Supplemental Materials

Methods

Transgenic mice

All animal procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at the Allen Institute for Brain Science. Transgenic mouse lines were generated using conventional and BAC transgenic, or knock-in strategies as previously described[1,2]. External sources included Cre lines generated as part of the NIH Neuroscience Blueprint Cre Driver Network (http://www.credrivermice.org) and the GENSAT project (http://gensat.org/), as well as individual labs. In transgenic lines with regulatable versions of Cre young adult tamoxifen-inducible mice (CreERT2) were treated with ~200 µl of tamoxifen solution (0.2 mg/g body weight) via oral gavage once per day for 5 consecutive days to activate Cre recombinase.

We used the transgenic mouse line Ai93, in which GCaMP6f expression is dependent on the activity of both Cre recombinase and the tetracycline controlled transactivator protein (tTA)[1]. Triple transgenic mice (Ai93, tTA, Cre) were generated by first crossing Ai93 mice with Camk2a-tTA mice, which preferentially express tTA in forebrain excitatory neurons[3]. Double transgenic mice were then crossed with a Cre driver line to generate mice in which GCaMP6f expression is induced in the specific populations of neurons that express both Cre and tTA. In a subset of mice, we alternatively leveraged the TIGRE2.0 transgenic platform that combines the expression of tTA and Gcamp6f in a single reporter line (Ai148(TIT2L-GC6f-ICL-tTA2)[4].

Cux2-CreERT2;Camk2a-tTA;Ai93(TITL-GCaMP6f) expression is regulated by the tamoxifen-inducible Cux2 promoter, induction of which results in Cre-mediated expression of GCaMP6f predominantly in superficial cortical layers 2, 3 and 4[5] (see Supplemental Figure 7). Both Emx1-IRES-Cre;Camk2a-tTA;Ai93 and Slc17a7-IRES2-Cre;Camk2a-tTA;Ai93 are pan-excitatory lines and show expression throughout all cortical layers[6,7]. SST-IRES-Cre;Ai148 exhibit GCaMP6f in somatostatin-expressing neurons[8]. VIP-IRES-Cre; Ai148 exhibit GCaMP6f in *Vip*-expressing cells by the endogenous promoter/enhancer elements of the vasoactive intestinal polypeptide locus[8]. Rorb-IRES2-Cre;Cam2a-tTA;Ai93 exhibit GCaMP6f in excitatory neurons in cortical layer 4 (dense patches) and layers 5,6 (sparse)[6]. Scnn1a-Tg3-Cre;Camk2a-tTA;Ai93 exhibit GCaMP6f in excitatory neurons in cortical layer 4 and in restricted areas within the cortex, in particular primary sensory cortices. Nr5a1-Cre;Camk2a-tTA;Ai93 exhibit GCaMP6f in excitatory neurons in cortical layer 4[9]. Rbp4-Cre;Camk2a-tTA;Ai93 exhibit GCaMP6f in excitatory neurons in cortical layer 5[10]. Fezf2-CreER;Ai148 exhibits GCaMP6f in subcerebral projection neurons in the layer 5 and 6[11]. Tlx3-Cre_PL56;Ai148 exhibits GCaMP6f primarily restricted to IT corticostriatal in the layer 5[10]. Ntsr1-Cre_GN220;Ai148 exhibit GCaMP6f in excitatory corticothalamic neurons in cortical layer 6[12].

We maintained all mice on a reverse 12-hour light cycle following surgery and throughout the duration of the experiment and performed all experiments during the dark cycle.

Cross platform registration

We developed an integrated suite of tools and procedures that leveraged surgical implant hardware, mouse behavior platforms, and imaging instruments. These tools provided a means to register data acquired between instruments and repeatedly target and record neurons in brain areas identified with intrinsic imaging. (1) Each animal was implanted with a stereotaxically-aligned headframe that provided a cranial window for brain imaging and permitted head fixation in a reproducible configuration (see *Surgery*). (2) All behavioral hardware components were custom-designed and assembled in house so that we could register the underlying geometry to a common coordinate system (see *Intrinsic Imaging*, *Two photon in vivo calcium imaging and* Supplementary Figure 1a). (3) All imaging datasets were registered to this common coordinate system using shared reticles clamped throughout the data collection pipeline (Supplementary Figure 1b). We maintained this coordinate system by monitoring the stability of the reticles on a weekly basis across all steps of the data collection pipeline. Any excessive deviations were flagged for further inspection (Supplementary Figure 1c). (4) We converted locations from the intrinsic imaging cortical map into stage coordinates on the two-photon imaging system so as to allow repetitive targeting of individual field of views.

Surgery

Prior to implantation, a 3D printed acrylic photopolymer microscope well (to facilitate the use of liquid immersion objectives) was glued to the titanium headframe with Loctite 406 using a jig to ensure uniform offset between the center of the well and the reference surfaces of the clamp (Supplemental Figure 2b).

Transgenic mice expressing GCaMP6f were weaned and genotyped at ~p21, and surgery was performed between p37 and p63. Surgical eligibility criteria included: 1) weight ≥19.5g (males) or ≥16.7g (females); 2) normal behavior and activity; and 3) healthy appearance and posture. A pre-operative injection of dexamethasone (3.2 mg/kg, S.C.) was administered 3h before surgery. Mice were initially anesthetized with 5% isoflurane (1-3 min) and placed in a stereotaxic frame (Model# 1900, Kopf, Tujunga, CA), and isoflurane levels were maintained at 1.5-2.5% for the duration of the surgery. An injection of carprofen (5-10 mg/kg, S.C.) was administered and an incision was made to remove skin, and the exposed skull was levelled with respect to pitch (bregma-lamda level), roll and yaw.

The stereotax was zeroed on lambda using a custom headframe holder equipped with a stylus affixed to a clamp-plate (Supplemental Figure 2c). The stylus was then replaced with the headframe, which was lowered and affixed to the skull with Metabond. Once dried, the mouse was placed in a custom clamp (Supplemental Figure 2d) to

position the skull at a rotated angle of 23 and pitch angle of 6, such that visual cortex was horizontal to facilitate the craniotomy. The craniotomy was centered at X = -2.8mm and Y = 1.3mm with respect to lambda (centered over the left mouse visual cortex). A circular piece of skull 5 mm in diameter was removed, and a durotomy was performed. A coverslip stack (two 5mm and one 7mm glass coverslip adhered together) was cemented in place with Vetbond[13]. Metabond cement was applied around the cranial window inside the well to secure the glass window. Post-surgical brain health was documented using a custom photo-documentation system (Supplemental Figure 2e) to acquire a spatially registered image of the cranial window. One, two, and seven days following surgery, animals were assessed for overall health (bright, alert and responsive), cranial window clarity and brain health.

Intrinsic Imaging

A retinotopic map was created using intrinsic signal imaging (ISI) in order to define visual area boundaries and target *in vivo* two-photon calcium imaging experiments to consistent retinotopic locations[14]. Mice were lightly anesthetized with 1-1.4% isoflurane administered with a somnosuite (model #715; Kent Scientific, CON). Vital signs were monitored with a Physiosuite (model # PS-MSTAT-RT; Kent Scientific). Eye drops (Lacri-Lube Lubricant Eye Ointment; Refresh) were applied to maintain hydration and clarity of eye during anesthesia. Mice were placed on a lab jack platform and headfixed for imaging normal to the cranial window.

The brain surface was illuminated with two independent LED lights: green (peak λ=527nm; FWHM=50nm; Cree Inc., C503B-GCN-CY0C0791) and red (peak λ=635nm and FWHM of 20nm; Avago Technologies, HLMP-EG08-Y2000) mounted on the optical lens. A pair of Nikon lenses lens (Nikon Nikkor 105mm f/2.8, Nikon Nikkor 35mm f/1.4), provided 3.0x magnification (M=105/35) onto an Andor Zyla 5.5 10tap sCMOS camera. A bandpass filter (Semrock; FF01-630/92nm) was used to only record reflected red light onto the brain.

A 24” monitor was positioned 10 cm from the right eye. The monitor was rotated 30° relative to the animal’s dorsoventral axis and tilted 70° off the horizon to ensure that the stimulus was perpendicular to the optic axis of the eye. The visual stimulus displayed was comprised of a 20° x 155° drifting bar containing a checkerboard pattern, with individual square sizes measuring 25º, that alternated black and white as it moved across a mean-luminance gray background. The bar moved in each of the four cardinal directions 10 times. The stimulus was warped spatially so that a spherical representation could be displayed on a flat monitor[15].

After defocusing from the surface vasculature (between 500 µm and 1500 µm along the optical axis), up to 10 independent ISI timeseries were acquired and used to measure the hemodynamic response to the visual stimulus. Averaged sign maps were produced from a minimum of 3 timeseries images for a combined minimum average of 30 stimulus sweeps in each direction[16].

The resulting ISI maps were automatically segmented by comparing the sign, location, size, and spatial relationships of the segmented areas against those compiled in an ISI-derived atlas of visual areas. A cost function, defined by the discrepancy between the properties of the matched areas, was minimized to identify the best match between visual areas in the experimental sign map and those in the atlas, resulting in an auto-segmented and annotated map for each experiment. Manual correction and editing of the results included merging and splitting of segmented and annotated areas to correct errors.

Finally, target maps were created to guide *in vivo* two-photon imaging location using the retinotopic map. Target locations were identified for each visual area, restricted to within 10° of the center of gaze.

Habituation

Following successful ISI mapping, mice spent two weeks being habituated to head fixation and visual stimulation. During the first week mice were handled and head fixed for progressively longer durations, ranging from 5 to 10 minutes. During the second week, mice were head fixed and presented with visual stimuli, starting for 10 minutes and progressing to 50 minutes of visual stimuli by the end of the week. During this week they were exposed to all of the stimuli used during data collection. Mice received a single 60 min habituation session on the two-photon microscope, during which they were head fixed under the objective and a stimulus was presented.

Two photon in vivo calcium imaging

Calcium imaging was performed using a two-photon-imaging instrument (either a Scientifica Vivoscope or a Nikon A1R MP+; the Nikon system was adapted to provide space to accommodate the behavior apparatus). Laser excitation was provided by a Ti:Sapphire laser (Chameleon Vision – Coherent) at 910 nm. Pre-compensation was set at ~10,000 fs2. Movies were recorded at 30Hz using resonant scanners over a 400 μm field of view. Temporal synchronization of all data-streams (calcium imaging, visual stimulation, body and eye tracking cameras) was achieved by recording all experimental clocks on a single NI PCI-6612 digital IO board at 100 kHz.

Mice were head-fixed on top of a rotating disk and free to walk at will. The disk was covered with a layer of removable foam (Super-Resilient Foam, 86375K242, McMaster) to alleviate motion-induced artifacts during imaging sessions. All two-photon imaging experiments were conducted under ambient red light to maintain the reversed day-night cycle. Data was initially obtained with the mouse eye centered both laterally and vertically on the stimulation screen and positioned 15 cm from the screen, with the screen parallel to the mouse's body. Later, the screen was moved to better fill the visual field. The normal distance of the screen from the eye remained at 15 cm, but the screen center moved to a position 118.6 mm lateral, 86.2 mm anterior and 31.6 mm dorsal to the right eye.

An experiment container consisted of three imaging sessions (60 min each) at a given field of view during which mice passively observed three different stimuli. One imaging session was performed per day, for a maximum of 16 sessions for each mouse.

On the first day of imaging at a new field of view, the ISI targeting map was used to select spatial coordinates. A comparison of superficial vessel patterns was used to verify the appropriate location by imaging over a field of view of ~800 μm using epi-fluorescence and blue light illumination. Once a cortical region was selected, the objective was shielded from stray light coming from the stimulation screen using opaque black tape. In two-photon imaging mode, the desired depth of imaging was set to record from a specific cortical depth. On subsequent imaging days, we returned to the same location by matching (1) the pattern of vessels in epi-fluorescence with (2) the pattern of vessels in two photon imaging and (3) the pattern of cellular labelling in two photon imaging at the previously recorded location.

Once a depth location was stabilized, a combination of PMT gain and laser power was selected to maximize laser power (based on a look-up table against depth) and dynamic range while avoiding pixel saturation (max number of saturated pixels <1000). The stimulation screen was clamped in position, and the experiment began. Two-photon movies (512x512 pixels, 30Hz), eye tracking (30Hz), and a side-view full body camera (30Hz) were recorded and continuously monitored. Recording sessions were 1 hour long but were interrupted if any of the following was observed: 1) mouse stress as shown by excessive secretion around the eye, nose bulge, and/or abnormal posture; 2) excessive pixel saturation (>1000 pixels) as reported in a continuously updated histogram; 3) loss of baseline intensity in excess of 20% caused by bleaching and/or loss of immersion water; 4) hardware failures causing a loss of data integrity. Immersion water was occasionally supplemented while imaging using a micropipette taped to the objective (Microfil MF28G67-5 WPI) and connected to a 5 ml syringe via an extension tubing. At the end of each experimental session, a z-stack of images (+/- 30 μm around imaging site, 0.1 μm step) was collected to evaluate cortical anatomy and evaluate z-drift during the course of experiment. Experiments with z-drift above 10μm over the course of the entire session were excluded. In addition, for each experimental area analyzed, a full-depth cortical z stack (~700 μm total depth, 5 μm step) was collected to document the imaging site location.

Visual Stimulation

Visual stimuli were generated using custom scripts written in PsychoPy[17,18] (Peirce, 2007, 2008) and were displayed using an ASUS PA248Q LCD monitor, with 1920 x 1200 pixels. Stimuli were presented monocularly, and the monitor was positioned 15 cm from the mouse's eye, and spanned 120° X 95° of visual space. Each monitor was gamma corrected and had a mean luminance of 50 cd/m^2. To account for the close viewing angle of the mouse, a spherical warping was applied to all stimuli to ensure that the apparent size, speed, and spatial frequency were constant across the monitor as seen from the mouse's perspective.

Visual stimuli included drifting gratings, static gratings, locally sparse noise, natural scenes and natural movies. These stimuli were distributed across three ~60 minute imaging sessions (Figure 1f). During session A the drifting gratings, natural movie one and natural movie three stimuli were presented. During session B the static gratings, natural scenes, and natural movie one were presented. During session C the locally sparse noise, natural movie one and natural move two were presented. In each session, the different stimuli were presented in segments of 5-13 minutes and interleaved with each other. In addition, at least 5 minutes of spontaneous activity were recorded in each session.

The drifting gratings stimulus consisted of a full field drifting sinusoidal grating at a single spatial frequency (0.04 cycles/degree) and contrast (80%). The grating was presented at 8 different directions (separated by 45°) and at 5 temporal frequencies (1, 2, 4, 8, 15 Hz). Each grating was presented for 2 seconds, followed by 1 second of mean luminance gray before the next grating. Each grating condition (direction & temporal frequency combination) was presented 15 times. Trials were randomized, with blank sweeps (i.e. mean luminance gray instead of grating) presented approximately once every 20 trials.

The static gratings stimulus consisted of a full field static sinusoidal grating at a single contrast (80%). The grating was presented at 6 different orientations (separated by 30°), 5 spatial frequencies (0.02, 0.04, 0.08, 0.16, 0.32 cycles/degree), and 4 phases (0, 0.25, 0.5, 0.75). The grating was presented for 0.25 seconds, with no inter-grating gray period. Each grating condition (orientation, spatial frequency, and phase) was presented ~50 times. Trials were randomized, with blank sweeps presented approximately once every 25 trials.

The natural scenes stimulus consisted of 118 natural images. Images were taken from the Berkeley Segmentation Dataset[19], the van Hateren Natural Image Dataset[20], and the McGill Calibrated Colour Image Database[21]. The images were presented in grayscale and were contrast normalized and resized to 1174 X 918 pixels. The images were presented for 0.25 seconds each, with no inter-image gray period. Each image was presented ~50 times. Trials were randomized, with blank sweeps approximately once every 100 images.

Three natural movie clips were used from the opening scene of the movie *Touch of Evil*[22]. Natural Movie One and Natural Movie Two were both 30 second clips while Natural Movie Three was a 120 second clip. All clips had been contrast normalized and were presented in grayscale at 30 fps. Each movie was presented 10 times with no inter-trial gray period. Natural Movie One was presented in each imaging session.

The locally sparse noise stimulus consisted of white and dark spots on a mean luminance gray background. Each spot was a square, 4.65° on a side. Each frame of the stimulus had ~11 spots on the monitor, with no two spots within 23° of each other, and was presented for 0.25 seconds. Each of the 16 x 28 spot locations was occupied by white and black spots a variable number of time (mean=115). For most of the collected data, this

stimulus was adapted such that half of it used 4.65° spots while the other half used 9.3° spots, with an exclusion zone of 46.5°.

Serial Two-Photon Tomography

Serial two-photon tomography was used to obtain a 3D image volume of coronal brain images for each specimen. This 3D volume enables spatial registration of each specimen's associated ISI and optical physiology data to the Allen Mouse Common Coordinate Framework (CCF). Methods for this procedure have been described in detail in whitepapers associated with the Allen Mouse Brain Connectivity Atlas in Documentation for that resource, and in the associated publication[23].

Mice were anesthetized with 5% isoflurane and intracardially perfused with 10 ml of saline (0.9% NaCl) followed by 50 ml of freshly prepared 4% paraformaldehyde (PFA) at a flow rate of 9 ml/min. Brains were rapidly dissected and post-fixed in 4% PFA at room temperature for 3-6 hours and overnight at 4°C. Brains were then rinsed briefly with PBS and stored in PBS with 0.02% sodium azide before proceeding to the next step. Agarose was used to embed the brain in a semisolid matrix for serial imaging. The brain was placed in a 4% oxidized agarose solution made by stirring 10 mM $NaIO_4$ in agarose, then transferring through 50 mM phosphate buffer and embedding at 60°C in a grid-lined embedding mold to standardize placement of the brain in an aligned coordinate space. The agarose block was then left at room temperature for 20 minutes to allow agarose to solidify, and then covalent interaction between the brain tissue and the agarose was promoted by placing the block in 0.2% sodium borohydride in 50 mM sodium borate buffer (pH 9.0) for 48 hours at 4°C. The agarose block was then mounted on a 1x3 glass slide using Loctite 404 glue and prepared immediately for serial imaging.

Multi-photon image acquisition was performed using a customized TissueCyte 1000 system (TissueVision, Cambridge, MA) coupled with an ultra-fast mode-locked Ti:Sapphire laser. First the mounted specimen was placed on the metal plate in the center of the cutting bath, which was filled with PBS with 0.02% sodium azide and placed onto the sample stage. A new vibratome blade was used for each specimen and aligned to be parallel to the leading edge of the specimen block. Next, the top surface of the specimen block was brought up to the level of the vibratome blade by adjusting the sample stage height. The z-stage was set to slice at 100 µm intervals. Specimens were oriented for image acquisition to occur from the caudal to the rostral end. The XY scan area consists of 221 tiles (17 rows x 13 columns). Each tile was imaged at a resolution of 0.3 µm/pixel or 0.8 µm/pixel. The specimen was illuminated with a 925 nm wavelength laser with a Zeiss 20x water immersion objective (NA = 1). A 560 nm dichroic (Chroma, Bellows Falls, VT) split the emission light, and a 500 nm dichroic (Chroma) further split the emission for a total of three channels. The 593/40 nm (Chroma), 520/35 nm (Semrock, Rochester, NY) and 447/60 nm emission filter (Chroma) were used for the Red, Green and Blue channels, respectively. In order to scan a full tissue section, individual tile images were acquired, and the entire stage (Physik Instrumente) was moved between each tile. After an entire section was imaged, the X and Y stages moved the specimen to the vibratome, which cut a 100

μm section and returned the specimen to the objective for imaging of the next section. The blade vibrated at 60 Hz and the stage moved toward blade at 0.5 mm/sec during cutting.

Post-mortem assessment of brain structure

Morphological and structural analysis of each experimental mouse's brain was performed following collection of the 2P serial imaging (TissueCyte) dataset (Supplemental Figure 6).

The following characteristics warranted an automatic failure of all datasets associated with the mouse: (1) Abnormal GCaMP6 expression pattern; (2) Necrotic brain tissue; (3) Compression of the contralateral cortex that resulted in disruption to the cortical laminar structure; (4) Compression of the ipsilateral cortex (caused by a skull growth) or adjacent to the cranial window.

The following characteristics may have resulted in, but did not warrant automatic, failure of the datasets associated with the mouse. (1) Compression of the contralateral cortex due to a skull growth; (2) Excessive compression of the cortex underneath the cranial window; (3) Abnormal or enlarged ventricles.

Image processing

For each two-photon imaging session, the image processing pipeline performed, in order, 1) spatial or temporal calibration specific to a particular microscope, 2) motion correction, 3) image normalization to minimize confounding random variations between sessions, 4) segmentation of connected shapes, and 5) classification of soma-like shapes from remaining clutter (Supplemental Figure 9). Once all the parameters were initially tuned, the pipeline ran fully automatically across all of the varied experimental conditions of the campaign without manual intervention.

The motion correction algorithm relied on phase correlation and only corrected for rigid translational errors. It performed the following steps. Each movie was partitioned into 400 consecutive frame blocks, representing 13.3 seconds of video. Each block was registered iteratively to its own average 3 times (Supplementary Figure 9a-b). A second stage of registration integrated the periodic average frames themselves into a single global average frame through 6 additional iterations (Supplementary Figure 9c). The global average frame served as the reference image for the final resampling of every raw frame in the video (Supplementary Figure 9d).

Each 13.3 second block was used to generate normalized periodic averages using the following steps. First, we subtracted the mean from the maximum projection to retain pixels from active cells (Supplementary Figure 9e-f-g). To select objects of the right size during segmentation, we convolved all periodic normalized averages with a 3x3 median filter and a 47x47 high-pass mean filter. We then normalized the histogram of all resulting frames (Supplementary Figure 9g-h).

All normalized periodic averages were then segmented using an adaptive threshold filter to create an initial estimate of binarized ROI masks of unconnected components (Supplemental Figure 9i). Given GCaMP6 lower expression in cell nucleli, good detections from somata tended to show bright outlines and dark interiors. We then performed a succession of morphological operations to fill closed holes and concaves shapes (Supplemental Figure 9j-k).

These initial ROI masks cluded shapes from multiple periods that were actually from a single cell. To further reduce the number of masks to putative individual cell somas, we computed a feature vector from each masks that included morphological attributes such as location, area, perimeter, and compactness, among others (Supplemental Figure 9l). A battery of heuristic decisions applied on these attributes allowed to combine, eliminate or maintain ROI (Supplemental Figure 9l-m). A final discrimination step, using a binary relevance classifier fed by experimental metadata (Cre, imaging depth) along with the previous morphological features, further filtered the global masks into the final ROIs used for trace extraction.

