candidate markers is demanding to conduct on large scales and relies on prior knowledge of cell types, which in many cases can be limited or inaccurate. Bulk transcriptomic profiling of discrete cell populations (e.g., phosphorylated ribosome capture, TRAP, and RiboTag) does not preserve single-cell information and heavily depends on single pre-selected markers (Heiman et al., 2008; Knight et al., 2012; Sanz et al., 2009).

By addressing these limitations, Act-seq allows us to reveal distinct IEG activation profiles in different cell types, and to identify neuronal subpopulations activated by a specific behavior and/or experience (discussed below), linking molecularly defined cell types to brain function.

### Induction of Distinct Sets of IEGs in Neuronal versus Non-neuronal Cells in the MeA

Previous studies have examined brain gene expression changes in response to seizure but did not differentiate between different cell types (Lacar et al., 2016; Sandberg et al., 2000). We found that seizure induced robust IEG activation not only in neurons but also in glial cells, endothelial cells, and mural cells. While previous studies reported changes in cell morphology, proliferation, and gene expression in non-neuronal cell types hours or days following seizure (Fabene et al., 2008; Luo et al., 2015; Rakhade and Jensen, 2009), our data suggest that early IEG induction events, which may potentially result from seizure-induced neural activity and/or excitotoxicity, could possibly be an acute indicator of cellular responses in non-neuronal cells. Importantly, we observed that different sets of IEG were induced in a cell-type-specific manner within the same brain area. This substantial difference in IEG induction between neuronal and non-neuronal cells and its underlying mechanisms are not well understood and may stem from differences in sensitivity to stimuli, signal transduction mechanism, transcriptional regulatory machinery, chromatin structure, and/or factors regulating RNA processing and stability (Fowler et al., 2011). How non-neuronal cells respond in physiological settings and what downstream cellular processes these different IEG activations may lead to would be of interest for future investigation.

Act-seq also allowed us to pinpoint specific MeA neuronal subpopulations that were preferentially activated by a behavior-

ally relevant stimulus, acute stress. Although the MeA has long been implicated in regulating stress response (Crane et al., 2005; Ons et al., 2010; Solomon et al., 2010), the molecular identity of the MeA neurons mediating this function has been a mystery. Following acute stress, we observed preferential IEG elevation in two specific neuronal subpopulations that expressed *Cck*. While other non-*Cck*<sup>+</sup> neurons might not necessarily be inactive, the *Cck*<sup>+</sup> neurons were preferentially activated relative to other subpopulations. Interestingly, the CCK neuropeptide has been suggested to modulate stress-related behaviors in a wide variety of paradigms (Griebel and Holsboer, 2012; Pérez de la Mora et al., 2007; Rotzinger and Vaccarino, 2003). The identification of the *Cck*<sup>+</sup> populations preferentially activated by acute stress now opens a door for future in-depth investigation of these neurons, such as imaging and manipulation of neuronal activity. The application of Act-seq to linking a specific behavior/experience to molecularly defined neuronal populations generates interesting and testable hypotheses on their functional roles and provides molecular handles for systematic morphological, electrophysiological, and functional interrogations.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODELS AND SUBJECT DETAILS
  - Mice
- METHODS DETAILS
  - Generation of Single-Cell Suspensions
  - Drop-Seq Procedure
  - Drop-Seq Read Alignment and Generation of Digital Expression Data
  - Dimensionality Reduction, Clustering, and Classification for Major Cell Types
  - Characterization of Neuronal and Astrocyte Subtypes
  - Weighted Gene Coexpression Network Analysis
  - Gene Ontology Analysis
  - Comparison of IEG Expression
  - Dual-Color Fluorescent *In Situ* Hybridization and Immunohistochemistry
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

## SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and seven tables and can be found with this article online at <https://doi.org/10.1016/j.neuron.2017.09.026>.

## AUTHOR CONTRIBUTIONS

Y.E.W., L.P., and W.H. designed and performed experiments and analyzed data. Y.Z. assisted in behavior analysis. X.L. provided expertise and input to the study and assisted in RNA sequencing. Y.E.W. and W.H. wrote the paper with inputs from all authors. W.H. supervised the entire study.

## ACKNOWLEDGMENTS

We thank L. Luo, F. Guo, M. Sofroniew, and P. Chen for helpful discussions and comments, and the UCLA TCGB for RNA sequencing service. The work was supported in part by the Whitehall Foundation (W.H.), a NARSAD Young Investigator grant (W.H.), and a Searle Scholars Award (W.H.).

