ALLEN INSTITUTE for BRAIN SCIENCE

Allen Cell Types Database

TECHNICAL WHITE PAPER: OVERVIEW

CELL CHARACTERIZATION DATA SETS

- **Electrophysiology:** Whole cell current clamp recordings made from identified, fluorescent Cre-positive neurons or nearby Cre-negative neurons in acute brain slices derived from adult mice. Whole cell current clamp recordings made from adult human neocortical neurons in brain slices derived from surgical specimens.
- **Morphology:** reconstruction-quality, 2D image stacks containing the complete structure of mouse and human neurons filled and recorded from *in vitro* slice preparations and 3D reconstructions of the dendrites and the initial part of the axon and/or the full axon of each neuron.
- **Histology:** Images of immunohistochemical staining from human brain slices using antibodies chosen to evaluate tissue integrity, cell distribution and histopathology. The panel includes Nissl, NeuN, SMI-32, GFAP, PVALB, Iba1 and Ki67
- **Transcriptomics:** single-cell RNA-sequencing with SMART-Seq v4 on fluorescent Cre-positive and Cre-negative cells enriched by FACS for mouse and fluorescent NeuN-positive (neuronal) or NeuN-negative (non-neuronal) nuclei enriched by FACS for human.
- **Generalized Leaky Integrate-and-Fire (GLIF) models:** a series of point neuron models of increasing complexity to reproduce the spiking behaviors of the recorded neurons. Starting with a leaky integrate and fire model, more complex models attempt to model variable spike threshold, afterspike currents, and threshold adaptation.
- **Biophysical-Perisomatic models:** compartmental model of neurons that account for the neural morphology and emulate electrophysiological responses by assuming biophysically detailed mechanisms for specific families of ionic conductances, with passive dendrites and active conductances at the soma.
- **Biophysical-All Active models:** compartmental model of neurons that account for the neural morphology and emulate electrophysiological responses by assuming biophysically detailed mechanisms for specific families of ionic conductances, with active conductances everywhere.

BACKGROUND

The constituent elements of mammalian cortex are neurons and their supporting actors - astrocytes, oligodendrocytes and microglia, as well as the vasculature. Given our interest in understanding perception and sensorimotor action that takes place in the 1 second time scale, we focus here on nerve cells, of which there are about 16 billion in the human cortex and 14 million in the mouse (Herculano-Houzel 2009). We hypothesize that these neurons can be classified into distinct types based on their dendritic morphology and axonal projections, their physiological and functional behavior and the genes that they express.

The idea that neuronal cell types have functional relevance is best documented in the vertebrate retina, where around 70 functionally distinct cell types have been reported and described using a combination of receptive field properties, cell morphology, cellular location within the retina, electrophysiology, connectivity and, more recently, gene expression. These neural components combine to process the rain of incoming photons into

action potentials that leave the retina along the optic nerve (Masland 2012). Since the earliest anatomical studies of Ramón y Cajal and Camillo Golgi, it has been evident that there is likewise a plethora of neural types in the cerebral cortex (Cauli *et al*., 1997, 2000; Toledo-Rodriguez *et al*., 2004; Sugino *et al*., 2006; Petilla *et al*., 2008; Thomson 2010; DeFelipe *et al*., 2013). The total number is unclear but could be as high as 1000 cell types, inferred from the tiling principle, the hypothesis that the dendritic footprint of any one cell type must cover the entire surface of cortex at least once (Crick 1994; Stevens 1998).

Yet despite much effort, there is no consensus on how many cortical cell types exist, whether they tile the cortical surface, how a given cell type varies across the cortical sheet and among different species, or even whether every cell in the adult neocortex can be unambiguously assigned to a unique cell class. One limiting factor has been combining the findings across individual studies and laboratories due to differences in preparation. An additional limitation has been the availability of tools for repeatedly targeting groups of neurons. With the recent proliferation of transgenic Cre mouse lines that label genetically-defined subsets of neurons, we now have the tools to reproducibly target and probe distinct neural populations across experiments in the mouse. These tools will also allow us to correlate neural class diversity with functionally described cell classes obtained in other Allen Institute projects.

One goal of the Allen Institute for Brain Science is to characterize, in a systematic and standardized manner, individual neurons in neocortical and thalamic regions of the mouse and adult human brain. This characterization includes: intrinsic electrophysiology properties, neuron morphology, transcriptomics profile, and/or parameters from multiple computational models incorporating one or more of these experimentally defined modalities. Through clustering and correlational analyses of these different modalities, we seek to characterize neuronal cell types, and to enable cross-species comparison.

The current main components of data generation and analysis include electrophysiology, morphology, transcriptomics, and modeling (**Figures 1** and **2**). In more detail, the four distinct components are: (i) an electrophysiological characterization of the biophysical and intrinsic firing properties of these cells, based on somatic patch recordings in slices, (ii) image data and morphological reconstructions of the dendritic tree and the initial part of the axon and/or the full axon within the slice based on biocytin fills through patch pipette, (iii) transcriptomics data obtained from single cells or nuclei enriched by various strategies and fluorescence-guided cell selection by FACS, and (iv) a variety of abstract point models (*generalized, leaky integrate-and-fire* or GLIF models) and more biophysically realistic active compartmental models of these neurons.