Neuropil Subtraction

To correct for contamination of the ROI calcium traces by surrounding neuropil, we modeled the measured fluorescence trace of each cell as $F_M = F_C + rF_N$, where F_M is the measured fluorescence trace, F_C is the unknown true ROI fluorescence trace, F_N is the fluorescence of the surrounding neuropil, and r is the contamination ratio. To estimate the contamination ratio for each ROI, we selected the value of r that minimized the cross-validated error, $E = \sum_t |F_C - F_M + rF_N|^2$, over four folds. This minimization was performed by computing the error over each fold with a fixed value of r, for a range of r values. For each fold, F_C was computed by minimizing the cost function $C = \sum_t |F_C - F_M + rF_N|^2 + \lambda |LF_C|^2$, where L is the discrete first derivative (to enforce smoothness of F_C) and λ is a penalty parameter we set to 0.05. After determination of r, we computed the true trace as $F_C = F_M - rF_N$, which is used in all subsequent analysis.

Demixing traces from overlapping ROIs

To avoid artificially correlating neurons' activity by averaging fluorescence over two spatially overlapping ROIs, we demixed the activity of all recorded ROIs. We used a model where every ROI had a trace distributed in some spatially heterogeneous, time-dependent fashion:

$$F_{it} = \sum_k W_{kit} T_{kt}$$

where *W* is a tensor containing time-dependent weighted masks: W_{kit} measures how much of neuron k's fluorescence is contained in pixel i at time t. T_{kt} is the fluorescence trace of neuron k at time t - this is what we want to estimate. F_{it} is the recorded fluorescence in pixel i at time t.

Importantly, this model applied to all ROIs, including those too small to be a neuron or otherwise filtered out. We filtered out duplicate ROIs (defined as two ROIs with >70% overlap) and ROIs that were the union of two other ROIs (any ROI where the union of any other two ROIs accounted for 70% of its area) before demixing and applied the remaining filtering criteria after demixing. Projecting the movie F onto the binary masks, A, reduced the dimensionality of the problem from 512x512 pixels to the number of ROIs:

$$\sum_i A_{ki} F_{it} = \sum_{k,i} A_{ki} W_{kit} T_{kt}$$

where A_{ki} is one if pixel i is in ROI k and zero otherwise–these are the ROI masks from segmentation, after filtering out duplicate and union ROIs. At a particular time point t, this yields the simple linear regression:

$$AF(t) = (AW^T(t))T(t)$$

where we estimated the weighted masks W by the projection of the recorded fluorescence F onto the binary ROI masks A. On every imaging frame t, we computed the linear least squares solution $\hat{T}$ to extract each ROI's trace value at that time point.

It was possible for ROIs to have negative or zero demixed traces $\hat{T}$. This occurred if there were union ROIs (one ROI composed of two neurons) or duplicate ROIs (two ROIs in the same location with approximately the same shape) that our initial detection missed. If this occurred, those ROIs and any that overlapped with them were removed from the experiment. This led to the loss of ~1% of ROIs.

ROI Matching

The field of view for each session, and the segmented ROI masks, were registered to each other using an affine transformation. To map cells, a bipartite graph matching algorithm was used to find correspondence of cells between sessions A and B, A and C, and B and C. The algorithm took cells in the pair-wise experiments as nodes, and the degree of spatial overlapping and closeness between cells in the two experiments as weight of edge of the nodes. By maximizing the summed weights of edges, the bipartite matching algorithm found the best matching between cells of the two experiments. Finally, a label combination process was applied to the matching results of A and B, A and C, and B and C, producing a unified label for all three experiments.

$\Delta F/F$

To calculate the $\Delta F/F$ for each fluorescence trace, we first calculate baseline fluorescence using a median filter of width 5401 samples (180 seconds). We then calculate the change in fluorescence relative to baseline fluorescence (ΔF), divided by baseline fluorescence (F). Due to preprocessing of the fluorescence traces, in rare cases the estimated baseline fluorescence can be very small or negative. To prevent spurious

$\Delta F/F$ estimates, we thus set the baseline as the maximum of the median filter estimated baseline and the standard deviation of the estimated noise of the fluorescence trace.

L0 penalized event detection

The measured fluorescence traces that result from calcium imaging are noisy and slow reflections of underlying neural spikes. To properly estimate the functional properties of neurons we wish to infer their instantaneous spiking activity based on their measured fluorescence traces. A recent paper demonstrated a fast algorithm for finding the exact L0-penalized solution for the event detection problem [Jewell, Witten, 2017] and showed that the L0-penalized solutions typically outperform L1-penalized solutions in the detection of spikes from fluorescence traces. This L0-penalized algorithm forms the basis for our event detection pipeline. As we show below, low firing rate activity such as isolated spikes is difficult to detect in calcium signals, so we refer to this as "event" detection.

For each $\Delta F/F$ trace we remove slow timescale shifts in the fluorescence using a median filter of width 101 samples (3.3 seconds). We then apply the L0-penalized algorithm to the corrected $\Delta F/F$ trace. The L0 algorithm has two hyperparameters: gamma and lambda. Gamma corresponds to the decay constant of the calcium indicator. We set gamma to be the decay constant obtained from jointly recorded optical and electrophysiology with the same genetic background and calcium indicator. Supplemental Figure 11 shows the extracted linear kernels for Emx1-Ai93 and Cux2-Ai93 from which gamma has been extracted by fitting the fluorescence decay with a single exponential. The rise time, amplitude, and shape of the extracted linear kernels are mainly a function of the genetically encoded calcium indicator (GCaMP6f) and appear to be largely independent of the specific promoter driving expression.

Lambda controls the strength of the L0 penalty. To estimate lambda, we first estimate the standard deviation of the intrinsic noise in the trace. We then set lambda to minimize the number of events smaller than two standard deviations of the noise distribution, while retaining at least one recovered event. The values for gamma and lambda as described above are then used to obtain a set of predicted neural events. We use these event traces as input for all of our various analyses.

To assess how the events detected using the above procedure relate to actual measured action potentials (spikes) that underlie the fluorescence time series, we performed L0 event detection on the calcium-related fluorescence of cells that had been imaged simultaneously with loose patch recordings. Since the true spike train is known for these data, we computed the expected probability of detecting an event, as well as the expected event magnitude, as a function of the number of spikes observed in a set of detection windows relevant to the pipeline data analyses (e.g. static gratings, natural scenes, and locally sparse noise templates are presented for 0.25 s each). Supplemental Figure 11cd show results for Emx1-Ai93 and Cux2-Ai93, respectively.

The error bars in the graphs represent twice the standard error of the mean across cells. The false positive rate, as measured by the probability of detecting an event given

no spike in the simultaneously recorded membrane potential in a window of given length, is generally low but increases with detection window length. Bursts within short time windows correspond to high instantaneous firing rates and result in detected events with high likelihood (e.g. >5 spikes within 100 ms result in events detected with > 80% chance). Bursts within longer windows can have lower instantaneous firing rates (i.e. longer average inter-spike intervals), and thus, for a given spike count, the average detection probability decreases with increasing detection window length. For similar reasons, expected event magnitudes tend to decrease for longer event detection windows as well. The size of the error bars, in turn, decreases with increasing window length because for longer windows, more instances of any given spike count were observed.

Overall, the calibration results are quite similar across the two juxtaposed promoters (Cux2 vs Emx1), which were compared for excitatory Layer 2/3 pyramidal cells. For both promoters, the chances of detecting an event in response to a single spike is negligible (<10%) for short windows, and not significantly higher than the false positive rate for longer windows. 50% event detection probability is reached for 4 spikes (Cux2) and 3 spikes (Emx1), respectively within windows shorter than 500 ms. The expected magnitude of the detected events is largely a monotonic function of the number of underlying neural spikes. However, the size of the error bars suggests that this information may not be entirely sufficient to precisely infer spike counts from event magnitudes. Moreover, for Emx1, this relationship seems to saturate earlier than for Cux2. These differences between Cux2 and Emx1 are minor but seem consistent with higher levels of GCaMP6f expression under the Emx1 promoter.

Analysis

All analysis was performed using custom scripts written in Python using NumPy, SciPy[24], Pandas[25] and Matplotlib[26].

Using the derived events, the average responses to each stimulus condition was computed for each cell in response to the drifting gratings, static gratings and natural scenes. From this, the preferred stimulus condition was identified as the stimulus condition that elicited the largest average response. For each trial of the stimulus, the neural activity of the cell was compared to a distribution of activity for that cell taken during the epoch of spontaneous activity, and a p-value computed. If at least 25% of the trials in response to a cell's preferred condition had a significant difference from the distribution of spontaneous activities, the cell was deemed to be responsive to that stimulus and included in further analyses.

Direction selectivity was computed from mean responses to drifting gratings, at the cell's preferred temporal frequency, as

$$gDSI = \frac{\sum R_\theta e^{i\theta}}{\sum R_\theta}$$

where θ is the direction of grating movement, and R_θ is the mean response to that direction of movement.

The temporal frequency tuning, at a cell's preferred orientation, was fit using either an exponential curve (if the peak temporal frequency was at either extreme) or a Gaussian curve (if the peak temporal frequency was at an intermediate value). The reported preferred temporal frequency was taken from these fits. The same was done for spatial frequency tuning, fit at the cell's preferred orientation and phase in response to the static gratings. In both cases, if a fit could not converge, a preferred frequency was not reported.

Spatial receptive fields were fit from the responses to locally sparse noise. Because more than one stimulus spot was present during a given trial, it is not possible to conclude (on a per-trial basis) the stimulus-response relationship between spot locations and responses. Therefore, a statistically significant co-occurrence of spot presentation and responses across trials defined the inclusion criteria for membership of a stimulus pixel in the receptive field. To begin, the stimulus was convolved with a spatial Gaussian (4.65° per sigma), to allow pooling of contributions to responses from nearby pixels. A p-value is was computed for each spot (black and white separately) by constructing a null distribution for the number of trials that a spot was present during responsive trials. This per-pixel null distribution was estimated by shuffling the identity of the responsive trials (n=10,000 shuffles), breaking the relationship between stimulus and response under the assumption of a background level of responsiveness independent of the stimulus. Statistical outliers (i.e. pixels present during events more often that can be accounted for by chance) were identified by computing a p-value for each spot relative to its null (shuffled) distribution. These p-values were then corrected for false discoveries using the Šidák multiple comparisons correction, and thresholded at p=0.05 to identify receptive field membership.

An additional test was performed to confirm the presence of a receptive field. We defined the null model that a neuron lacking a receptive field has equal probability of producing a response regardless of the location and luminance (i.e. black or white) of the spots displayed on the screen on any given trial. In contrast, a neuron has a statistically significant receptive field if there is a deviation beyond chance from responsive trials occurring with equal probability across all spot locations and luminances. Therefore, chi-square tests for independence were performed for each neuron to quantify the dependence of responsive trials on the stimulus.

An assumption of the chi-square test is that the response of the neuron on a given trial can only be attributed to a single spot; that is, only a single stimulus spot is presented on each trial. Although multiple non-gray spots appeared on the screen during each trial, the exclusion region of the locally sparse noise stimulus, by construction, prevented two non-gray pixels within a 23° radius (for the 4.65° spot size) or 46° radius (for the 9.3° spot size) of one another from being presented on the same trial. Leveraging this structure in the stimulus, chi-square tests were performed on patches in visual space small enough to ensure that two or more non-gray pixels were rarely presented on the same trial, but large enough to ensure that the patch completely contains the receptive field and includes visual space outside of the receptive field in order for the chi-square test to detect the dependence of neuron responses on spot locations. We chose 32.2°x32.2° patches for 4.65° spots and 64.4°x64.4°patches for the 9.3° spot LSN (i.e. 7x7 grid of spot locations in each case). For each neuron, multiple chi-square tests were performed on such patches to

tile the entire stimulus monitor and the p-values from these tests were then corrected using the Šidák method to account for the multiple comparisons. If the p-value for any patch on the stimulus monitor yielded was significant ($p<0.05$) after multiple comparison correction, the neuron was considered to have a receptive field.

Finally, if a neuron was found to have a receptive field, the spots that were identified for receptive field membership were fit with a two-dimensional Gaussian distribution, with orientation, azimuth/elevation, and x/y standard deviation serving as degrees of freedom for the optimization. On and Off subunits (eg. white and black spots) were fit separately. Subunit area was defined as the 1.5 standard deviation ellipse under this fit gaussian, measured in units of squared visual degrees. Up to two On and Off subunits were fit. The dimensionless overlap index is computed as the area of overlap of the two-dimensional gaussians fit to each On and Off subunits, normalized by their geometric means

$$OI = \frac{A_{on} \cap A_{off}}{\sqrt{A_{on} A_{off}}}$$

The total areas of the receptive field was computed as the sum of all subunit areas, correcting for overlap.

Lifetime sparseness was computed using the definition in Vinje and Gallant[27].

$$S_L = \frac{1 - \frac{1}{N} \frac{(\sum_i r_i)^2}{\sum_i r_i^2}}{1 - \frac{1}{N}}$$

where *N* is the number of stimulus conditions and r_i is the response of the neuron to stimulus condition i averaged across trials. Population sparseness was computed with the same metric, but where N is the number of neurons and r_i is average response vector of neuron i to all stimulus conditions.

Reliability was defined as CC_{max}^2, where CC_{max} is the expected correlation between the sample trial averaged response and the true (unmeasured) mean response. We follow the computation from Schoppe, *et al.*[28]:

$$\frac{1}{CC_{max}} = \sqrt{1 - \frac{1}{N}\left(\frac{\left(1 - \frac{1}{N}\right) \times \sum_{n=1}^{N} Var(R_n)}{Var(\sum_{n=1}^{N} R_n) - \sum_{n=1}^{N} Var(R_n)}\right)}$$

where N is the number of trials and R_n is the time series of the response on the nth trial. For R_n we use the trace of extracted event magnitudes at 30Hz, smoothed with a Gaussian window of width 0.25s.

We computed “noise” and “signal” correlations in the population responses. Signal correlations were computed as the Pearson correlation between the trial-averaged stimulus responses of pairs of neurons. Noise correlations were computed as the Pearson correlation of the single-trial stimulus responses for a pair of neurons and a given stimulus,

and then averaged over stimuli. For natural movies, we computed the noise and signal correlations of the binned event counts in non-overlapping 10 frame windows. We computed “spontaneous correlations” as the Pearson correlation of the detected event trains during the periods of spontaneous activity recording.

Decoding

We used non-parametric (K-neighbors) classifiers to decode the visual stimulus identity (e.g. the natural scene number, within the natural scene responses) from the population vector of single-trial responses, using the Pearson correlation distance between response vectors. We report the performance on the held-out data from five-fold cross-validation. On each cross-validation fold, we performed an inner-round of 2-fold cross-validation to choose the number of neighbors from eight logarithmically spaced options (1, 2, 4, 7, 14 and 27). To examine how decoding performance depended on the number of neurons, we took a range of sample sizes. For each sample size, we uniformly chose five samples of neurons. To compare decoding performance between behavioral conditions, we used stimulus conditions (e.g. individual natural scenes) with at least five repetitions in each behavioral condition. We used the same number of repetitions (uniformly sampled) for each stimulus in each behavioral condition.

3D Gabor Wavelet Model

We model the response properties of each neuron as a quadratic function of the input pixels. Each neuron is modeled as a sparse linear combination of linear and quadratic basis functions. Similar techniques have been successfully used to model neurons and voxels in visual cortex both in animals and humans[29–31]. Our basis functions are a pyramid of linear and quadratic 3D Gabor wavelet filters that tile the stimulus at a variety scales, directions and temporal frequencies (see Figure 6a. The parameters that generate the set of filters were adapted and scaled to the tuning properties of mouse visual cortex. We estimate weights for 10 time-lags for each basis function to enable fitting of the temporal kernel. The weighted sum of the basis functions is passed to a parameterized soft-plus nonlinearity to ensure the model has only positive outputs. This quadratic model is thus akin to a highly regularized STA/STC analysis, but adapted to fit the full spatio-temporal receptive field using stimuli from the data set.

We estimated a sparse combination of basis functions for each neuron using a variant of threshold gradient descent[32]. Threshold gradient descent works just like gradient descent, except only basis functions whose gradients have magnitudes larger than some fractional threshold, t, of the largest gradient magnitude have their weights updated. All weights start at 0 and the descent is terminated using early stopping, i.e. it continues only if generalization performance on a stopping set improves. The threshold parameter, which can range from 0 to 1, controls the sparsity of the solution, with larger values returning sparse solutions. All models were fit with a threshold value of 0.8.

We modified the original algorithm in three ways to improve its performance on our data. First it was changed so that it updated the weights at all time lags for any basis function that exceeded threshold. This allowed the temporal kernel for each basis function to vary smoothly in time, rather than being sparse. Second, we maintained an active-set containing all weights that had ever been updated. At each iteration, any basis function whose gradient exceeded the threshold had its weight added to the active set, and then all weights in the active set were updated. This prevented oscillations during fitting. Third we used an adaptive step size to speed up training. The step size increased by a factor of 1.2 at each iteration if generalization to the stopping set improved, and decreased by a factor of 0.5 if generalization worsened[33].

All models were fit and tested in a nested six-fold cross-validation framework. We split the data into six sets each containing many 50 sample long continuous blocks from throughout the dataset. For each fold of cross validation, one set was used as a test set to evaluate the model and the remaining five were used to train a model. A model was trained by starting with five separate models, each trained on a different combination of four of the five training sets, with the remaining set functioning as the stopping set for that model. After fitting, the five models were averaged together to create a single model for making predictions on the test set. Reported model performance is the average performance on the test set across the six folds. Separate models were fit for the natural stimuli (movies and scenes) and artificial stimuli (drifting gratings, static gratings and locally sparse noise).

We define “simple” cells as neurons that exhibit linear receptive fields and “complex” cells as neurons that exhibit quadratic receptive fields. We assess the degree of “complexity” of each neuron by constructing a metric that is the ratio of total excitatory weight on quadratic basis functions to the total excitatory weight on all basis functions. A model that only has weights on quadratic basis functions would be 1 on this index, while a model that only has weights on linear basis functions would be 0 on this index.

The linear component of each model can be visualized directly, similar to a spike triggered average, as it is simply weights on pixels. To visualize the quadratic component, the weighted quadratic basis functions must first be converted into an equivalent matrix of second order terms. This matrix can then be analyzed by eigenvector decomposition, similar to spike triggered covariance.

SSM Analysis for comparison with convolutional neural networks

The comparison of representations to those of convolutional neural networks is performed via a similarity-of-similarity matrix analysis. To compare two representations, e.g. a model layer with the responses from Cux2, layer4, V1, we compute the Spearman correlation between the similarity matrices for each representation. The similarity matrix for a representation is computed by taking the Pearson correlation of the response across neurons or units in that representation for two images from the natural scene stimuli, including the grey screen condition, resulting in a 119x119 matrix for each representation. For the CNN, the response is just the set of outputs of a particular model layer. For the

neural responses, we use the trial average of the integrated event magnitude for the interval during which a particular image was on screen. We compare only to the pooling layers of VGG16 in the main text. A hierarchical clustering of the layers using the one minus the similarity matrix to compute a distance generates six clusters one of which consists of the input layer alone. The remaining clusters are the layers between and including the successive pooling layers. Images were resized to (50,64) before CNN responses were computed. Several different sizes from (50,64) to (400,512) in powers of two did not yield substantially different results.

Statistical significance is determined by generating a null distribution for the SSM correlation by creating 100 shuffles of the image labels for both representations and computing the SSM correlation from that. The distribution of shuffles was fit to a Student's T distribution and a p value was computed from that fit. Significance was considered at the 0.005 level, with a Bonferroni correction for the number of comparisons across model layers, cortical areas, cortical layers, and Cre lines.

Population Overlap

Population overlap quantifies the extent to which the population of neurons that responds to one stimulus set (e.g. locally-sparse noise) overlaps with the population of neurons that responds to another stimulus set (e.g. static gratings). For each pair of stimulus sets, the population overlap is defined as the number of neurons in an experiment that were responsive to both stimulus sets divided by the number of neurons that were successfully imaged for the entirety of both stimulus sets.

To place the observed population overlap in context, we defined an index of population overlap in which 0.0 indicates the population overlap of statistically independent populations, 1.0 indicates maximally-overlapping populations, and -1.0 indicates minimally-overlapping populations. If f_1, f_2, f_{12} indicate the fraction of cells responsive to stimulus set 1, stimulus set 2, and both stimulus sets 1 and 2, respectively, the population overlap is calculated as follows.

$$f_{12}^{min} = \max(0.0, f_1 + f_2 - 1.0)$$
$$f_{12}^{max} = \min(f_1, f_2)$$
$$f_{12}^{chance} = f_1 f_2$$

$$PO = \begin{cases} \dfrac{f_{12} - f_{12}^{chance}}{f_{12}^{chance} - f_{12}^{min}}, & f_{12} < f_{12}^{chance} \\ \dfrac{f_{12} - f_{12}^{chance}}{f_{12}^{max} - f_{12}^{chance}}, & f_{12} \geq f_{12}^{chance} \end{cases}$$

Statistically independent populations have the property that whether or not a neuron is responsive to one stimulus set provides no information about whether or not that neuron is responsive to the other stimulus set. By definition, the population overlap of statistically

independent populations is equal to the product of the marginal fractions of neurons responsive to each stimulus set alone. An observed overlap greater than that expected for statistically independent populations indicates that a neuron that is responsive to one stimulus set is more likely than not to also be responsive to the other stimulus set. By contrast, an observed overlap less than that expected for statistically independent populations indicates that a neuron that is responsive to one stimulus set is more likely to be non-responsive than responsive to the other stimulus set.

The fraction of cells responsive to either stimulus set alone also constrains the range of possible values of population overlap. Maximum overlap occurs when all neurons responsive to one stimulus set are also responsive to the other stimulus set, which implies that the overlap can be no greater than the lesser of the two marginal fractions responsive. Conversely, minimum overlap occurs when the fewest neurons responsive to one stimulus set are also responsive to the other stimulus set, which implies that the overlap must be greater than zero when the sum of the marginal fractions responsive is greater than 1.0.

691 **References**

692

693 1. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization
694 system for the whole mouse brain. *Nat. Neurosci.* **13,** 133–140 (2010).

695 2. Madisen, L. *et al.* Transgenic mice for intersectional targeting of neural sensors and
696 effectors with high specificity and performance. *Neuron* **85,** 942–958 (2015).

697 3. Mayford, M. *et al.* Control of memory formation through regulated expression of a
698 CaMK11 transgene. *Science (80-.).* **274 IS-,** 1678-1683 EP- (1996).