Received: May 22, 2017

Revised: July 27, 2017

Accepted: September 15, 2017

Published: October 11, 2017

## REFERENCES

- Bensaude, O. (2011). Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? *Transcription* 2, 103–108.
- Bergan, J.F., Ben-Shaul, Y., and Dulac, C. (2014). Sex-specific processing of social cues in the medial amygdala. *eLife* 3, e02743.
- Bian, X., Yanagawa, Y., Chen, W.R., and Luo, M. (2008). Cortical-like functional organization of the pheromone-processing circuits in the medial amygdala. *J. Neurophysiol.* 99, 77–86.
- Blondel, V.D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech.* 2008, P10008.
- Bosmann, H.B. (1971). Mechanism of cellular drug resistance. *Nature* 233, 566–569.
- Choi, G.B., Dong, H.W., Murphy, A.J., Valenzuela, D.M., Yancopoulos, G.D., Swanson, L.W., and Anderson, D.J. (2005). Lhx6 delineates a pathway mediating innate reproductive behaviors from the amygdala to the hypothalamus. *Neuron* 46, 647–660.
- Clarke, L.E., and Barres, B.A. (2013). Emerging roles of astrocytes in neural circuit development. *Nat. Rev. Neurosci.* 14, 311–321.
- Covey, E., and Carter, M. (2015). *Basic Electrophysiological Methods* (Oxford University Press).
- Crane, J.W., French, K.R., and Buller, K.M. (2005). Patterns of neuronal activation in the rat brain and spinal cord in response to increasing durations of restraint stress. *Stress* 8, 199–211.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Ebner, K., and Singewald, N. (2006). The role of substance P in stress and anxiety responses. *Amino Acids* 31, 251–272.
- Fabene, P.F., Navarro Mora, G., Martinello, M., Rossi, B., Merigo, F., Ottoboni, L., Bach, S., Angiari, S., Benati, D., Chakir, A., et al. (2008). A role for leukocyte-endothelial adhesion mechanisms in epilepsy. *Nat. Med.* 14, 1377–1383.
- Fowler, T., Sen, R., and Roy, A.L. (2011). Regulation of primary response genes. *Mol. Cell* 44, 348–360.
- Griebel, G., and Holsboer, F. (2012). Neuropeptide receptor ligands as drugs for psychiatric diseases: the end of the beginning? *Nat. Rev. Drug Discov.* 11, 462–478.
- Guenther, C.J., Miyamichi, K., Yang, H.H., Heller, H.C., and Luo, L. (2013). Permanent genetic access to transiently active neurons via TRAP: targeted recombination in active populations. *Neuron* 78, 773–784.
- Habib, N., Basu, A., Avraham-Davidi, I., Burks, T., Choudhury, S.R., Aguet, F., Gelfand, E., Ardlie, K., Weitz, D.A., Rozenblatt-Rosen, O., et al. (2017). DroNc-Seq: Deciphering cell types in human archived brain tissues by massively-parallel single nucleus RNA-seq. *bioRxiv*, <https://doi.org/10.1101/115196>.
- Heiman, M., Schaefer, A., Gong, S., Peterson, J.D., Day, M., Ramsey, K.E., Suárez-Fariñas, M., Schwarz, C., Stephan, D.A., Surmeier, D.J., et al. (2008). A translational profiling approach for the molecular characterization of CNS cell types. *Cell* 135, 738–748.
- Hong, W., Kim, D.-W., and Anderson, D.J. (2014). Antagonistic control of social versus repetitive self-grooming behaviors by separable amygdala neuronal subsets. *Cell* 158, 1348–1361.
- Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.
- Isogai, Y., Si, S., Pont-Lezica, L., Tan, T., Kapoor, V., Murthy, V.N., and Dulac, C. (2011). Molecular organization of vomeronasal chemoreception. *Nature* 478, 241–245.
- Janak, P.H., and Tye, K.M. (2015). From circuits to behaviour in the amygdala. *Nature* 517, 284–292.
- Kawashima, T., Okuno, H., and Bito, H. (2014). A new era for functional labeling of neurons: activity-dependent promoters have come of age. *Front. Neural Circuits* 8, 37.
- Keshavarzi, S., Sullivan, R.K.P., Ianno, D.J., and Sah, P. (2014). Functional properties and projections of neurons in the medial amygdala. *J. Neurosci.* 34, 8699–8715.
- Knight, Z.A., Tan, K., Birsoy, K., Schmidt, S., Garrison, J.L., Wysocki, R.W., Emiliano, A., Ekstrand, M.I., and Friedman, J.M. (2012). Molecular profiling of activated neurons by phosphorylated ribosome capture. *Cell* 151, 1126–1137.
- Kołodziejczyk, A.A., Kim, J.K., Svensson, V., Marioni, J.C., and Teichmann, S.A. (2015). The technology and biology of single-cell RNA sequencing. *Mol. Cell* 58, 610–620.
- Lacar, B., Linker, S.B., Jaeger, B.N., Krishnaswami, S., Barron, J., Kelder, M., Parylak, S., Paquola, A., Venepally, P., Novotny, M., et al. (2016). Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat. Commun.* 7, 11022.
- Lake, B.B., Codeluppi, S., Yung, Y.C., Gao, D., Chun, J., Kharchenko, P.V., Linnarsson, S., and Zhang, K. (2017). A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA. *Sci. Rep.* 7, 6031.
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559.
- Langfelder, P., Zhang, B., and Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 24, 719–720.
- Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445, 168–176.
- Li, C.-L., Li, K.-C., Wu, D., Chen, Y., Luo, H., Zhao, J.-R., Wang, S.-S., Sun, M.-M., Lu, Y.-J., Zhong, Y.-Q., et al. (2016). Somatosensory neuron types identified by high-coverage single-cell RNA-sequencing and functional heterogeneity. *Cell Res.* 26, 83–102.
- Lin, L.C., and Sibille, E. (2015). Somatostatin, neuronal vulnerability and behavioral emotionality. *Mol. Psychiatry* 20, 377–387.
- Luo, Y., Hu, Q., Zhang, Q., Hong, S., Tang, X., Cheng, L., and Jiang, L. (2015). Alterations in hippocampal myelin and oligodendrocyte precursor cells during epileptogenesis. *Brain Res.* 1627, 154–164.
- Macosko, E.Z., Basu, A., Satija, R., Nemes, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 161, 1202–1214.
- McDavid, A., Finak, G., Chattopadhyay, P.K., Dominguez, M., Lamoreaux, L., Ma, S.S., Roederer, M., and Gottardo, R. (2013). Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. *Bioinformatics* 29, 461–467.
- Ons, S., Rotllant, D., Marin-Blasco, I.J., and Armario, A. (2010). Immediate-early gene response to repeated immobilization: Fos protein and arc mRNA levels appear to be less sensitive than c-fos mRNA to adaptation. *Eur. J. Neurosci.* 31, 2043–2052.
- Pérez de la Mora, M., Hernández-Gómez, A.M., Arizmendi-García, Y., Jacobsen, K.X., Lara-García, D., Flores-Gracia, C., Crespo-Ramírez, M., Gallegos-Cari, A., Nuche-Bricaire, A., and Fuxe, K. (2007). Role of the amygdaloid cholecystokinin (CCK)/gastrin-2 receptors and terminal networks in