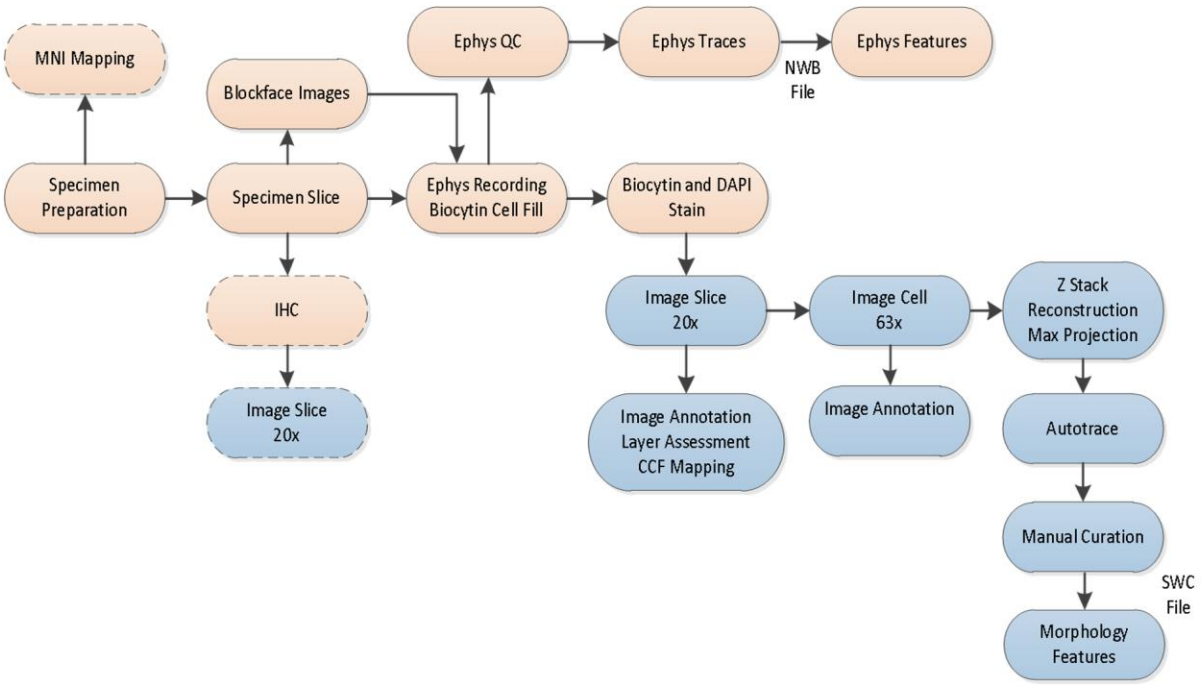

Figure 1. Overview of the electrophysiology and morphology data generation workflow.
Tissue enters the processing workflow from a transgenic mouse or human surgical specimen. The brain specimen is then dissected and sliced, followed by electrophysiological recording and filling of the target cell of interest with biocytin. Biocytin-filled cells are then imaged and morphologicallly reconstructed. Elements of the workflow specific to human surgical cases are indicated with dashed borders.

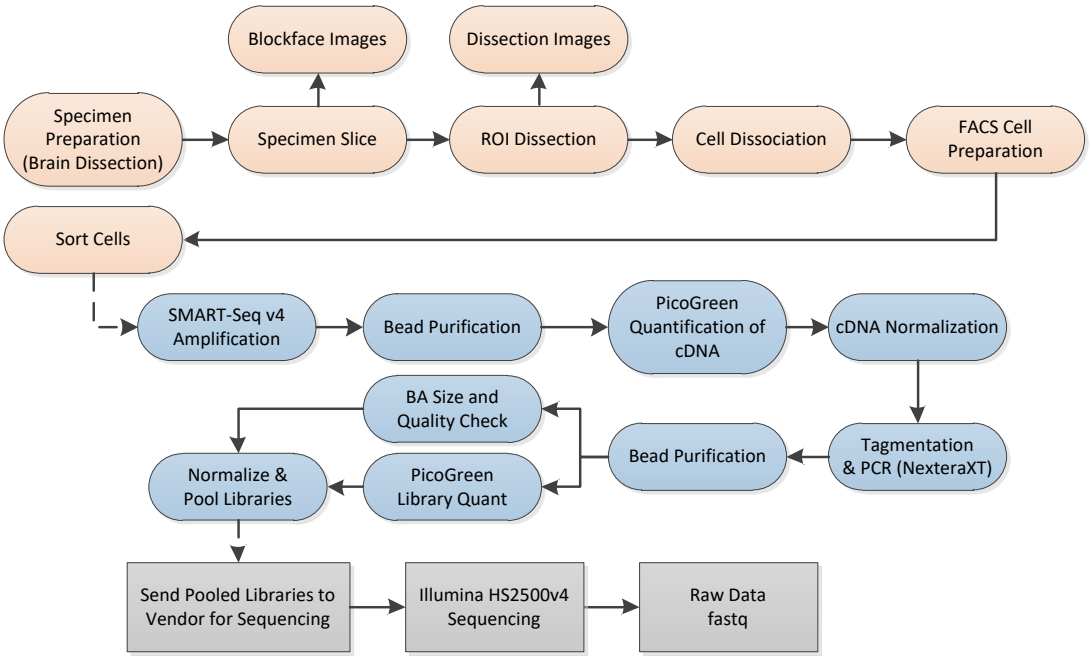

Figure 2. Overview of the workflow for RNA-Seq data generation.
The main components of the workflow for transcriptional data generation include brain dissection, region of interest (ROI) dissection, cell sorting, RNA purification, amplification and library preparation for RNA sequencing (RNA-Seq). Quality control steps are integrated into the data generation workflow.