699 4. Daigle, T. L. *et al.* A suite of transgenic driver and reporter mouse lines with
700 enhanced brain cell type targeting and functionality. (2017). doi:10.1101/224881

701 5. Franco, S. J. *et al.* Fate-restricted neural progenitors in the mammalian cerebral
702 cortex. *Science (80-.).* **337,** 746–749 (2012).

703 6. Harris, J. A. *et al.* Anatomical characterization of Cre driver mice for neural circuit
704 mapping and manipulation. *Front. Neural Circuits* **8,** 1–16 (2014).

705 7. Gorski, J. a *et al.* Cortical excitatory neurons and glia, but not GABAergic neurons,
706 are produced in the Emx1-expressing lineage. *J. Neurosci.* **22,** 6309–6314 (2002).

707 8. Taniguchi, H. *et al.* A Resource of Cre Driver Lines for Genetic Targeting of
708 GABAergic Neurons in Cerebral Cortex. *Neuron* **71,** 995–1013 (2011).

709 9. Dhillon, H. *et al.* Leptin directly activates SF1 neurons in the VMH, and this action
710 by leptin is required for normal body-weight homeostasis. *Neuron* **49,** 191–203
711 (2006).

712 10. Gerfen, C. R., Paletzki, R. & Heintz, N. GENSAT BAC cre-recombinase driver lines
713 to study the functional organization of cerebral cortical and basal ganglia circuits.
714 *Neuron* **80,** 1368–1383 (2013).

715 11. Guo, C. *et al.* Fezf2 expression identifies a multipotent progenitor for neocortical
716 projection neurons, astrocytes, and oligodendrocytes. *Neuron* **80,** 1167–1174
717 (2013).

718 12. Gong, S. *et al.* Targeting Cre Recombinase to Specific Neuron Populations with
719 Bacterial Artificial Chromosome Constructs. *J. Neurosci.* **27,** 9817–9823 (2007).

720 13. Goldey, G. J. *et al.* Removable cranial windows for long-term imaging in awake
721 mice. *Nat. Protoc.* **9,** 2515–2538 (2014).

722 14. Kalatsky, V. A. & Stryker, M. P. New paradigm for optical imaging: Temporally
723 encoded maps of intrinsic signal. *Neuron* **38,** 529–545 (2003).

724 15. Marshel, J. H., Garrett, M. E., Nauhaus, I. & Callaway, E. M. Functional
725 specialization of seven mouse visual cortical areas. *Neuron* **72,** 1040–1054 (2011).

726 16. Garrett, M. E., Nauhaus, I., Marshel, J. H. & Callaway, E. M. Topography and Areal
727 Organization of Mouse Visual Cortex. *J. Neurosci.* **34,** 12587–12600 (2014).

728 17. Peirce, J. W. Generating Stimuli for Neuroscience Using PsychoPy. *Front.*
729 *Neuroinform.* **2,** 10 (2008).

730 18. Peirce, J. W. PsychoPy-Psychophysics software in Python. *J. Neurosci. Methods*
731 **162,** 8–13 (2007).

19. Martin, D., Fowlkes, C., Tal, D. & Malik, J. A database of human segmented natural images and its application to\nevaluating segmentation algorithms and measuring ecological statistics. *Proc. Eighth IEEE Int. Conf. Comput. Vision. ICCV 2001* **2,** 416–423 (2001).

20. van Hateren, J. H. & van der Schaaf, a. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. Biol. Sci.* **265,** 359–366 (1998).

21. Olmos, A. & Kingdom, F. A. A. A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception* **33,** 1463–1473 (2004).

22. Welles, O. *Touch of Evil*. (Universal - International, 1958).

23. Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508,** 207–214 (2014).

24. Jones, E., Oliphant, T., Peterson, P. & Others. SciPy.org. *SciPy: Open source scientific tools for Python2* (2001).

25. McKinney, W. & Team, P. D. Pandas - Powerful Python Data Analysis Toolkit. *Pandas - Powerful Python Data Anal. Toolkit* (2015).

26. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* (2007). doi:10.1109/MCSE.2007.55

27. Vinje, W. E. & Gallant, J. L. Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. *Science (80-.).* **287,** 1273–1276 (2000).

28. Schoppe, O., Harper, N. S., Willmore, B. D. B., King, A. J. & Schnupp, J. W. H. Measuring the Performance of Neural Models. *Front. Comput. Neurosci.* **10,** 1–11 (2016).

29. Kay, K. N., Naselaris, T., Prenger, R. J. & Gallant, J. L. Identifying natural images from human brain activity. *Nature* (2008). doi:10.1038/nature06713

30. Nishimoto, S. *et al.* Reconstructing visual experiences from brain activity evoked by natural movies. *Curr. Biol.* (2011). doi:10.1016/j.cub.2011.08.031

31. Willmore, B. D. B., Prenger, R. J. & Gallant, J. L. Neural Representation of Natural Images in Visual Area V2. *J. Neurosci.* (2010). doi:10.1523/JNEUROSCI.4099-09.2010

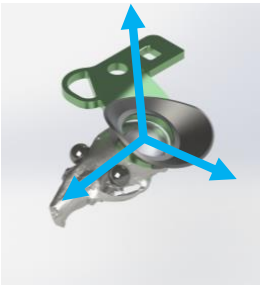
32. Friedman, J. H. & Popescu, B. E. Predictive learning via rule ensembles. *Ann. Appl. Stat.* (2008). doi:10.1214/07-AOAS148

33. Riedmiller and Braun - 1992 - RPROP - A Fast Adaptive Learning Algorithm.

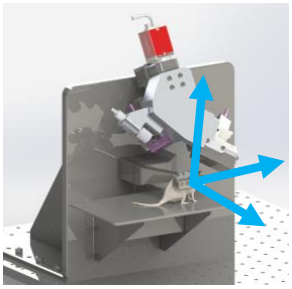
Supplemental Figure 1

a

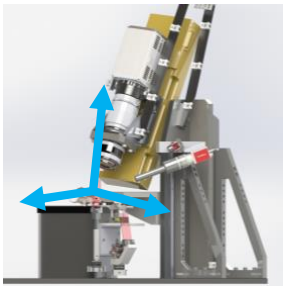
Headframe



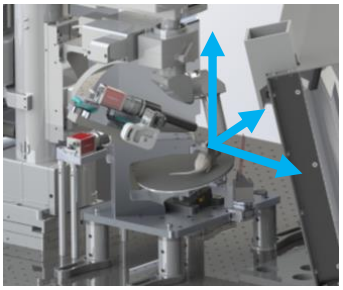
Surgery



ISI



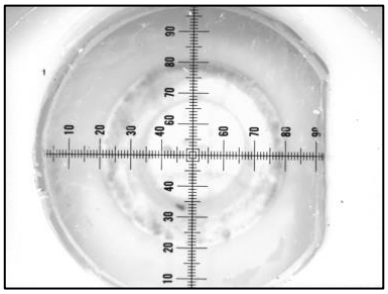
Physiology



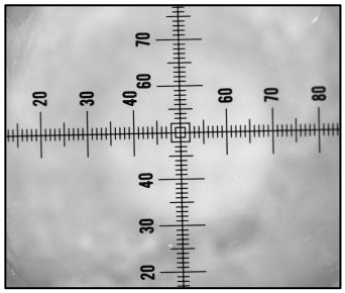
b

Headframe Reticle

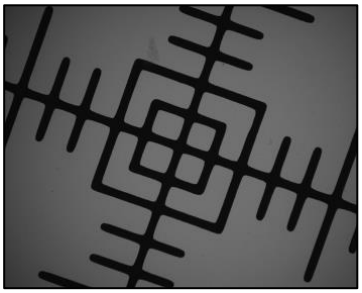
Surgery



ISI

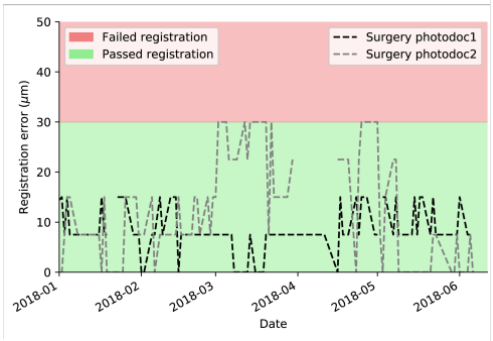


Physiology

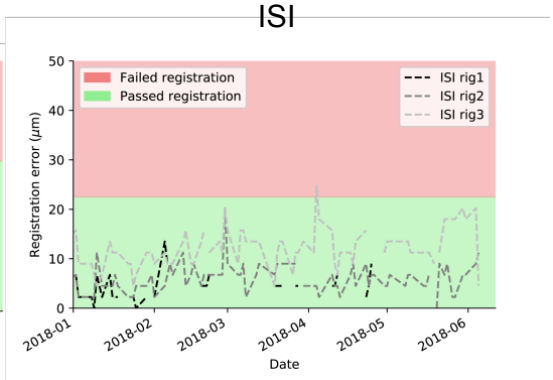


c

Surgery



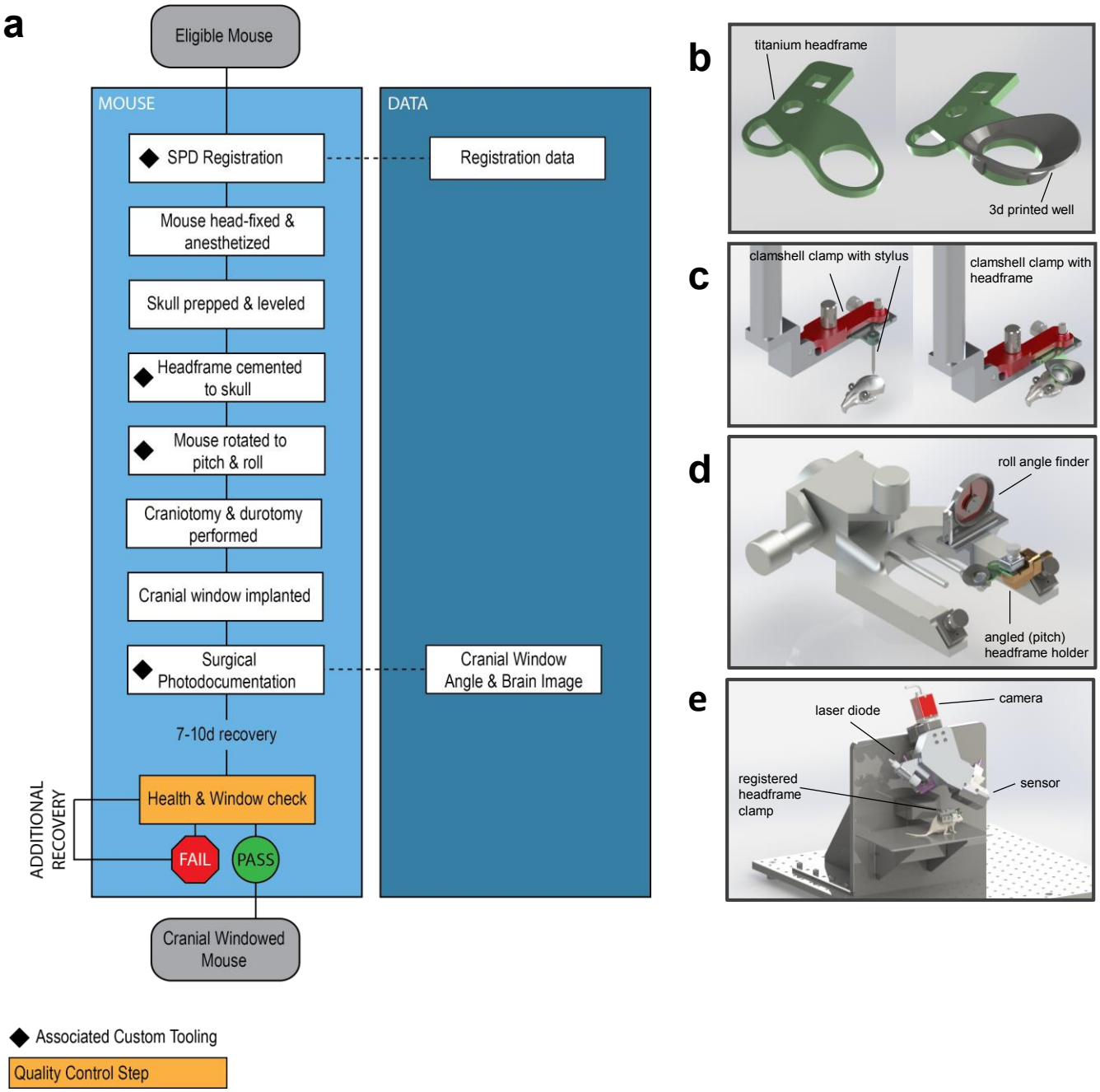
ISI



Supplemental Figure 1: Cross-platform registration for pipeline data collection

All data collection platforms are precisely designed and built around the headframe geometry and coordinate space (a). A registration artifact (a headframe with reticle glued inside the well) defines the imaging coordinate system (b). All experimental systems (including surgical photo-documentation, ISI, and 2P optical physiology microscopes) are calibrated so they produce data that is referenced within the imaging coordinate system. Registration integrity is monitored by routine measurement of a registration artifact (6 months of data shown in (c)). If the registration parameters exceed tight control limits the system is re-calibrated.

Supplemental Figure 2

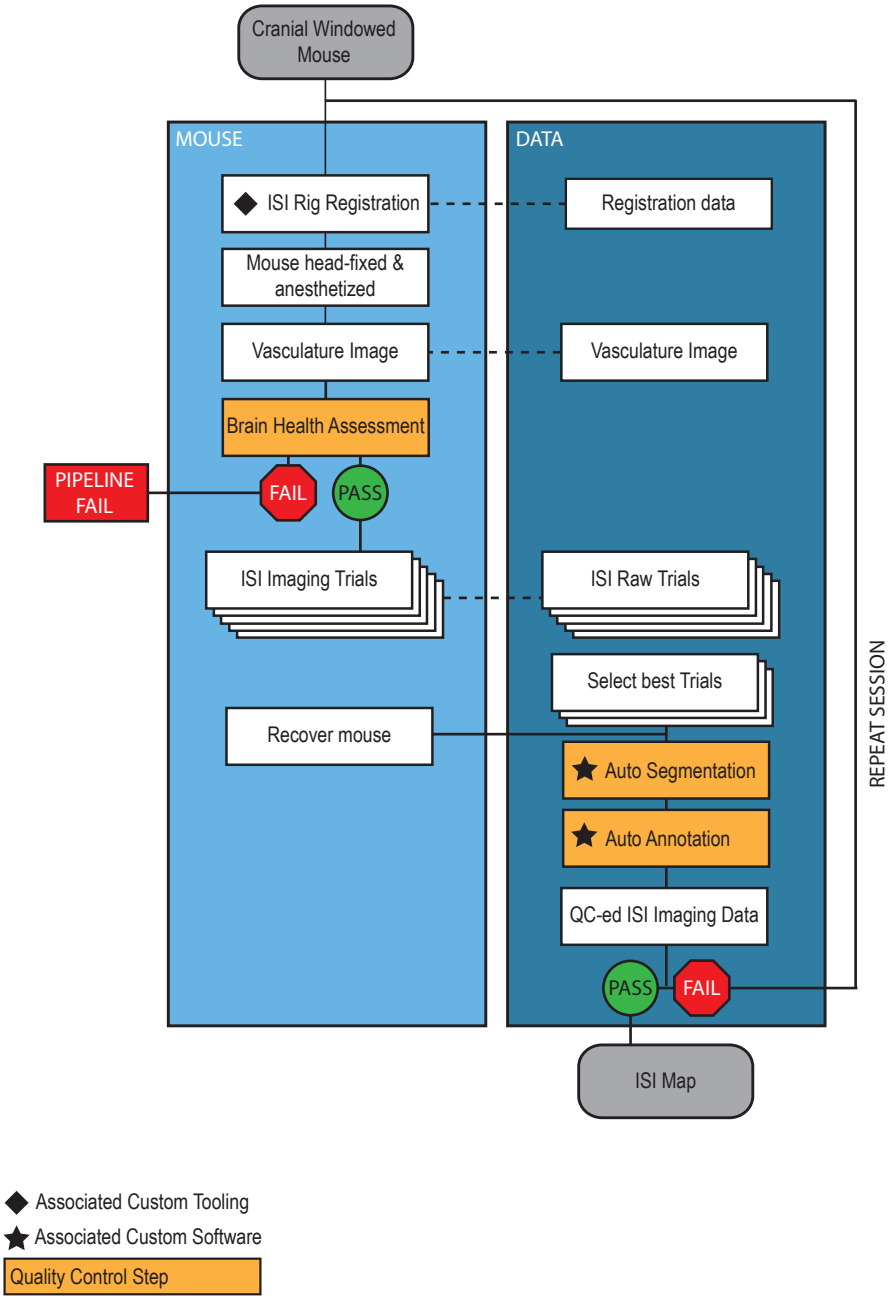


Supplemental Figure 2: A standardized workflow for headframe and cranial window placement

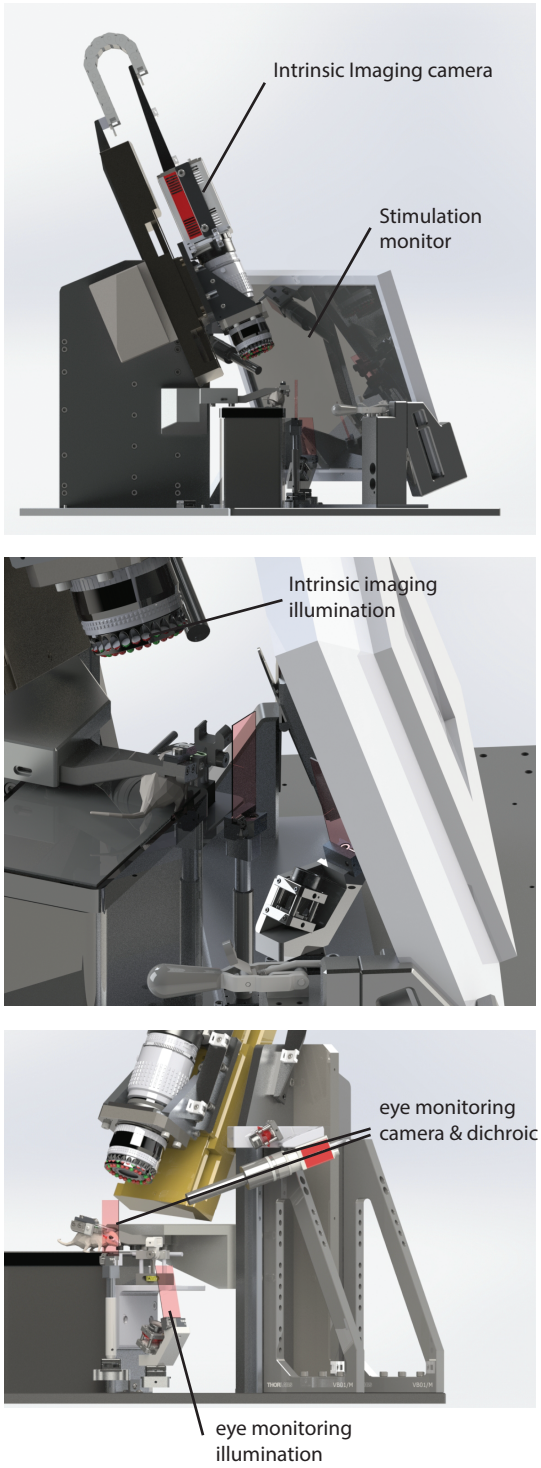
The cranial window surgery consists of procedural and data collection steps (a). The surgery is standardized across mice through precise placement of a registerable titanium headframe with objective well (b). Reproducible placement of the headframe is achieved using a suite of custom tooling that ensures precise placement of the headframe with respect to lambda (c), reproducible cranial window angles across mice (d), and quality control tracking of surgical outcome, cranial window angle and brain health (e).