- the modulation of anxiety in the rat. Effects of CCK-4 and CCK-8S on anxiety-like behaviour and [3H]GABA release. *Eur. J. Neurosci.* **26**, 3614–3630.
- Poulin, J.-F., Tasic, B., Hjerling-Leffler, J., Trimarchi, J.M., and Awatramani, R. (2016). Disentangling neural cell diversity using single-cell transcriptomics. *Nat. Neurosci.* **19**, 1131–1141.
- Rakhade, S.N., and Jensen, F.E. (2009). Epileptogenesis in the immature brain: emerging mechanisms. *Nat. Rev. Neurol.* **5**, 380–391.
- Reichmann, F., and Holzer, P. (2016). Neuropeptide Y: A stressful review. *Neuropeptides* **55**, 99–109.
- Rogge, G., Jones, D., Hubert, G.W., Lin, Y., and Kuhar, M.J. (2008). CART peptides: regulators of body weight, reward and other functions. *Nat. Rev. Neurosci.* **9**, 747–758.
- Romanov, R.A., Zeisel, A., Bakker, J., Girach, F., Hellysaz, A., Tomer, R., Alpár, A., Mulder, J., Clotman, F., Keimpema, E., et al. (2017). Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. *Nat. Neurosci.* **20**, 176–188.
- Rotzinger, S., and Vaccarino, F.J. (2003). Cholecystokinin receptor subtypes: role in the modulation of anxiety-related and reward-related behaviours in animal models. *J. Psychiatry Neurosci.* **28**, 171–181.
- Sandberg, R., Yasuda, R., Pankratz, D.G., Carter, T.A., Del Rio, J.A., Wodicka, L., Mayford, M., Lockhart, D.J., and Barlow, C. (2000). Regional and strain-specific gene expression mapping in the adult mouse brain. *Proc. Natl. Acad. Sci. USA* **97**, 11038–11043.
- Sanz, E., Yang, L., Su, T., Morris, D.R., McKnight, G.S., and Amieux, P.S. (2009). Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proc. Natl. Acad. Sci. USA* **106**, 13939–13944.
- Shekhar, K., Brodin, P., Davis, M.M., and Chakraborty, A.K. (2014). Automatic classification of cellular expression by nonlinear stochastic embedding (ACCENSE). *Proc. Natl. Acad. Sci. USA* **111**, 202–207.
- Shekhar, K., Lapan, S.W., Whitney, I.E., Tran, N.M., Macosko, E.Z., Kowalczyk, M., Adiconis, X., Levin, J.Z., Nemes, J., Goldman, M., et al. (2016). Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. *Cell* **166**, 1308–1323.e30.
- Sofroniew, M.V. (2015). Astrocyte barriers to neurotoxic inflammation. *Nat. Rev. Neurosci.* **16**, 249–263.
- Solomon, M.B., Jones, K., Packard, B.A., and Herman, J.P. (2010). The medial amygdala modulates body weight but not neuroendocrine responses to chronic stress. *J. Neuroendocrinol.* **22**, 13–23.
- Swanson, L.W. (2000). Cerebral hemisphere regulation of motivated behavior. *Brain Res.* **886**, 113–164.
- Swanson, L.W., and Petrovich, G.D. (1998). What is the amygdala? *Trends Neurosci.* **21**, 323–331.
- Tasic, B., Menon, V., Nguyen, T.N., Kim, T.-K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346.
- Ting, J.T., Daigle, T.L., Chen, Q., and Feng, G. (2014). Acute brain slice methods for adult and aging animals: application of targeted patch clamp analysis and optogenetics. *Methods Mol. Biol.* **1183**, 221–242.
- Tonegawa, S., Pignatelli, M., Roy, D.S., and Ryan, T.J. (2015). Memory engram storage and retrieval. *Curr. Opin. Neurobiol.* **35**, 101–109.
- Unger, E.K., Burke, K.J.J., Jr., Yang, C.F., Bender, K.J., Fuller, P.M., and Shah, N.M. (2015). Medial amygdalar aromatase neurons regulate aggression in both sexes. *Cell Rep.* **10**, 453–462.
- van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605.
- Xu, X., Coats, J.K., Yang, C.F., Wang, A., Ahmed, O.M., Alvarado, M., Izumi, T., and Shah, N.M. (2012). Modular genetic control of sexually dimorphic behaviors. *Cell* **148**, 596–607.
- Yount, G.L., Ponsalle, P., and White, J.D. (1994). Pentylentetrazole-induced seizures stimulate transcription of early and late response genes. *Brain Res. Mol. Brain Res.* **21**, 219–224.
- Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnerberg, P., La Manno, G., Jureus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142.
- Zeng, H., and Sanes, J.R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. *Nat. Rev. Neurosci.* **18**, 530–546.
- Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O’Keefe, S., Phatnani, H.P., Guarnieri, P., Caneda, C., Ruderisch, N., et al. (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* **34**, 11929–11947.
- Zirlinger, M., Kreiman, G., and Anderson, D.J. (2001). Amygdala-enriched genes identified by microarray technology are restricted to specific amygdaloid subnuclei. *Proc. Natl. Acad. Sci. USA* **98**, 5270–5275.