Specimen selection: Mice
Interneuron-specific or layer-specific Cre driver lines for electrophysiology, morphology and modeling (Ephys column in **Table 1**), or excitatory, pan-neuronal or inhibitory Cre driver lines for transcriptomics (RNA-Seq column in **Table 1**) are combined with the universally expressing Cre-dependent responder line, Ai14 or Ai65 (see **Table 2**) (Madisen *et al*., 2010; Madisen *et al*., 2015). Mice are group-housed (5 per cage) in micro ventilated cages with a 12 h light/dark cycle. Purina Lab diet 5001 mouse food and water are given *ad libitum*. The Cre mice are backcrossed to C57BL/6J mice to minimize genetic variance.

Specimen selection: Human
This database of cell types includes diverse experimental data derived from cells of the adult human brain. The human brain tissue from both postmortem and neurosurgical origin were made available through the generosity of tissue donors. Human tissue from neurosurgical specimens included the collection of paired histological sections for cytochemical context. Detailed staining protocols for Nissl, NeuN, SMI-32, GFAP, PVALB, Iba1 and Ki67 are provided in the **Morphology Overview** of the Documentation section. For information on human specimen case qualification and a complete table of individual donor profiles, see the **Case Qualification and Donor Profiles** summary.

Electrophysiology
Electrophysiological data is collected from both mouse and human tissue. Detailed protocols for slice preparation are included in the **Electrophysiology Overview** in the Documentation. For mouse, acute brain slices are prepared from male or female animals between the ages of P45 to P70. Slices (350 µm) are sectioned using a vibrating microtome: each slice is imaged to aid in brain region identification. Whole cell current clamp recordings are made from tdTomato-positive neurons; in some cases, nearby tdTomato-negative neurons are recorded from, for comparative purposes. Human surgical specimens are transported from the hospital site to the Allen Institute and sliced at 350 µm using a vibrating microtome. Whole cell current clamp recordings are performed, sampling from neurons across all layers, using the distance from the pial surface and transitions in cell size and density as a guide to laminar position.

A standardized set of elecrophysiological stimuli are applied to both mouse and human neurons. Stimulation waveforms are designed to: 1) interrogate intrinsic membrane properties that contribute to the input/output function of neurons, 2) understand aspects of neural response properties *in vivo*, and 3) construct and test computational models of varying complexity emulating the neural response to stereotyped stimuli. Stimulus sets are divided into 2 groups: a fixed ‘Core 1’ sequence designed to establish baseline properties, which is applied to every neuron, and a more flexible ‘Core 2’ sequence that can be tailored to evaluate hypotheses based on cell type or class. Electrophysiology recordings include metadata such as electrode resistance, tight seal resistance, series resistance, solution bath temperature and amplifier settings on a sweep by sweep basis. Data qualification is based on cell-wide and sweep-based criteria, and other considerations described in the Overview document. For detailed descriptions of electrophysiology experiments, see the **Electrophysiology Overview** in the Documentation tab.

Morphology
Cell shape and structure are critical parameters for evaluating cell class. Images of the recorded cells are thus processed and quantitated. These data from cell reconstructions are used for biophysical modeling and neuron classification efforts. Three-dimensional (3D) reconstructions of dendrites and the initial segment and/or full axon are generated for a subset of mouse and human neurons that have both electrophysiological data and qualilty biocytin staining. 3D reconstructions are generated from a digital image stack. Data that pass quality control steps enter a Vaa3D-based (Peng *et al*. 2010) image processing and reconstruction analysis workflow. The subsequent reconstructions are then manually corrected and curated using Vaa3D tools and the Mozak extension (https://www.mozak.science/). Neuron reconstructions are then used for qualitative and quantitative morphological analyses. For detailed information on morphological reconstruction, see **Morphology Overview** in the Documentation tab.

Transcriptomics

RNA sequencing (RNA-Seq) can provide a transcriptomic profile for each cell. RNA transcripts are isolated from either whole cells or nuclei. For mouse, acute brain slices are prepared from male or female animals between the ages of P45 to P70. Slices (350 µm) are sectioned using a vibrating microtome: each slice is imaged to aid in brain region identification. Regions of interest are microdissected and dissected tissue pieces are treated with protease and subsequently triturated. From this cell suspension, single cells are isolated by fluorescence-activated cell sorting (FACS). After sorting, cDNA amplification and library construction is performed using SMART-Seq v4 (Clontech) and Nextera XT (Illumina) kits. Single cell libraries are sequenced on HiSeq (Illumina) to generate 50 base-pair paired-end reads. Raw read (fastq) files are aligned to the mm10 mouse genome sequence (Genome Reference Consortium, 2011) with the RefSeq transcriptome version GRCm38.p3 (current as of 01/15/2016). Transcriptome alignment is performed using RNA-Seq by Expectation-Maximization RSEM (Li *et al.*, 2011). Reads that did not map to the transcriptome are then aligned to the mm10 genome sequence using Bowtie with default settings (Langmead et al., 2009). Reads that mapped to neither the transcriptome with RSEM or to the genome with Bowtie are mapped against the ERCC sequences. After alignment, quality control is performed, followed by clustering the single cell data into groups using a consensus approach based on two iterative clustering techniques - iterative weighted gene co-expression network analysis (WGCNA) (as described in Tasic *et al.*, 2016) and an iterative version of Seurat (as described in Macosko *et al*., 2015). For detailed information on RNA-Seq data generation from mouse, see the **Transcriptomics Overview** in the Documentation tab.