Supplemental Figure 3

a



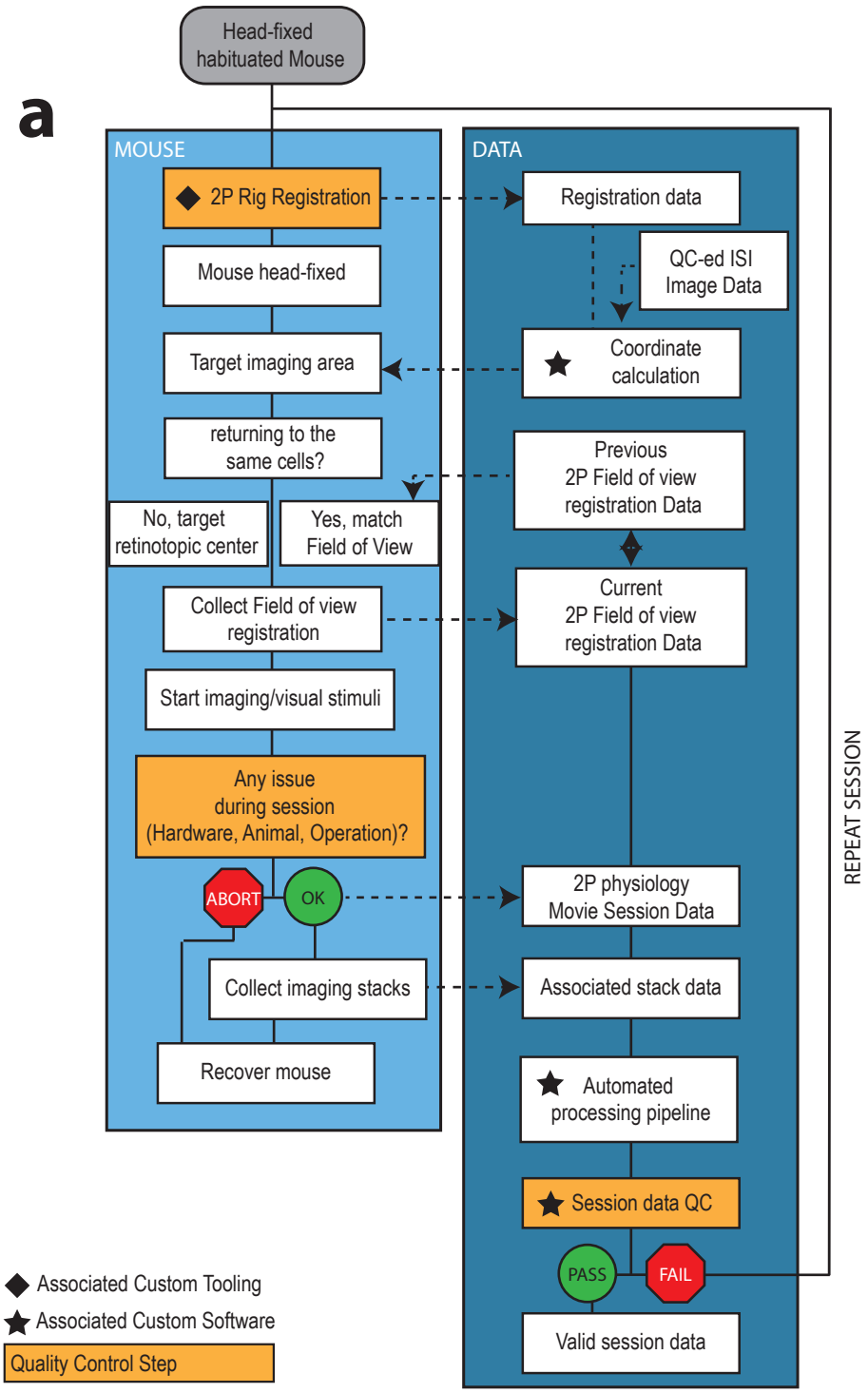
b

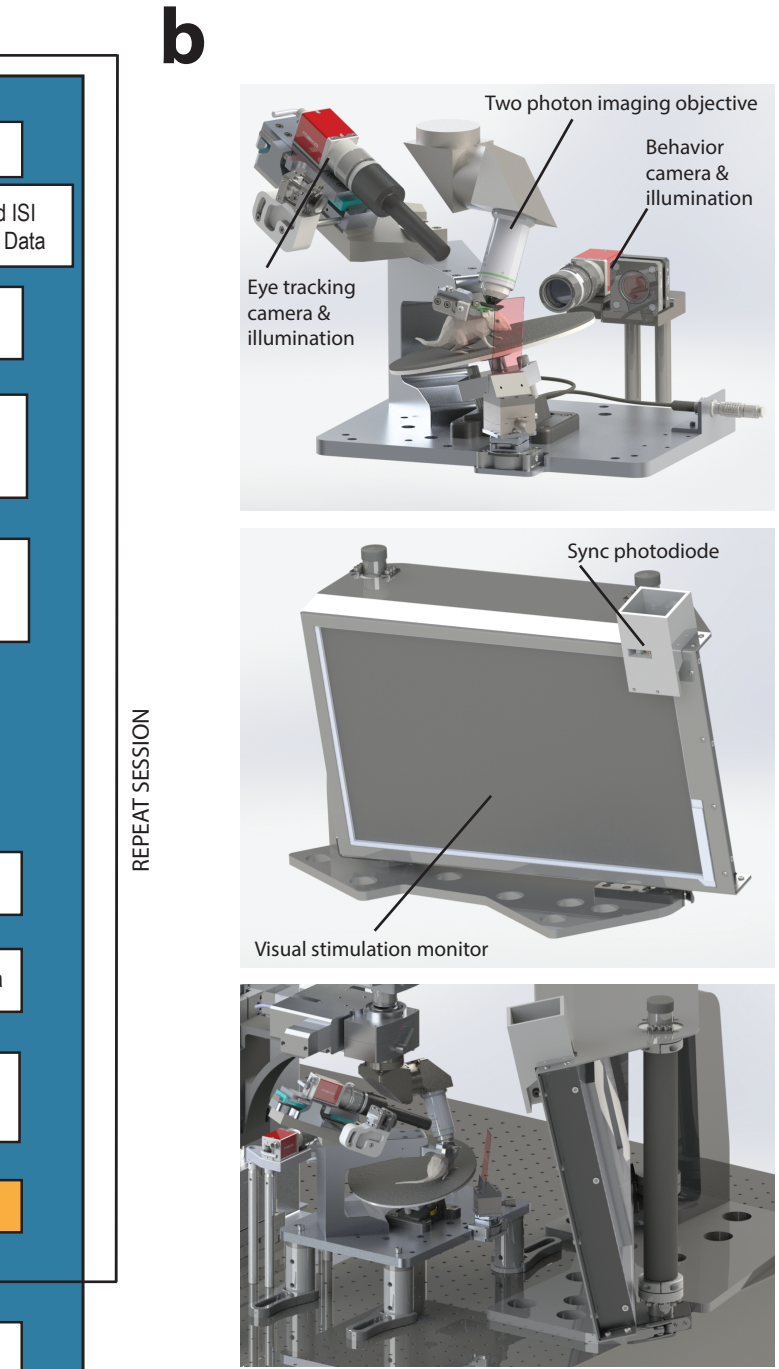


Supplemental Figure 3: A standardized intrinsic imaging experimental workflow to map the visual cortex at scale

(a) The intrinsic imaging consists of procedural and data collection steps. Data is standardized using custom tools (b) and data quality metrics and control procedure. (b) The Visual cortex is illuminated via custom LED rings positioned around the imaging objective (middle) and the fluctuation is light reflection is imaged using an imaging camera (top). Periodic stimuli presented on the stimulation screen allowed to map individual brain areas. (bottom) A set of camera allowed to monitor eye position in anesthetized mice.

Supplemental Figure 4

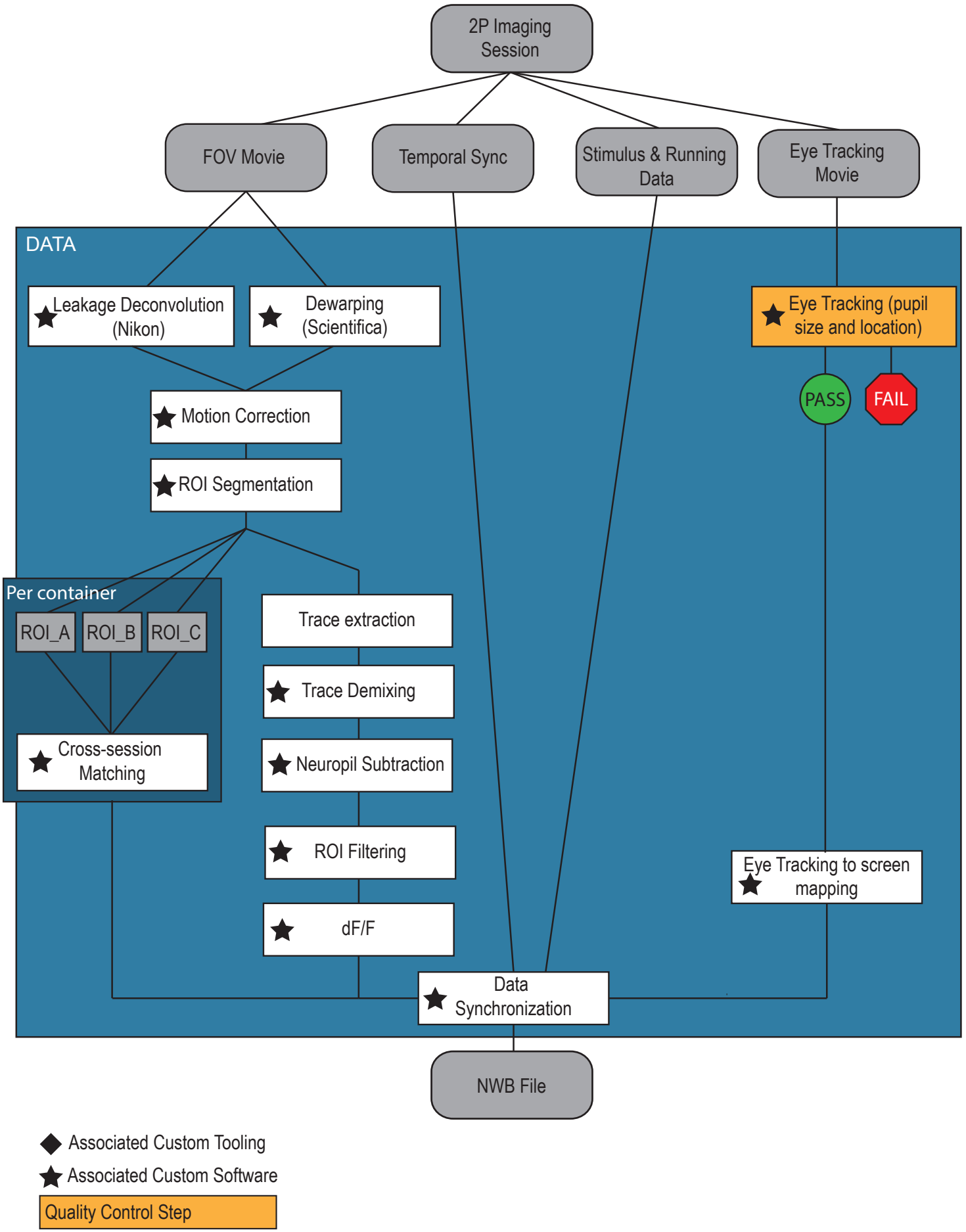




Supplemental Figure 4: A standardized 2-photon calcium imaging experimental workflow

(a) Two photon calcium imaging consists of procedural and data collection steps standardized using a formal experimental workflow. (b-top) Mice are head-fixed using custom behavioral hardware and monitored thanks to behavior and eye tracking cameras. (b-bottom) Visual stimuli are presented on a stimulation screen positioned reliably from session to session. (b-middle) Visual stimulation timing is monitored using a photodiode positioned on top of the screen for every experiment. The experimental workflow integrated tightly experimental procedures with QC metrics and any experimental that do not meet our standardized criteria (see Supplemental Figure 8) is re-attempted.

Supplemental Figure 5

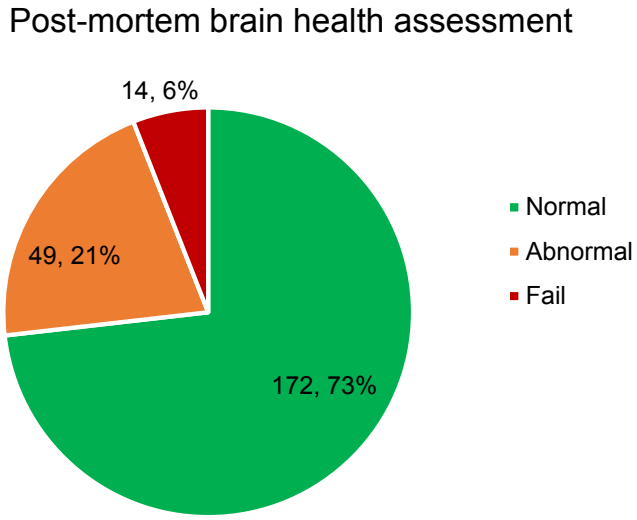


Supplemental Figure 5: Image processing workflow

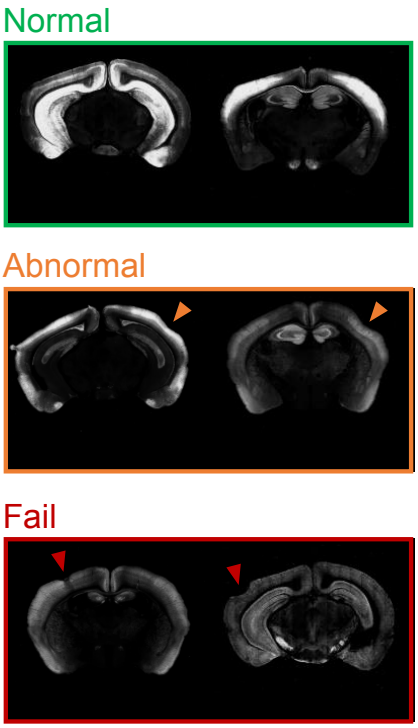
In vivo 2-photon imaging data is processed using a standardized pipeline. Calcium movies are motion corrected and segmented using an automated segmentation algorithm. Within each session, the traces are extracted from the identified ROI, and overlapping ROIs demixed. Across all three imaging sessions in a single experiment, the segmented ROIs are matched across sessions. Following the matching step, ROIs are filtered to select only somatic masks, neuropil contamination is subtracted, and $\Delta F/F$ is computed within each session. These traces are aligned to the stimulus, running, and eye tracking data using the temporal sync that was recorded during the experiment. The eye tracking movie is processed (see Supplemental Figure 12) and aligned to visual space.

Supplemental Figure 6

a



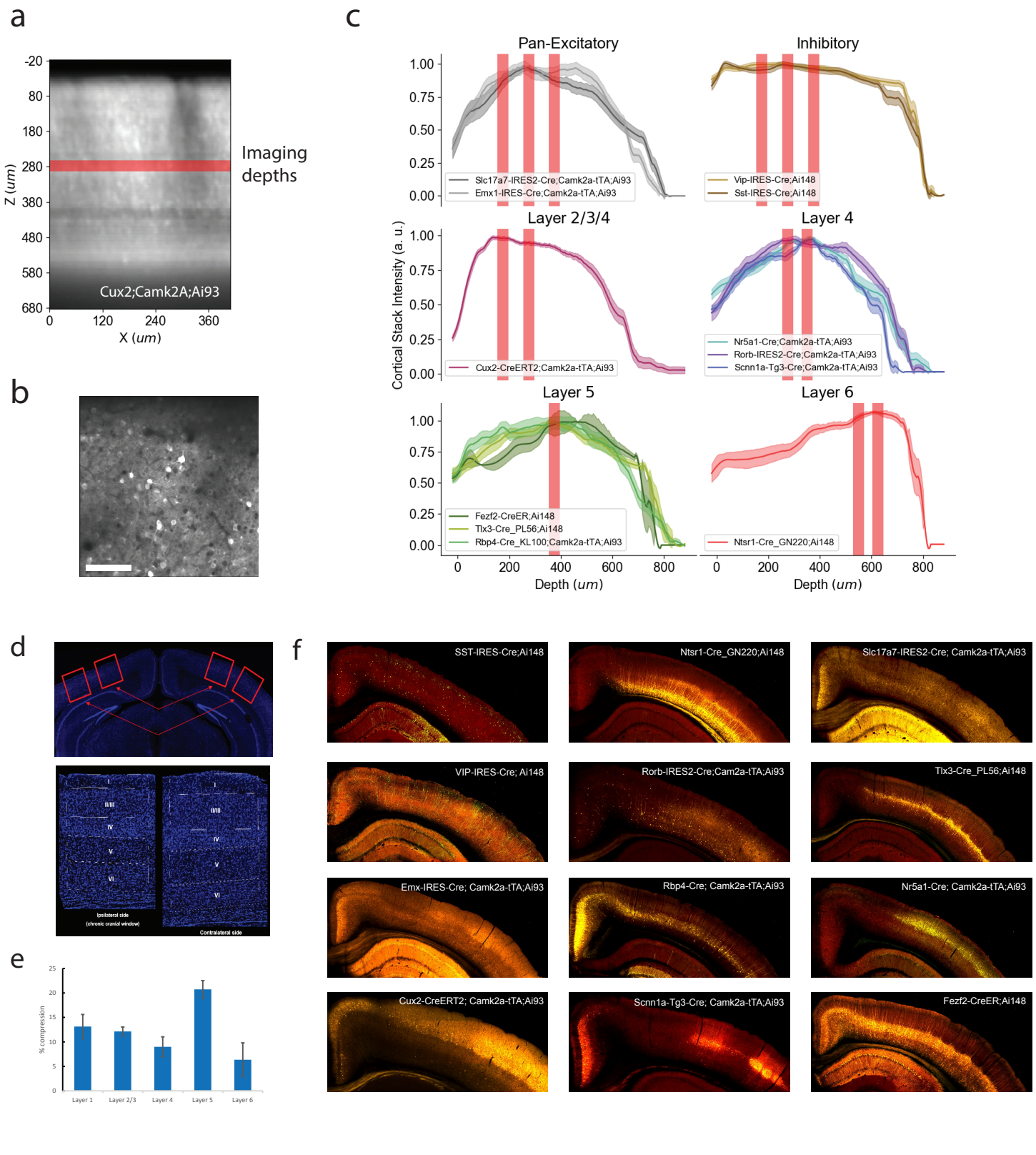
b



Supplemental Figure 6: Brain health assessment

(a) We observed brain abnormalities in approximately 27% of the 235 mice that completed the experimental workflow. Abnormalities were identified during post-mortem histological analysis of the tissue (examples shown in (b)) and had no observable behavioral manifestation during the life of the mouse. Additionally, analysis of a random subset of datasets obtained from mice with abnormalities revealed no differences in physiological responses. In most cases the abnormalities either occurred in the contralateral hemisphere or did not result in observable disruption to the cortical laminar structure (Abnormal). In rarer cases, the abnormalities 1) occurred under or adjacent to the cranial window, 2) disrupted the cortical layers, or 3) resulted in necrosis. In these cases, experimental datasets from these mice were excluded from analysis (Fail). (b) Examples of Normal, Abnormal and Failed brains.

Supplemental Figure 7



Supplemental Figure 7: Quantification of cortical layer expression levels across all transgenic mice for mice implanted with cortical glass windows

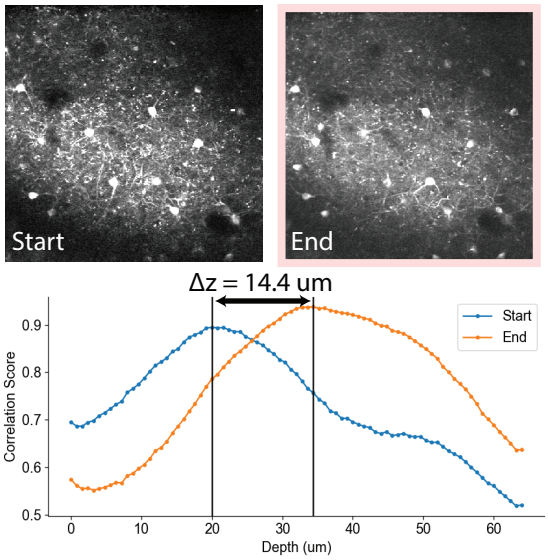
(a) Side view projection (XZ) of a single cortical stack acquired using two photon imaging. All experimental session were associated with such a cortical stack. Red bar denotes imaging depth shown in (b). Scale bar in (b) = 100 μm. (c) Distribution of labelling intensity across all cortical stacks acquired in V1 for all imaged cre-lines. Red bars denotes imaging depths used in each line. (d) Brain compression was quantified using DAPI staining and confocal imaging of entire coronal sections. Cortical layers were annotated and their thickness measured for both the ipsi-lateral and contra-lateral side. (e) Distribution of cortical thickness ratio for all 6 cortical layers between the ipsi-lateral and contra-lateral side. (f) Example expression distribution of all cre-lines as shown on coronal sections imaged with serial two photon imaging.

Supplemental Figure 8

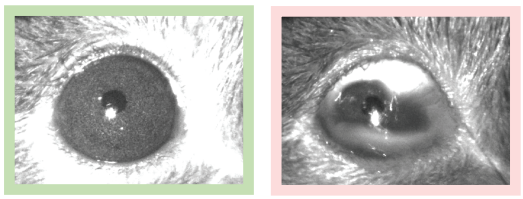
a

QC evaluation	Metric(s)	Threshold(s)	Notes
z-drift	Z distance between start and end of physiological movie	10 um	Measured using cross-correlation between movie plane and a local z-stack
Animal stress	Presence of foam over the eye, general animal stress	Foam covering pupil	Evaluated using eye tracking movie, body posture movie and general handling
Cell matching	Subjective evaluation of matched field of view between consecutive session	~ more than 50 % cells matching	
Wheel rotation failure	Wheel rotation uncorrelated with behavior movie; Presence of discontinuous spikes in wheel data		
Interictal events	Width and size of whole field calcium events	size larger than 30 % ΔF/F, width between 100 and 300 ms	
Laser damage	Number of discontinuous cellular compartments appearing between sessions	>0	
Bleaching and loss of signal	Average fluorescence from start to end of session	20% drop in baseline intensity	
Abnormal gcamp expression	Depth profile of gcamp	Any abnormal pattern	Evaluated by comparing to reference z stack data
Calcium physiology syncing	Number of dropped frames	1 for Nikon,60 for Scientifica	
Eye tracking syncing	Number of dropped frames	100 frames	
Visual stimulus syncing	Number of long frames	60 frames	Frame duration is plotted over time and number of abrnormal frames are tracked
Saturation	Number of satured pixels	1000 pixels	
Laser alignement stability	Power at objective (mW)	>10% change	Trigger maintenance or evaluation of rig

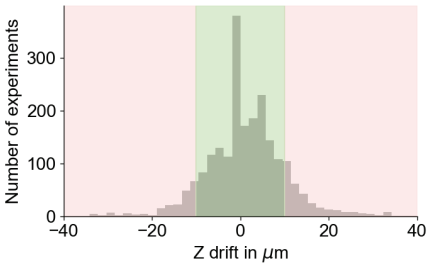
b



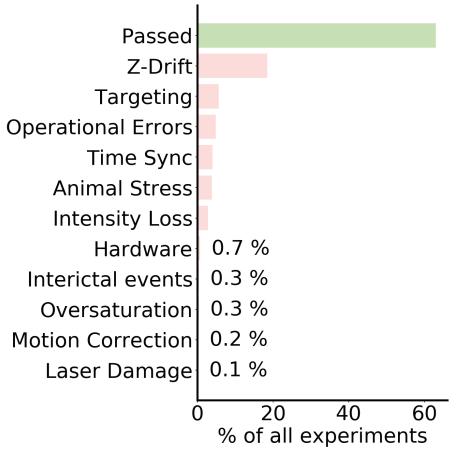
c



d



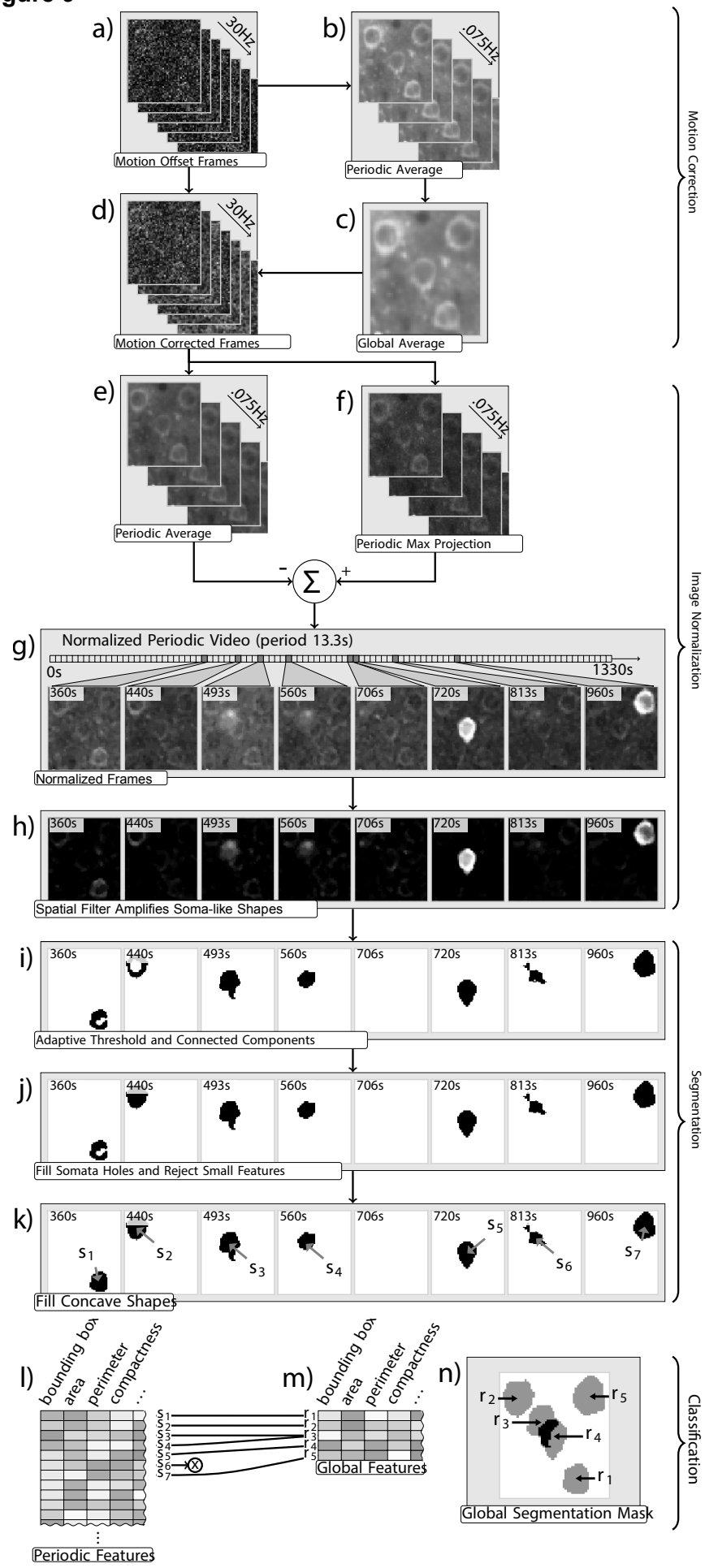
e



Supplemental Figure 8: 2-photon quality control metrics used for Quality Control (QC)

(a) Table of metrics used to validate experiment collected in two photon imaging. (b) Example images at the onset and end of a session showcasing excessive z-drift (>10 μm). Bottom panel in (b) shows the correlation score of all images in a local 60um z-stack with the onset and end image shown at the top. The peak was used as a measure of the actual z-position. (c) Example eye tracking data showing passing (green) and failing (red) signs of stress. (d) Distribution of z-drift across both passed and failed experiments. Green and red areas indicate failure threshold. (e) Distribution of all experimental failures for all individual sessions collected on the 2 photon imaging pipeline.