## STAR★METHODS

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Sheep anti-fluorescein-POD	Sigma-Aldrich	Cat# 11426346910, RRID: AB_840257
Sheep anti-DIG-AP	Sigma-Aldrich	Cat# 11093274910, RRID: AB_514497
Rabbit anti-DNP-Alexa488	Molecular Probes	Cat# A-11097, RRID: AB_2314332
Mouse anti-NeuN	Thermo Fisher Scientific	Cat# MAB377MI
Donkey anti-Mouse IgG (H+L)-Alexa488	Thermo Fisher Scientific	Cat# A-21202, RRID: AB_141607
<b>Chemicals, Peptides, and Recombinant Proteins</b>		
Pentylentetrazole	Sigma-Aldrich	Cat# P6500
Pronase	Sigma-Aldrich	Cat# P6911
Fetal bovine serum	Thermo Fisher Scientific	Cat# 10437028
BSA	Sigma-Aldrich	Cat# A8806
N-methyl-D-glucamine	Sigma-Aldrich	M2004
Actinomycin D	Sigma-Aldrich	Cat# A1410
<b>Critical Commercial Assays</b>		
Nextera XT DNA Library Prep Kit	Illumina	Cat# FC-131-1024
DIG RNA Labeling Mix	Sigma-Aldrich	Cat# 11277073910
Fluorescein RNA Labeling Mix	Sigma-Aldrich	Cat# 11685619910
TSA Plus DNP (HRP) System	PerkinElmer	Cat# NEL747A001KT
HNPP Fluorescent Detection Set	Sigma-Aldrich	Cat# 11758888001
<b>Deposited Data</b>		
Raw and processed scRNA-seq data	Gene Expression Omnibus	GEO: GSE103976
<b>Experimental Models: Organisms/Strains</b>		
Mouse: C57BL/6J	The Jackson Laboratory	Cat# 000664, RRID: IMSR_JAX:000664
<b>Software and Algorithms</b>		
Drop-seq_tools (v1.12)	Macosko et al., 2015	<a href="http://mccarrolllab.com/dropseq/">http://mccarrolllab.com/dropseq/</a>
STAR (v2.5.0a)	Dobin et al., 2013	<a href="https://github.com/alexdobin/STAR">https://github.com/alexdobin/STAR</a>
Clustering and differential expression between cell types	Shekhar et al., 2016	<a href="https://github.com/broadinstitute/BipolarCell2016">https://github.com/broadinstitute/BipolarCell2016</a>
t-distributed stochastic neighbor embedding	van der Maaten and Hinton, 2008	<a href="https://lvdmaaten.github.io/tsne/">https://lvdmaaten.github.io/tsne/</a>
Weighted gene coexpression network analysis	Langfelder and Horvath, 2008	R package WGCNA
BackSPIN	Zeisel et al., 2015	<a href="https://github.com/linnarsson-lab/BackSPIN">https://github.com/linnarsson-lab/BackSPIN</a>
Database for Annotation, Visualization and Integrated Discovery (DAVID)	Huang et al., 2009	<a href="https://david.ncifcrf.gov">https://david.ncifcrf.gov</a>
Likelihood-ratio test	McDavid et al., 2013	<a href="http://satijalab.org/seurat/">http://satijalab.org/seurat/</a> ; RRID: SCR_007322
ImageJ	NIH	<a href="https://imagej.nih.gov/ij/index.html">https://imagej.nih.gov/ij/index.html</a> ; RRID: SCR_003070
<b>Other</b>		
Drop-seq beads	ChemGenes	Cat# MACOSKO-2011-10
Drop-seq reagents	Macosko et al., 2015	<a href="http://mccarrolllab.com/dropseq/">http://mccarrolllab.com/dropseq/</a>

## CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Weizhe Hong ([whong@ucla.edu](mailto:whong@ucla.edu)).

## EXPERIMENTAL MODELS AND SUBJECT DETAILS

### Mice

All animal care and experimental procedures followed National Institutes of Health guidelines, and were approved by the UCLA Institutional Animal Care and Use Committee. Mice were housed with a reverse light cycle (12 hr light/12 hr dark) and *ad libitum* access to food and water. Single-cell RNA-seq, *in situ* hybridization, and immunohistochemistry experiments were conducted using adult (8–10 weeks) male C57BL/6J mice purchased from Jackson Laboratories. Mice were single housed for 2 days before experiments to minimize background IEG expression in the amygdala due to interactions with other mice. For pentylenetetrazole (PTZ) treatment, mice were injected intraperitoneally with PTZ at 50 mg/kg body weight and sacrificed 1 hr after injection. The mice displayed seizure-like behaviors ~5 min following PTZ administration. For saline injection, mice were injected intraperitoneally with 250  $\mu$ L 0.9% (w/v) NaCl solution and sacrificed 1 hr after injection. Acute immobilization was performed by taping the four limbs of the mice to a plastic board. The mice were restrained for 45 min, returned to their home cages, and sacrificed 15 min later (1 hr following the start of the treatment).

## METHODS DETAILS

### Generation of Single-Cell Suspensions

#### Comparison of different single-cell dissociation procedures (Figures 4 and 5)

Individual mice were anesthetized in an isoflurane chamber and decapitated, and the brain was immediately extracted and placed in fresh ice-cold artificial cerebrospinal fluid (ACSF) containing 124 mM NaCl, 2.5 mM KCl, 1.2 mM  $\text{NaH}_2\text{PO}_4$ , 24 mM  $\text{NaHCO}_3$ , 5 mM HEPES, 13 mM glucose, 2 mM  $\text{MgSO}_4$ , and 2 mM  $\text{CaCl}_2$ , bubbled with a carbogen gas (95%  $\text{O}_2$  and 5%  $\text{CO}_2$ ) and with a pH of 7.3–7.4. The brain was sectioned in ice-cold ACSF on a vibratome (Leica VT1200) into 300  $\mu$ m slices.