Human brain cells are transcriptionally profiled from postmortem or surgically resected tissue specimens. For postmortem brain processing, tissue is embedded in alginate, and coronal slices are cut at approximately 0.5-1.0 cm intervals and quick-frozen. Targeted areas for study (ie, middle temporal gyrus) are resected and sectioned at 500 µm on a vibrating microtome. Tissue procured from neurosurgical sources is acquired through a procurement workflow, and sliced at 350 µm on a vibrating microtome and stained. Individual cortical layers are microdissected using a fluoresecent Nissl stain. Tissue is homogenized to liberate nuclei, and single nuclei are captured with FACS by gating on DAPI and NeuN to isolate either NeuN-positive (neuronal) or NeuN-negative (non-neuronal) events. After sorting, cDNA amplification and library construction is performed using SMART-Seq v4 (Clontech) and Nextera XT (Illumina) kits. Single cell libraries are sequenced on HiSeq (Illumina) to generate 50 base-pair paired-end reads. Raw read (fastq) files are aligned to the GRCh38 human genome sequence (Genome Reference Consortium, 2011) with the RefSeq transcriptome version GRCh38.p2 (assembly date 12/5/2014). Transcriptome alignment is performed with the Spliced Transcripts Alignment to a Reference (STAR) protocol (Dobin *et al*., 2013) using default settings. Reads that did not map to the genome were then aligned to synthetic constructs sequences and the *E.coli* genome (version ASM584v2). After alignment, quality control is performed. For detailed information on RNA-Seq data generation from human tissue, see the **Transcriptomics Overview** and the **Case Qualification and Donor Profiles** summary in the Documentation tab.

GLIF Models

Rich data acquired from detailed electrophysiology can be utilized to created models of cell activity. Some of the electrophysiological stimuli protocols are designed to provide specific input for neuronal modeling. Several types of models of different levels of complexity aim to capture or predict a neuron's response to these stimuli. For simulations of neural networks, there is a tradeoff between the size of the network that can be simulated and the complexity of the model used for individual neurons. Generalized leaky integrate-and-fire (GLIF) models reduce synaptic integration occurring in the extended dendritic tree to a simple linear process; the output is compared to a single firing threshold. If the threshold is exceeded, a spike is generated and the internal variables are reset. A series of models of increasing complexity aim to improve reproducibility of the spiking behavior from the recorded neurons. Starting with a leaky integrate-and-fire model, three mechanisms are added: afterspike currents, reset rules in which membrane potential and threshold after a spike depend on the electrical state prior to the spike, and voltage dependent changes in threshold (Koch 1999; Pozzorini *et al*. 2015). Parameters are tuned to optimize the reproduced spike times based on a training noise stimulus. More precisely, the likelihood is maximized for a model neuron to reproduce exactly the observed spike train by using intrinsic noise. The model performance is subsequently evaluated on a test stimulus: for different time scales, the fraction of the variance of the neuronal response, which is explained by the model, is computed. For more information, see the **Neuronal Models: GLIF** overview in the Documentation tab.

Biophysical Models

The combination of experimental data from cortical neurons (electrophysiology and morphology) can be utilized to construct biophysically detailed models of individual cells (Koch 1999). To create biophysically detailed compartmental models, neural morphology and and electrophysiological responses are emulated in two different methods: (1) by assuming active conductances at the soma, with the rest of the neuron remaining passive, and (2) by assuming active conductances along the entirety of the neural morphology (all-active). Models are constructed for neurons that have high quality electrophysiological recordings and intact dendritic morphological reconstructions. For the first method, a genetic optimization algorithm is used, taking advantage of repeated single-cell simulations with the NEURON program (Hines and Carnevale, 1997) and directed evolution of biophysical parameters, following the approach of Segev and colleagues (Druckmann *et al*., 2007; Hay *et al*., 2011). For the second method, a genetic optimization algorithm run (Blue Gene Q supercomputer, Lugano, Switzerland) for the development of all-active models with algorithms, workflows, and analyses (Druckmann *et al*., 2007; Hay *et al*., 2011), adapted for experimental protocols also used in collaboration with the Blue Brain Project. Model performance is evaluated by comparing the model and experimental values of specific electrophysiological features computed from the somatic voltage response, to somatic current injection (including frequency of firing in response to a step current injection, average action potential peak values, average width of action potentials, and several others). For detailed methods, see the **Neuronal Models: Biophysical - Perisomatic** and **Neuronal Models: Biophysical - All Active** overviews in the Documentation tab.

Table 1. Cre transgenic lines that are the source of single cells for the Allen Cell Types Database.