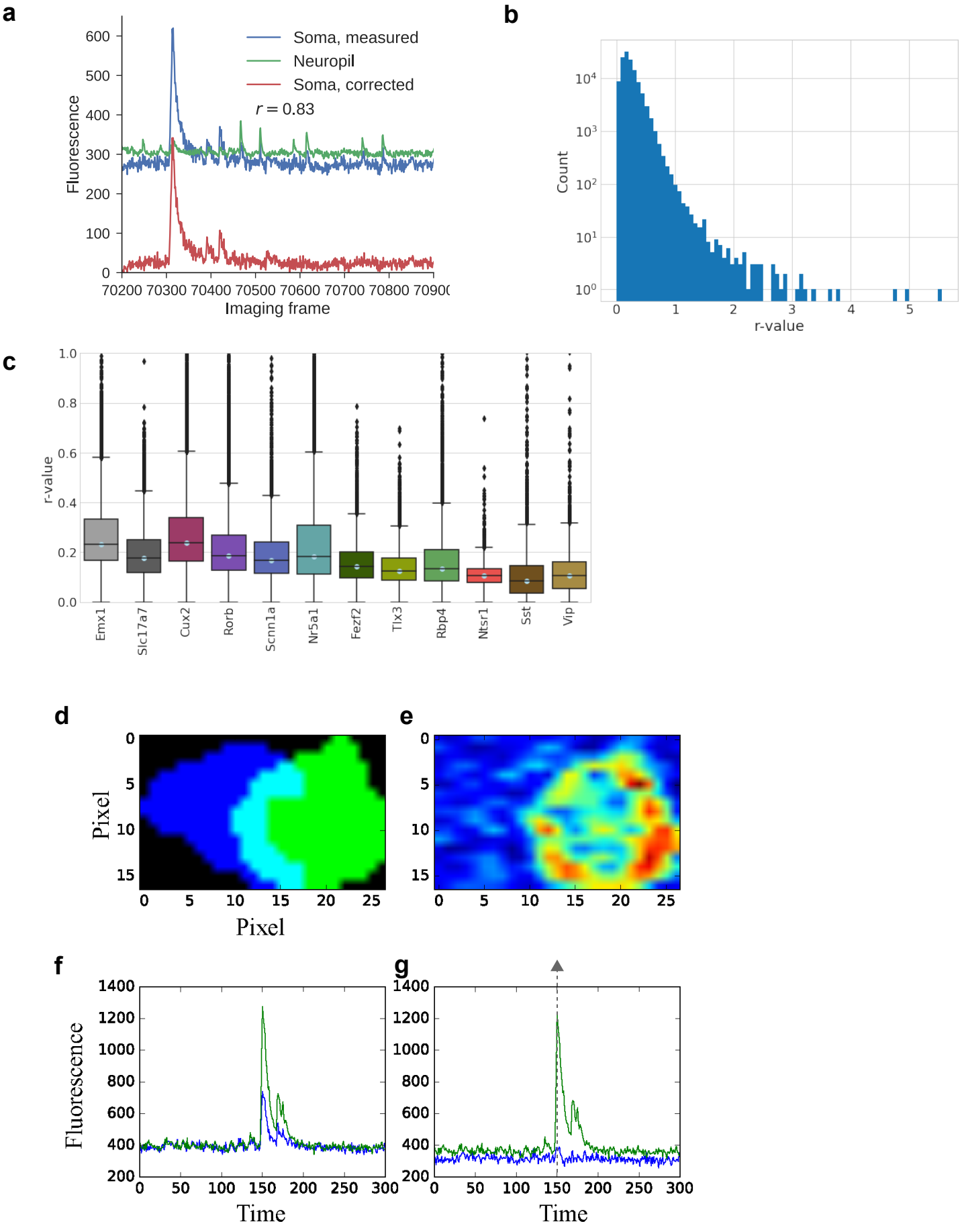
Supplemental Figure 9



Supplemental Figure 9: Image processing pipeline

(a) Uncorrected movie. (b) Periodic average of 400-frame set. (c) Global average frame computed by registering the periodic averages together. (d) Motion corrected movie, aligned to the global average, used for downstream processing. (e) Periodic average frame from full motion corrected movie subtracted from (f) periodic maximum projection frame creates (g) normalized periodic frames. Periodic frames after (h) spatial filtering, (i) adaptive thresholding, and morphological operations (j, k). Features of each ROI mask identified in each period are computed (l). Heuristic decisions combine duplicate masks and label non-somatic masks for filtering (m), yielding the final ROI masks (n).

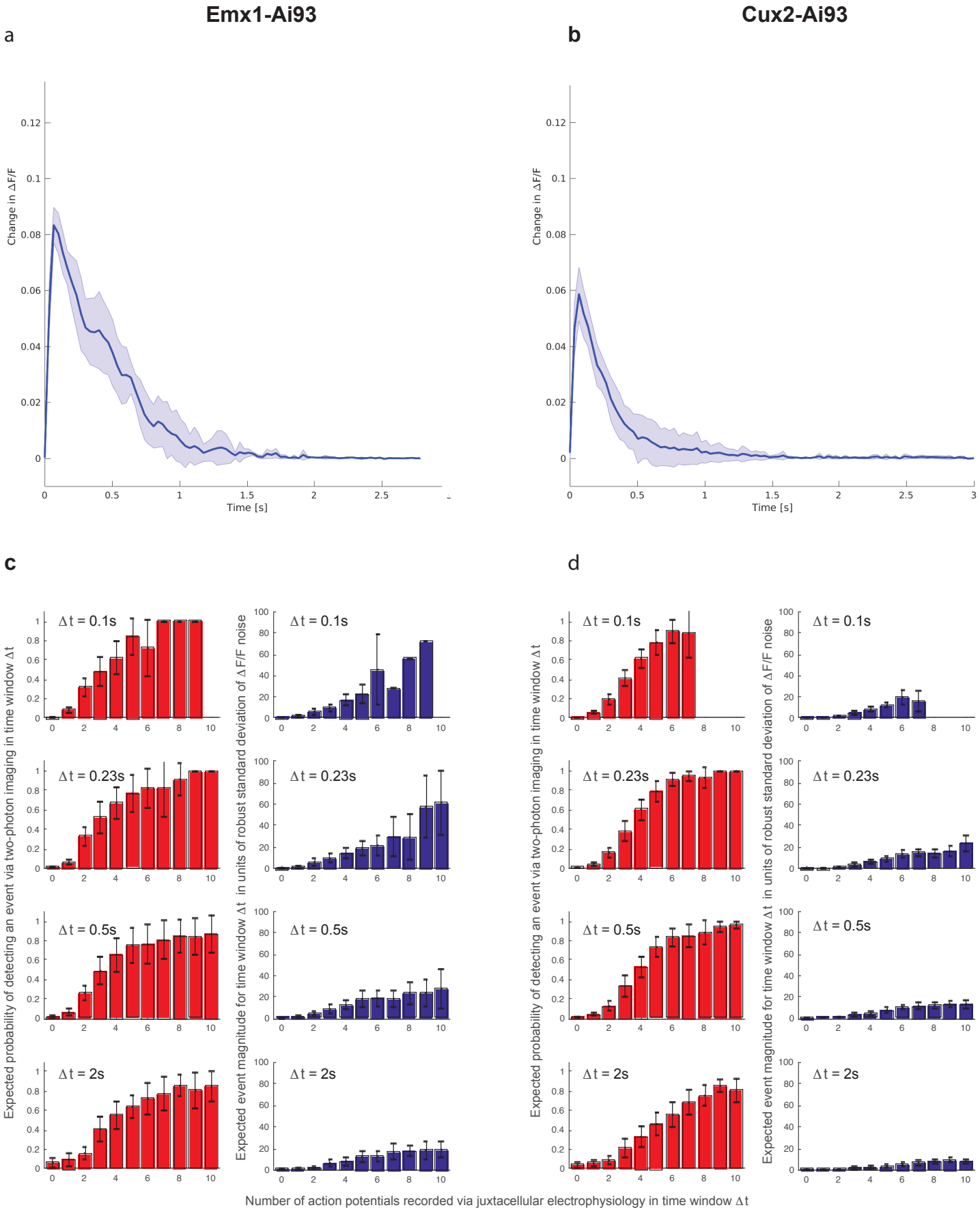
Supplemental Figure 10



Supplemental Figure 10: Neuropil subtraction and demixing

(a) Example of neuropil subtraction showing somatic fluorescence, neuropil fluorescence and the corrected trace. (b) Histogram of computed r-values for all cells in the dataset. (c) Distributions of r-values for each Cre line. (d) Example of two overlapping ROIs (blue and green, with overlapping pixels in cyan). (e) Non-uniform distribution of fluorescence across a ROI, used for demixing. (f) Raw fluorescence traces from each ROI, centered on the time point in e. (g) Demixed traces for each ROI, centered on the time point in e.

Supplemental Figure 11

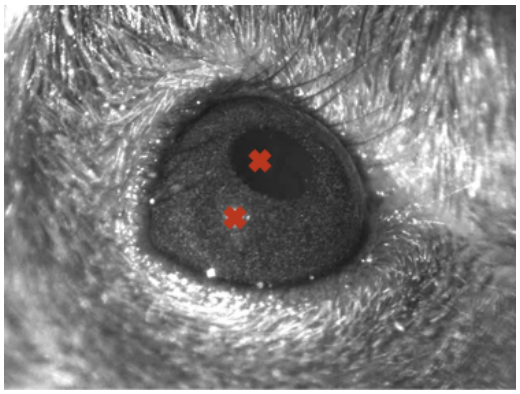


Supplemental Figure 11: Evaluation of Event Detection

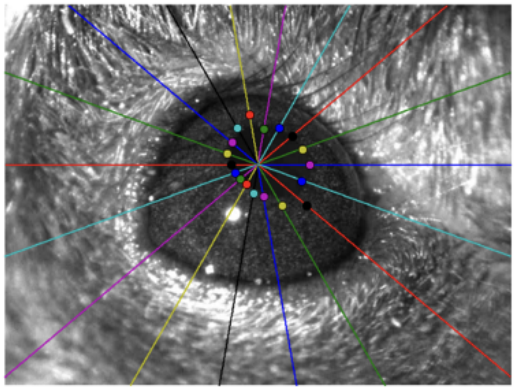
(a) Single-spike calcium response extracted from calcium imaging data acquired simultaneously with juxtacellular electrophysiology *in vivo* (15 cells in Emx1-Ai93 mice). (b) Single-spike calcium response extracted from calcium imaging data acquired simultaneously with juxtacellular electrophysiology *in vivo* (20 cells in Cux2-Ai93 mice). (c) The red bar graphs indicate the expected probability of detecting an event via L0 event detection (y-axis) as a function of the number of action potentials juxtacellularly recorded (x-axis) in a given time window (subplot title) based on 15 cells in Emx1-Ai93 mice. The blue bar graphs indicate the expected event magnitude reported by the L0 event detection (y-axis) as a function of the number of action potentials juxtacellularly recorded (x-axis) in a given time window (subplot title) based on 15 cells in Emx1-Ai93 mice. (d) As in c but based on 20 cells in Cux2-Ai93 mice.

Supplemental Figure 12

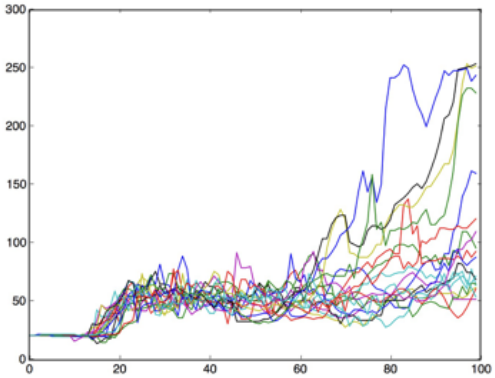
a



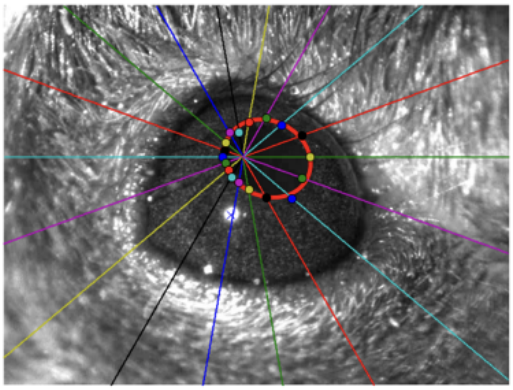
b



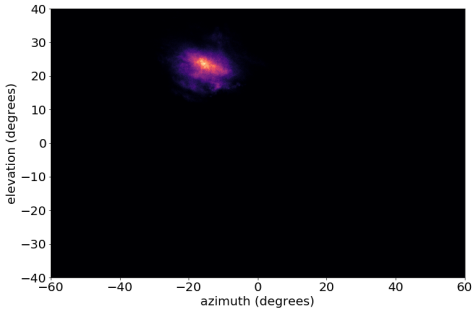
c



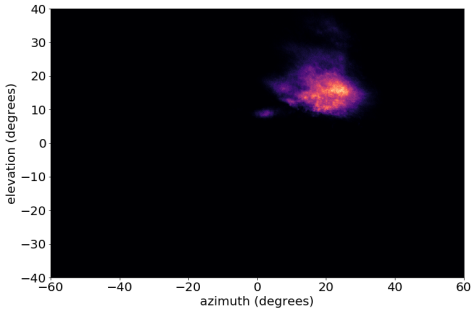
d



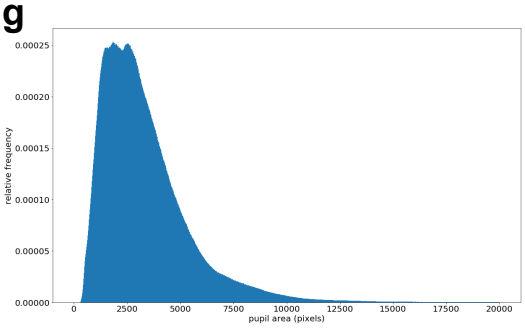
e



f



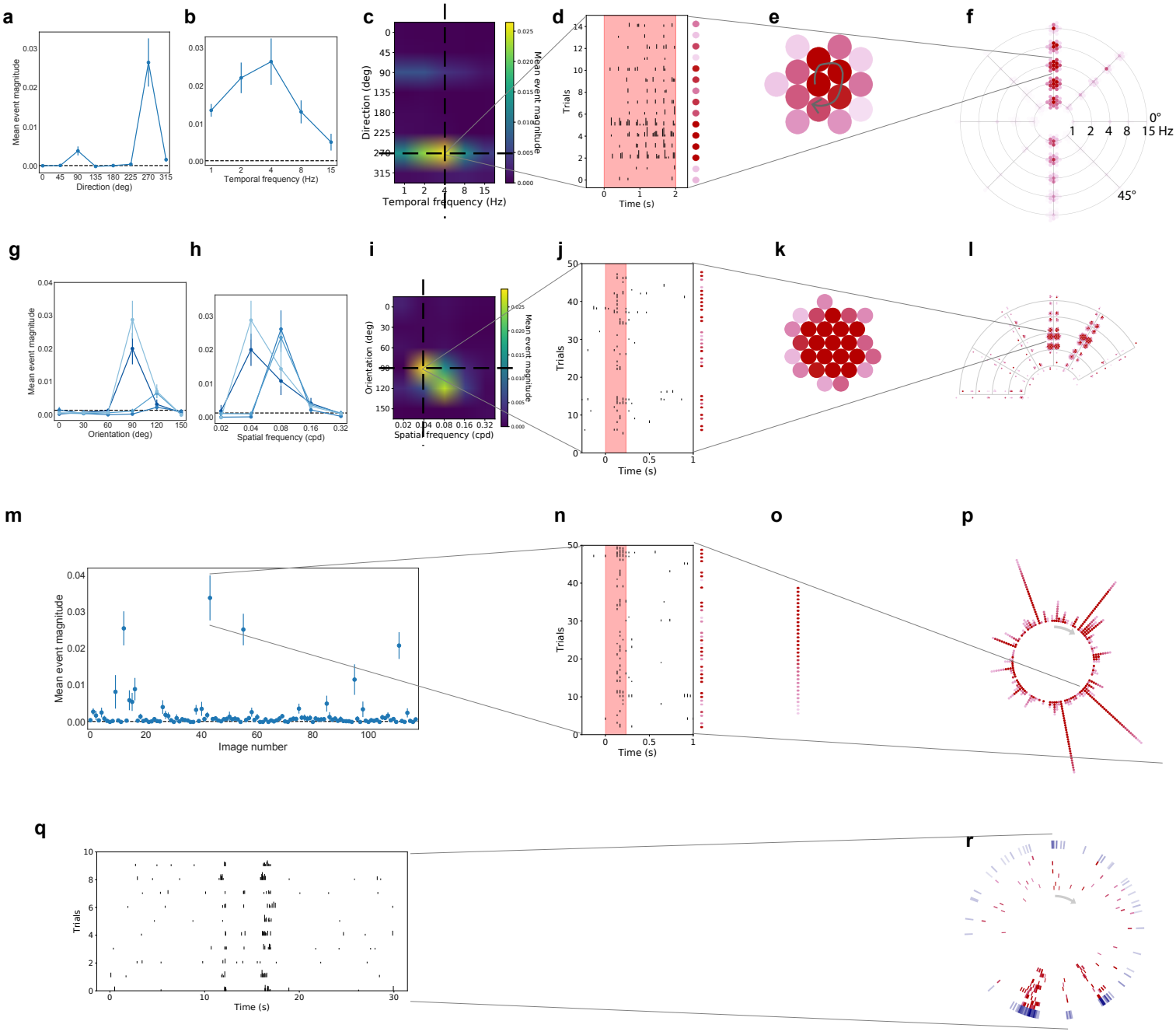
g



Supplemental Figure 12: Eye tracking and characterization

(a) Seed points are selected by finding points of maximal correlation with bright or dark spots. (b) Rays are traced from these points and boundaries are detected by thresholding against a proportion of the mean of the initial portion of the ray. (c) Example luminance progression along rays. (d) A ransac algorithm is used to fit an ellipse to the detected boundary points. (e) Eye position for the first monitor position. (f) Eye position for the second monitor position. (g) Histogram of pupil area.

Supplemental Figure 13



Supplemental Figure 13: Response visualizations

Conventional tuning curves for drifting grating responses for one cell. (a) Direction tuning plotted at the preferred temporal frequency (4 Hz) (mean ± sem). Dotted line represents the mean response to the blank sweep. (b) Temporal frequency tuning plotted at the preferred grating direction (270°). (c) Heatmap of the direction and temporal frequency responses for cell, showing any possible interaction of direction and temporal frequency. (d) All 15 trials at the preferred direction and temporal frequency, 2 second grating presentation is indicated by pink shading. The mean event magnitude is represented by intensity of the dot to the right of the trial. (e) All trials are clustered, with the strongest response in the center and weaker responses on the outside. (f) Clusters are plotted on a "Star plot". Arms indicated the direction of grating motion, arcs indicate the temporal frequency of the grating, with the lowest in the center and the highest at the outside. Clusters of red dots are located at the intersection and arms and arcs, representing the trial responses at that condition.

Tuning curves for static gratings for one cell. (g) Orientation tuning plotted at the preferred spatial frequency for each of the four phases. Dotted line represents the mean response to the blank sweep. (h) Spatial frequency tuning plotted at the preferred orientation for each of the four phases. (i) Heatmap of the orientation and spatial frequency at the preferred phase (j) All trials at the preferred orientation, spatial frequency and phase, the 250 ms grating presentation is indicated by pink shading. The mean event magnitude is represented by the intensity of the dot to the right of the trial. (k) All trials are clustered, with the strongest response in the center and weaker responses on the outside. (l) Clusters are placed on a "Fan plot". Arms represent the orientation and arcs represent the spatial frequency of the grating. At each intersection, there are four lobes of clustered dots, one for each phase at that grating condition.