For the “conventional” dissociation method, the slices of interest were immediately transferred to a Petri dish containing ice-cold, carbogen-bubbled ACSF. An amygdala region (containing part of the basal lateral, basal medial, central, medial, and cortical amygdala nuclei) was microdissected under a dissecting microscope from three consecutive 300  $\mu$ m sections (~Bregma  $-1.06$  mm – Bregma  $-1.94$  mm) according to Paxinos and Franklin’s *the Mouse Brain in Stereotaxic Coordinates*. The dissected tissue was cut into small pieces < 1 mm in each dimension and transferred to a Petri dish containing 1 mg/mL pronase (Sigma-Aldrich, Cat# P6911) in carbogen-bubbled ACSF. Pronase digestion was performed at 34°C for 20 min in a chamber continuously aerated with carbogen. Following digestion, the pronase solution was exchanged with ice-cold, carbogen-bubbled ACSF containing 1% fetal bovine serum. The tissue pieces were then dissociated in a microcentrifuge tube by gentle trituration through Pasteur pipettes with polished tip openings of 600  $\mu$ m, 300  $\mu$ m, and 150  $\mu$ m diameter. Following trituration, the cells were filtered through a 20  $\mu$ m filter to eliminate cell clumps, pelleted, washed with and resuspended in ice-cold, carbogen-bubbled ACSF + 0.01% BSA.

For the Act-seq method, the brain was sectioned as described above, and the slices of interest were immediately transferred to a Petri dish with an ice-cold, carbogen-bubbled recovery solution containing 93 mM N-methyl-D-glucamine, 2.5 mM KCl, 1.2 mM  $\text{NaH}_2\text{PO}_4$ , 30 mM  $\text{NaHCO}_3$ , 20 mM HEPES, 25 mM glucose, 10 mM  $\text{MgSO}_4$ , 0.5 mM  $\text{CaCl}_2$ , 5 mM sodium ascorbate, 2 mM thiourea, and 3 mM sodium pyruvate with a pH of 7.3–7.4, and incubated for 15 min on ice in a chamber continuously aerated with carbogen. This “protective recovery” step has been shown to substantially enhance neuronal survival during acute brain slice preparation from adult mice by reducing cellular stress, metabolic demand, and excitotoxicity (Covey and Carter, 2015; Ting et al., 2014). We found that in addition to improving cell survival, this procedure alone also led to a modest decrease in dissociation-induced IEG expression, but had minimal effects on further suppressing IEG induction in the presence of ActD (data not shown). Therefore, this step is beneficial for the processing of adult brain tissues, but is not required for the suppression of IEG induction during dissociation in the presence of ActD. The slices were then transferred to a Petri dish containing ice-cold, carbogen-bubbled ACSF and the amygdala was dissected as described above. Pronase digestion was performed at 34°C for 20 min or at 22°C or 11°C for 30 min in a chamber continuously aerated with carbogen. The subsequent dissociation steps were performed as described above in the conventional method. Actinomycin D (Sigma-Aldrich, Cat# A1410) was added at three steps during the dissociation process: to the “protective recovery” solution at 45  $\mu$ M (for 34°C digestion) or 30  $\mu$ M (for 22°C or 11°C digestion), to the pronase solution at 45  $\mu$ M (for 34°C digestion) or 15  $\mu$ M (for 22°C or 11°C digestion), and during trituration at 3  $\mu$ M.

Our results indicated that at 34°C, ActD alone was effective in inhibiting dissociation-induced IEG expression in all cell types except for endothelial and mural cells (Figure 4G). This allows for its ready incorporation into established single-cell dissociation protocols using different proteases at room temperature or higher temperature. In addition, we observed that pronase, which is isolated from *Streptomyces griseus* (a bacterial species that normally lives in a wide range of temperature in the soil), exhibits partial protease activity at low temperature. Combining ActD application with lower digestion temperature could block artificial IEG induction in all major cell types (Figures 4H and 4I). We note that caution needs to be exercised to make sure that the tissue is effectively dissociated at lower temperature, as protease activity decreases with temperature; a reduced amount of tissue, smaller sizes of tissue chunks, longer dissociation time, and/or extra trituration may be needed to ensure thorough cell dissociation.

The Act-seq protocol takes ~2 hr from animal anesthesia to droplet generation, including ~2 min for animal anesthesia, ~3 min for brain extraction, ~10 min for vibratome sectioning, ~15 min for protective recovery, ~15 min for amygdala dissection, ~20–30 min for protease digestion, ~10 min for cell trituration, ~20 min for cell washing and centrifugation, and ~25 min for droplet generation.

#### **Single-cell dissociation from the MeA for cell type classification and IEG analysis (Figures 1, 2, 3, 6, and 7)**

The MeA was dissected from three consecutive 300  $\mu$ m sections (~Bregma –1.06 mm - Bregma –1.94 mm) encompassing most of the rostro-caudal extension of the MeA, including MeAad, MeAav, MeApd, and MeApv, according to *Paxinos and Franklin's the Mouse Brain in Stereotaxic Coordinates* using the optic tract as the primary landmark. Single cells were dissociated using Act-seq as described above with pronase digestion at 11°C for 30 min.

#### **Drop-Seq Procedure**

Drop-seq was performed largely as described previously (Macosko et al., 2015). Cells were processed for Drop-seq within ~15 min after collection. Briefly, cells were diluted to a final droplet occupancy of ~5%, and barcoded beads were diluted to a final droplet occupancy of ~10%. Flow rates of 3 mL/hr were used for each aqueous suspension and 15 mL/hr for the droplet generation oil. Droplets generated from ~2 mL of aqueous volume (1 mL cells and 1 mL beads) were broken, beads were harvested, and hybridized RNA was reverse transcribed. Populations of ~3000 beads (~150 STAMPs [single-cell transcriptomes attached to microparticles]) were separately amplified with 14 PCR cycles. Single-cell sequencing libraries were generated by tagmentation using Nextera XT (Illumina, Cat# FC-131-1024) and sequenced on the Illumina NextSeq500 sequencer with 20 bp for read 1 and 60–63 bp for read 2.