Driver Line (Alias)	Originating Lab or Donating Investigator	Public Repository (Stock #)	Full Strain Name	Expression Pattern Summary	Usage	
					Ephys	RNA-Seq
Calb1-IRES2-Cre	Allen Institute for Brain Science	The Jackson Laboratory (028532)	B6;129S-*Calb1*$^{tm2.1(cre)Hze}$/J	Cre expression is enriched in layers 2/3 and 4 of cortex; and in restricted populations within hippocampal formation, striatum, thalamus, hypothalamus, midbrain, pons, medulla, and cerebellum.		✓
Chat-IRES-Cre	Bradford Lowell	The Jackson Laboratory (006410)	B6;129S6-*Chat*$^{tm1(cre)Lowl}$/J	Expressed in cholinergic neurons: restricted populations in medulla, pons, thalamus, midbrain; and in scattered cells in striatum, basal forebrain, and very sparse population of cells in cortex.	✓	✓
Chrna2-Cre_OE25	Nathaniel Heintz and Charles Gerfen	MMRRC (036502)	STOCK Tg(Chrna2-cre)OE25Gsat/Mmucd	Enriched in cortical layer 5. Expressed in restricted populations throughout brain, including olfactory areas, hippocampus, lateral septal complex, pallidum, thalamus, hypothalamus, midbrain, hindbrain, cerebellum.	✓	✓
Chrnb3-Cre_SM93	Nathaniel Heintz and Charles Gerfen	MMRRC (036469)	STOCK Tg(Chrnb3-cre)SM93Gsat/Mmucd			✓
Crh-IRES-Cre_ZJH	Z. Josh Huang	The Jackson Laboratory (012704)	B6(Cg)-*Crh*$^{tm1(cre)Zjh}$/J	Scattered expression in cortex. Enriched expression in piriform cortex, amygdala, hypothalamus, thalamus.		✓
Ctgf-2A-dgCre	Allen Institute for Brain Science	The Jackson Laboratory (028535)	B6.Cg-*Ctgf*$^{tm1.1(folA/cre)Hze}$/J	Cre expression is selective for layer 6b of cortex and in restricted populations within cortical subplate.	✓	✓
Cux2-CreERT2	Ulrich Mueller	MMRRC (032779)	B6(Cg)-*Cux2*$^{tm3.1(cre/ERT2)Mull}$/Mmmh	Enriched in cortical layers 2/3/4, thalamus, midbrain, pons, medulla and cerebellum.	✓	✓
Esr2-IRES2-Cre	Allen Institute for Brain Science	The Jackson Laboratory (030158)	B6;129S-*Esr2*$^{tm1.1(cre)Hze}$/J	Scattered expression in cortical layer 6. Restricted expression in the hypothalamus and amygdala.	✓	✓
Esr2-IRES2-Cre-neo	Allen Institute for Brain Science				✓	
Etv1-CreERT2	Z. Josh Huang	The Jackson Laboratory (013048)	B6(Cg)-*Etv1*$^{tm1.1(cre/ERT2)Zjh}$/J	Enriched in neocortical layer 5, and restricted populations in cerebellum, thalamus, hippocampus, piriform cortex, cortical subplate (amygdala).		✓

Driver Line (Alias)	Originating Lab or Donating Investigator	Public Repository (Stock #)	Full Strain Name	Expression Pattern Summary	Usage	
					Ephys	RNA-Seq
Gad2-IRES-Cre	Z. Josh Huang	The Jackson Laboratory (010802)	STOCK *Gad2*$^{tm2(cre)Zjh}$/J	Specific to GABAergic neurons. Enriched in striatum, piriform cortex, and in restricted populations in thalamus, hypothalamus, cerebellum, olfactory areas, and GABAergic interneurons of cortex.	✓	✓
Glt25d2-Cre_NF107	Nathaniel Heintz and Charles Gerfen	MMRRC (036504)	STOCK Tg(Colgalt2-cre)NF107Gsat/Mmucd	Reporter expression displays laminar enrichment in rostral cortex and in restricted populations within cortical subplate.	✓	✓
Gng7-Cre_KH71	Nathaniel Heintz and Charles Gerfen	MMRRC (031181)	STOCK Tg(Gng7-cre)KH71Gsat/Mmcd	Enriched in striatum and neocortical layers 2 and 5, and in restricted populations in medulla, pons, midbrain, hippocampus, hypothalamus, thalamus, and olfactory areas.		✓
Htr3a-Cre_NO152	Nathaniel Heintz and Charles Gerfen	MMRRC (036680)	STOCK Tg(Htr3a-cre)NO152Gsat/Mmucd	Reporter expression detected in subset of cortical interneurons. Enrichment is also detected in restricted populations in olfactory areas, pallidum, hypothalamus, pons, medulla, and cerebellum.	✓	✓
Ndnf-IRES2-dgCre	Allen Institute for Brain Science	The Jackson Laboratory (028536)	B6.Cg-*Ndnf*$^{tm1.1(folA/cre)Hze}$/J	Cre expression is restricted to layer 1 of cortex and in restricted populations in olfactory areas, hippocampal formation, striatum, midbrain.	✓	✓
Nkx2-1-CreERTT2	Z. Josh Huang	The Jackson Laboratory (014552)	STOCK *Nkx2-1*$^{tm1.1(cre/ERT2)Zjh}$/J	Strong scattered cells in restricted populations throughout the brain including striatum, cortical subplate (amygdala), hippocampus, and hypothalamus. Expressed in choroid plexus.	✓	
Nos1-CreERT2	Z. Josh Huang	The Jackson Laboratory (014541)	B6;129S-*Nos1*$^{tm1.1(cre/ERT2)Zjh}$/J	Expressed in molecular and granular layer in cerebellum. Strong expression in amygdala and olfactory bulb, particularly the accessory olfactory bulb, granular layer. Expressed in scattered cell populations in other brain regions.	✓	✓
Nr5a1-Cre (Sf1-Cre)	Bradford Lowell	The Jackson Laboratory (006364)	FVB-Tg(Nr5a1-cre)2Lowl/J	Expressed in restricted populations within the hypothalamus (ventromedial hypothalamus), and in cortical layer 4.	✓	✓
Ntsr1-Cre_GN220	Nathaniel Heintz and Charles Gerfen	MMRRC (030648)	B6.FVB(Cg)-Tg(Ntsr1-cre)GN220Gsat/Mmucd	Specific to cortical layer 6 neurons.	✓	✓
Oxtr-2A-Cre	Allen Institute for Brain Science			Cre expression is enriched in restricted populations within hippocampal formation, cortical subplate, striatum, thalamus, hypothalamus, and medulla. Sparse, scattered expression in cortex.	✓	✓
Pdyn-T2A-CreERT2	Allen Institute for Brain Science	The Jackson Laboratory (030197)	B6;129S-*Pdyn*$^{tm1.1(cre/ERT2)Hze}$/J	Sparse expression in cortical layer 5. Restricted expression in populations of the anterior hypothalamus and amygdala, and scattered expression in the posterior hypothalamus and amygdala.		✓
Penk-IRES2-Cre-neo	Allen Institute for Brain Science				✓	✓
Pvalb-2A-Dre	Allen Institute for Brain Science	The Jackson Laboratory (021190)	B6.Cg-*Pvalb*$^{tm3.1(dreo)Hze}$/J	Dre expression scattered throughout many brain regions including cortex, hippocampal formation, pons, medulla; and in restricted populations in pallidum, thalamus, midbrain, and cerebellum.	✓	✓
Pvalb-IRES-Cre	Silvia Arber	The Jackson Laboratory (008069)	B6;129P2-*Pvalb*$^{tm1(cre)Arbr}$/J	Expressed in restricted and/or sparse populations within the cerebellum, medulla, pons, midbrain, cortex, hippocampus, thalamus, and striatum.	✓	✓
Pvalb-2A-CreERT2	Allen Institute for Brain Science	The Jackson Laboratory (021189)	B6.Cg-*Pvalb*$^{tm1.1(cre/ERT2)Hze}$/J	Scattered expression throughout cortex. Enriched in restricted populations in cerebellum, medulla, pons, pallidum, thalamus.	✓	
Pvalb-2A-FlpO	Allen Institute for Brain Science	The Jackson Laboratory (022730)	B6.Cg-*Pvalb*$^{tm4.1(FLPo)Hze}$/J	Scattered expression throughout cortex. Enriched in restricted populations in cerebellum, medulla, pons, pallidum, thalamus.	✓	