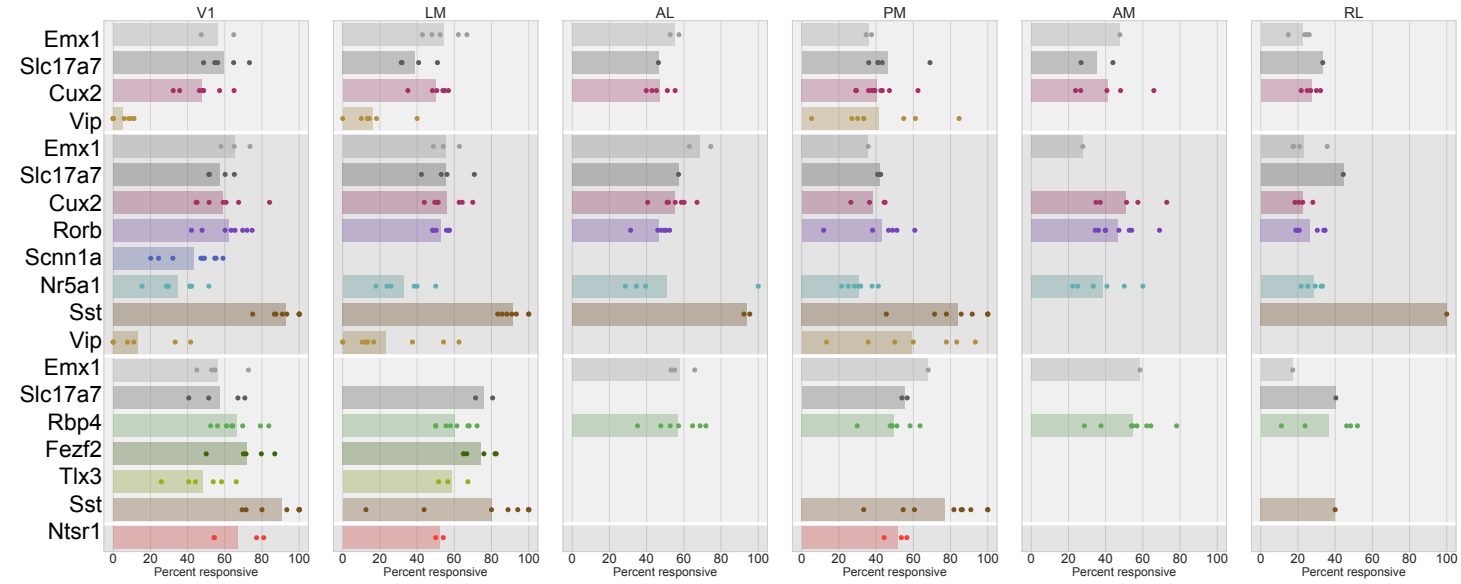
Responses to natural scenes. (m) Responses to each image presented, mean ± sem. Dotted line represents the mean response to the blank sweep. (n) All trials of the image which elicited the largest mean response, the 250 ms image presentation is indicated by pink shading. The mean event magnitude is represented by the intensity of the dot to the right of the trial. Trials are sorted (o) and are plotted on a "Corona plot" (p). Each ray represents the response to one image, with the strongest response on the inside and weaker responses at the outside.

Responses to natural movies (q) Responses of one cell's response to each trial of the natural movie. (r) Responses are plotted on a "Track plot". Each red ring represents the activity of the cell to one trial, proceed clockwise from the top of the track. The outer blue track represents the mean response across all ten trials.

Suppplemental Figure 14

Drifting gratings

Cre	V1			LM			AL			PM			AM			RL		
	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent
Emx1 - layer 2/3	461	269	58%	769	450	59%	328	181	55%	293	106	36%	105	50	48%	694	154	22%
Emx1 - layer 4	969	639	66%	677	362	53%	591	411	70%	168	60	36%	83	23	28%	1019	240	24%
Emx1 - layer 5	745	414	56%				316	186	59%	75	51	68%	53	31	58%	250	43	17%
Slc17a7 - layer 2/3	868	510	59%	437	184	42%	41	19	46%	342	155	45%	138	51	37%	45	15	33%
Slc17a7 - layer 4	932	526	56%	678	362	53%	182	104	57%	265	111	42%				115	51	44%
Slc17a7 - layer 5	766	435	57%	102	79	77%				159	89	56%				52	21	40%
Cux2 - layer 2/3	1721	810	47%	850	431	51%	723	332	46%	1036	421	41%	352	149	42%	438	124	28%
Cux2 - layer 4	1815	1069	59%	1540	845	55%	1500	822	55%	849	325	38%	726	383	53%	524	119	23%
Rorb	1617	1010	62%	767	399	52%	794	383	48%	411	202	49%	516	252	49%	1202	302	25%
Scnn1a	1200	522	44%															
Nr5a1	528	188	36%	242	75	31%	144	55	38%	203	66	33%	110	39	35%	854	254	30%
Rbp4	369	249	67%	342	208	61%	301	178	59%	432	210	49%	244	137	56%	97	39	40%
Fezf2	335	247	74%	550	405	74%												
Tlx3	840	400	48%	684	405	59%												
Ntsr1	244	149	61%	109	56	51%				232	116	50%						
Sst - layer 4	101	94	93%	108	101	94%	35	33	94%	112	96	86%				15	15	100%
Sst - layer 5	201	159	79%	100	74	74%				122	90	74%				5	2	40%
Vip - layer 2/3	116	6	5%	92	12	13%				144	59	41%						
Vip - layer 4	61	10	16%	105	27	26%				82	47	57%						



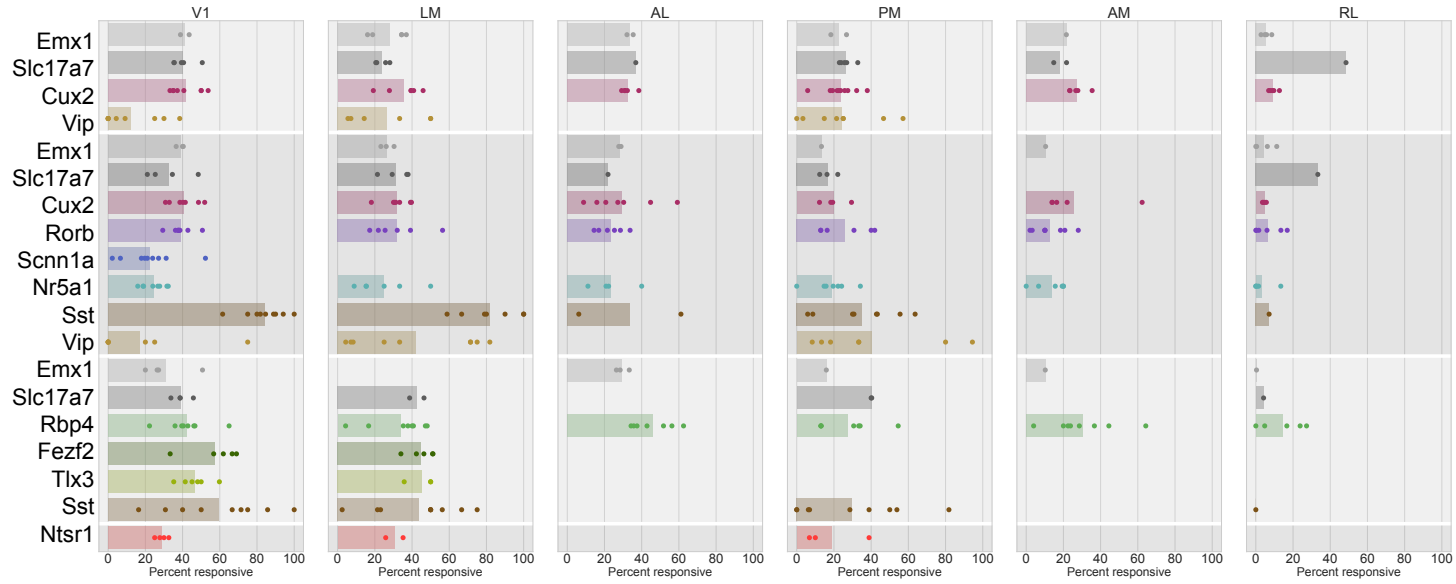
Supplemental Figure 14: Responsiveness to drifting gratings

(a) Table summarizing the numbers of cells imaged for each Cre line, layer, area combination in response to drifting grating stimulus and the number, and percent, of cells that were responsive to the drifting grating stimulus. (b) Strip plots of the percent of cells responsive to the drifting grating stimulus for each experiment.

Suppplemental Figure 15

Static gratings

Cre	V1			LM			AL			PM			AM			RL		
	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent
Emx1 - layer 2/3	430	174	40%	721	210	29%	340	116	34%	251	58	23%	88	19	22%	697	37	5%
Emx1 - layer 4	729	284	39%	634	162	26%	543	154	28%	188	25	13%	134	14	10%	966	42	4%
Emx1 - layer 5	684	202	30%				233	69	30%	44	7	16%	39	4	10%	272	1	0%
Slc17a7 - layer 2/3	767	302	39%	438	102	23%	38	14	37%	338	89	26%	137	26	19%	33	16	48%
Slc17a7 - layer 4	828	267	32%	663	199	30%	242	53	22%	192	32	17%				63	21	33%
Slc17a7 - layer 5	695	275	40%	90	37	41%				169	68	40%				48	2	4%
Cux2 - layer 2/3	1716	698	41%	851	302	35%	664	212	32%	944	221	23%	355	102	29%	412	41	10%
Cux2 - layer 4	1806	720	40%	1566	509	33%	1211	352	29%	771	149	19%	554	155	28%	491	23	5%
Rorb	1404	537	38%	753	217	29%	791	180	23%	341	76	22%	388	49	13%	1149	61	5%
Scnn1a	1114	263	24%															
Nr5a1	458	118	26%	232	40	17%	83	15	18%	228	45	20%	123	19	15%	854	27	3%
Rbp4	304	129	42%	378	122	32%	285	122	43%	411	94	23%	212	60	28%	89	14	16%
Fezf2	304	177	58%	582	266	46%												
Tlx3	829	382	46%	612	279	46%												
Ntsr1	216	61	28%	143	42	29%				196	22	11%						
Sst - layer 4	111	93	84%	86	70	81%	115	17	15%	178	43	24%				14	1	7%
Sst - layer 5	126	55	44%	166	40	24%				187	37	20%				7	0	0%
Vip - layer 2/3	113	13	12%	82	20	24%				136	29	21%						
Vip - layer 4	52	12	23%	104	40	38%				87	37	43%						



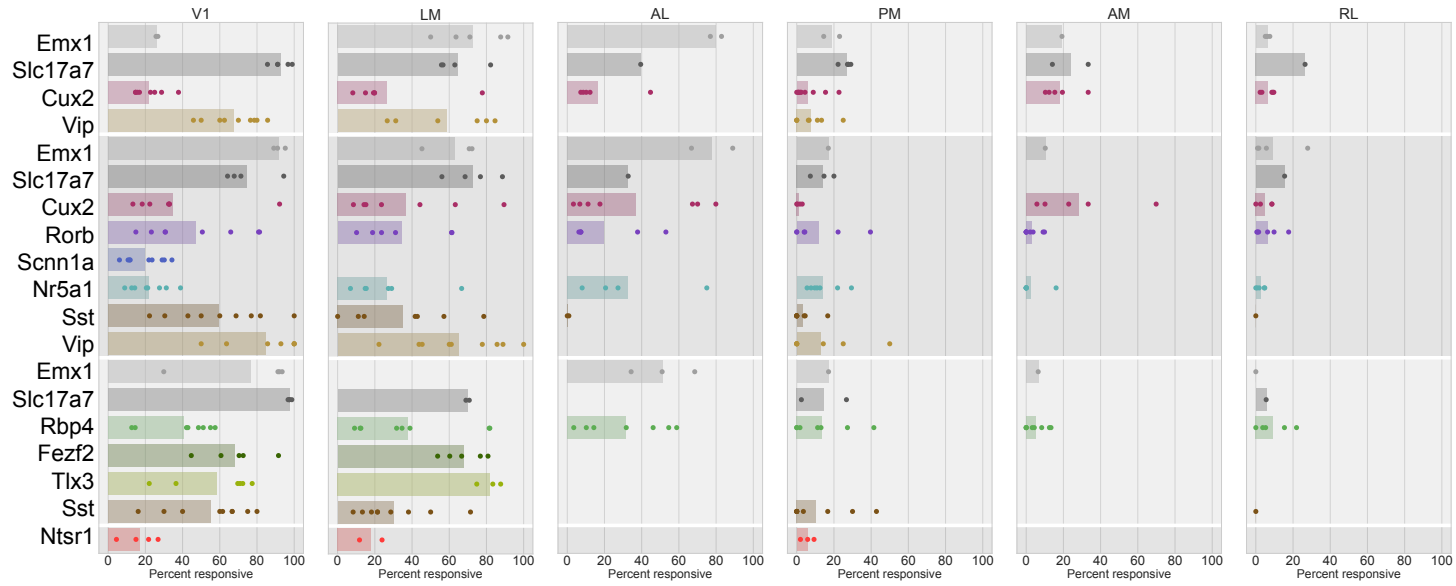
Supplemental Figure 15: Responsiveness to static gratings

(a) Table summarizing the numbers of cells imaged for each Cre line, layer, area combination in response to static grating stimulus and the number, and percent, of cells that were responsive to the static grating stimulus. (b) Strip plots of the percent of cells responsive to the static grating stimulus for each experiment.

Suppplemental Figure 16

Locally Sparse Noise

Cre	V1			LM			AL			PM			AM			RL		
	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent
Emx1 - layer 2/3	407	107	26%	654	516	79%	316	255	81%	270	51	19%	78	15	19%	759	47	6%
Emx1 - layer 4	855	785	92%	584	398	68%	557	424	76%	206	35	17%	117	12	10%	1088	100	9%
Emx1 - layer 5	640	475	74%				275	148	54%	76	13	17%	31	2	6%	238	0	0%
Slc17a7 - layer 2/3	873	812	93%	427	300	70%	38	15	39%	344	92	27%	149	36	24%	34	9	26%
Slc17a7 - layer 4	818	621	76%	634	486	77%	229	75	33%	207	27	13%				84	13	15%
Slc17a7 - layer 5	790	772	98%	94	66	70%				156	32	21%				36	2	6%
Cux2 - layer 2/3	1391	300	22%	902	209	23%	656	74	11%	845	33	4%	320	51	16%	425	29	7%
Cux2 - layer 4	1732	575	33%	1417	542	38%	1165	415	36%	759	9	1%	535	169	32%	541	36	7%
Rorb	1414	646	46%	783	262	33%	753	156	21%	397	38	10%	404	16	4%	1214	59	5%
Scnn1a	1176	226	19%															
Nr5a1	427	86	20%	253	49	19%	85	18	21%	234	34	15%	115	9	8%	805	17	2%
Rbp4	353	128	36%	437	145	33%	309	72	23%	382	37	10%	237	14	6%	95	11	12%
Fezf2	280	198	71%	604	413	68%												
Tlx3	874	497	57%	629	511	81%												
Ntsr1	196	31	16%	135	24	18%				203	11	5%						
Sst - layer 4	107	63	59%	78	32	41%	132	1	1%	188	3	2%				25	0	0%
Sst - layer 5	88	38	43%	103	27	26%				103	9	9%				7	0	0%
Vip - layer 2/3	120	81	68%	79	44	56%				126	8	6%						
Vip - layer 4	55	46	84%	94	57	61%				77	6	8%						



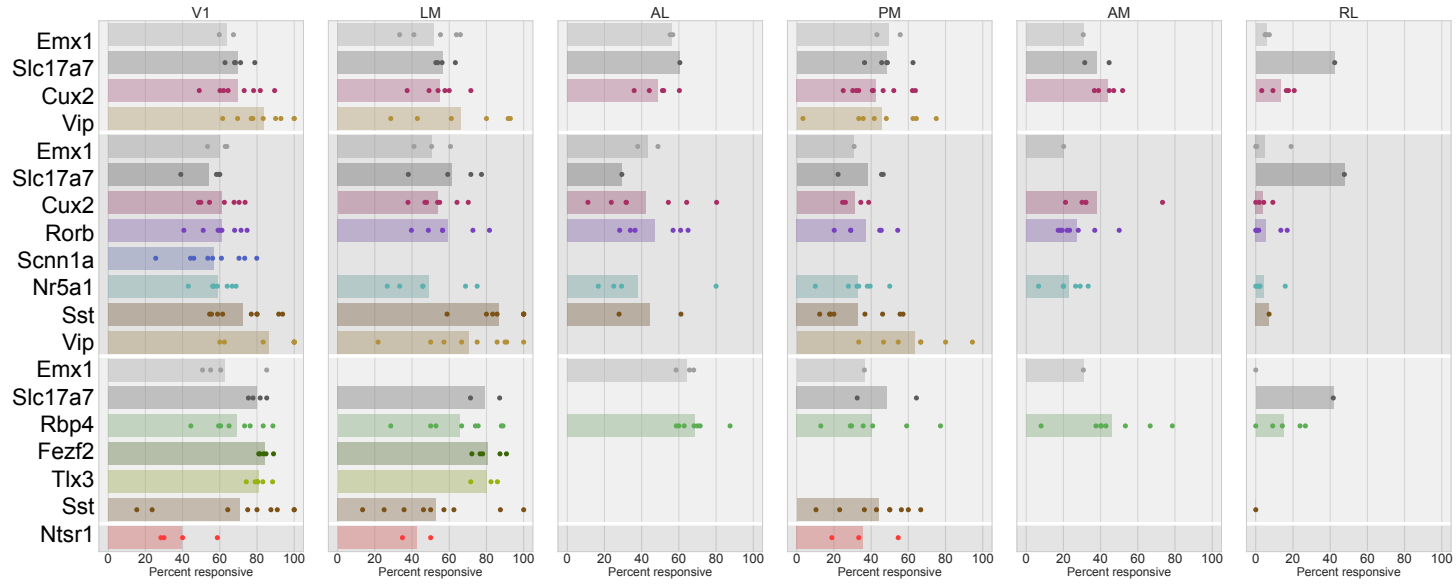
Supplemental Figure 16: Responsiveness to locally sparse noise

(a) Table summarizing the numbers of cells imaged for each Cre line, layer, area combination in response to locally sparse noise stimulus and the number, and percent, of cells that were responsive to the locally sparse noise stimulus. (b) Strip plots of the percent of cells responsive to the locally sparse noise stimulus for each experiment.

Suppplemental Figure 17

Natural Scenes

Cre	V1			LM			AL			PM			AM			RL		
	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent
Emx1 - layer 2/3	430	268	62%	721	393	55%	340	190	56%	251	126	50%	88	27	31%	697	41	6%
Emx1 - layer 4	729	437	60%	634	305	48%	543	232	43%	188	58	31%	134	27	20%	966	51	5%
Emx1 - layer 5	684	422	62%				233	151	65%	44	16	36%	39	12	31%	272	0	0%
Slc17a7 - layer 2/3	767	532	69%	438	247	56%	38	23	61%	338	160	47%	137	54	39%	33	14	42%
Slc17a7 - layer 4	828	445	54%	663	367	55%	242	71	29%	192	69	36%				63	30	48%
Slc17a7 - layer 5	695	545	78%	90	74	82%				169	96	57%				48	20	42%
Cux2 - layer 2/3	1716	1197	70%	851	464	55%	664	323	49%	944	396	42%	355	161	45%	412	53	13%
Cux2 - layer 4	1806	1098	61%	1566	843	54%	1211	511	42%	771	241	31%	554	229	41%	491	19	4%
Rorb	1404	843	60%	753	438	58%	791	370	47%	341	116	34%	388	87	22%	1149	46	4%
Scnn1a	1114	648	58%															
Nr5a1	458	259	57%	232	92	40%	83	23	28%	228	80	35%	123	30	24%	854	35	4%
Rbp4	304	216	71%	378	230	61%	285	188	66%	411	141	34%	212	93	44%	89	17	19%
Fezf2	304	254	84%	582	472	81%												
Tlx3	829	673	81%	612	492	80%												
Ntsr1	216	79	37%	143	58	41%				196	71	36%						
Sst - layer 4	111	81	73%	86	75	87%	115	38	33%	178	43	24%				14	1	7%
Sst - layer 5	126	64	51%	166	55	33%				187	56	30%				7	0	0%
Vip - layer 2/3	113	92	81%	82	53	65%				136	54	40%						
Vip - layer 4	52	45	87%	104	69	66%				87	56	64%						



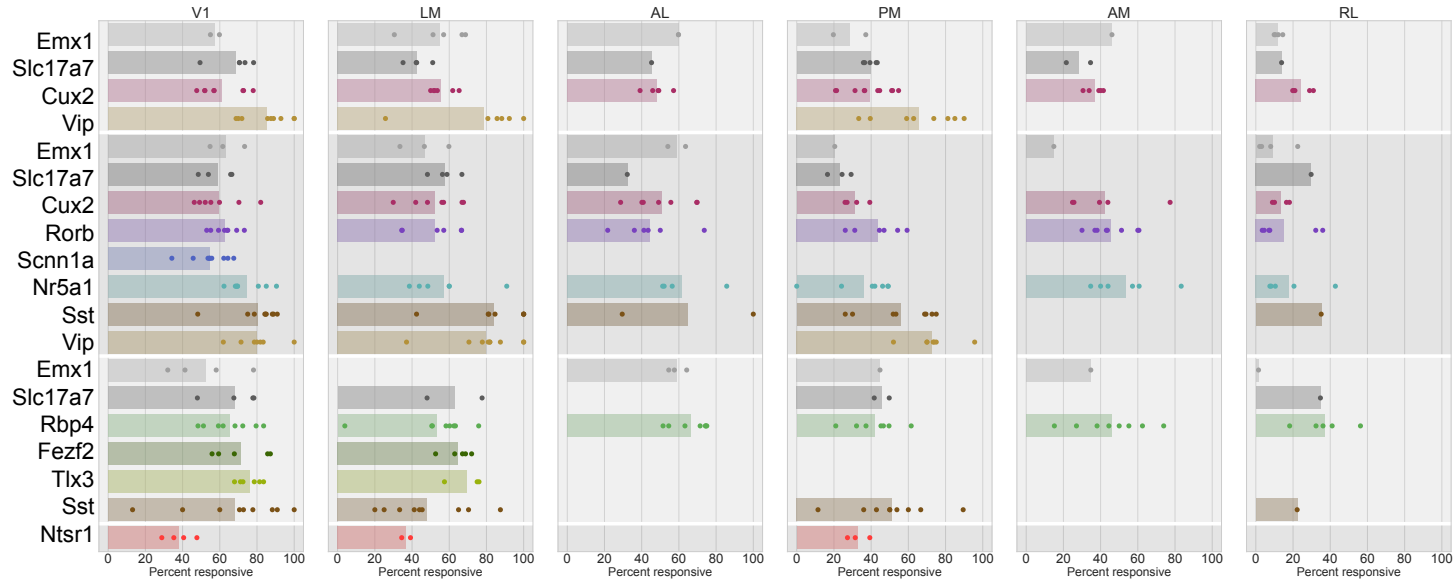
Supplemental Figure 17: Responsiveness to natural scenes

(a) Table summarizing the numbers of cells imaged for each Cre line, layer, area combination in response to the natural scenes stimulus and the number, and percent, of cells that were responsive to the natural scenes stimulus. (b) Strip plots of the percent of cells responsive to the natural scenes stimulus for each experiment.