#### **Drop-Seq Read Alignment and Generation of Digital Expression Data**

Paired-end sequence reads were processed as described before (Macosko et al., 2015; Shekhar et al., 2016). Briefly, bases 1–12 and 13–20 in read 1 were used to infer the cell of origin (cell barcode) and molecule of origin (unique molecular identifier or UMI), respectively. Read pairs were filtered to remove those with a cell barcode or UMI base quality of less than 10. Read 2 was trimmed at the 5' end to remove any portion of the SMART adaptor sequence and at the 3' end to remove polyA tails of 6 bases or longer, and then aligned to the mouse genome (mm10) using STAR v2.5.0a (Dobin et al., 2013) with default settings. Reads uniquely mapped to exonic regions of genes were recorded and grouped by cell barcode. Correction for barcode synthesis errors was performed as previously described (Shekhar et al., 2016). A digital expression matrix containing the number of UMIs in each gene within each cell was then assembled, with UMIs within an edit distance of 1 collapsed.

#### **Dimensionality Reduction, Clustering, and Classification for Major Cell Types**

MeA cells with > 450 genes and > 700 transcripts were used for the analysis of major cell types, resulting in a dataset of 20,679 cells. Mitochondrial RNAs, tRNAs, and rRNAs were excluded. Transcript counts were normalized to library size as described before (Shekhar et al., 2016). Raw transcript count of each gene in each cell was divided by the total transcript number in that cell and then timed by the median of total transcript numbers in all cells in the dataset. Dimensionality reduction and clustering analysis were then performed using log-transformed expression data ( $\ln(\text{normalized counts} + 1)$ ).

Principal component analysis (PCA) was performed on the whole expression data matrix (excluding immediate early genes [IEGs; Table S3]) using the `fast.pcomp` function in the `gmodels` R package. PCs 1–50 (Figure S1) were then used as input for t-distributed stochastic neighbor embedding (tSNE) to generate a two-dimensional non-linear embedding of the cells, with 2000 iterations and the perplexity parameter set to 30 as previously described (Shekhar et al., 2014; van der Maaten and Hinton, 2008). To group cells into transcriptionally similar clusters, we implemented a Louvain–Jaccard graph clustering algorithm (Blondel et al., 2008; Shekhar et al., 2016), using the scores of the cells along PCs 1–50 and with the number of nearest neighbors set to 30. This resulted in 29 clusters in our data.

To assign each cluster to a major cell type, we identified top genes differentially upregulated in each cluster. We performed comparison between every single cluster and all the other clusters using a previously described non-parametric binomial test (Shekhar et al., 2016), which evaluates whether a gene  $g$  is upregulated in population A with reference to population B by computing the probability of detecting  $g$  in A with the observed frequency or higher, given the detection frequency of  $g$  in B.  $p$  values were FDR adjusted using Benjamini–Hochberg correction. Differentially expressed genes that were detected in at least 10% of the cells in the first cluster and have a fold change of at least 2 and an FDR < 0.05 were then ranked by  $\ln(\text{fold change}) \times \text{proportion of positive cells within the first cluster}$  to identify top genes enriched in each cluster. We then assigned each cluster to a major cell type (neurons, astrocytes, microglia, oligodendrocyte precursor cells, oligodendrocytes, endothelial cells, and mural cells) based on the presence/absence of known cell type markers within its top upregulated genes. We next compared each major cell type to the other cell types using the binomial test (Shekhar et al., 2016) to identify cell-type-specific genes, and ranked the genes as described above. The top 50 markers for each major cell type were listed in Table S1.

For the clustering of MeA major cell types and further clustering of MeA neurons and astrocytes (Figures 1, 2, and 3), we pooled cells from mice in the control, seizure, and acute stress groups as well as additional male C57BL/6J mice (8–10 weeks) to increase the number of cells used in the analysis. Cells from the additional mice were only used in cell type clustering analysis (Figures 1, 2, and 3), and were not used for any IEG differential expression analysis (Figures 6 and 7). Of note, different experimental conditions did not cause significant changes in the expression of major cell-type markers, neuronal/astrocyte subtype markers, or genes used for neuronal clustering (Figures 6D, 7E, 7J, Figures S6M, S6N, and S7).

For the classification of amygdala cells sequenced for the comparison of different single-cell dissociation methods (Figures 4 and 5), cells with > 500 genes (4556 cells) were used for the analysis. PCA, tSNE, Louvain-Jaccard clustering and cell type assignment were performed as described above.