Driver Line (Alias)	Originating Lab or Donating Investigator	Public Repository (Stock #)	Full Strain Name	Expression Pattern Summary	Usage	
					Ephys	RNA-Seq
Rasgrf2-T2A-dgFlpO	Allen Institute for Brain Science	The Jackson Laboratory (029589)	B6.Cg-*Rasgrf2*$^{tm2.1(folA/flpo)Hze}$/J	Scattered expression in upper cortical layers. Enrichment in specific populations of the pons and hypothalamus.		✓
Rbp4-Cre_KL100	Nathaniel Heintz and Charles Gerfen	MMRRC (031125)	STOCK Tg(Rbp4-cre)KL100Gsat/Mmucd	Enriched in cortical layer 5 and dentate gyrus.	✓	✓
Rorb-IRES2-Cre	Allen Institute for Brain Science	The Jackson Laboratory (023526)	B6;129S-*Rorb*$^{tm1.1(cre)Hze}$/J	Strong expression in zonal layer of superior colliculus and thalamus subregions. Dense, patchy expression in layer 4 and sparse expression in layers 5, 6 in cortex. Expressed in trigeminal nucleus and small patches of cells in cerebellum.	✓	
Rorb-IRES2-Cre-neo	Allen Institute for Brain Science					✓
Rorb-P2A-FlpO	Allen Institute for Brain Science			Expression in layer 4 and layer 5 of the cortex, pontine gray, thalamus, arcuate hypothalamic nucleus, and cortical amygdala.		✓
Scnn1a-Tg2-Cre	Allen Institute for Brain Science	The Jackson Laboratory (009112)	B6;C3-Tg(Scnn1a-cre)2Aibs/J	Reporter expression in sparse and/or restricted regions of cortex (layer 4), thalamus, midbrain, medulla, pons, and cerebellum.	✓	✓
Scnn1a-Tg3-Cre	Allen Institute for Brain Science	The Jackson Laboratory (009613)	B6;C3-Tg(Scnn1a-cre)3Aibs/J	Enriched in cortical layer 4 and in restricted populations within cortex, thalamus, and in cerebellum.	✓	✓
Sim1-Cre	Bradford Lowell	The Jackson Laboratory (006451)	B6.FVB(129X1)-Tg(Sim1-cre)1Lowl/J	Enriched in restricted populations in cerebellum, midbrain, hypothalamus, hippocampus (CA1), cortical subplate (amygdala), thalamus; area-dependent enrichment in cortex layers 2/3, 4, 6b.		✓
Sim1-Cre_KJ18	Nathaniel Heintz and Charles Gerfen	MMRRC (031742)	STOCK Tg(Sim1-cre)KJ18Gsat/Mmucd	Enriched in restricted populations in layer 5 of cortex, striatum (amygdala), hypothalamus. Sparse, scattered expression elsewhere in brain.	✓	✓
Slc17a6-IRES-Cre (VGLUT2-ires-Cre)	Bradford Lowell	The Jackson Laboratory (016963)	STOCK *Slc17a6*$^{tm2(cre)Lowl}$/J	Widespread expression throughout most of brain, except very sparse expression in striatum and restricted populations within cerebellum, medulla, and pons.	✓	✓
Slc17a7-IRES2-Cre	Allen Institute for Brain Science	The Jackson Laboratory (023527)	B6;129S-*Slc17a7*$^{tm1.1(cre)Hze}$/J	Strong expression throughout cortex, olfactory bulb, anterior olfactory nuclei. Scattered expression in striatum, hippocampus. Enriched in restricted populations in pons, superior colliculus, anterodorsal nucleus of thalamus.		✓
Slc17a8-IRES2-Cre	Allen Institute for Brain Science	The Jackson Laboratory (028534)	B6;129S-*Slc17a8*$^{tm1.1(cre)Hze}$/J	Scattered expression in many brain regions. Expresssion is enriched in layers 5 and 6 of cortex; and in restricted populations within pallidum, thalamus, midbrain, pons, medulla, and cerebellum.		✓
Slc32a1-IRES-Cre (VGAT-ires-Cre)	Bradford Lowell	The Jackson Laboratory (016962)	STOCK *Slc32a1*$^{tm2(cre)Lowl}$/J	Specific to GABAergic neurons. Enriched in striatum and in restricted populations in thalamus, hypothalamus, cerebellum, olfactory areas, and GABAergic interneurons of the cortex.		✓
Slc32a1-2A-FlpO	Allen Institute for Brain Science	The Jackson Laboratory (029591)	B6.Cg-*Slc32a1*$^{tm1.1(flpo)Hze}$/J	Widespread, scattered expression throughout the brain. Expressed in restricted populations within thalamus, hypothalamus, midbrain, pons, and medulla.	✓	
Snap25-IRES2-Cre	Allen Institute for Brain Science	The Jackson Laboratory (023525)	B6;129S-*Snap25*$^{tm2.1(cre)Hze}$/J	Strong widespread expression throughout the brain.		✓
Sst-IRES-Cre	Z. Josh Huang	The Jackson Laboratory (013044)	STOCK *Sst*$^{tm2.1(cre)Zjh}$/J	Strong scattered expression throughout brain. Localized areas of enrichment include restricted populations in thalamus, amygdala, midbrain, hindbrain, cortical subplate, Purkinje cell layer.	✓	✓