Suppplemental Figure 18

Natural Movies

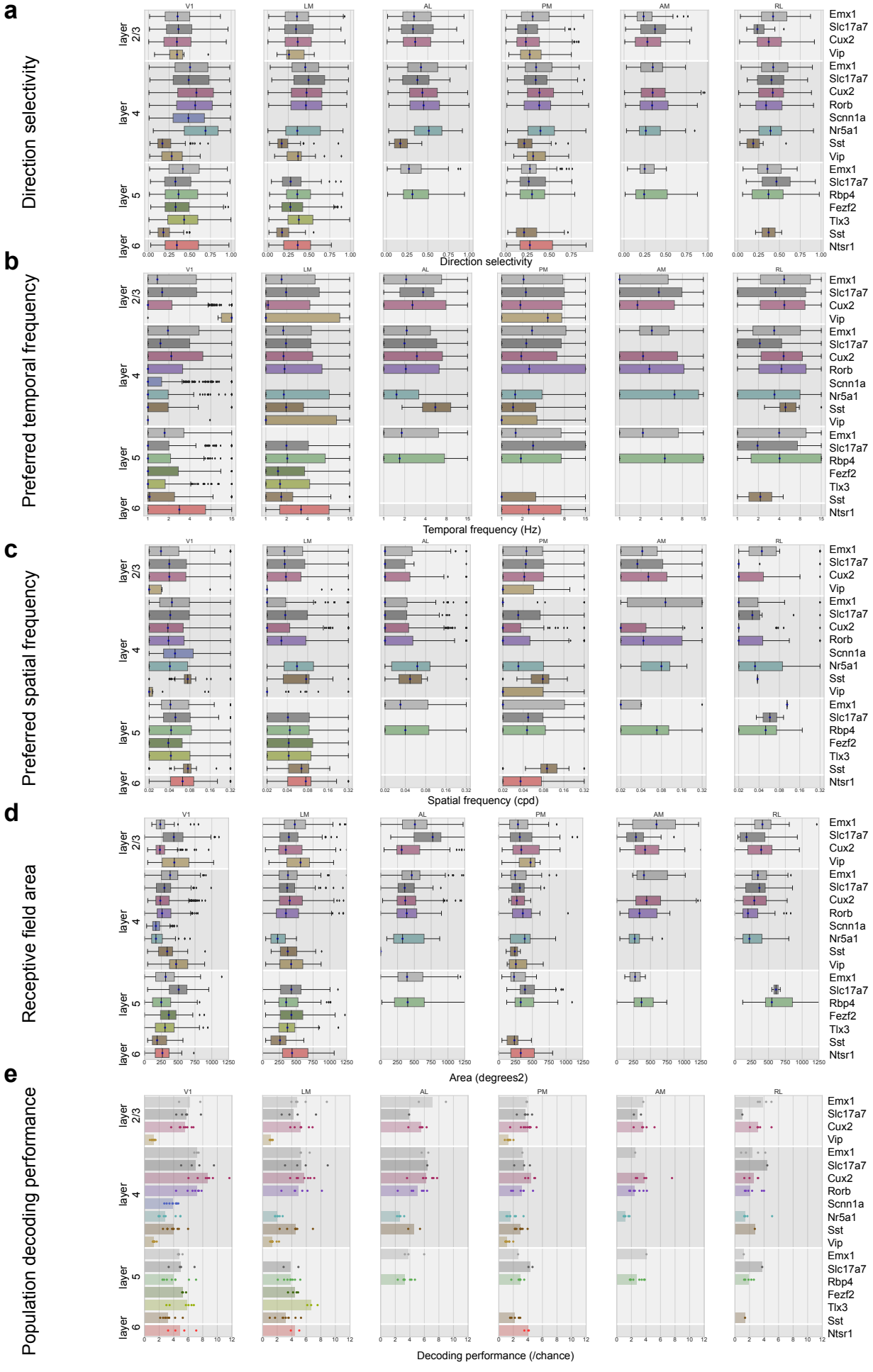
Cre	V1			LM			AL			PM			AM			RL		
	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent	total	responsive	percent
Emx1 - layer 2/3	695	402	58%	1122	628	56%	512	306	60%	422	123	29%	167	77	46%	1080	127	12%
Emx1 - layer 4	1302	812	62%	976	459	47%	838	486	58%	308	63	20%	195	29	15%	1516	132	9%
Emx1 - layer 5	1099	530	48%				437	257	59%	105	47	45%	95	33	35%	415	6	1%
Slc17a7 - layer 2/3	1284	874	68%	749	309	41%	64	29	45%	581	232	40%	235	68	29%	65	9	14%
Slc17a7 - layer 4	1288	761	59%	969	541	56%	310	100	32%	376	88	23%				185	55	30%
Slc17a7 - layer 5	1162	784	67%	146	98	67%				245	117	48%				72	25	35%
Cux2 - layer 2/3	2533	1530	60%	1341	761	57%	1084	538	50%	1615	628	39%	580	220	38%	630	156	25%
Cux2 - layer 4	2635	1532	58%	2504	1317	53%	1953	978	50%	1372	428	31%	1031	444	43%	740	97	13%
Rorb	2218	1372	62%	1191	642	54%	1242	543	44%	593	239	40%	735	322	44%	1757	221	13%
Scnn1a	1873	1047	56%															
Nr5a1	702	497	71%	416	206	50%	172	95	55%	331	128	39%	171	76	44%	1318	266	20%
Rbp4	531	367	69%	640	299	47%	490	298	61%	590	262	44%	355	168	47%	136	51	38%
Fezf2	490	353	72%	981	643	66%												
Tlx3	1181	892	76%	946	659	70%												
Ntsr1	331	118	36%	210	78	37%				330	100	30%						
Sst - layer 4	159	124	78%	152	121	80%	200	76	38%	310	142	46%				37	13	35%
Sst - layer 5	290	103	36%	261	100	38%				298	105	35%				9	2	22%
Vip - layer 2/3	157	130	83%	126	89	71%				198	116	59%						
Vip - layer 4	90	72	80%	154	114	74%				122	89	73%						



Supplemental Figure 18: Responsiveness to natural movies.

(a) Table summarizing the numbers of cells imaged for each Cre line, layer, area combination in response to any of the natural movie stimuli and the number, and percent, of cells that were responsive to the natural movie stimuli. (b) Strip plots of the percent of cells responsive to the natural movie stimuli for each experiment.

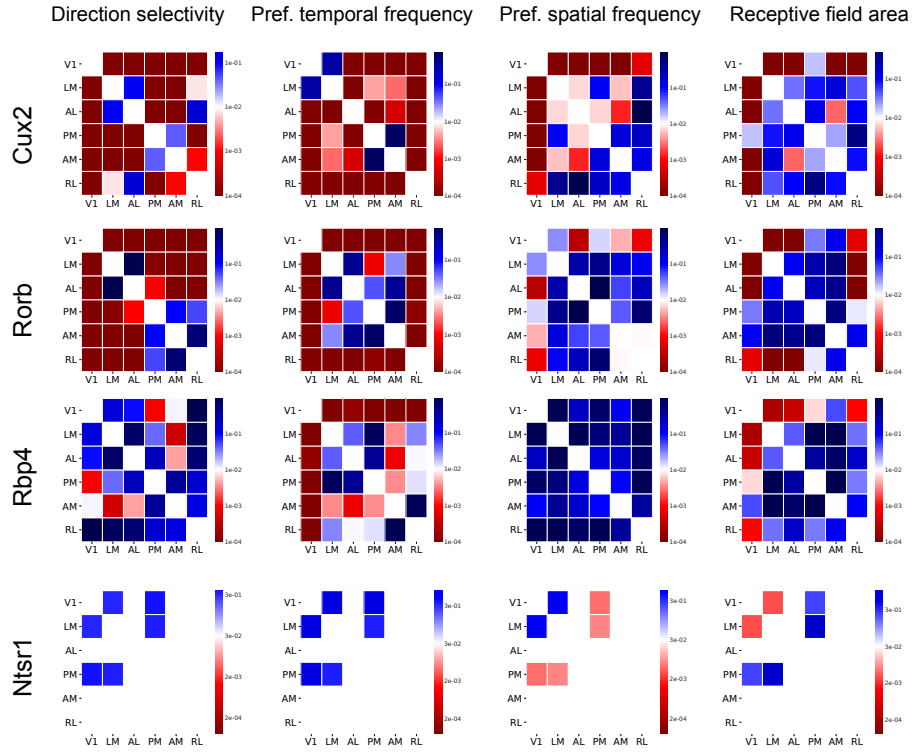
Supplemental Figure 19

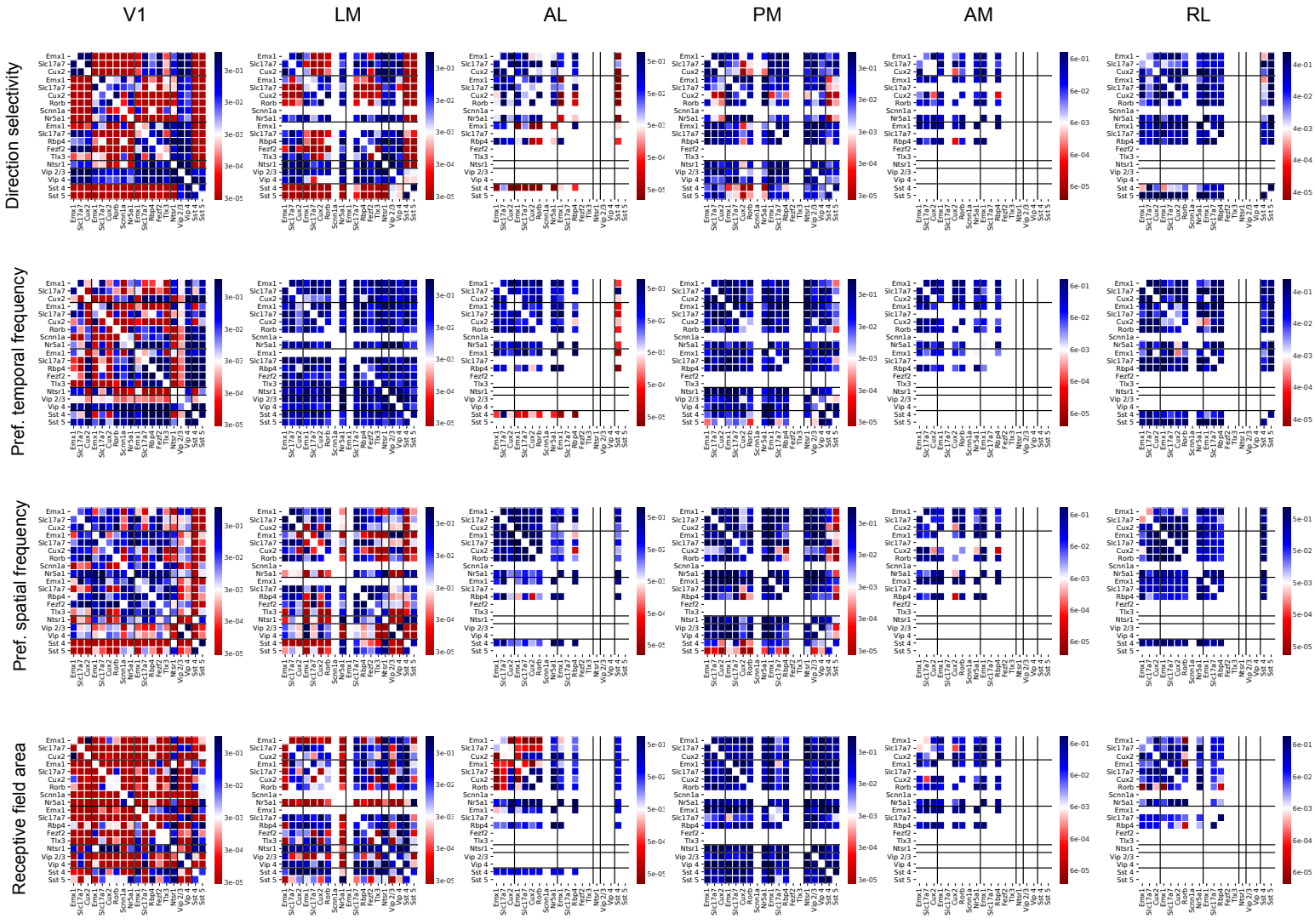


Supplemental Figure 19: Response characterizations across all areas

Box or strip plots representing the distribution of values for (a) direction selectivity, (b) preferred temporal frequency, (c) preferred spatial frequency, (d) receptive field area, and (e) population decoding of grating direction for each Cre line and layer across all six visual areas.

Supplemental Figure 20

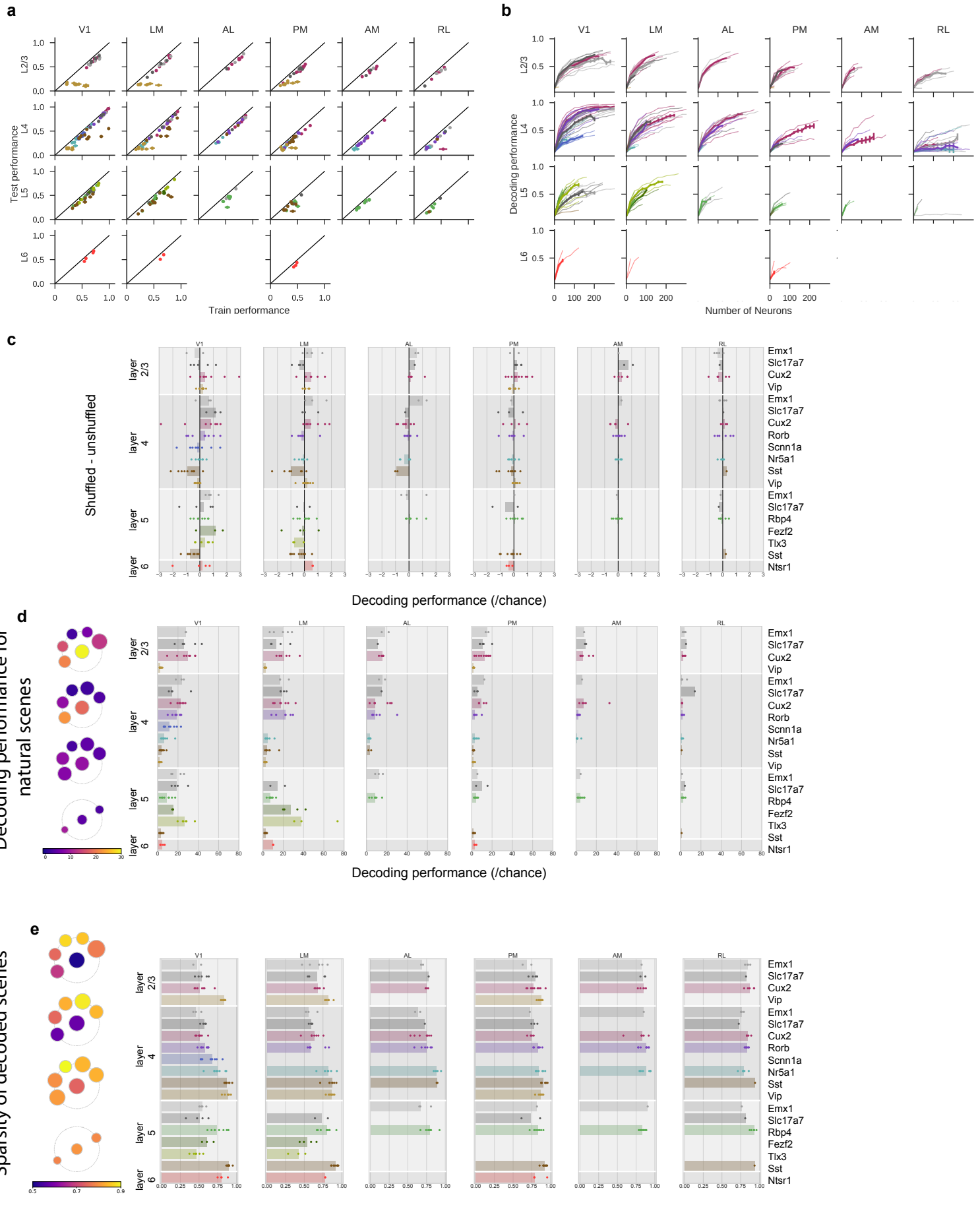




Supplemental Figure 20: Statistical tests for single cell response metrics

(a) Heatmaps of p-values of pairwise comparisons for each Cre line across areas using the Kolmogorov-Smirnov (KS) test with a Bonferroni correction for the number of comparisons. The heatmap is centered on the significance criteria. E.g. for comparisons across all six visual areas, $p<0.01$ (=0.05/5) is significant. For comparisons across only 3 visual areas (eg. for Ntsr1) $p<0.025$ is significant. (b) Heatmaps of p-values for pairwise comparison for each Cre-line and layer combination within each visual area, using KS test with a Bonferroni correction for the number of comparisons. The colormap for the p-values is centered at the significance criteria.

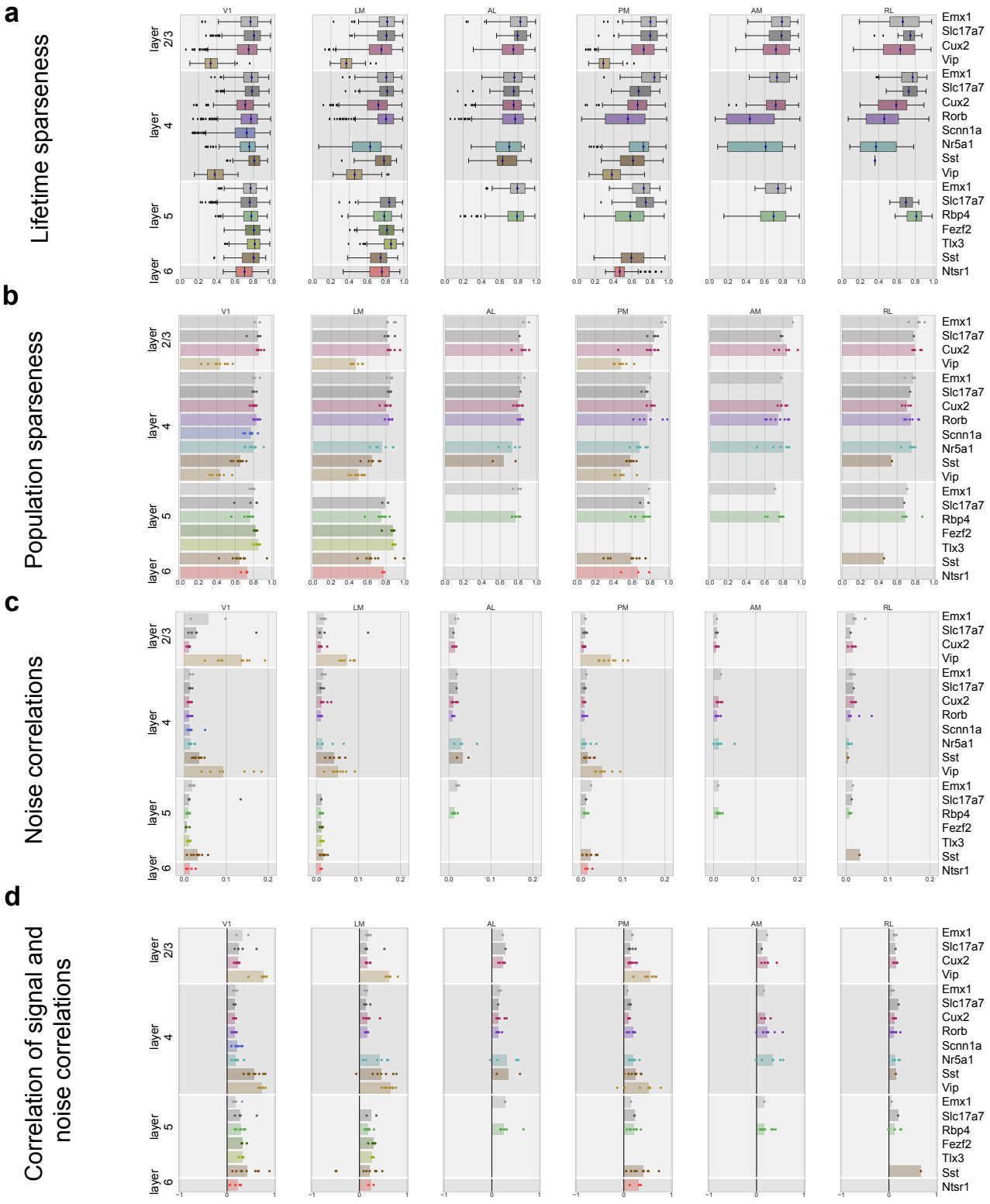
Supplemental Figure 21



Supplemental Figure 21: Population decoding

(a) Test vs train performance for K-Nearest Neighbor decoding of drifting grating direction for each experiment, across all areas, layers and Cre lines. For all experiments, other than Vip, the experiments are close to unity, reflecting good generalization. Error bars represent the standard error of the mean across five-fold cross-validation. (b) Dependence of decoding performance on the number of cells included in the decoding, for all areas and Cre lines. (c) Strip plots of the difference between the decoding performance when trials are shuffled, destroying noise correlations, and trials are kept intact. For excitatory cells in V1, the median differences are small but positive, indicating that noise correlations hurt decoding performance. For Sst cells, noise correlations help decoding performance. (d) Pawplot and strip plots summarizing decoding performance for natural scenes, relative to chance. (e) Pawplot and strip plots summarizing the sparsity of natural scene decoding, reflecting the sparsity of the distribution of decoding performance across natural scenes in a given experiment.