### Characterization of Neuronal and Astrocyte Subtypes

To identify neuronal subtypes, cells that were classified as neurons in the major cell-type analysis were used for further clustering. Similar to previous studies (Romanov et al., 2017; Zeisel et al., 2015), we first selected genes that were enriched in neurons (excluding IEGs [Table S3]), in order to reduce noise. We used the binomial test (Shekhar et al., 2016) to compare neurons to all the other cell types combined and selected 3237 genes with fold change  $\geq 1.3$ , FDR < 0.05, and expressed in  $\geq 0.3\%$  neurons. To cluster the neurons in a way that is more functionally relevant, we selected genes that were directly related to neuronal function, including genes encoding neurotransmitter metabolism enzymes and transporters, neuropeptides, neurotransmitter and neuropeptide receptors, other endogenous ligands and receptors, and ion channels/solute carriers (263 genes; Table S2). We then used these genes and neurons with > 600 transcripts of the 3237 neuronally enriched genes (2444 neurons) for further clustering analysis, employing a bi-clustering algorithm, BackSPIN (Zeisel et al., 2015), which simultaneously partitions genes and cells. In the original BackSPIN algorithm, the cells are split into two groups in each round, and each gene is assigned to one group of the cells and not taken into account when splitting the other group of the cells in the subsequent rounds. Considering that one gene might contribute to the differentiation between cell subgroups at multiple division points, we implemented the BackSPIN algorithm in a way that took every gene into account at every split. This was done by running the BackSPIN algorithm with only one split and with all the genes as input, dividing the cells into two groups, and then applying the one-split BackSPIN algorithm recursively to each group of cells, using all genes as input in each round. We also added the sum of *Gad1*, *Gad2*, and *Slc32a1* as a feature during clustering to increase the robustness of clustering of GABAergic neurons. Other parameters of BackSPIN were set to default values. This resulted in 16 neuronal subclusters. Subclusters 1-8 and 13-16 largely corresponded with the N1 neuronal cluster in the tSNE map for all MeA cell types, and subclusters 9-12 mainly corresponded with N2-N5 (Figure S4A). Despite this overall correspondence, the division between neuronal subclusters identified from BackSPIN clustering did not exactly follow the boundaries in the all-cell tSNE map. The two-dimensional tSNE map was primarily used as a visualization tool that displayed the distribution of single-cell transcriptomes only in the first two dimensions of the tSNE space, whereas additional dimensions were not easily visualizable. The BackSPIN clustering of neurons using neuronally enriched genes has higher sensitivity, but boundaries between neuronal subclusters defined this way may not be clearly visible in the first two dimensions of the tSNE space of all cell types. We further used the binomial test (Shekhar et al., 2016) to identify subtype-enriched genes (fold change  $\geq 1.5$ , FDR < 0.05, detected in  $\geq 10\%$  of neurons in a subtype) among the 3237 neuronally enriched genes by comparing each neuronal subtype with all the other subtypes and ranking the genes by  $\ln(\text{fold change}) \times \text{proportion of positive cells in that subtype}$ .

For the subclustering of astrocytes, cells that were classified as astrocytes in the major cell type analysis were used (5932 cells). We first selected astrocyte enriched genes (excluding IEGs [Table S3]) using the binomial test (Shekhar et al., 2016). Astrocytes were compared to all the other cell types combined and 3576 genes with fold change  $\geq 1.3$ , FDR < 0.05, and expressed in  $\geq 0.3\%$  astrocytes were selected. These genes were then used for PCA analysis, and tSNE was performed using PCs 1-20 as input, with 2000 iterations and the perplexity parameter set to 30. The three astrocyte subclusters were identified through the Louvain-Jaccard clustering algorithm (Blondel et al., 2008; Shekhar et al., 2016), using the scores of the cells along PCs 1-20 and with the number of nearest neighbors set to 20. To identify genes specifically enriched in each subtype, we used the binomial test (Shekhar et al., 2016) to compare each subtype with the other two combined and ranked the genes as described above.

It should be noted that our subclustering analysis of MeA neurons represents one computational approach, among many others (reviewed in Poulin et al., 2016), for classifying neuronal subtypes. Also, tissue dissociation, microfluidic cell sorting, and RNA capture and amplification may lead to different efficiency in detecting different cell types. The molecular map of the MeA could be further refined as scRNA-seq techniques continue to improve and by combining scRNA-seq with additional morphological, connective, and electrophysiological studies.

### Weighted Gene Coexpression Network Analysis

The Weighted Gene Coexpression Network Analysis (WGCNA) R package (Langfelder and Horvath, 2008) was used for building signed coexpression networks for the major cell types. Cells with  $\geq 1200$  genes and 1500 transcripts (4446 cells in total) were used in the analysis. Mural cells were excluded due to small cell number. We selected 5000 most variable genes using the `feature_selection` function in the BackSPIN algorithm (Zeisel et al., 2015) for network construction. Biweight midcorrelation was first used to calculate pairwise correlations between genes. Next, pairwise topological overlap was calculated with a power of 9 based on a fit to scale-free topology. Coexpression modules comprised of positively correlated genes with high topological overlap were then identified using the `cutreeDynamic` function in the `dynamicTreeCut` R package (Langfelder et al., 2008), with the following parameters: `method = "hybrid,"` `deepSplit = 2,` `pamStage = F,` `minClusterSize = 30`. The expression of each module was summarized by the module eigengene (ME, defined as the first principle component of all genes in a module). Modules whose eigengenes were highly correlated were further merged using the `mergeCloseModules` function in the WGCNA R package. Pearson correlations between MEs and different major cell types (treated as binary numeric variables) were calculated. p values were FDR adjusted using Benjamini-Hochberg correction.

### Gene Ontology Analysis

Gene ontology (GO) analysis for gene coexpression modules was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang et al., 2009). All genes used for the WGCNA analysis were used as background. Representative top enriched GO terms (of the biological process category) with FDR < 0.15 were shown.

### Comparison of IEG Expression

To compare the expression of individual IEGs between different single-cell dissociation methods (Figures 4A–4E, 4J, 4L, 4N), we performed one-sided Fisher's exact test to evaluate whether the IEGs were expressed in higher proportions of cells in the conventional dissociation method compared to the modified methods. We also performed a likelihood-ratio test designed for single-cell data (Macosko et al., 2015; McDavid et al., 2013), which simultaneously tests for changes in mean expression and in the percentage of expressed cells. The statistical significance of the results was largely identical to that derived from Fisher's exact test (data not shown).