Driver Line (Alias)	Originating Lab or Donating Investigator	Public Repository (Stock #)	Full Strain Name	Expression Pattern Summary	Usage	
					Ephys	RNA-Seq
Sst-IRES-FlpO	Z. Josh Huang	The Jackson Laboratory (028579)	*Sst*$^{tm3.1(flpo)Zjh}$/J		✓	✓
Tac1-IRES2-Cre	Allen Institute for Brain Science	The Jackson Laboratory (021877)	B6;129S-*Tac1*$^{tm1.1(cre)Hze}$/J	Scattered expression in caudate, septum, hypothalamus, midbrain, hindbrain, cerebellum. Dense expression in accessory olfactory bulb and anterior olfactory nucleus, thalamus, VMH and midbrain structures (superior colliculus). Cre expression pattern is similar but more sparse.		✓
Th-Cre_FI172	Nathaniel Heintz and Charles Gerfen	MMRRC (031029)	B6.FVB(Cg)-Tg(Th-cre)FI172Gsat/Mmucd	Enriched in restricted populations in olfactory areas, cortical amygdalar area, hypothalamus, thalamus, midbrain, pons, medulla, cerebellum. Sparse, scattered expression in cerebral nuclei.		✓
Tlx3-Cre_PL56	Nathaniel Heintz and Charles Gerfen	MMRRC (036547)	STOCK Tg(Tlx3-cre)PL56Gsat/Mmucd	Reporter expression enriched in layer 5a of cortex and in restricted populations of pons and medulla.	✓	✓
Trib2-2A-CreERT2	Allen Institute for Brain Science	The Jackson Laboratory (022865)	B6.Cg-*Trib2*$^{tm1.1(cre/ERT2)Hze}$/J	Enriched in superficial layer 5 neurons of cortex. In scattered cells in cortex and throughout brain. Widespread in vasculature.		✓
Vip-IRES-Cre	Z. Josh Huang	The Jackson Laboratory (010908)	STOCK *Vip*$^{tm1(cre)Zjh}$/J	Strong scattered expression throughout brain. Enriched in superficial cortical layers and restricted populations in hindbrain and midbrain.	✓	✓
Vipr2-IRES2-Cre	Allen Institute for Brain Science			In subcortical areas, this Cre line has expression in arcuate hypothalamic nucleus, dorsal lateral geniculate nucleus, ventral nucleus of the thalamus, bed nuclei of the stria terminalis, and scattered expression in superior colliculus. In the cortex, there is scattered expression in neurons and non-neuronal cells.	✓	✓
Vipr2-IRES2-Cre-neo	Allen Institute for Brain Science			Expression is mostly restricted to subcortical areas, with enrichment in the geniculate complex of the thalamus, arcuate hypothalamic nucleus, and central amygdalar nucleus.	✓	

Table 2. Reporter line crossed to Cre transgenic lines.

Reporter Line	Expressed Gene	Originating Lab or Donating Investigator	Public Repository (Stock #)	Full Strain Name	Expression Pattern Summary
Ai14	tdTomato	Allen Institute for Brain Science	The Jackson Laboratory (007914)	B6.Cg-*Gt(ROSA)26Sor*$^{tm14(CAG-tdTomato)Hze}$/J	tdTomato fluorescent protein is expressed in cytoplasm.
Ai65(RCFL-tdT)	tdTomato	Allen Institute for Brain Science	The Jackson Laboratory (021875)	B6;129S-*Gt(ROSA)26Sor*$^{tm65.1(CAG-tdTomato)Hze}$/J	Expression of tdTomato fluorescent protein is dependent on both Cre and Flp recombinase activity.
Ai66(RCRL-tdT)	tdTomato	Allen Institute for Brain Science	The Jackson Laboratory (021876)	B6;129S-*Gt(ROSA)26Sor*$^{tm66.1(CAG-tdTomato)Hze}$/J	Expression of tdTomato fluorescent protein is dependent on both Cre and Dre recombinase activity.
Snap25-LSL-2A-GFP	EGFP	Allen Institute for Brain Science	The Jackson Laboratory (021879)	B6.Cg-*Snap25*$^{tm1.1Hze}$/J	Expression of EGFP fluorescent protein is limited to neurons.