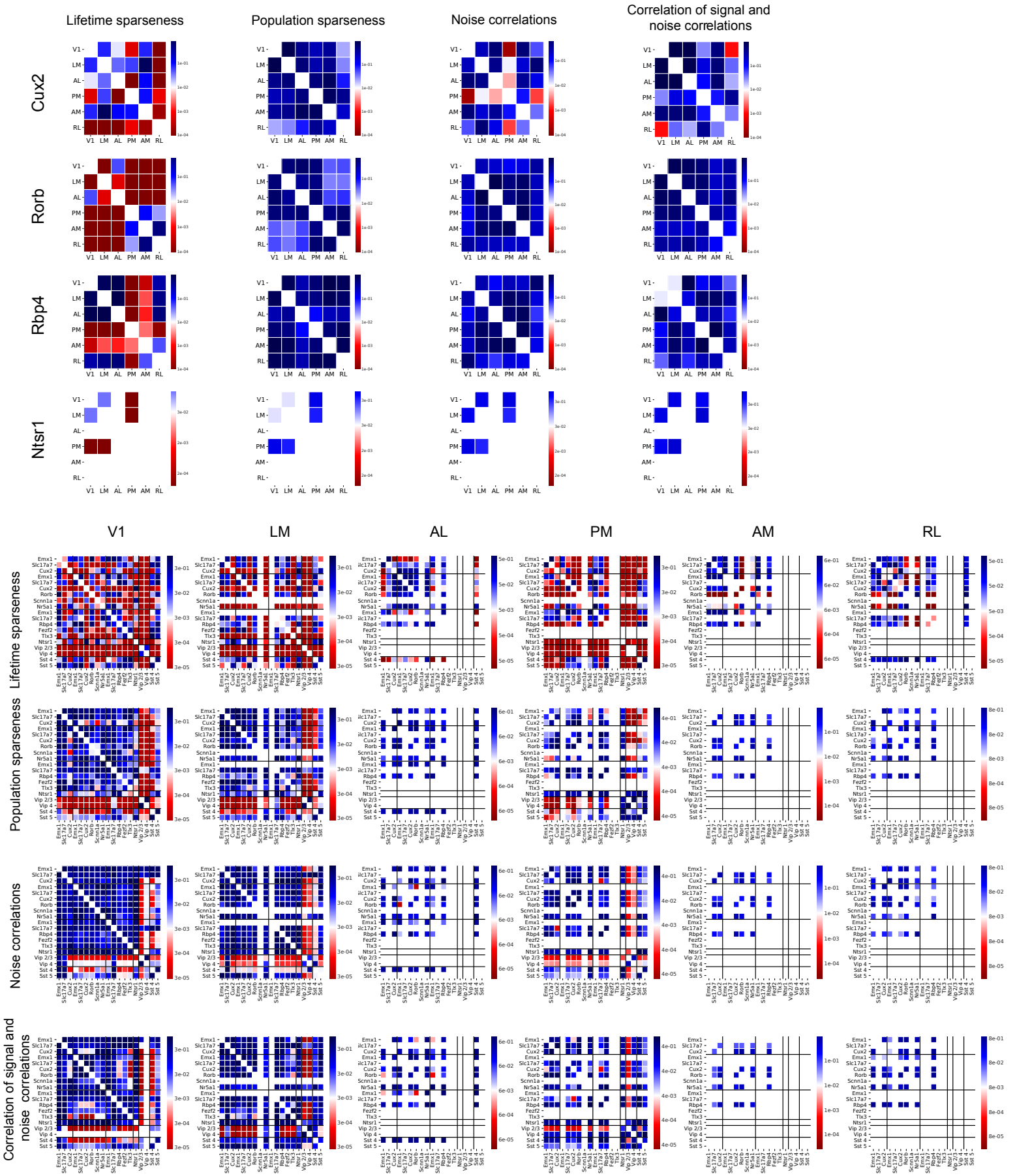
Supplemental Figure 22



Supplemental Figure 22: Sparsity characterization across all areas

Box or strip plots representing the distribution of values for (a) lifetime sparseness, (b) population sparseness, (c) noise correlations and (d) correlation of signal correlations and noise correlations for each Cre line and layer across all six visual areas.

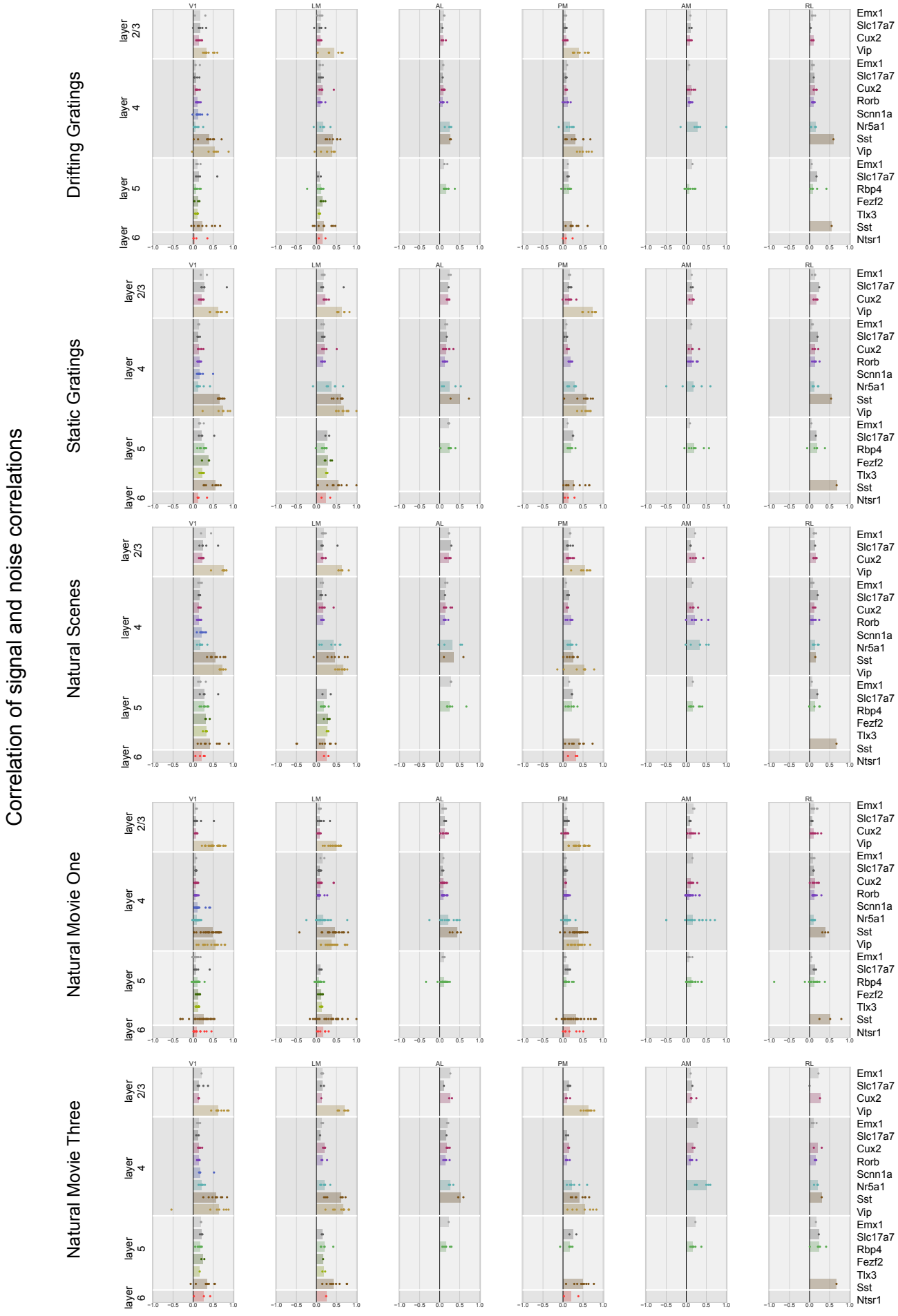
Supplemental Figure 23



Supplemental Figure 23: Sparsity statistics

(a) Heatmaps of p-values of pairwise comparisons for each Cre line across areas. For single cell metrics (e.g. lifetime sparseness) the Kolmogorov-Smirnov (KS) test was used with a Bonferroni correction for the number of comparisons. For population metrics (e.g. population sparseness) a t-test was used with a Bonferroni correction for the number of comparisons. The heatmap is centered on the significance criteria. E.g. for comparisons across all six visual areas, $p<0.01$ (=0.05/5) is significant. For comparisons across only 3 visual areas (eg. for Ntsr1) $p<0.025$ is significant. (b) Heatmaps of p-values for pairwise comparison for each Cre-line and layer combination within each visual area with a Bonferroni correction for the number of comparisons. The colormap for the p-values is centered at the significance criteria.

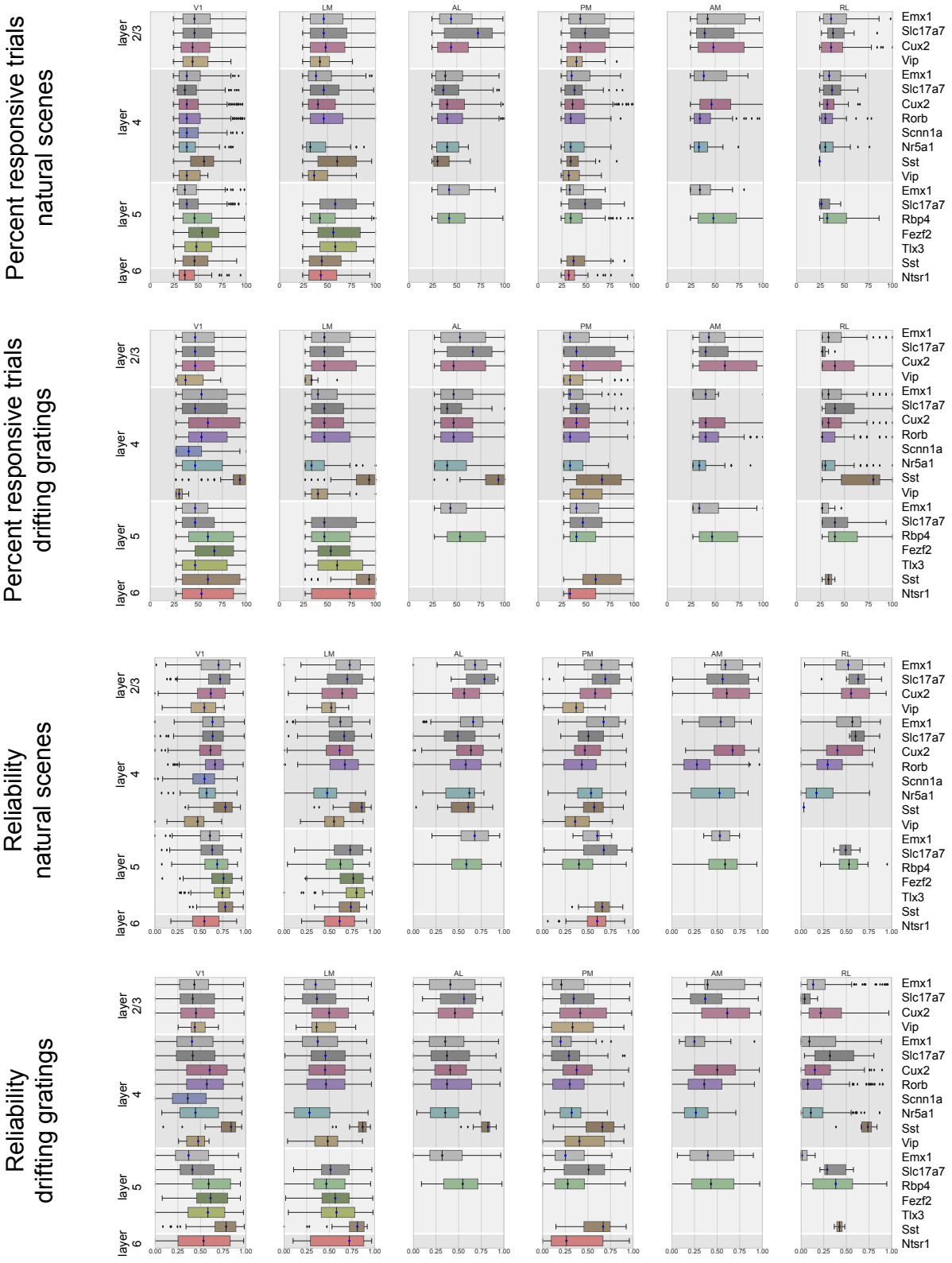
Supplemental Figure 24



Supplemental Figure 24: Correlations of signal and noise correlations

Strip plots for the correlations of signal and noise correlations for each Cre line and layer across all six visual areas, for all stimuli.

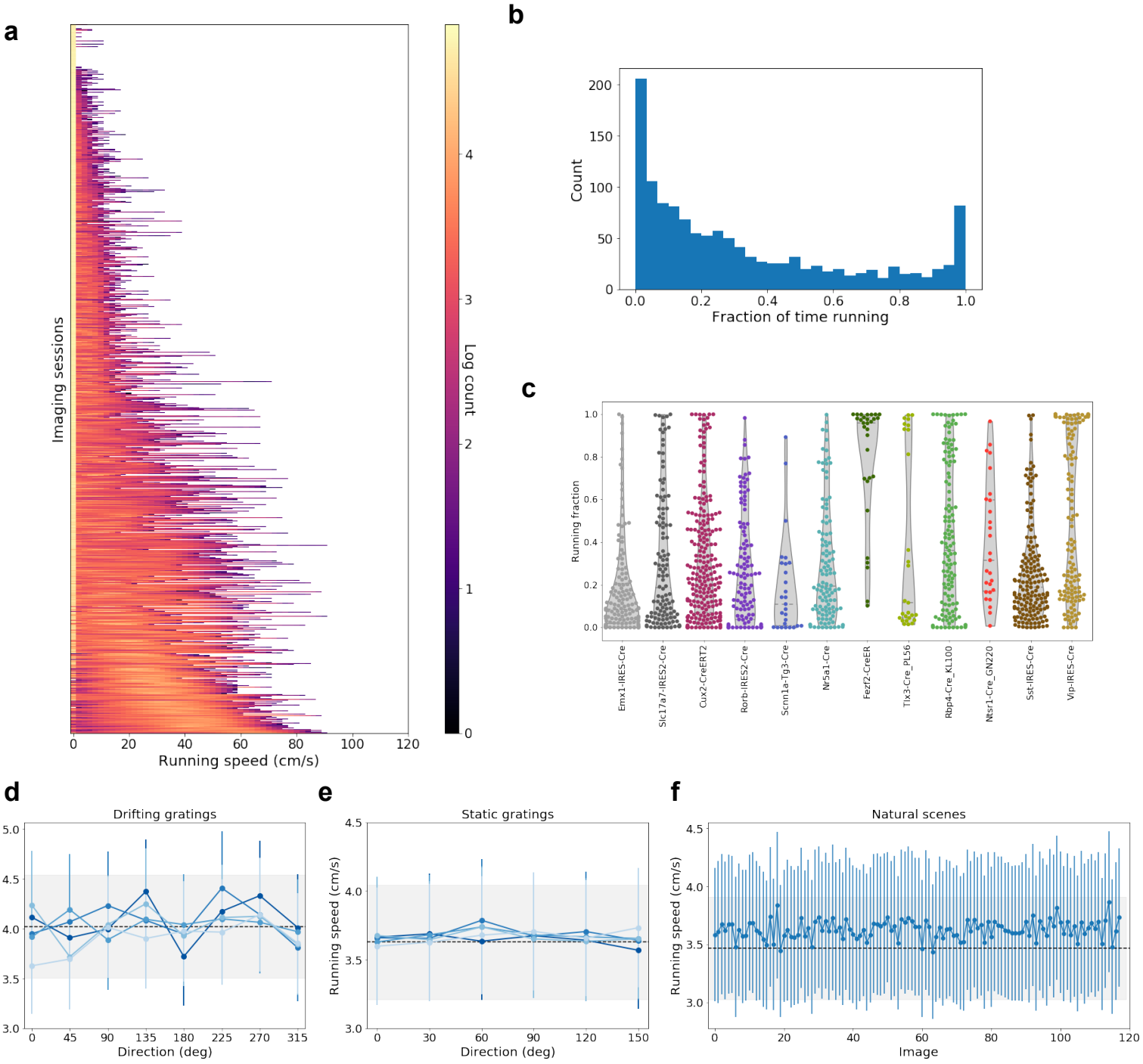
Supplemental Figure 25



Supplemental Figure 25: Reliability

Box plots representing the distribution of values for (a) percent responsive trials for natural scenes, (b) percent responsive trials for drifting gratings, (c) reliability for natural scenes and (d) reliability for drifting gratings for each Cre line and layer across all six visual areas.

Supplemental Figure 26



Supplemental Figure 26: Running characterization

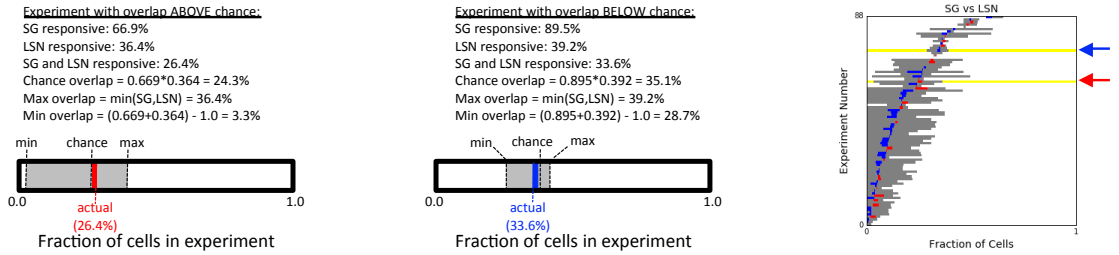
(a) Heatmap of running speed distributions for all imaging sessions, ordered by the mean running speed. (b) Histogram of the fraction of time the mouse is running (>1 cm/s) for all imaging sessions. (c) Distribution of the fraction of time the mouse is running for each Cre line. Mice from some Cre lines exhibit more running (eg. Fezf2). (d) Mean running speed for grating direction for drifting gratings at each temporal frequency, with blank sweep indicated by dashed line. (e) Mean running speed for grating orientation for static gratings at each spatial frequency, with blank sweep indicated by dashed line. (f) Mean running speed for natural scenes.

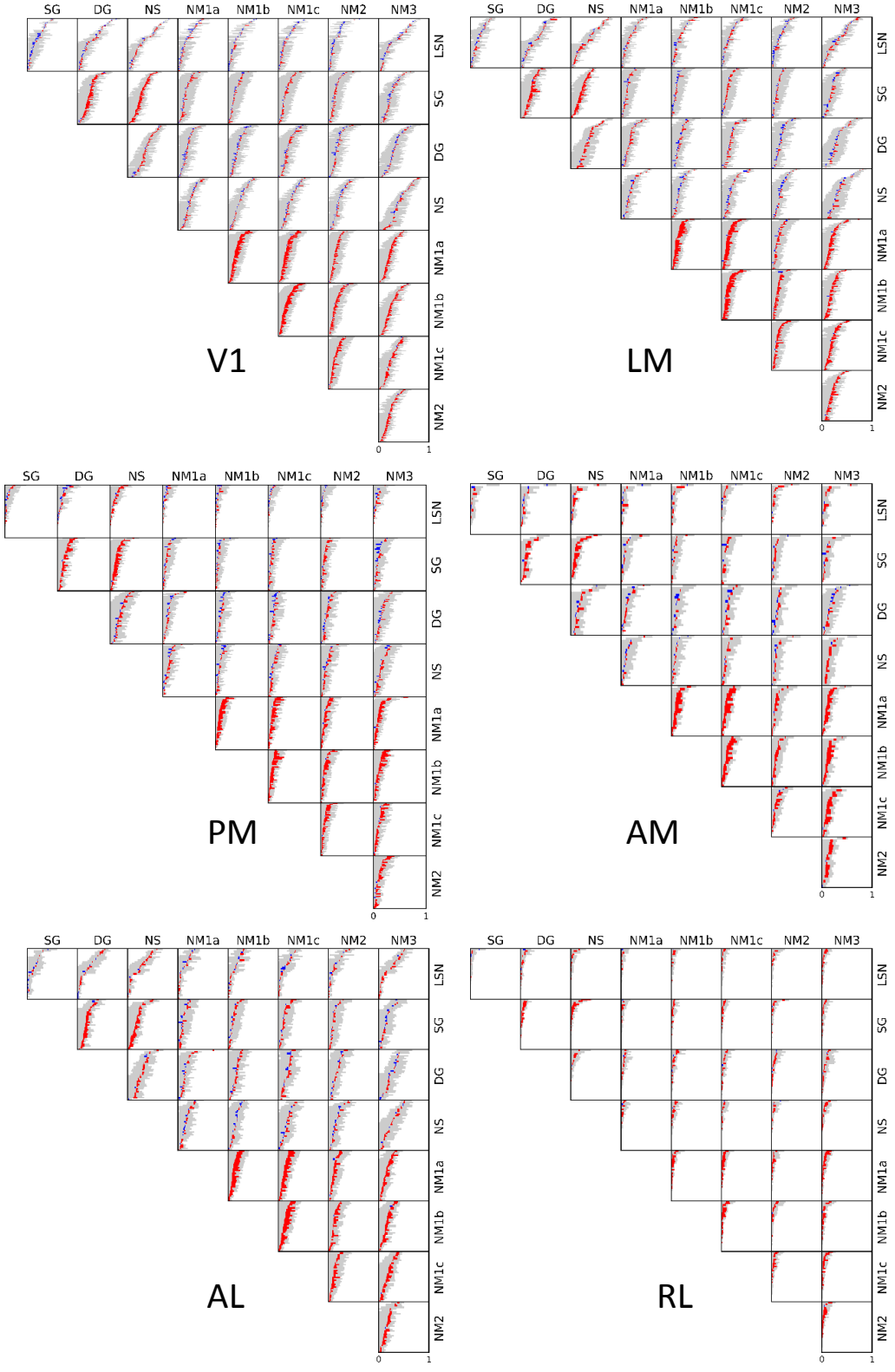
Supplemental Figure 27

Supplemental Figure 27: Model performance across all areas

(a) Box plots representing the distribution of model performance (r) across all areas. (b) Box plots representing the distribution of the complexity ration across all areas.

Supplemental Figure 28





Supplemental Figure 28: Responses to different stimuli are largely independent.

For each visual area, the amount of overlap of responsive cells for each pairwise comparison of stimuli. Top, comparison of static gratings and locally sparse noise in V1, highlighting two examples. For each experiment, the gray bar indicates the range of possible overlap given the percent of cells that respond to each stimulus. Colored bar indicates the actual overlap relative to chance. Red bar reflects above chance overlap. Blue bar reflects below chance overlap. Below, overlap comparisons for each stimulus pair for each visual area. Stimulus abbreviations: DG: drifting gratings, SG: static gratings, LSN: locally sparse noise, NS: natural scenes, NM: natural movie. Natural movie 1 is repeat in each imaging session (NM1a, NM1b, NM1c).