To compare the proportion of IEG expressing cells in different cell types under different dissociation conditions (Figures 4F–4I; Table S4), we first selected IEGs showing the most prominent upregulation in each cell type when comparing the conventional dissociation procedure and Act-seq with 11°C digestion. We defined IEG-positive cells as those expressing any of the selected IEGs (neurons, *Fos*, *Fosb*, and *Junb*; astrocytes, *Fosb*, *Egr1*, *Junb*, and *Nr4a1*; microglia, *Dusp1*, *Egr1*, and *Fos*; OPCs, *Fos*, *Fosb*, and *Junb*; oligodendrocytes, *Fos*, *Fosb*, and *Egr1*; endothelial cells and mural cells, *Fos* and *Fosb*). We then examined whether the proportions of IEG-positive cells were higher in the conventional dissociation method compared to the modified methods using one-sided Fisher's exact test.

For the evaluation of seizure-induced IEG upregulation using either the conventional method or Act-seq (Figures 6A<sub>1</sub>–6B<sub>6</sub>), we used cells dissociated from PTZ-injected animals to compare with cells dissociated from saline-injected control animals. 8071 cells with > 550 genes and > 700 transcripts were used for this analysis. We have performed control experiments in both saline-injected animals and naive animals. In Act-seq, we found no IEG significantly differentially expressed (fold change > 2, FDR < 0.05, two-sided Fisher's exact test) between these two groups (Figures S6A–S6F). In further in-depth analysis of IEG upregulation in major cell types using Act-seq (Figures 6C<sub>1</sub>–6F<sub>6</sub>), we combined both saline-injected and naive animals as one control group to increase statistical power. 9822 cells with > 550 genes and > 700 transcripts were used for this analysis. For the analysis of gene expression changes following acute stress (Figures 7B–7E), naive animals were used as control.

For the evaluation of IEG upregulation in neuronal subtypes in response to seizure and acute stress (Figures 7A–7E), we used the combination of four top upregulated IEGs identified in seizure (*Fos*, *Fosb*, *Egr4*, and *Nr4a3*; Figures 6C<sub>1</sub>, 6E<sub>1</sub>, 6F<sub>1</sub>; Table S6) as an indicator of activation, and defined cells expressing any of these genes as IEG-positive cells. We then used the Fisher's exact test and the likelihood ratio test (Macosko et al., 2015; McDavid et al., 2013) to compare the proportion of IEG-positive cells and the summed expression level of the selected IEGs within single cells (Figures 7A–7C). For the analysis of IEG upregulation in astrocyte subtypes following seizure (Figures 7I–7M), astrocytes with > 600 genes and > 700 transcripts were used (4694 cells; Table S6).

For the comparison of cell type markers and other non-IEG genes between different experimental conditions, two-sided Fisher's exact test was used.

### Dual-Color Fluorescent *In Situ* Hybridization and Immunohistochemistry

Dual-color fluorescent *in situ* hybridization (FISH) experiments were performed largely as described (Li et al., 2016) with some modifications. Briefly, antisense probes were generated using cDNA templates derived from brain slices containing the MeA. The sequences of the probes were the same as those used by the Allen Brain Atlas (<http://www.alleninstitute.org>). Probes were labeled with digoxigenin (Sigma-Aldrich Cat# 11277073910) or fluorescein (Sigma-Aldrich Cat# 11685619910). Brains were fixed with 4% paraformaldehyde and dehydrated in 15% sucrose and 30% sucrose at 4°C overnight consecutively. 20 μm sections were hybridized with both probes for 16 hr at 65°C. After washing and blocking, the sections were incubated with anti-fluorescein-POD (Sigma-Aldrich Cat# 11426346910, 1:5000) at 4°C overnight. The sections were then sequentially incubated with TSA Plus (DNP, PerkinElmer Cat# NEL747A001KT, 1:1000) for 13 min, anti-DNP-POD (PerkinElmer Cat# NEL747A001KT, 1:300) for 45min, and TSA Plus (DNP, PerkinElmer Cat# NEL747A001KT, 1:1000) for 13 min. Subsequently, the sections were incubated with anti-DIG-AP (Sigma-Aldrich Cat# 11093274910, 1:1000) and anti-DNP Alexa488 (Molecular Probe Cat# A-11097, 1:500) at 4°C overnight. Detection of the digoxigenin-labeled probe was performed with HNPP/FR (Sigma-Aldrich Cat# 11758888001, both at 1:100). For NeuN immunohistochemistry, brain sections were fixed with 4% PFA for 20 min. After washing and blocking, the sections were incubated with anti-NeuN (Thermo Fisher Scientific, Cat# MAB377MI, 1:500) overnight at 4°C, followed by secondary antibody (anti-mouse IgG-Alexa488, Thermo Fisher Scientific, Cat# A-21202, 1:500) at room temperature for 2 hr. Images were acquired using Zeiss LSM 880 and quantified using ImageJ.

### QUANTIFICATION AND STATISTICAL ANALYSIS

Specifics on the statistical methodologies and software used for various analyses are described in the corresponding sections in RESULTS, figure legends, METHODS DETAILS, and supplemental tables. Comparison of gene expression between cell types or experimental conditions was performed using the Fisher's exact test, binomial test (Shekhar et al., 2016), and/or likelihood-ratio

test (Macosko et al., 2015; McDavid et al., 2013) as indicated. Overlap between different cell populations (Figures 7D, 7F–7H) was assessed using the Fisher's exact test. p values were adjusted using Benjamini-Hochberg correction or Bonferroni correction as indicated. Bar plots with error bars show mean  $\pm$  SD. Boxplots show the first and third quartiles and the median. Whiskers in the boxplots encompass data points within 1.5-fold interquartile range of the lower and upper quartiles.

#### **DATA AND SOFTWARE AVAILABILITY**

Gene expression data are available in the Gene Expression Omnibus, GEO: GSE103976. The data and code that support the findings of this study are available upon request from the corresponding author.