REFERENCES

Cauli B, Audinat E, Lambolez B, Angulo MC, Ropert N, Tsuzuki K, Hestrin S, Rossier J (1997) Molecular and physiological diversity of cortical nonpyramidal cells. *Journal of Neuroscience* 15:3894-3906.

Cauli B, Porter JT, Tsuzuki K, Lambolez B, Rossier J, Quenet B, Audinat E (2000) Classification of fusiform neocortical interneurons based on unsupervised clustering. *Proceedings of the National Academy of Sciences USA* 97:6144-6149.

Crick FHC (1994) *The Astonishing Hypothesis: The Scientific Search for the Soul*. Simon & Schuster, New York.

DeFelipe J, López-Cruz PL, Benavides-Piccione R, Bielza C, Larrañaga P, Anderson S, Burkhalter A, Cauli B, Fairén A, Feldmeyer D, Fishell G, Fitzpatrick D, Freund TF, González-Burgos G, Hestrin S, Hill S, Hof PR, Huang J, Jones EG, Kawaguchi Y, Kisvárday Z, Kubota Y, Lewis DA, Marín O, Markram H, McBain CJ, Meyer HS, Monyer H, Nelson SB, Rockland K, Rossier J, Rubenstein JL, Rudy B, Scanziani M, Shepherd GM, Sherwood CC, Staiger JF, Tamás G, Thomson A, Wang Y, Yuste R, Ascoli GA (2013) New insights into the classification and nomenclature of cortical GABAergic interneurons. *Nature Reviews Neuroscience* 14:202-216.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1):15-21.

Druckmann S, Banitt Y, Gidon A, Schürmann F, Markram H, Segev I (2007) A novel multiple objective optimization framework for constraining conductance-based neuron models by experimental data. *Frontiers in Neuroscience* 1:7–18.

Hay E, Hill S, Schürmann F, Markram H, Segev I (2011) Models of neocortical layer 5b pyramidal cells capturing a wide range of dendritic and perisomatic active properties. *PLoS Computational Biology* 7:e1002107.

Herculano-Houzel S (2009) The human brain in numbers: a linearly scaled-up primate brain. *Frontiers in Human Neuroscience* 3: doi:10.3389/neuro.09.031.2009.

Hines ML, Carnevale NT (1997) The NEURON simulation environment. *Neural Computation* 9:1179–1209.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* 10:R25.

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323.

Koch C (1999) *Biophysics of Computation: Information Processing in Single Neurons.* Oxford University Press, New York.

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 161:1202-1214.

Madisen L, Zwingman TA, Sunkin SM, Oh SW, Zariwala HA, Gu H, Ng LL, Palmiter RD, Hawrylycz MJ, Jones AR, Lein ES, Zeng H (2010) A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neuroscience* 13:133-140.

Madisen L, Garner A, Shimaoka D, Chuong A, Klapoetke N, Li L, van der Bourg A, Niino Y, Egolf L, Monetti C, Gu H, Mills M, Cheng A, Tasic B, Nguyen T, Sunkin S, Benucci A, Nagy A, Niyawaki A, Helmchen F, Empson RM, Knöpfel T, Boyden E, Reid C, Carandini M, Zeng H (2015) Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. *Neuron* 85:942-958.

Masland, RH (2012) The neuronal organization of the retina. *Neuron* 76**:**266-280.

Peng H, Ruan Z, Long F, Simpson JH, Myers EW (2010) V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets. *Nature Biotechnology* 28:348-353.

Petilla Interneuron Nomenclature Group, Ascoli GA, Alonso-Nanclares L, Anderson SA, Barrionuevo G, Benavides-Piccione R, Burkhalter A, Buzsáki G, Cauli B, Defelipe J, Fairén A, Feldmeyer D, Fishell G, Fregnac Y, Freund TF, Gardner D, Gardner EP, Goldberg JH, Helmstaedter M, Hestrin S, Karube F, Kisvárday ZF, Lambolez B, Lewis DA, Marin O, Markram H, Muñoz A, Packer A, Petersen CC, Rockland KS, Rossier J, Rudy B, Somogyi P, Staiger JF, Tamas G, Thomson AM, Toledo-Rodriguez M, Wang Y, West DC, Yuste R (2008) Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nature Reviews Neuroscience* 9:557-568.

Pozzorini C, Mensi S, Hagens O, Naud R, Koch C, Gerstner W (2015) Automated high-throughput characterization of single neurons by means of simplified spiking models. *PLoS Computational Biology* 11(6):e1004275.

Stevens CF (1998) Neuronal diversity: Too many cell types for comfort? *Current Biology* 8:R708-R710.

Sugino K, Hempel CM, Miller MN, Hattox AM, Shapiro P, Wu C, Huang ZJ, Nelson SB (2006) Molecular taxonomy of major neuronal classes in the adult mouse forebrain. *Nature Neuroscience* 9**:**99-107.

Tasic B, Menon V, Nguyen TN, Kim TK, Jarsky T, Yao Z, Levi B, Gray LT, Sorensen SA, Dolbeare T, Bertagnolli D, Goldy J, Shapovalova N, Parry S, Lee C, Smith K, Bernard A, Madisen L, Sunkin SM, Hawrylycz M, Koch C, Zeng H (2016) Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nature Neuroscience* 19:335-346.

Thomson AM (2010) Neocortical layer 6, a review. *Frontiers in Neuroanatomy* 4:13.

Toledo-Rodriguez M, Blumenfeld B, Wu C, Luo J, Attali B, Goodman P, Markram H (2004) Correlation maps allow neuronal electrical properties to be predicted from single-cell gene expression profiles in rat neocortex. *Cerebral Cortex* 14:1310-1327.